Essays in Financial Technology: Banking Efficiency and Application of Machine Learning Models in Supply Chain Finance and Credit Risk Assessment.

Yixuan Li

Submitted in fulfilment of the requirements of the Degree of Doctor of Philosophy

Adam Smith Business School

College of Social Sciences

University of Glasgow

August 2023

Abstract

The financial landscape is undergoing a significant transformation, driven by technological innovations that are reshaping traditional banking practices. This thesis examines the evolving relationship between financial technology (FinTech) and banking, specifically addressing the credit risk aspects within the domains of Supply Chain Finance (SCF) and peer-to-peer (P2P) lending.

FinTech has experienced rapid growth and innovation over the past decade. It encompasses a wide range of technologies and services that aim to enhance and streamline financial processes, disrupt traditional banking models, and offer new solutions to consumers and businesses. The status of FinTech and banking is assessed through an extensive review of the current literature and empirical data. Accordingly, FinTech development has significantly impacted the financial landscape, driving innovation, competition, and customer expectations while it has exposed inefficiencies within traditional banking, it has also compelled banks to evolve and embrace technological advancements. The impact of FinTech on traditional banking models, customer behaviours, and market competition is aimed to be explored. This investigation highlights the challenges and opportunities that arise as FinTech disrupts and reshapes the banking sector, emphasizing its potential to enhance efficiency, accessibility, and customer experiences. As Chapter 3 focuses on an empirical analysis of the impact of FinTech on the operating efficiency of commercial banks in China.

Further, in the context of credit risk, the thesis focuses on SCF and P2P lending, two prominent areas influenced by FinTech innovation. SCF has witnessed substantial transformation with the infusion of FinTech solutions. Digital platforms have streamlined the flow of funds within complex supply networks, enhancing the liquidity of suppliers and optimizing working capital for buyers. However, this transformation introduces new credit risk challenges. As suppliers' financial data becomes more accessible, the need for accurate risk assessment and predictive modelling becomes paramount. The integration of big data analytics, machine learning, and artificial intelligence (AI) holds the promise of refining credit risk evaluation by offering real-time insights into supplier financial health, thereby improving lending decisions and reducing defaults.

Similarly, P2P lending has redefined the borrowing and lending landscape, enabling direct connections between individual borrowers and lenders. While P2P lending platforms offer

speed, convenience, and access to credit for previously underserved segments, they also grapple with credit risk concerns. Evaluating the creditworthiness of individual borrowers without sufficient credit history demands innovative risk assessment methodologies. The emergence of data issues, such as imbalanced data issues, feature selection, and data processing, presents challenges in building accurate credit risk profiles for P2P lending participants. FinTech solutions play a pivotal role in creating and implementing these alternative risk assessment models. Note that, few studies in the literature investigate the benchmark of the advanced method of solving the credit risk assessment in emerging financial services.

This thesis aims to address this research gap by evaluating the effectiveness of credit risk assessment models in these FinTech-driven contexts, considering both traditional methodologies and novel data-driven approaches. Chapter 4 investigates the credit risk assessment issue in Digital Supply Chain Finance (DSCF) with the Machine Learning approach and Chapter 5 emphasises the issue of data imbalance of credit risk assessment in P2P Lending.

By addressing these gaps and issues, this thesis aims to contribute to the broader discourse on FinTech's role in shaping the future of banking. The findings have implications for financial institutions, policymakers, and regulators seeking to harness the benefits of FinTech while mitigating associated risks. Ultimately, this study offers insights into navigating the evolving landscape of credit risk in SCF and P2P lending within the context of an increasingly technology-driven financial ecosystem.

Content

List of Tables

List of Figures

Acknowledgements

First and foremost, I would like to express my deep and sincere gratitude to my research supervisors, Professor Charalampos Stasinakis and Dr Wee Meng Yeo, for their patience, motivation, enthusiasm, and immense knowledge. Their guidance helped me in all the time of research and writing of the thesis. Working and studying under their guidance was a great privilege and honour.

I also sincerely thank Dr Filipa Da Silva Fernandesd for sharing her knowledge and suggestive comments on my research. Meantime, I want to give a special thanks to all Adam Smith Business School staff for their instant help in my daily life and work.

I am very thankful to my partner Yaxiong Wu, for his accompany, understanding and continuing support to complete this research work. I thank my friends, including Jiawen Liang, Wanquan Zhang, Xiaotong Sun, Yu Guo and many others for bringing joy and helping in my research life.

Last but not least, I am extremely grateful to my parents Xun Li and Tong He, my grandparents Yunyuan He and Naihong Chen, my uncle and aunt Gang He and Yin Ban, and other family members, for their unconditional love, caring, and encouragement and sacrifices for educating and supporting me for my future.

Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed name: Yixuan Li

Signature:

Abbreviations

ANN	Artificial Neural Network
BM	Bookmaker Informedness
CCB	City Commercial Bank
CS	Cost Sensitive
DEA	Data Envelopment Analysis
DFII	Digital Financial Inclusion Index
DMU	Decision Making Unit
DSCF	Digital Supply Chain Finance
DT	Decision Tree
EFFCH	Efficiency Change
ERP	Enterprise Resource Planning
FD	Financial Development
FI	FinTech Index
FinTech	Financial Technology
IHT	Instance Hardness Threshold
JOSB	Joint-stock Commercial Bank
KNN	K-Nearest Neighbours
LR	Logistic Regression
MCC	Matthews Correlation Coefficient
MLP	Multi-layer Perceptron
MPI	Malmquist Productivity Index
NB	Naïve Bayes
PURE	Pure Technical Efficiency Change
RF	Random Forest
SCALE	Scale Efficiency Change
SCF	Supply Chain Finance
SFA	Stochastic Frontier Analysis
SME	Small and Medium Enterprise
SMOTE	Synthetic Minority Oversampling Technique
SOCB	State-owned Commercial Bank
SVM	Support Vector Machine
TECHCH	Technology Change
TFPCH	Total Factor Productivity Change
XGBoost	Extreme Gradient Boosting

Chapter 1 Introduction

1.1 Background and Motivations

FinTech is a technology-driven financial innovation which refers to a broad range of technological innovations that aim to disrupt and improve various financial activities, including banking, payments, investments, insurance, lending, and wealth management. These innovations leverage software, applications, algorithms, and data analytics to provide novel solutions that aim to enhance user experiences, reduce costs, and increase accessibility.

In the continuous evolution and upgrading of technologies such as cloud computing, AI, big data, and blockchain, increasingly powerful computing capabilities have enabled the entire lifecycle of data creation, storage, use, sharing, archiving and destruction to be enhanced in quality and efficiency, which has brought transformative opportunities for the financial industry and is highly dependent on data and information. In terms of global strategic requirements for the development of FinTech, it is estimated that more than 170 countries worldwide have issued national FinTech and digitalisation strategies[1]. Among them, the financial industry is often treated as the first echelon of digitalisation due to its information-intensive nature, playing the role of a "leader" in digital transformation. Taking AI as an example, in the past three years, 60 countries and regions, accounting for 90% of global GDP, have formulated AI policies and strategies, and the use of AI to facilitate digital transformation and intelligent operations in the financial sector has become a common strategic goal for all countries[2]. In terms of industry trends, major global banks have increased their investment in FinTech. in 2019 BNP Paribas invested 90.56% of its profits in technology, JP Morgan Chase invested 81.2% of its net profits in technology, and Bank of America, Citibank, and State Street invested 30% to 60% of its net profits in technology. In terms of results orientation, during the period 2017-2020, UBS's net profit grew by 160.87% per annum, Morgan Stanley's net profit grew by 27.64% per annum and Deutsche Bank, despite having a negative net profit in 2019, invested an average of €3 billion per annum in technology in 2019-2020 and achieved a goal of turning a net profit from negative to positive in 2020[3], demonstrating the positive effects of the bank's strong focus on FinTech (see Figure

[1] https://www.ibanet.org/document?id=FinTech-legal-frameworks-23
[2] https://e.huawei.com/cn/material/enterprise/85edb14ae4d9452cbde0777ab9c5d8e7
[3] http://www.caict.ac.cn/kxyj/qwfb/bps/202009/P020200918520670741842.pdf

1.1). The comparison of IT expenditure between global banks and technology companies, as shown in the chart, further demonstrates the drive for technological upgrades and growing competition for services in the financial sector.

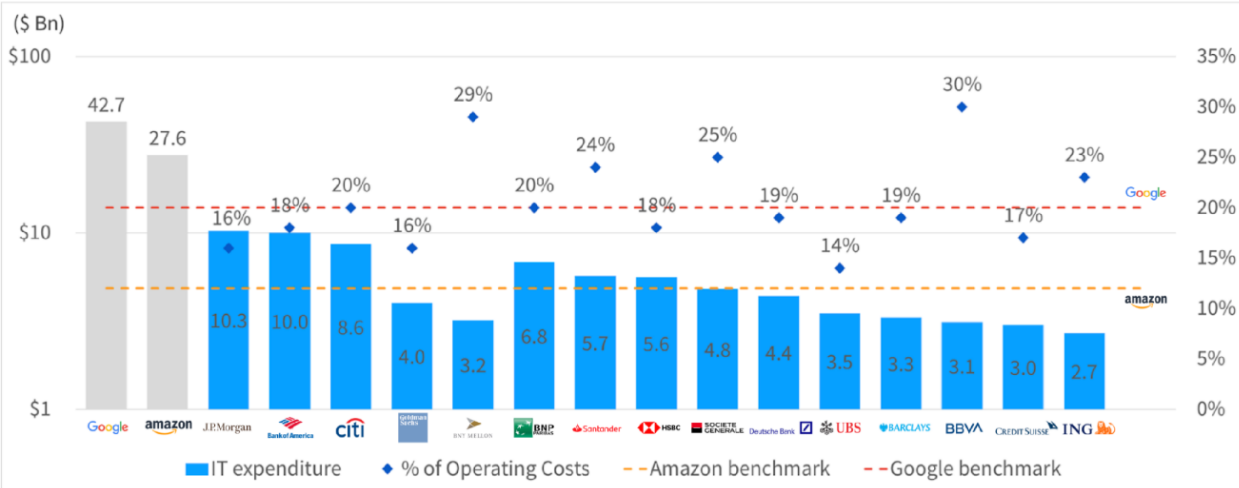


Figure 1.1 Global Banks and tech companies – IT expenditure 2020[4]

FinTech is injecting new momentum into the transformation and development of commercial banks. On the one hand, FinTech leads a significant shift in the service concept and mindset of traditional commercial banks, thus helping to enhance their ability to innovate in financial services. Based on the technological support of FinTech, traditional commercial banks re-engineer their data ecosystem, further integrate diverse digital channels, enhance data collection, processing, and application, and bring a better experience to customers (Arslanian and Fischer, 2019). On the other hand, FinTech facilitates commercial banks to reduce operating costs and improve efficiency. Moreover, FinTech assists banks in improving traditional credit assessment systems, effectively alleviating information asymmetries, reducing the cost of a single operation, and greatly increasing the efficiency of lending while expanding their service coverage (Leong et al., 2017).

However, the rapid development of FinTech has imposed a series of fundamental, institutional, and disruptive impacts on the banking industry that should not be overlooked, mainly in terms of directly squeezing the development space of commercial banks' major businesses such as payment and settlement, deposit management and finance. Many banks are also facing a serious issue of disintermediation (Navaretti et al., 2018). This has led the

[4] https://www.cdotrends.com/story/15968/4-trends-driving-record-investment-technology-global-banks

essay to further examine the empirical evidence of the impact of FinTech development on the efficiency of commercial banks.

The application of FinTech has strongly driven disruptive innovation and reshaping of financial services, nevertheless, challenges and opportunities coexist, and the development of innovative financial services has also created more risk exposure (Mahalle et al., 2021). As the most important business of financial institutions, the credit business has given rise to more innovative businesses based on the needs of different customer groups, such as supply chain finance (SCF) and P2P lending. The SCF starts from the supply chain industry chain, consolidating logistics, capital chain, information flow, and other information with FinTech, and building an integrated financial system and risk assessment system for core enterprises and upstream and downstream enterprises, aiming to quickly respond to the comprehensive needs of industry chain enterprises for settlement, financing, and management (Liu, 2020). Compared with traditional bank credit, the biggest innovation of SCF is the change in its credit model (Moretto et al., 2019). The traditional credit model of banks is based on financial information such as balance sheets, and credit decisions are made based on an evaluation of the enterprises' own situation, but this is not the case with SCF. Under the financing model of SCF, banks have diluted the financial analysis of the enterprises, and no longer emphasise the scale of the industry in which the enterprise is located, the value of its fixed assets, financial indicators, guarantee methods, etc (Wang and Wu, 2021). Instead, the real background of the enterprise's trade and the strength and credit level of the core enterprises in the supply chain are more focused, i.e., banks assess the credit status of the entire supply chain (Chen et al., 2019). The change has inevitably required financial institutions to re-establish a credit rating system for the credit risk assessment of SMEs in SCF. The development of digitalisation has further strengthened the establishment of information-sharing platforms, such as the application of enterprise resource planning (ERP) systems, which provide more sources of information for credit assessment in digital supply chain finance (DSCF) and increase the requirements for the effectiveness of credit risk assessment models. In addition, P2P platforms are online service that mediates debt contracts between lenders and borrowers. In the traditional lending market, the information collected by banks is usually privatised, whereas P2P lending activities leave behind a large amount of information and transaction data of the participants (Chen et al., 2020), which allows the lending platform to collect more data to identify credit risks. However, how to build relevant credit risk assessment models and collate the existing credit data has become a crucial concern of current research.

The manual credit risk assessment is highly subjective and arbitrary, and the process is cumbersome and costly in terms of human and material resources (Byanjankar et al., 2015). Solving credit assessment problems and thus reducing risk through FinTech approaches is becoming a necessity for the development of financial services. Figure 1.2 below shows that according to the report, AI applications in risk management account for 32% of the financial services sector, and machine learning as an important part of AI has become an important tool for credit risk identification. Under this circumstance, machine learning can effectively capture the non-linear characteristics of data in credit assessment and identify default risk points with supervised learning, by building relevant models, acquiring sufficient data, and evaluating the output results with metrics (Van Gestel et al., 2005). In terms of identifying the defaults on lending, machine learning algorithms apply a classification approach to separate default data from non-default samples, which provides a more capable fit the features and is stronger at non-linear and cross-sectional features. Secondly, machine learning algorithms offer more robust fitting effectiveness through the ensemble models (Mosavi et al., 2018), such as XGBoost and LightGBM. Essentially, the complexity of the model is increased and the ability of the model to identify fixed patterns is enhanced. Thirdly, machine learning is also more responsive to iterative forms of data and the existence of data processing issues (Saker, 2021). Overall, machine learning unlocks the value of vast amounts of data.

Industry	Human Resource	Manufacturing	Market and sales	Product and/or Service Development	Risk	Service Operations	Strategy and Corporate Finance	Supply-chain Management
All industries	9%	12%	20%	23%	13%	25%	9%	13%
Automotive and Assembly	11%	26%	20%	15%	4%	18%	6%	17%
Business, Legal, and Professional Services	14%	8%	28%	15%	13%	26%	8%	13%
Consumer Goods/Retail	2%	18%	22%	17%	1%	15%	4%	18%
Financial Services	10%	4%	24%	20%	32%	40%	13%	8%
Healthcare Systems/Pharma and Medical Products	9%	11%	14%	29%	13%	17%	12%	9%
High Tech/Telecom	12%	11%	28%	45%	16%	34%	10%	16%

% of Respondents (Function)

Source: McKinsey & Company, 2021[5]

[5] https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf

Figure 1.2 AI adoption by industry and function, 2021

Both sides of the coin exist, machine learning approaches also have limitations. Firstly, machine learning is a data-driven training model, it requires a high-quality and wide range of training data. Data issues like Imbalanced data require further analysis to avoid errors and over-fitting. Secondly, machine learning suffers from many constraints in its application due to its lack of explainability. Deep learning, as exemplified by MLP, is like a black box (Rai, 2020), giving input and getting output as the result, but the rationale behind the result and the reliability of the decision are not explained. However, for financial institutions, causal explanations for decision-making are essential. For instance, in practice, credit risk assessment is usually based on the financial institution's risk-taking capacity and risk appetite (Epetimehin, 2013). When the decision-making of a financial institution is conservative, the credit assessment focuses more on identifying potential defaulters, whereas when the decision-making of a financial institution is more yield-oriented, the credit assessment focuses more on expanding more low-risk borrowers. This has led to further research into the use of machine learning for predictive analysis of credit risk assessment.

Overall, while the role of FinTech in banking development has been studied, and credit risk assessment has made strides with machine learning, challenges of systematic investigation and gaps in the specific verification persist. Addressing regulatory issues based on the specific country and era, promoting financial inclusion, optimizing operations, and improving credit risk models through high-quality data and interpretability are crucial steps toward bridging these gaps and advancing the financial industry.

1.2 Thesis Structure

The objective of this thesis is to construct a research framework for exploring the evolution of FinTech alongside banking, and to delve into the particular procedure of evaluating credit risks in SCF and P2P lending, all rooted in the underlying motivations. This thesis contains 7 chapters. Chapter 2 provides a review of FinTech developments, a description of advanced financial services, and a summary of machine learning methods. Three chapters (Chapter 3,4,5) display empirical applications. The final Chapter 6 presents a summary. Limitations of the study and future work are also explained in this section. The overall structure of the article is illustrated in Figure 1.3 below. The introduction of empirical studies is also briefly described as follows:

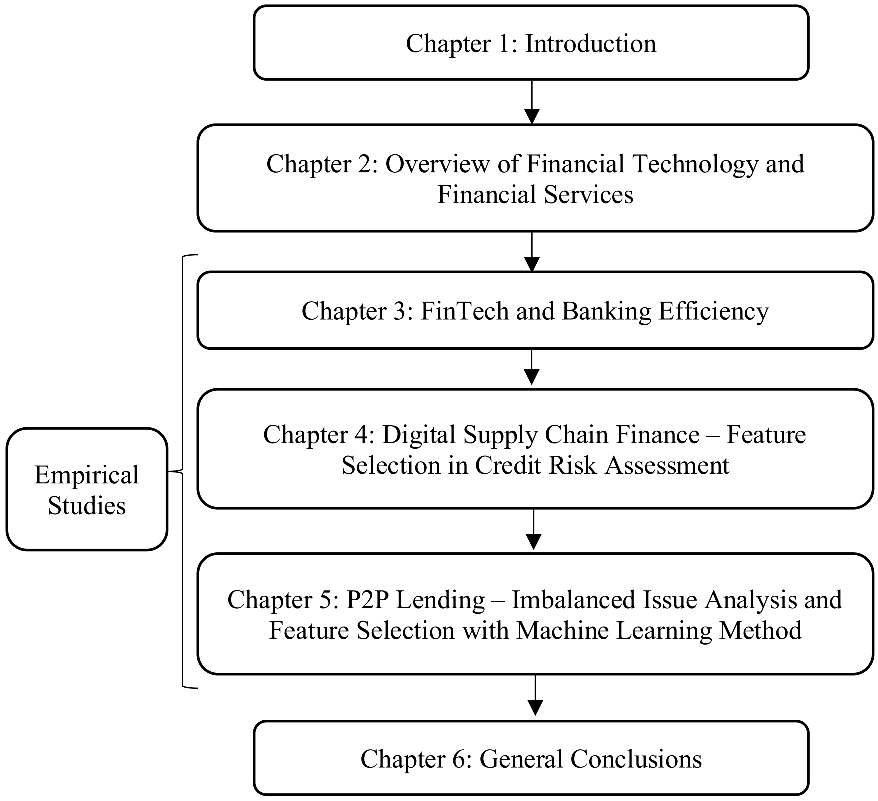


Figure 1.3. Overview of thesis structure

In Chapter 3, the impact of FinTech development on the efficiency of commercial banks in China is investigated by applying a two-stage Double Bootstrapped truncated regression. FinTech development in China has grown rapidly in recent years, with a wide range of physical application scenarios, and commercial bank efficiency is widely influenced. In terms of bank ownership and regional financial development, hypotheses based on different dimensions were formulated to detect the impact of FinTech on the efficiency of commercial banks in China. The DEA-Malmquist approach was used to estimate the dynamic efficiency change of 101 commercial banks for the period 2011-2020. Further, according to Simar & Wilson (2000), truncated regression models were used to regress DEA-Malmquist results to address inconsistent results due to changes in returns to scale. A double bootstrapped truncated regression is generated to obtain the bias-corrected score and examine whether FinTech development positively affects the efficiency changes of Chinese commercial banks. The overall results suggest that the FinTech development has a positive influence on Chinese commercial banks' efficiency. In addition, the positive impact of a higher level of FinTech development on the efficiency of CCBs is greater than that of SOCBs and JSCBs. Finally, FinTech has had a higher positive impact on CCBs in more financially developed regions than those in less financially developed regions.

In Chapter 4, a hybrid XGBoost-MLP method is implemented in the credit risk assessment of DSCF. The operational processes of DSCF as a new type of financial service are presented. Meanwhile, digitalisation features are incorporated into the set of features for credit risk assessment. The growing non-linear data issues make traditional credit assessment methods inefficient. In this context, machine learning methods are applied to credit risk assessment. At first, to avoid subjective arbitrariness in data processing, XGBoost is proposed for feature selection. Using feature importance ranking, this chapter filters out the set of features that make the evaluation optimal. Afterwards, LR, KNN, NB, RF, DT, SVM and MLP are used as classifiers. The whole sample is evaluated by an in-sample test and out-of-sample test to avoid overfitting. The results demonstrate the effectiveness of the hybrid XGBoost-MLP model for DSCF credit evaluation and the importance of including digitalisation features in SCF credit risk assessment.

In Chapter 5, a comprehensive investigation that combines feature selection with imbalanced learning methods is conducted for credit risk assessment of P2P lending. The credit assessment of P2P lending as exemplified by the Lending Club is considered by scholars to have a more serious imbalanced data problem, but the existing literature has ignored the importance of feature selection in dealing with the imbalanced data problem. Accordingly, the full sample of 151 original features was divided into two comparison sets, one with 17 features selected from the existing literature as the selected feature set, and the other with 87 features processed by data missing and data leakage only as our complete comparison feature set. For the selected feature set, XGBoost, DT and LR are used as baseline classifiers and imbalance learning methods by Bagging, Easy-Ensemble, IHT, Tomek links, SMOTE and CS are used for model evaluation. For comparison feature sets, the features are first selected by XGBoost and then subjected to the same model evaluation. By comparing the two feature sets, the results suggested that the performance of credit risk assessment is poor when the features are incomplete and that imbalanced learning methods are not effective in addressing the inefficiency of the evaluation results. Moreover, an analysis of the interactions between the degree of data imbalance and feature selection is performed. As a result, the more complete the data features are and the better the credit risk assessment performance is and the less the data imbalance problem affects them. In addition, the Easy-Ensemble method enables us to solve the problem to a great extent and obtain optimal results when there are limitations in the completeness of the features and imbalances in the data.

1.3 Contributions

This thesis contributed to the FinTech development and machine learning methods in financial applications mainly from banking efficiency, and credit risk assessment in DSCF and P2P. In addition, each chapter has separate outputs for research motivation, technical application, and empirical analysis.

Chapter 3 examines the impact of FinTech on the efficiency of commercial banks in China, a topic of both academic and practical significance. From an academic perspective, this paper enriches the research on FinTech and the influence of FinTech on commercial banks. In recent years, the world has witnessed a surge of research on FinTech and its effect on commercial banks, but the research is not yet systematic, and the research methods are not yet rich. In studies on FinTech, for example, it is difficult to conduct in-depth discussions as the definition of the connotation of FinTech has no standard in much research. Meanwhile, most studies believe that the development of FinTech hurts the traditional financial industry, thus ignoring the current growing cooperation. This chapter analyses the impact of FinTech development on commercial banks from the perspective of their efficiency which generates valuable implications for the subsequent development and application of FinTech in financial services. Further, this paper aims to overcome these problems as follows: Firstly, defining the meaning of FinTech based on a comprehensive review of previous research findings. Secondly, identifying the specific mechanisms through which FinTech affects the efficiency of banks, and constructing an econometric model based on data from listed banks to test the proposed hypotheses. The goal is to analyse the impact of FinTech on the efficiency of commercial banks with different ownership levels. This provides a closer look at the current state of cooperation and competition between different types of banks and FinTech startups and provides an empirical analysis of the future practice of commercial banks. Nevertheless, the degree of regional financial development is considered in a study related to the impact of FinTech development, about city commercial banks. This also clarifies the mutually reinforcing relationship between FinTech development and the degree of financial development for commercial banks, expanding the significance of research combining local financial development and FinTech. From a practical perspective, this chapter provides insights into the transformation of Chinese commercial banks' operations in the context of FinTech development. With the increasing share of FinTech applications in the financial sector, the development of FinTech has a growing impact on the overall operation of commercial banks, which in turn greatly affects the financial system (Vives, 2017). This chapter also clarifies the mechanism and extent of the impact of FinTech on the efficiency of banks and proposes countermeasures to accelerate the transformation of Chinese commercial banks in the FinTech environment as an example, which is beneficial for modern commercial banks to adopt more targeted and effective transformation measures.

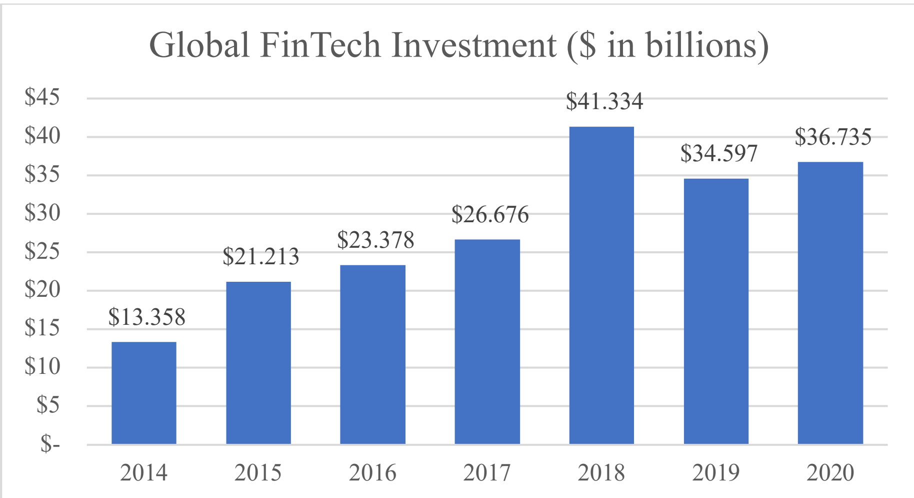
Unlike the previous chapters, Chapters 4 and 5 focus more on machine learning methods to study practical models for credit risk assessment of specific financial businesses and to produce better model solutions. In Chapter 4, the process of DSCF, a new type of financing business, is summarised systematically in detail, and the findings contribute to improving the theoretical framework of DSCF in influencing corporate investment, financing behaviour, and enhancing the impact of machine learning training models in credit risk assessment. The current research on SCF is based on SMEs' perspectives to explore the mitigation effects of SCF and the investment and financing decisions of enterprises through games, while digital-based SCF risk assessment has not received sufficient attention and lacks a corresponding theoretical guidance framework. Thus, based on manually collated data, this chapter adds new digital features to the training model for credit risk assessment. Using listed companies as the target of the study, it systematically investigates the business processes and risks of DSCF, provides theoretical explanations, completes the theory of SCF, and obtains a more effective feature set for credit risk assessment through empirical evidence. The conclusion of the study provides a reference for banks and other financial institutions to work closely with core enterprises, use DSCF to innovate financial supply methods, reduce credit risks, broaden long-tail customers, and promote profitable expansion and business upgrades.

Chapter 5 focuses on the problems of imbalanced data and subjective and arbitrary feature selection in existing credit risk assessments and identifies the difficulties and solutions in data analysis and processing. Using P2P lending data as the research object for in-depth analysis, this chapter draws on previous literature to extract key characteristics related to credit risk for comparison and analysis and implement machine learning methods to build a comprehensive assessment model for individual credit risk, the findings of which help to improve the credit risk avoidance capability of P2P lending. Further, by comparing and validating existing credit assessment models, the conclusions of this paper make clearer the relationship between the imbalance data issue and feature selection in the credit risk assessment, which provides more effective theoretical support for the data processing. Finally, the identification of key influencing features provides a reference for the selection of indicators in the subsequent credit risk assessment and also provides practical implications for the credit audit of P2P lending services and their following development.

Chapter 2 Overview of Financial Technology and Financial Services

2.1 Introduction

FinTech is a sweeping, technology-driven financial innovation that has seen the deep integration of representative information technologies such as big data, AI, cloud computing, blockchain and mobile internet with finance, resulting in a series of new financial industries such as mobile payment, big data credit, smart investment advisory, financial cloud, crypto-digital currency, and insurance technology. FinTech-related investment in the market is also on the rise (see Figure 2.1). Financial innovation is common, nor is it always beneficial to economic development, but is only worthy of encouragement and support if it serves the real economy. Therefore, the role of FinTech in serving the real economy has become the focus of scholars. FinTech is a "disruptive innovation" to traditional finance, which has changed the established financial solutions (Anshari et al., 2020) and consumption patterns (Feng and Zhang, 2021). Applying big data technology to expand data sources, FinTech facilitates financing for small and micro enterprises and other market players and promotes industrial development and upgrading (Alt et al., 2018). The risk control capabilities of financial institutions are improved by FinTech, which helps financial institutions acquire and retain quality customers (Xiang et al., 2017). Increasing innovative business models reduces the cost of financial services (Leong and Sung, 2018).



Source: CB INSIGHT CORNERSTONE ADVISOR[6]

Figure 2.1 Global FinTech Investment ($ in billions)

[6] https://www.forbes.com/sites/ronshevlin/2020/09/01/nobody-wants-an-online-bank-and-other-wisdom-from-investing-100-billion-in-FinTech/

FinTech is the fusion of finance and technology, but technological progress is always a double-edged sword with both advantages and disadvantages. The impact of FinTech on the banking industry shares the same internal logic following the theory of innovation destruction (Schumpeter J., 2000). From the perspective of creation, FinTech is conducive to reducing transaction costs and improving the efficiency of resource allocation in the banking industry, helping to bank financial institutions to break through the traditional business scope to obtain new profit growth points. From the perspective of disruption, the emergence and development of new financial business models based on FinTech has broken the established interest pattern of the banking industry and has had a strong impact on the mindset and business space of the traditional banking industry through a service approach that is closer to modern customer needs and more efficient and convenient.

In recent years, the function of FinTech to promote the transformation and upgrading of the traditional financial industry has gradually emerged, especially after big breakthroughs in underlying technologies, i.e., AI (Rasiwala and Kohli, 2021). With such technologies, a new wave of integration between traditional financial services such as payments, financing, wealth management and FinTech is rapidly developing, which drive the financial industry to improve the efficiency of financial services, provide better and more targeted products and services, and enhance the inclusiveness of the financial system (Philippon T., 2016). FinTech is providing a potential opportunity for traditional financial sectors. As Buchak et al. (2018) suggest, the relationship between shadow banking and FinTech in the US concluded that FinTech has played an important role in the rapid development of shadow banking. In contrast to shadow banks that do not use FinTech, shadow banks that use FinTech allocate credit resources to borrowers with better credit and participate more actively in the refinancing market, and FinTech improves the overall resource allocation efficiency of the financial market. Omarini (2020) also claims that under the wave of FinTech sweeping the world, traditional financial institutions can get a win-win situation by cooperating with FinTech companies. FinTech companies need to access customer data and payment systems through traditional financial institutions, and as long as they adopt the right response strategies, the creative effects of FinTech on traditional financial institutions will be greater than the "destructive effects".

FinTech is driving the improving financial services, it is creating new types of financial risks and challenging the existing financial regulatory system. DSCF as a new FinTech service is an important means of corporate finance, originating from the management of capital flows in supply chain organisations, aiming to optimise the allocation of funds in the organisation through the synergistic operation of logistics, information flow and capital flow, reduce the overall costs of the system and achieve the best overall value of the supply chain (Du et al., 2020). Traditional SCF is more of a financial lending based on interpersonal trust, which not only requires long-term cooperation between trading parties but also relies on the core parties' guarantee of the SMEs' rights (Moretto et al., 2019). The inherent weaknesses of SMEs and the goal of maximising the interests of the core enterprises greatly limit the effectiveness of SCF. In the face of the swift ascent of the digital economy, industrial enterprises are intensifying their focus on adopting novel technologies and merging internal systems (Gomber et al., 2018). This is aimed at enhancing the amalgamation and utilization of enterprise data. Concurrently, a growing number of commercial banks are leveraging digital technology and FinTech to spearhead the strategic evolution of their operations. They are engaging in wide-ranging collaborations and strategic alignment within the supply chain ecosystem. This entails extracting and refining data from the operational contexts of enterprises, closely tied with panoramic data (He et al., 2020). These efforts centre on the functioning of businesses across procurement, manufacturing, and marketing domains. The company's credit rating is based on the consistency and validity of the company's behavioural data in procurement, manufacturing, and sales.

P2P lending, meanwhile, regarded as a new FinTech service for personal financial needs, is a transaction model that enables peer-to-peer lending over the internet, with the transaction venue provided by an online platform (Havrylchyk and Verdier, 2018). The purpose of P2P lending is to provide loans to groups that have difficulty borrowing from large traditional banks. This has enabled the digitalisation, informatisation and dynamisation of lending behaviour and the effective integration of online and offline financial resources, which has had a significant impact on the development of SMEs and the expansion of individual spending power. However, amid the ongoing sequence of technological advancements, the credit risks associated with P2P lending operations have evolved into intricate and diverse challenges. P2P lending frequently involves SMEs and average consumers lacking robust financial stability within the lending framework (Wang et al., 2016; Bavoso, 2020). This not only augments the financial and spending capabilities of these segments, but also introduces complexities in accurately appraising the creditworthiness of individual borrowers and establishing more stringent limitations for borrowers within the digital realm. The issue of credit risk in P2P lending is therefore very prominent and deserves attention.

Risk management in new financial businesses is one of the core innovations in FinTech. As mobile internet and big data technologies advance, individuals are gaining access to an immense volume of novel data. This data is marked by its extensive size, intricate nature, and lack of structure, presenting a substantial challenge for data analysts. At the same time, the reality is that the collected independent and dependent variables often do not satisfy a simple linear relationship with each other, thus making traditional risk assessment models even more limited. These challenges are where machine learning excels. Machine learning combines the disciplines of statistics, computer technology and probability theory to solve the problem of automatically building computational models through experience and is at the heart of AI and data science (Jordan & Mitchell, 2015).

Machine learning approaches, which are becoming mainstream in AI, are based on a data-driven approach to building models that recognise and learn specific patterns and laws, with 'learning' rather than established 'rules' at their core. Algorithms in the field of machine learning are often less constrained by assumptions about the data, which makes them flexible, efficient, and highly accurate. Machine learning has the following characteristics: firstly, it is a multi-disciplinary subject, involving probability theory, statistics, algorithmic complexity theory, computational theory and other disciplines; secondly, as the word suggests, it is "the machine learning itself", which differs from traditional programming based on explicit rule-based algorithms. It operates under human-set rules and summarised knowledge and instead investigates how to recognise patterns or acquire knowledge from data without human intervention. (Bell, 2022). Naturally, machine learning is a systematic set of components that includes the problem to be solved, the data, the model, the optimisation algorithm, and the processes of validation and testing (Weichert et al., 2019). Machine learning is not a completely new concept in historical terms, as it has been developed in theory and practice since the 1950s and has undergone several methodological and paradigm shifts over the past 70 years (Shinde and Shah, 2018). In the 21st century, machine learning methods have made great breakthroughs in areas such as knowledge learning and pattern recognition. In addition to computer engineering applications, machine learning applications in the social sciences, especially financial fields, have also begun to increase (Athey, 2018).

The existing literature has studied the impact of FinTech on the banking industry in three areas: operational efficiency, business innovation and risk management. The motivation behind this study is rooted in the quest to shed light on the intricate relationship between scholars working in the field of FinTech and traditional financial institutions. This thesis aims to explore the dynamic interactions and collaborations that shape the evolution of both domains, and how they influence each other in terms of knowledge exchange, technological advancements, and the integration of innovative solutions into traditional financial practices. By exploring the development logic of FinTech, sorting out the background, evolutionary stages, basic driving technologies and primary business models of FinTech, and revealing the differences and connections between FinTech and traditional finance, this research clarifies the development logic of FinTech and its connotation. Moreover, the research endeavours to systematically analyse the role of machine learning in revolutionizing traditional financial businesses. Machine learning techniques have proven to be particularly transformative, offering new possibilities for data analysis, risk assessment, fraud detection, customer service, and investment strategies. This thesis seeks to comprehensively examine the adoption of machine learning algorithms by traditional financial institutions, their benefits, and challenges encountered during integration.

2.2 Financial Services Powered by FinTech

FinTech, financial innovation enabled by technology, (FSB 2016) relies on high technologies such as big data, cloud computing, blockchain and AI to gradually open new paths for SME finance. Scholars agree that FinTech innovates business models and services and improves the quality of information dissemination and tracking (Wonglimpiyarant et al., 2017). Farboodi and Veldkamp (2017) point out that FinTech measures information, trading strategies and market efficiency through price information and market liquidity. One of the crucial functions of FinTech is to provide access to finance for companies, and lending platforms and the smart settlement and credit rating technologies behind them all reflect specific applications of FinTech for corporate finance (Berg et al., 2022). FinTech lending is an application that has been developed with debt financing as a fundamental point. The Basel Committee on Banking Supervision classifies FinTech business models and considers that FinTech activities are classified into depository and financing services, payment and clearing services, market infrastructure services and investment management services (see Table 2.1). The birth of FinTech as a product of the coupling of finance and technology did not happen overnight (Arner et al., 2015). Alt and Puschmann (2012) argue that advances in information technology, changes in consumer behaviour, the lack of traditional formal banking services and the existence of regulatory arbitrage have contributed to the emergence of FinTech.

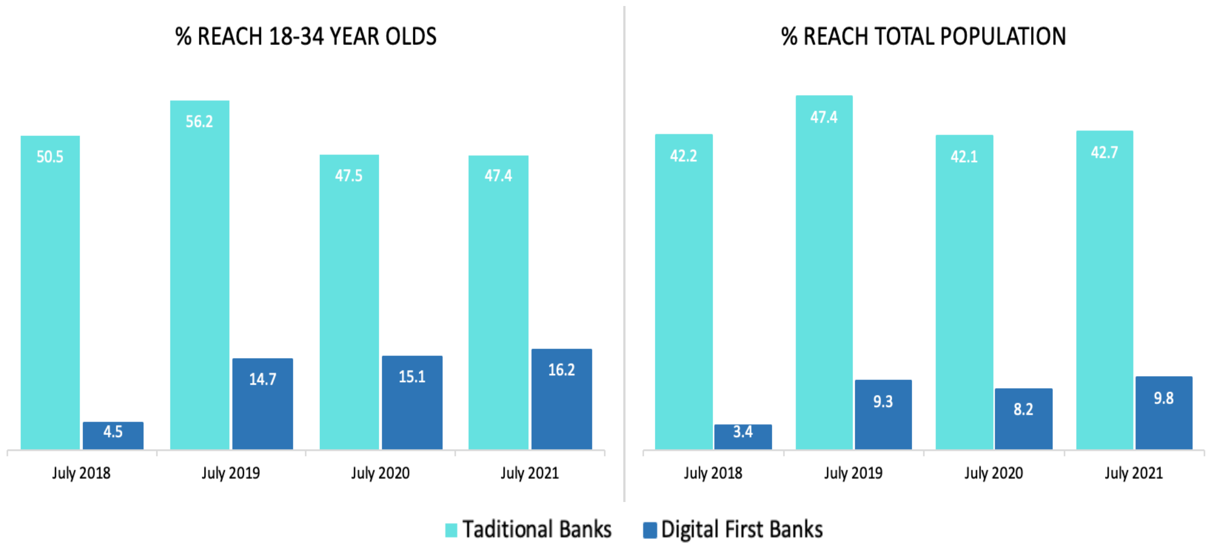
Table 2.1 Sectional Innovations by FinTech

Depository and Financing Services	Payment, Clearing and Settlement Services	Market Infrastructure Services	Investment Management Services
Lending Marketplaces	Cryptocurrency	AI	Robot Advisor
Mobile Banking	P2P Transfers	Big Data Management	E-Trading
Credit Rating	Mobile Wallet	Digital certificates	High Frequency Trading
Crowdfunding		Blockchain	Copy-Trading
		Cloud Computing	

Source: https://www.bis.org/bcbs/publ/d415.pdf

2.2.1 Banking

FinTech is evolving rapidly and is often presented in the news as 'revolutionary', with 'digital weapons' that will destroy industry barriers and traditional financial institutions. The question of whether and to what extent FinTechs will replace traditional financial institutions such as banks remains a major point of research (Navaretti et al., 2018). The existing research suggests that there are two broad impact mechanisms of FinTech development on the operations of commercial banks, namely positive technology spillover effects and negative market crowding-out effects. As Chuen and Teo (2015) illustrated that the marginal cost of financial services decreases and the marginal benefit increases with the continuous development and application of FinTech, and the economic efficiency of banks benefits from customer connection behaviour and the resulting economies of scale and positive externalities, resulting in the high network effect. FinTech companies are often regarded to be ahead of traditional commercial banks in terms of digital technology, and technological knowledge itself has positive externalities (Gazel and Schwienbacher, 2021). As new market players, FinTech provides society with more efficient financial services and uses more comprehensive data to reduce transaction costs and improve decision-making, increasing financial inclusion (Zetsche et al., 2017). Alternatively, Li et al. (2022) suggested that there is a negative market crowding out effect because the development of FinTech does pose a significant challenge to the banking industry, with emerging financial sectors such as internet finance companies and third-party payment platforms crowding out commercial banks' liability, intermediate and asset businesses. There are surveys showing that younger users are becoming more inclined to use digital banking and less inclined to use traditional banking (see Figure 2.2). Numerous studies have compared the characteristics, advantages, and disadvantages of FinTech and traditional banks. However, a research gap exists in understanding the specific factors that drive consumers to choose FinTech services over traditional banks and vice versa. Moreover, there is a research gap in understanding how FinTech affects small and medium-sized banks, which might face unique challenges and opportunities in the evolving financial landscape. At the same time, internet finance has had an impact on the marketisation of deposit rates, affecting commercial banks' profits and increasing their risks. In the face of the huge challenges posed by FinTech, traditional commercial banks have also started to take the initiative in using FinTech.



Source: https://www.comscore.com/Insights/Blog/FinTech-and-traditional-banks-competition-or-symbiosis

Figure 2.2 Traditional Banks vs Digital Banks.

There are two new models offered by FinTech for traditional banks to improve access to finance. Firstly, big data credit, where institutions can better screen the creditworthiness of borrowers by accessing big data information that banks do not have, and then grant loans. Secondly, P2P lending, where lenders and borrowers are matched directly via the Internet. The issues of information asymmetries and data complexity are addressed on this basis. The non-traditional information is a proxy for credit information, acts as a signal, and is often considered soft credit information. For example, physical appearance (Duarte et al., 2012), ethnicity and gender (Pope & Sydnor, 2011), financial status (Donnelly et al., 2009), social connections (Freedman & Jin, 2008), social networks (Freedman & Jin, 2017), borrowing statements and investor rationality.

2.2.2 Supply Chain Finance

SCF is a specialised area of commercial banks' credit business and an important financing channel for enterprises based on big data, especially SMEs. SCF exists in a complete industrial chain, where the enterprises in the chain use the credit of the core enterprises and real transactions as the background and use some non-cash liquid assets such as accounts receivable, prepaid accounts, inventory, etc. as the guarantee for repayment. The information flow, logistics, and capital flow provide the foundations for the service in the supply chain. This financing mode breaks through the traditional enterprise financing process, which only focuses on the enterprise's own financial and operational status and reduces the financing risks caused by information asymmetry by improving the transparency of information of enterprises in the supply chain. With the development of FinTech, commercial banks are rapidly developing their electronic banking channels, while enterprises are also expanding the application of technological tools such as Enterprise Resource Planning (ERP), Supply Chain Management (SCM), and Money Management System (TMS), which gives SCF the basis for networking. Combining the resource-based concept with open innovation, commercial banks have two complementary models, from the perspective of external resources which are used to meet customer demand, and from the perspective of internal sources, where their resources are shared with the outside world, for example by forming strategic alliances to achieve integration (Gianiodis et al, 2014). In contrast, commercial banks use digital technology to empower SCF and develop DSCF, which streamlines processes, reduces costs, and enables them to serve more customers at lower costs (Caniato et al., 2016). Innovative technologies such as blockchain are also considered key elements to address existing difficulties and can further drive SCF to achieve multi-level credit penetration, combining SCF with recent technologies and driving SCF into the era of digital visualisation.

The risk management capability of SCF has also been rapidly enhanced with the innovation of business models and the in-depth application of FinTech. In the development of DSCF, the construction of a data-based risk control system for SCF has played a key role. Big data technology, the booming development of e-commerce platforms, and the trend of networked transactions are improving the accessibility of risk control, as well as the supply chain transaction data that can be obtained with low cost, high efficiency, and high credibility. In addition, regarding data acquisition, enterprise production and operation data and transaction data are collected in real-time through the Internet. In terms of data storage, cloud computing, and distributed databases accommodate massive amounts of data. Further, data processing, natural language understanding, image recognition, data mining, and machine learning technologies are becoming increasingly mature, and the application of AI technology improves the efficiency of information processing and partially or even completely replaces expert decision-making. While applying FinTech to SCF, the biggest challenge facing SCF is that the traditional credit system has lost its effectiveness, while a new credit system has not yet been established (Song et al., 2023). The heterogeneity of business contexts, irregular data forms, and information silos in the existing system make it difficult to form industry-level supply chain information, the authenticity of which is questionable, making the establishment of a new credit risk management industry one of the issues worth studying.

2.2.3 P2P Lending

P2P lending services are a complement to traditional commercial bank credit, providing a new financing channel for SMEs and individuals to borrow money due to the "no credit, no collateral" business model adopted by existing financial institutions due to the imperfection of central bank credit collection (Havrylchyk and Verdier, 2018). Freedman and Jin (2008), by examining the social network nature of P2P lending, suggest that online lending has the advantage of lower interest rates and hence lower lending costs compared to traditional financial intermediaries. However, due to the nature and scope of business of online lending institutions, there are not enough funds to meet the needs of borrowers and lenders, and there is vicious competition in the industry under the practice of "zero interest" for online lending. P2P lending has also been seen as an alternative to traditional financial institutions. All loans on the Lending Club platforms are made through a partnership with WebBank, and loans are still granted through commercial banks, which act as intermediaries between the lenders and borrowers, so Lending Club executives do not see the rise of online lending platforms as competition to commercial banks. However, De Roure et al. (2022) argue that the emergence of the P2P lending model, with its high transparency and simplicity of procedures, has attracted more investors to invest through P2P platforms, which has dispersed the customers of commercial banks. While Tang (2019) constructed a theoretical model to identify the substitute role of P2P lending to banks. Nevertheless, with the numerous incidents of P2P lending platforms plunging into thunder, risk management in P2P lending has become a principal issue that needs to be addressed urgently, and the development and application of a series of FinTech technologies such as AI and big data have revived the confidence of P2P platforms.

Regarding the risks of P2P lending, P2P lending creates a creditor-debtor relationship just like ordinary loans, and therefore P2P lending cannot ignore credit risk. The credit risk factor for P2P lending is higher than for bonds and bank loans, due to the lack of collateral and credit references of P2P borrowers, or even having a poor credit history (Emekter et al., 2015). This is compounded by the fact that P2P lending has higher interest rates, which puts more pressure on borrowers to finance themselves than loans from banks, and therefore increases the risk of default. The adverse selection risk faced by institutions increases with higher interest rates, and with higher interest rates, the credit risk of borrowers increases accordingly (Stiglitz and Weiss, 1981). In addition, investors also are exposed to significant risks when moral hazard happened in response to the lure of high profits.

For P2P credit risk management, the massive voluntary information posted by lenders is unverifiable. Neither the lender nor the platform could screen the information for authenticity and cannot be used as a credible source of information. However, lenders are precisely susceptible to being influenced by this information and making potentially incorrect decisions (Michels, 2012). Thus, with the emergence of substantial amounts of data, researchers have begun to focus on the economic impact of big data. Goldfarb & Tucker (2019) found that the advantages of big data have led to cost reductions in data storage, computation, and data transmission. In addition, Begenau et al. (2018) and Farboodi et al. (2019) analysed the relationship between big data and firm size and growth and they found that data analysis improves investors' forecasts and reduces equity uncertainty, which, in turn, lowers the firm's cost of capital and increases the skewness of the firm size distribution due to the larger data generation and higher investment in active experimentation by large firms. Moreover, Farboodi & Veldkamp (2020) explored how access to big data affects market effectiveness and illustrated that the advancing data processing technology ensures the continued processing of both historical and future data in the long run, with data resolving investment risk while also introducing new risk factors. The proliferation of methods for analysing big data provides good methodological support for its application in the financial sector, as summarised by Varian (2014) and Mullainathan and Spiess (2017). Empirically, Hua and Huang (2021) use data from Ant Financial Services to find that big data credit can increase the volume and diversity of goods and improve merchant service levels. This finding not only validates the effectiveness of big data credit but also indirectly reflects that FinTech promotes real economic development channels. As such, AI methods for processing big data are being developed to effectively mine and process all kinds of data on the Internet to find valuable information thereby helping investors make effective decisions.

2.3 Financial Business and Artificial Intelligence Applications

The new businesses that have emerged from FinTech have also created more complex risk management needs. The four representative technologies of FinTech have different depths of application in risk management scenarios, with different areas of focus and some crossover. Cloud computing technology brings breakthroughs in computing power and speed for massive data; big data risk control technology is mainly applied in the field of credit risk management in internet finance, solving the problem of information asymmetry; AI risk control technology is based on big data, mainly solving the problem of optimising risk control models; blockchain is mainly applied in the field of technical security in operational risk management such as payment clearing.

Regarding the AI industry, its fervour is rising globally, and the market is expanding. Accordingly, global AI investment and financing reached $9.35 billion by 2021, an increase of 38% from 2020[7]. Meanwhile, according to the 2021 AI Index Report released by Stanford University (see Figure 2.3), global Al company funding continued to converge on leading start-ups starting in 2018, and the number of new companies formed each year continued to decline starting in 2018, but Al funding amounts continued to maintain an upward trend[8]. Building models from massive amounts of data, machine learning, and AI will optimise business decisions, provide bespoke services, and improve risk management. According to McKinsey Global Institute, the application of these technologies is expected to create more than $250 billion in value for the banking industry[9].

[7] https://www.statista.com/statistics/941137/ai-investment-and-funding-worldwide/
[8] https://aiindex.stanford.edu/ report/
[9] https://www.mckinsey.com/industries/financial-services/our-insights/banking-matters/a-new-era-of-divergence

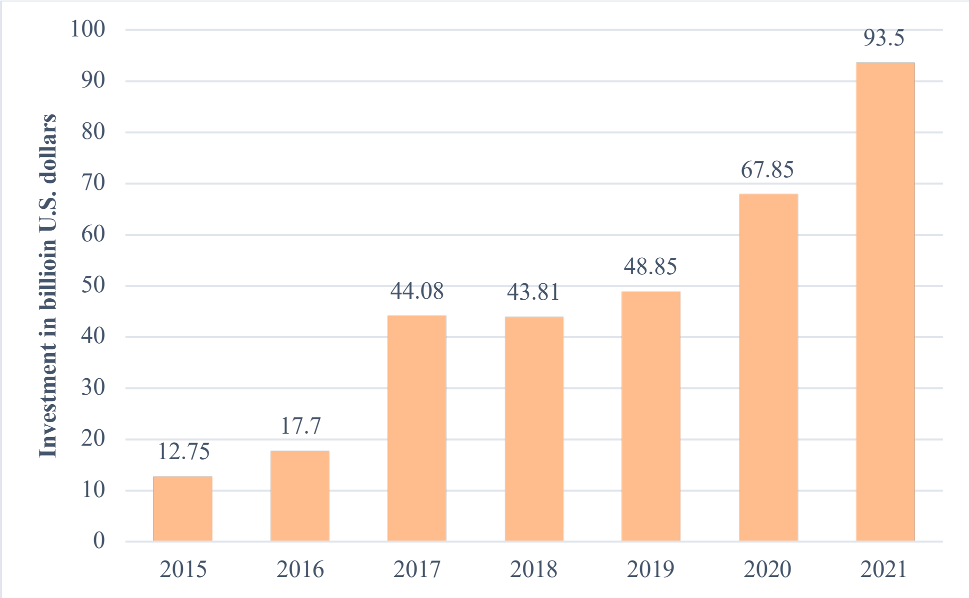


Figure 2.3 Global total corporate investment in AI from 2015-2021.

AI in the financial services industry empowers risk and service operations. In the financial sector, AI is implemented in three parts (see Figure 2.4): the infrastructure layer, the core technology layer, and the extended application layer. The infrastructure layer provides basic hardware and software facilities, including computing hardware (AI chips, sensors, etc.), computing system technology (big data, blockchain technology, etc.), and big data technology (data collection, data storage, etc.). The core technology layer includes algorithm theory, development framework, and common application technologies, relying on computing platforms and data resources for mass recognition training and machine learning modeling, including natural language processing, graphic recognition, and deep learning, which are also the most important aspects of AI services for financial services (Jia et al., 2018). Technology giants Google, IBM, Amazon, Apple, Alibaba, and Baidu are all deeply involved in this layer. The extended application layer addresses practical issues, where AI technologies provide products, services, and solutions for the financial industry, with commercialisation at its core. In the extended application layer, companies are integrating AI technology into their products and services to transform and upgrade their traditional business, for the data-intensive industry of finance. In parallel with rapid advances in machine learning techniques, there is a proliferation of diverse, high-frequency, voluminous, structured, or unstructured data at a granular level. Collecting and training in big data has undoubtedly led to an increase in the scope and capabilities of data-driven machine learning models. For financial services, machine learning advances risk management forward in time, further predicting the cycle of financial risk and the degree of harm through big data mining, and intelligently monitoring risky transactions and irregularities.

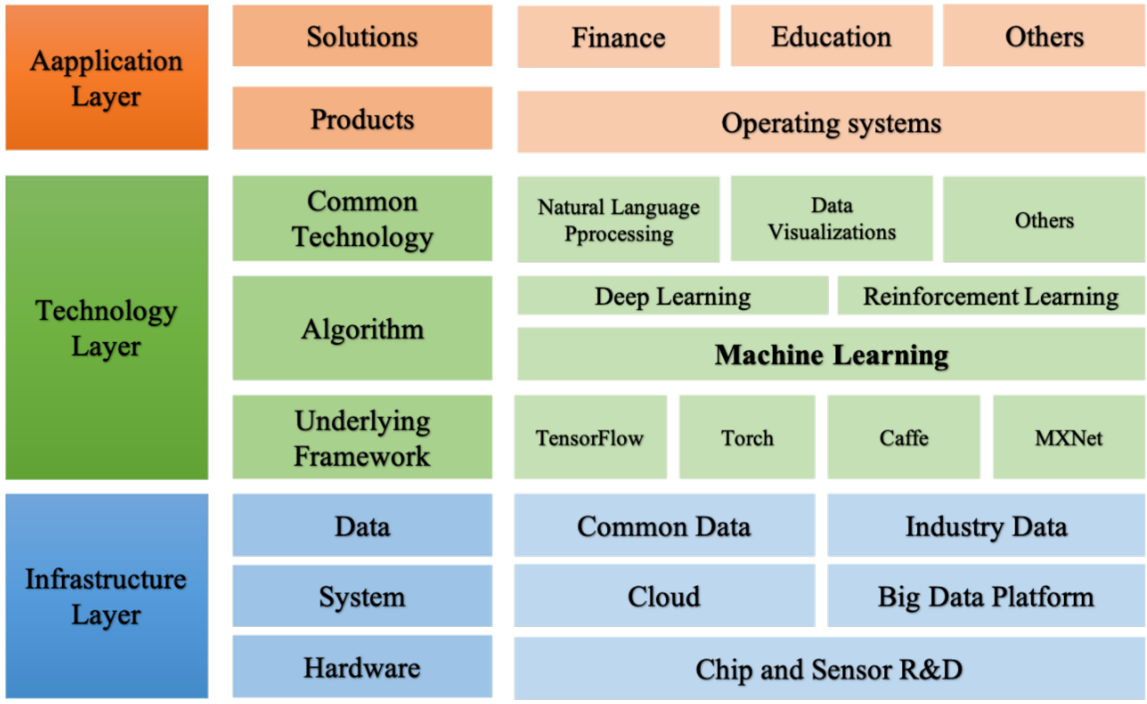


Figure 2.4 Implementation process of AI

However, the application of machine learning models also amplifies certain gaps that are present in traditional models and relevant research. Many researchers discuss the application of machine learning models in various financial tasks and highlight the advantages and disadvantages of using such models in comparison to traditional methods. Lee et al. (2019) reviewed and sorted the machine learning application in risk management of banking. As well as the study of Goodell et al. (2021) which provide the statistical finding for AI and machine learning research and usage in the financial field. Moreover, increasing model complexity and improving explainability is also regarded as significant research point when machine learning models are applied to different industries or applications. In the current study, Machine learning models are typically based on large-scale unstructured datasets (e.g., natural language, image, speech information, etc.) and are built using new software packages and specific computing infrastructures (Ma et al., 2014). Nevertheless, the model's complexity does not mean an overly complex response should be adopted. As shown in Figure 2.5 below, the McKinsey Risk Dynamics model risk validation and management team, has adapted the model validation framework and practice approach so that the bank's existing traditional model validation framework is fully capable of effectively managing the risks associated with machine learning models. Despite this, current research still has some gaps in developing hybrid models that standardize evaluation processes and integrate domain expertise.

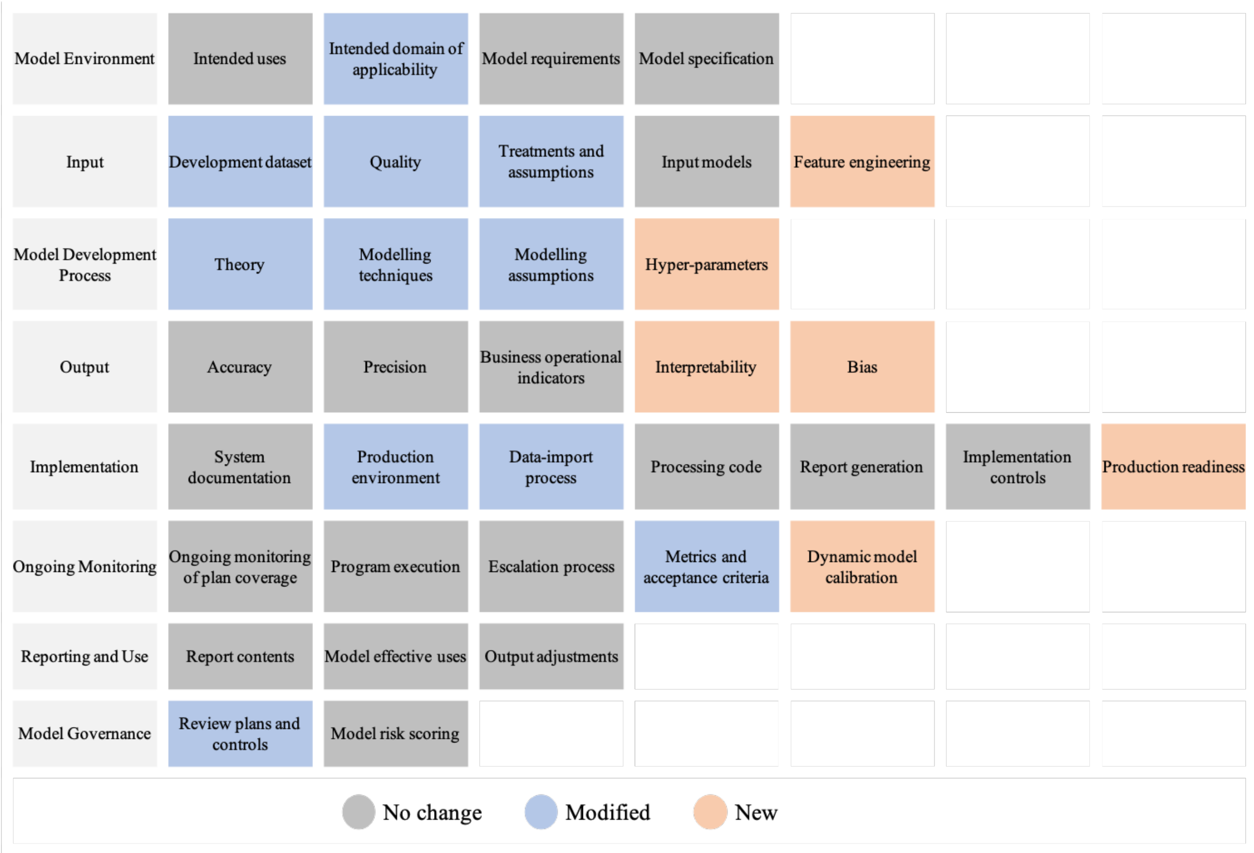


Figure 2.5 Machine Learning updated framework in risk management.

2.4 Machine Learning in Credit Risk Assessment

As the backbone of AI, the Machine learning approach allows the computer to identify the relationship between the data and model by learning the characteristics of the sample through advanced accuracy optimisation. Machine learning approaches are divided into two categories according to whether the data has labels: supervised and unsupervised learning. The supervised learning focus on the classification and regression targets where classification algorithms are used for discrete data distribution prediction and regression algorithms are used for continuous data distribution prediction (Nasteski, 2017). Unsupervised learning includes clustering and anomaly detection. Clustering algorithms scatter data into different groups according to similarity, and anomaly detection algorithms calculate the outlier degree of each data. From the perspective of credit risk assessment, the goal of identifying default groups is regarded as a binary classification target. On this basis, the modeling design of machine learning include the following five processes as Figure 2.6 displayed: at first, data import and data pre-processing, such as data cleaning, and normalisation. The second step is feature engineering, including feature selection and extraction, filtering, and transformation with prepared data. In the third step, model training aims to select an appropriate algorithm and perform model training on the data, and implement parameter tuning. Then, the model evaluation and the model export, deployment, and application are processed finally.

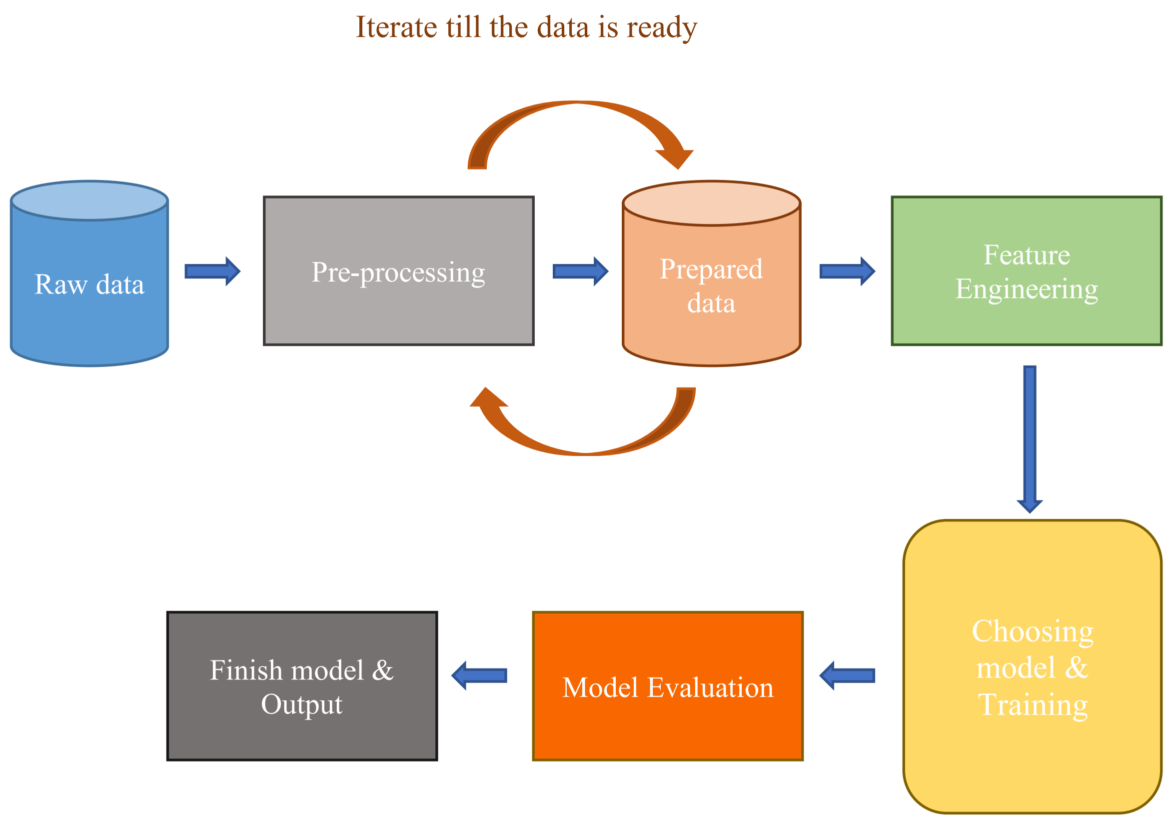


Figure 2.6 Overall process of machine learning design.

As computing and IT become more intergraded into daily life, the collection, collation, and processing of data relating to credit risk have become more feasible than ever before. Due to this, there has been a surge in the demand for data analysis and data classification, making statistical learning, machine learning, and ensemble learning emerge as prominent subjects of research. These different approaches are often cross-fertilised in current research methods.

According to Galindo and Tamayo (2000), traditional statistical methods are dominated by linear, quadratic, and LR, while further modern statistical methods include a range of methods such as KNN. LR was proposed by Wiginton in 1989 and is frequently used for credit risk assessment. It determines the probability of default by assuming that the applicant has both performance and default after taking out a loan and then doing an LR analysis based on the sample information data. This algorithm does not require the variables to follow a normal distribution and is more suitable for dealing with binary classification problems. Alternatively, the KNN method was introduced for the credit risk assessment (Henley and Hand, 1996). Applying a smoothing function to relate the category labels and sample attributes of the evaluation data, KNN uses a discriminant function to analyse the number of samples belonging to each category label among the samples of each category label. Obviously, statistical models often rely on specific assumptions, such as linearity, normality, and independence of observations. However, in credit risk assessment, these assumptions may not hold true for all data sets, leading to potential biases and inaccuracies in model predictions. Addressing the violation of assumptions and developing robust statistical models is essential. Meanwhile, credit risk assessment datasets can be high-dimensional, with numerous variables representing various aspects of the borrower's financial profile (Oreski and Oreski, 2014).

Thus, the machine learning approaches such as neural networks, support vector machines (SVM), and decision trees (DT) are used for credit risk assessment (Bhatore et al., 2020). Since the capability of dealing with non-linear and the efficiency of prediction, the DT algorithm is used in the most basic classification problems, such as a probabilistic decision algorithm mapped in feature space and category space. The DT method primarily classifies many influencing factors that affect the borrower's credit, and through the introduction of entropy indicators to filter out the key classification factors for default and non-default situations, the DT method is highly operable and enables to better eliminate data noise. However, this model is more sensitive to data analysis and more prone to overfitting during the classification of single trees. In turn, neural networks models take the attribute variables of the sample data as input data by constructing multiple hidden layers and then pass them through layer by layer to obtain the output results. The neural networks algorithm model can be backpropagated if the output results are far from the error margin. During the iterative forward and backward propagation process, the parameter values of the individual models, i.e., thresholds and weights, are continuously adjusted and the procedure is terminated when the results of the output layer reach the set error range. Different neural network techniques are used to solve the problem of credit risk assessment, such as probabilistic neural networks, artificial neural networks, and hybrid neural networks. Furthermore, SVM as a linear classifier based on a feature space is introduced to credit scoring (Baesens et al., 2003). The algorithmic process of SVM is based on the construction of an approach by assigning weights to variables. SVMs are designed based on a high-dimensional space using kernel functions, which improves the classification accuracy of the predictive model compared to logistic algorithms, which are more suitable for binary classification problems. While the SVM is slow to train, it is not only inefficient when applied to large samples of data for classification. Then, single machine learning can no longer effectively meet the needs of prediction because of richer, more voluminous, and more imbalanced forms of data (Zhu et al., 2017).

Ensemble learnings are considered improved predictive solutions (Saqi and Rokach, 2018). By iteratively learning sample data, multiple weak classifiers are used as base classifiers, and then the base classifiers are combined to improve the prediction accuracy and generalisation ability of the model. Each weak classifier is differentiated by the addition of a factor perturbation mechanism, and each better weak classifier is combined into a strong classifier with high prediction accuracy. Currently, two common ensemble methods are Bagging and Boosting, which are used to obtain the final strong classifier by combining different classification algorithms to improve the accuracy of a single classifier and to improve the generalisation ability of the classifier.

In Bagging, a bootstrap approach is used to obtain N datasets from the overall dataset by taking a put-back sample, learning a model on each dataset, and using the output of the N models to obtain the final prediction, specifically as the classification problem represented by credit risk assessment uses N models to predict votes. Random Forest (RF), for example, is a form of Bagging, and it was confirmed that it outperforms a single classification DT model by composing multiple DTs, one of which is more accurate for a particular attribute and category of data (Breiman, 2001). The final discriminatory result is obtained by analysing the results of multiple DTs.

Boosting methods, the representative example of Boosting is AdaBoost (Adaptive boosting), which learns a series of weak classifiers and combines them into a strong classifier by weighting the same sample training set differently, then training the weak classifiers with this data and finally constructing a DT from these underlying aspects. It has the advantage of being highly tolerant of noise and can analyse large sample classes for weighting while computing in parallel to obtain the final integrated classifier. During each iteration, the sample weight is determined by whether it is correctly classified by the weak classifier, with its weight being increased if the current sample data is incorrectly classified, and decreased if the opposite is true, the advantage of AdaBoost includes less overfitting during the classification process and high prediction accuracy. Further, the Gradient Boosting Decision Tree (GBDT) was proposed as a Boosting method to improve the accuracy of a single model. The Gradient Boosting Decision Tree (GBDT) algorithm is trained in each round because of the residuals of the previous round, and it assigns an initial value to the input samples at the beginning of the prediction. Moreover, the Extreme Gradient Boosting (XGBboost) proposed by Tianqi Chen et al. in 2014 is an improvement of the GBDT algorithm, which solves both the regression and classification problems. The problem of inefficient classification by GBDT is solved by parallel computation.

Overall, ensemble methods offer several advantages, such as improved accuracy, robustness, and the ability to handle high-dimensional data. However, they also face challenges, such as increased computational complexity and potential overfitting, which need to be carefully addressed in credit risk assessment scenarios. Meanwhile, Florez-Lopez and Ramon-Jeronimo (2015) claimed that the current research related to the applications of ensemble methods often creates complex models that lack interpretability. In credit risk assessment, where transparency and justifiability are essential, the lack of interpretability can be a significant limitation. Moreover, credit risk datasets are commonly imbalanced, with a small fraction of default cases compared to non-default cases (Chen et al., 2016). Thus, the modifications to address the inherent bias in such datasets and produce balanced and fair credit risk assessments still exists in gaps. Finally, the emergence of new industries and novel business models introduces unique challenges to credit risk assessment. For instance, industries like FinTech start-ups, digital SCF services, and P2P lending platforms may have limited historical data or unconventional risk profiles, requiring adaptable ensemble techniques to model credit risk effectively. This critical literature review identifies the current research and gaps in the application of FinTech development and machine learning in credit risk assessment. The review also emphasizes the importance of adapting ensemble techniques to meet the challenges posed by emerging industries. By addressing these gaps and considering the evolving industry landscape, researchers and practitioners can develop more accurate and robust ensemble-based credit risk assessment models, contributing to enhanced risk management in the financial sector.

Chapter 3 FinTech and Banking Efficiency

3.1 Introduction

FinTech has become a driving factor for innovation across firms and financial institutions. This innovation varies across different dimensions, such as operational improvements, cost-cutting alternative applications, access to a larger pool of customers, and integration of more popular products and services, inter alia (Nguyen et al., 2022). This new trend in financial services combines digital innovations, technology-enabled business models, and big data analytics that can facilitate improved financial decision-making (Soni et al., 2022), enforce sustainable solutions (Chueca Vergara and Ferruz Agudo, 2021) and improve firms' risk profile (Austin and Dunham, 2022). All of the above concludes that there is strategic value for firms from FinTech innovation, as clearly explored by Chen et al. (2019).

However, the picture is not so clear when it comes to evaluating the importance of FinTech innovation in the banking sector. Although the literature theoretically links FinTech with the disruption of financial services and the potential disintermediation effect, empirical analysis showcasing the effects of this new trend in the operations of banks is scarce. There are different views in the literature linking financial innovation with banking diversity (Berger, 2003), resource allocation (Hartman et al., 2001), and innovation fragility and risk analysis (González et al., 2016; Zhao et al., 2022). Nonetheless, not enough attention is paid to the direct effects of FinTech. This is crucial, especially under the prism of Thakor (2020), who explains that when it comes to commercial banks, FinTech promotes the technological upgrade of their financial products and services, but also enhances overall competition in the sector.

The motivation for this study stems from the above situations of modern interaction of technology innovation and banking services. Our main goal is attempting to quantify the effects and assess the importance of FinTech for banking efficiency, focusing on the Chinese banking system and particularly Chinese commercial banks. China has seen FinTech emerging in the past few years. In 2021 it was reported that China's total FinTech investment and financing reached 12.67 billion Yuan in the first quarter, a year-on-year growth rate of nearly 130%. Traditional financial institutions have become significantly more proactive in the FinTech space, with large State-owned Commercial Banks (SOCBs) increasing their technology investment by 34.54% in 2020, well above their overall revenue growth rate of

4.44%. In line with this argument, Gorjón Rivas (2018) suggests that more than 25% of global users of FinTech work in China which has a large FinTech market where the Chinese use their mobile phones for business transactions. Proportion-wise this corresponds to 87.3% of the population, compared to 43.2% in the US.[10] In addition, China contains eleven times as many people in the US who use their mobile phones for payment transactions. This is consistent with KPMG's report (2018) that in 2018, four out of ten FinTech companies in the world were Chinese firms. Consequently, Chinese commercial banks have experienced an upsurge in FinTech-related advances to promote growth and economic stability.

This work is related to three different strands of literature. First, the empirical and theoretical framework is built that looks at the factors affecting the efficiency of banks. These confirm that bank-specific and country-specific factors influence efficiency levels (Casu et al., 2004; Ataullah and Le, 2006) but the new emerging external effect of technology has not been considered within the banking sector. A second relevant strand of literature emphasises the important role of efficiency measures and specific methodologies when dealing with banks (Fethi and Pasiouras, 2010). The majority of the studies use non-parametric methods, such as Data Envelopment Analysis (DEA) (Ferrier and Lovell, 1990; Zheng et al., 2003). Parametric methods represented by the Stochastic Frontier Approach (SFA) are also common (Bauer et al., 1993; Kwan and Eisenbeis, 1996; Clark and Siems, 2002). The SFA approach defines the functional form of the cost, profit, or production function and allows for the inclusion of the inefficiency factor in the error term, but it suffers from a pre-determined form of the optimal efficiency frontier function, which may lead to bias in the efficiency ridge (Berger and Humphrey, 1991). DEA is a linear method used to estimate the efficiency of a decision unit and is based directly on a firm-specific dataset, rather than a specific functional form to define the production frontier, while its assumptions ignore the effects of random error. A third related line of work is the literature on the financial development of emerging economies. Prior literature has shown that there is a positive relationship between local financial development and bank efficiency (Claessens and Laeven, 2003; Chen et al., 2020). Financial development has proven to be an important determinant

[10] The relevant statistics are available at
http://www.caict.ac.cn/english/research/whitepapers/202112/t20211224_394512.html and
https://www.businessofapps.com/data/mobile-payments-app-market/.

of FinTech development, and the regional nature of financial development is also influential (Laidroo and Avarmaa, 2020).

This paper adds value to the literature in many ways. First, the extent to which FinTech has an impact on the efficiency of commercial Chinese banks is tested. Previous studies have shown that the banking sector is considered to play a significant role in the development of any financial system, especially those dependent on banks (Heffernan, 2005). However, very few studies explore the relationship between the banking industry and FinTech. Phan et al. (2019) find that FinTech has a negative impact on Indonesian banks by constructing the FinTech index with the counts of FinTech enterprises. Iman (2019) also finds through a series of case studies that the development of FinTech in Indonesia has increased banking competition, while Le et al. (2021) analyse the impact of FinTech credit on bank efficiency using banking data from 80 countries for a period between 2013 and 2017. The authors confirm that although there is a negative relationship between FinTech credit and bank efficiency, higher FinTech credit promotes bank efficiency. Lee et al. (2021) confirm that the FinTech innovations improve the cost efficiency of commercial banks by using principal component analysis to construct the Fintech index. Meanwhile, Wang et al. (2021) apply factor analysis to build the indicator that represents the FinTech development and they find the positive influence of FinTech in commercial banks' efficiency. Taking this one step further, the relationship is explored by applying a Digital Financial Inclusion Index (DFII) (Stage, I.I., 2019),[11] which is used in research to objectively reflect China's FinTech development (Liu et al., 2021; Wang et al., 2022). Given the index's large and robust data sources, this indicator deeply reflects the development of financial digitisation in all regions within China, and serves as a proxy indicator for FinTech adoption, enabling the study to analyse the impact of FinTech development on banks from more perspectives (Yang et al., 2022).

Second, an analysis focusing on CCBs is presented. There is evidence of an upsurge in Commercial banks learning and cooperating with FinTech in China.[12] In China, these banks dominate funds coming from the local area (Zhang et al., 2012). All CCBs are based in

[11] The DFII index is obtained from the Ant Financial Service Group database. Information on this database can be found at: https://idf.pku.edu.cn/docs/20210421101507614920.pdf
[12] Source: https://www.centralbanking.com/central-banks/economics/7835346/the-irreversible-rise-of-FinTech-in-china.https://www.centralbanking.com/central-banks/economics/7835346/the-irreversible-rise-of-FinTech-in-china

central cities, i.e., provincial or prefecture-level cities. Before 2006, when the China Banking Regulatory Commission allowed CCBs to set up branches in other provinces, they were required to operate within their own administrative regions (Sun et al., 2013). In other words, their development and operations are closely related to the local economy. As of the end of 2021, there were 125 CCBs with total assets of 45.1 trillion RMB, accounting for 13.1% of the industry. According to Ferri (2009), CCBs are a new force among Chinese commercial banks. Compared with other major commercial banks classifications in China, SOCBs and Joint-stocked Commercial Banks (JSCBs), the scale of CCBs is small, but their continuous development promotes and supports the balance of the local economy and the operation of small and medium-sized enterprises.

Third, a two-stage bootstrap-DEA-Malmquist truncated regression method is considered. DEA-Malmquist approaches are commonly used when firm-level characteristics are explored, especially in regressions incorporating efficiency estimates as dependent variables. They provide estimates of the productivity of multi-input and output decision-making units (DMUs), but they have long criticised the accuracy and robustness. In this two-stage setup, a DEA-Malmquist framework is a relatively simple way to estimate the productivity of multi-input and output DMUs (Camanho and Dyson, 2006). Then, bootstrapped truncated regressions incorporating the Simar and Wilson (2007) method are used to eliminate bias in the estimates (Odeck, 2009; Fernandes et al., 2018). Wang et al. (2021) study the impact of FinTech development on the efficiency of commercial banks by obtaining the total factor productivity through DEA-Malmquist. However, they do not apply the same DEA-truncated regression framework leading to biased estimates, while the commercial banks' total factor productivity and its decomposers are specifically analysed in this chapter. This provides a more robust and complete approach.

Fourth, this chapter also considers that the development of FinTech at city-level with different regional financial development status should have a differential impact on CCBs. The impact of regional financial development is of great significance to the business development of CCBs in China and should be considered in our paper (Marcelin and Mathur, 2014). Chinese CCBs operate mainly in certain regions to which they are affiliated. They have a wider coverage of local businesses and are more closely connected to the local economic ecology, thereby making it easier for them to obtain policy support from local governments and regulatory authorities. CCBs are also an important contributor to supporting the construction of local enterprises and the development of local finance (Chen et al., 2020).

This analysis is based on a sample of 101 Chinese commercial banks sourced from the CSMAR database, over the period 2011-2020. The FinTech development index, i.e., DFII, is obtained from the Institute of Digital Finance, Peking University, with yearly data for a period between 2011 and 2020. This study provides an interesting set of findings by employing the two-stage bootstrap-DEA-Malmquist truncated regression. One of the first thing is to show that there is a positive relationship between FinTech development and the efficiency of Chinese commercial banks. Additionally, higher levels of FinTech development also contribute to the efficiency of Chinese commercial banks. Comparing the extent to which the efficiency of different commercial banks' affiliations is affected, it is find that CCBs are more positively affected. There is more potential for the application and development of FinTech in CCBs that are smaller in original size and have a single source of financing. Further, we involve the regional financial development and find, surprisingly, that CCBs located in regions with higher levels of financial development are better able to receive the positive impact of FinTech. This demonstrates that regions in China with a more developed regional financial development have more capacity and scope to practice pushing FinTech to work in CCBs. It also serves as a reminder that less financially developed regions should apply FinTech in conjunction with their local level of financial development to maximise their advantages.

The remainder of the paper is structured as follows. The theoretical background of Chinese commercial banks and the hypotheses are presented in Section 3.2. Section 3.3 includes the methodology applied, while Section 3.4 summarises the data description. The empirical results are presented in Section 3.5, while several robustness checks are given in Section 3.6. Finally, concluding remarks are provided in Section 3.7.

3.2 Theoretical Background and Hypotheses Development

3.2.1 Background of Chinese Banking System and Commercial Banks

The Chinese commercial banking system has undergone intensive innovative and institutional reforms since 1978 (Qian, 2000). China's banking industry has gradually developed from a single banking system to a more diversified multi-level banking system (Lin and Zhang, 2009; Cousin, 2011). Initially, the banking system was comprised of only one commercial bank (i.e., the People’s Bank of China), whereas by 1996 four major commercial banks played a crucial role in its development (Fu and Heffernan, 2009). During this period, China's commercial banking business gradually improved with the requirements of economic development and financial system reform. With China's reform and opening up, JSCBs were gradually established in the 1980s and 1990s as a kind of joint-stock bank with a mixture of state-owned investment and private equity participation. Under the guidance of the open-door policy, JSCBs operate entirely in a corporate manner in accordance with market norms, and they are relatively flexible and modern banks, with the customer as the guiding force and with the customer's needs as the top management philosophy (Craig, 2005). Until the gradual establishment of JSCBs and CCBs in 1986, China Minsheng Bank was the first national JSCB invested mainly by private enterprises. Since then, the rapid development of JSCB has realized its nationwide operation. CCBs are merged, reorganized, and transformed into urban cooperative banks from various urban credit cooperatives in order to resolve their existing risks and prevent and control incremental risks. Finally, the accession of China to the World Trade Organization (WTO) in 2001 allowed the country to develop and modernise its commercial banking system (Berger et al., 2009; Zha et al., 2016).

Modern Chinese commercial banks are divided into SOCBs, JSCBs and CCBs, in accordance with their functions and politics (Zha et al., 2016). SOCBs are wholly owned by the state, currently represented by the five largest banks i.e., Bank of China, Agricultural Bank of China, Industrial and Commercial Bank of China, China Construction Bank and Bank of Communications (Wang et al., 2014). These banks were established or restructured from the pre-reform era and continue to play a pivotal role in China's financial landscape. JSCBs which represent a more market-oriented segment of China's banking system allow non-state shares to participate, giving them more customised operations and greater flexibility. Currently, there are 12 joint-stock companies in China, of which four are private and eight are state-owned (Fu and Heffernan, 2009). CCBs are also a type of JSCB, but they are also a special group due to the capital of CCBs comes from the local area. All CCBs are divided into provinces or specific priority cities, and their development and operations are closely linked to the local economy (Zhang et al., 2012). At the end of 2018, there were 134 CCBs. They play a crucial role in supporting the development of small and medium-sized enterprises (SMEs) and fostering economic growth at the local level. China's banking system, with its diverse categories of banks, reflects a blend of state control and market-oriented approaches. While SOCBs dominate the sector, JSCBs and CCBs contribute to the overall dynamism of the financial system.

3.2.2 The FinTech development in China

The development of FinTech in China has also experienced three main stages due to the reforming of China's commercial banks (Hua and Huang, 2021). In particular, the financial informatisation stage (2005-2010), the Internet finance integration stage (2011-2015), and the FinTech and Digitisation stage (2016 up until today). The first stage was marked by the development of data aggregators and online banking. The financial system combined modern information technology with the aim of reorganising the traditional financial industry. In fact, China's WTO membership created a new market competitive landscape. It established new customer relationship management and financial product innovation, and it strengthened internal banking information. The second stage was a period of rapid development for online banking in China (Chen et al., 2014). The establishment of PPDAI[13] in 2007 became a landmark event in the development of FinTech in China, as it marked the real inclusion of FinTech into the core business of finance (Kong and Loubere, 2021).[14] Individual and corporate internet banking transactions began to grow rapidly. Consequently, banks entered a new development phase where the internet and finance were formally combined.[15]

Finally, the FinTech and Digitalisation stage is characterised by state-of-the-art technology, such as AI, quantitative trading, risk management by big data and cloud services. These are applied to the core industries of the financial sector. Overall, FinTech has made its way into the Chinese banking business, not only as a transformation tool in the industry but also as a strategic goal in Chinese national financial policy.[16]

3.2.3 FinTech, Financial Development and Banking Efficiency

Research on the efficiency of commercial banks in China has been active since the late 1990s, and there has been a diversity of approaches. Ariff and Luc (2008) calculate the cost efficiency and profit efficiency of 28 Chinese commercial banks for the period 1995-2004

---

[13] PPDAI is a leading online consumer finance marketplace in China. Launched in 2007, the company is the first online consumer finance marketplace in China connecting borrowers and investors (https://ir.finvgroup.com/2019-09-26-PPDAI-Group-Inc-to-Hold-Annual-General-Meeting-on-November-5-2019).

[14] Source: https://www.sec.gov/Archives/edgar/data/1691445/000119312517309953/d285990df1.htm

[15] Source: https://www.chinatechnews.com/2011/05/27/13382-twenty-seven-chinese-companies-gain-third-party-payment-licenses

[16] Source: http://www.chinadaily.com.cn/a/201908/22/WS5d5e5ed7a310cf3e35567595.html

using a DEA-based non-parametric approach. Further, Jiang et al. (2009) measure the efficiency of Chinese commercial banks from 1995 to 2005 using the SFA approach and analyse the impact of governance change on their efficiency. Matthews et al. (2007) further investigate the state-inefficiency and non-performing loans of Chinese commercial banks by using the non-parametric bootstrapping method, which is based on Simar and Wilson's (2000) approach and was more stable than the single DEA approach. Lin et al. (2009) also follow the SFA approach to measure the X-efficiency of the Chinese banking system and conclude that smaller banks are more efficient. See and He (2015), on the other hand, use the double bootstrap DEA method to measure the efficiency of 17 banks in China and analyse the determinants of bank technical efficiency in China, drawing on the methodology of Simar and Wilson (2007).

This paper argues that there is a link between Chinese commercial banks' efficiency levels and FinTech. Our motivation stems from two main aspects. First, FinTech can help commercial banks to explore data in the context of the digital age, extracting new information from any technological improvement. As a result, banks can achieve savings in a cost-efficiency way (Zhao et al., 2019). Previous studies have provided evidence that internet technology can improve the channels and services of commercial banks, enhancing their overall efficiency levels (Bons et al., 2012; Hoehle et al., 2012). In a similar vein, Chen et al. (2017) note the scale of traditional banks actively undergoing FinTech-driven transformation, with evidence that commercial banks have become more reliant on those that have implemented FinTech initiatives. Later, China Minsheng Bank and Alibaba Group signed a strategic pact to offer new finance services, a move that slightly perturbed China's state-run banking giants. The Industrial and Commercial Bank of China set up a FinTech Research Institute aimed at driving innovation of the bank's core businesses with information technology. In 2020, Everbright Bank announced an investment in a financial innovation incubator to conduct research on FinTech development.[17]

[17] Sources for the above are provided at the following reports: http://en.srcb.com/latestnews/-qiavExmOBOvytuXkRLRk.shtml , http://en.srcb.com/latestnews/-qiavExmOBOvytuXkRLRk.shtml, https://www.chinabankingnews.com/2019/11/05/icbc-announces-officially-launch-of-FinTech-research-academy/, https://www.chinabankingnews.com/2020/04/21/china-everbright-bank-invests-500-million-yuan-in-FinTech-innovation-incubator/

Second, the use of FinTech also allows for a more sustainable financial inclusion setting. Regions in China that are comprised of less concentrated traditional bank branches should benefit more from FinTech services (Hua and Huang, 2021). This is consistent with the view that internet finance blurs the boundaries among financial institutions and further eliminates barriers, encouraging traditional banks to achieve technological innovation and improve bank efficiency (Allen et al., 2002).

Based on the above, this chapter argues that these technological updates and services should have a positive impact on the efficiency of commercial banks in China. As such, the first testable hypothesis is as follows:

Hypothesis 1: Higher levels of FinTech development have a positive impact on CB's efficiency levels.

The role of FinTech in the different commercial banks' efficiency levels is also tested. CCBs differ from SOCBs and JSCBs in terms of resource scale, property rights structure and business modeling (Lin and Zhang, 2009). CCBs are characterised by a short establishment time, a small number of outlets, and customers that are small and medium-sized privatised firms. These banks benefit from fast information transmission due to their small asset scale and concentration in specific business areas (Ferri, 2009). As a result, CCBs can quickly absorb FinTech technological changes by forming external cooperative relationships with FinTech start-ups (Li, 2020). While SOCBs comprise wider customer groups, large assets, and a higher national reputation, they suffer from high-level government jurisdiction, i.e., a lack of incentive mechanisms (Zhang, 1998). This prevents large SOCBs from absorbing in a timely matter advanced technology in FinTech (García-Herrero et al., 2009). Hence, the lack of an adhocracy system restrains these banks from investing more efficiently in innovative activities (Firth et al., 2008). Finally, the JSCBs consist of more flexible operating mechanisms, complete management systems, and diversified products. However, the literature shows that these banks' positioning is not clear cut, as there is no fixed source of customer groups, giving the joint-stock system banks lower profit efficiency than CCBs (Jiang et al., 2013). Moreover, JSCBs lack policy support from local governments, which makes them benefit less from any local FinTech policy. It is argued that the positive impact of FinTech should exert a different effect on the efficiency levels of CCBs, SOCBs and

JSCBs. The development of FinTech poses a challenge to SOCBs and JSCBs; that is, whether nationwide business services can quickly respond to the transformation and upgrading of FinTech and implement it effectively. By contrast, the smaller size and the regional concentration of CCBs have become advantageous by absorbing the positive impact of FinTech development. Based on this argument, we propose the following hypothesis:

Hypothesis 2: Higher levels of FinTech development have a stronger positive impact on the efficiency of CCBs than on their SOCBs and JSCBs counterparts.

One of the features of CCBs is that they usually perform their business within the boundaries of the city or province where they are located, i.e., one-city-one bank (Ferri, 2009). Sun et al. (2013) show that CCBs operating regionally have higher efficiency levels due to the presence of strategic investors. However, the efficiency of CCBs is negatively related to the local region's economic development, due to CCBs' governance structure and the local officials' promotion system. This provides us with a unique opportunity to explore how the regional institutional environment affects local commercial bank efficiency. Previous literature denotes that improvement of the financial market can gradually promote the development of the local economy, therefore increasing the level of financial development (Levine, 1999; Law et al., 2013). However, the development of FinTech has also created both positive and negative aspects for CCBs. While the development of FinTech can improve internal operations efficiency, it can increase competition among commercial banks in a certain region. The more developed the economy of the cities, the more intense the banking competition would be, as there is a higher likelihood of the presence of large commercial banks, limiting the innovative capacity of FinTech and leading to relatively lower efficiency of CCBs (Yao et al., 2008). On the contrary, in cities with low economic development levels, CCBs are the main local financial institutions, and their profit and return efficiency will be relatively high. In addition, local economic development is often directly correlated to the degree of regional financial development (Cheng and Degryse, 2010). Studies on financial development in banking have focused on the overall financial development of the country and neglected the influence of uneven regional financial development from the perspective of CCBs (Wu et al., 2007). Based on this, this chapter argues that the relationship between FinTech and efficiency should be higher for CCBs in less regional financially developed areas.

Hypothesis 3: FinTech exerts a positive and stronger effect on CCBs in less financially developed regions than in their more financially developed counterparts.

3.3 Methodology

To test the main hypotheses, this study takes into consideration a non-parametric method based on a technical efficiency theory first developed by Farrell (1957), that utilising a DEA approach. There are some differences in the stability and accuracy of DEA's estimated efficiency. According to Simar and Wilson (2000), the DEA model has inconsistent efficiency results when the return to scale is assumed to be constant and/or variable. Hence, the DEA efficiency estimates are biased and should not be used directly in a regression framework, as in the work of Staub et al. (2010) and Chortareas et al. (2013). To alleviate these concerns, the hypotheses are tested by constructing a two-stage model. A two-stage model with a DEA-Malmquist process and a double bootstrapped truncated regression are generated.

3.3.1 First Stage: Banking Efficiency Estimations

The DEA-Malmquist method is a method of measuring Total Factor Productivity (TFP) changes (Färe et al., 1994). It combines the efficiency measuring method by Farrell (1957) and Caves et al. (1982). The main feature of this method is that panel data analysis, among multiple DMUs, can be performed without providing information on the price of the element. A variety of DEA models have been developed to measure efficiency and capacity in different ways. These largely fall within the category of input-oriented or output-oriented models.

The input-oriented method measures the optimised efficiency when the output quantity is fixed. Conversely, the output-oriented method aims to measure the optimised efficiency when the input quantity is fixed under the same circumstances. Most scholars exploring banks' efficiency consider the input-oriented model as the preferred model, as banks tend to control input quantity (Fethi and Pasiouras, 2010). However, in our case, CCBs are smaller in scale than national banks, which means that this level of control is not ensured. As such, we apply the output-oriented method as it is more reflective of banks that focus on growth in specific areas (Paradi and Schaffnit, 2004).

According to the output-orientation DEA-Malmquist method, the production technology P^t that converts input into output during time period $t = 1 \dots T$ can be expressed as $P^t = \{(x^t, y^t): x^t can\ produce\ y^t\}$, in which we assume that there is a transformation of inputs, $x^t \in R_+^N$, into outputs, $y^t \in R_+^M$. Then, the distance function D_o can be defined by the reciprocal of the largest proportional changes of output y^t, while the input x^t is given and can be presented as follows: $D_o^t(x_i^t, y_i^t) = \inf\{\delta: (x_i^t, y_i^t/\delta) \in P^t\}, i = 1, \dots, n_t$. Among them, we use n banks as DMU, o presents the output orientation, and δ represents the output-oriented efficiency index. If y_i^t is a component of P^t, the function value will be less than or equal to one, as $D_o^t(x_i^t, y_i^t) \leq 1$; if y_i^t is on the frontier of P^t, then the function value is equal to one, $D_o^t(x_i^t, y_i^t) > 1$, which means the improvement of efficiency.

On this basis, taking the period t as a reference, the Malmquist productivity index (MPI) which uses M to define for output-orientation from period t to period $t + 1$ is:

$$M_o^t(x_i^t, y_i^t, x_i^{t+1}, y_i^{t+1}) = \frac{D_o^t(x_i^{t+1}, y_i^{t+1})}{D_o^t(x_i^t, y_i^t)} \tag{1}$$

Then, taking $t + 1$ as the reference, the MPI is presented as:

$$M_o^t(x_i^{t+1}, y_i^t, x_i^{t+1}, y_i^{t+1}) = \frac{D_o^{t+1}(x_i^{t+1}, y_i^{t+1})}{D_o^{t+1}(x_i^t, y_i^t)} \tag{2}$$

This study considers MPI as the geometric average of the above two indices:

$$M_o(x_i^{t+1}, y_i^t, x_i^{t+1}, y_i^{t+1}) = \left[\frac{D_o^t(x_i^{t+1}, y_i^{t+1})}{D_o^t(x_i^t, y_i^t)} * \frac{D_o^{t+1}(x_i^{t+1}, y_i^{t+1})}{D_o^{t+1}(x_i^t, y_i^t)}\right]^{\frac{1}{2}} \tag{3}$$

Accordingly, the M_o, which also represents the total factor productivity change (TFPCH), could be decomposed into the product of efficiency change (EFFCH) and technology change (TECHCH):

$$M_o(x_i^{t+1}, y_i^t, x_i^{t+1}, y_i^{t+1}) = \frac{D_o^t(x_i^{t+1}, y_i^{t+1})}{D_o^t(x_i^t, y_i^t)} * \left[\frac{D_o^t(x_i^{t+1}, y_i^{t+1})}{D_o^{t+1}(x_i^{t+1}, y_i^{t+1})} * \frac{D_o^t(x_i^{t+1}, y_i^{t+1})}{D_o^{t+1}(x_i^t, y_i^t)}\right]^{\frac{1}{2}} \tag{4}$$

$$= EFFCH * TECHCH \tag{5}$$

$$EFFCH = \frac{D_o^t(x_i^{t+1}, y_i^{t+1})}{D_o^t(x_i^t, y_i^t)} \tag{6}$$

$$TECHCH = \left[\frac{D_o^t(x_i^{t+1}, y_i^{t+1})}{D_o^{t+1}(x_i^{t+1}, y_i^{t+1})} * \frac{D_o^t(x_i^{t+1}, y_i^{t+1})}{D_o^{t+1}(x_i^t, y_i^t)}\right]^{\frac{1}{2}} \tag{7}$$

According to Färe et al. (1994), the efficiency change index can be decomposed into pure technical efficiency change index (PECH) and scale efficiency change index (SECH); in other words, the technical efficiency improvement index based on the constant returns to scale (CRS) method is decomposed into PECH and SECH under the variable returns to scale (VRS) method.

$$PECH = \frac{D_{ov}^t(x_i^{t+1}, y_i^{t+1})}{D_{ov}^t(x_i^t, y_i^t)} \tag{8}$$

$$SECH = \frac{D_{ov}^t(x_i^{t+1}, y_i^{t+1})}{D_{ov}^{t+1}(x_i^{t+1}, y_i^{t+1})} * \frac{D_{oc}^t(x_i^{t+1}, y_i^{t+1})}{D_{oc}^{t+1}(x_i^t, y_i^t)} \tag{9}$$

In the above formula, the added subscripts $\boldsymbol{v}$ and $\boldsymbol{c}$ correspond to the technical efficiency change index of VRS and CRS, respectively.

3.3.2 Second Stage: Double Bootstrapped Truncated Regression

At this stage, in view of the error problem of the DEA-Malmquist method, we introduce the bootstrap method to correct bias. The revision and improvement of the traditional DEA-Malmquist model by Simar and Wilson (1999) made the results more robust. To test our hypotheses, the efficiency changes are the dependent variable in the second stage. This study is interested in an output-oriented model, thus, the specific model is built as follows:

$$\hat{S}_i^t = \alpha + \beta z_i^t + \epsilon_i \geq 1 \tag{10}$$

Where $\boldsymbol{\hat{S}_i^t}$ represents the Malmquist efficiency score and its decomposition scores, $\boldsymbol{z_i}$ is the vector of explanatory variables, $\boldsymbol{\alpha}$ is a constant, $\boldsymbol{\beta}$ is a vector of parameters to be estimated and $\boldsymbol{\epsilon_i}$ is an error term that is independent of $\boldsymbol{z_i}$, $\boldsymbol{\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)}$. We implement a bootstrap method to avoid the interference of random error accumulation and other factors in the model. Repeated sampling and estimation are applied and aim to achieve the highest convergence of the estimated parameters to the real ones. To achieve this, the experiment follows the steps of the bootstrap truncated regression model of Simar & Wilson (2007):

Step 1. Using the original data in P^t for each period $t(t = 1, \dots, T)$, we compute the $\hat{S}_i^t$ using the Malmquist model (3)-(9).

Step 2. Obtain the estimates $\hat{\alpha}$, $\hat{\beta}$ and $\widehat{\sigma_\epsilon^2}$ of α, β and σ_ϵ^2 in the truncated regression of $\hat{S}_i^t$ on z_i^t through applying maximum likelihood based on the (10).

Step 3. Loop over 3.1-3.4 L_1 times to obtain n set of bias-corrected estimates $B_i^t = \{\hat{S}_{ib}^{t*}\}_{b=1}^{L_1}$:

Step 3.1. Draw $\hat{\epsilon}_{ib}^t$ from $\mathcal{N}(0, \sigma_\epsilon^2)$ distributed on the left truncation of $\left(1 - \hat{\alpha} - z_i^t \hat{\beta}\right)$.

Step 3.2. Compute the $\hat{S}_{ib}^{t*} = \hat{\alpha} + \hat{\beta} z_i^t + \hat{\epsilon}_{ib}^t$ again for each $i = 1, .., n^t$.

Step 3.3. Set $x_{ib}^{t*} = x_i^t$, $y_{ib}^{t*} = \left(\frac{S_i^t}{\widehat{E_{ib}^{t*}}}\right) * y_i^t$ and $z_{ib}^{t*} = z_i^t$ for $i = 1, \dots, n^t$.

Step 3.4. Estimate the $\widehat{S_{ib}^{t*}}$ using the bootstrapped y_{jb}^{t*} and x_{jb}^{t*} instead of y_j^t and x_j^t.

Step 4. Compute the bias-corrected scores $\hat{\hat{S}}_i^t$ for each $j = 1, \dots, n^t$, which is defined by $\hat{\hat{S}}_i^t = \hat{S}^{t_i} - BIAS\left(\hat{S}^{t_i}\right)$ where $BIAS\left(\widehat{S_i^t}\right)$ represent the bootstrap estimator of bias in B_i^t obtained in step 3.

Step 5. Again, employ the maximum likelihood method to estimate the truncated regression of $\hat{\hat{S}}_i^t$ on z_i^t yielding new estimators of regression which defined as $\hat{\hat{\alpha}}$, $\hat{\hat{\beta}}$ and $\widehat{\widehat{\sigma_\epsilon}}$.

Step 6. Loop over 6.1-6.3 L_2 times to obtain a set of bootstrapped estimates $\{(\widehat{\alpha^*}, \hat{\beta}^*, \widehat{\sigma_\epsilon^*})_b\}_{b=1}^{L_2}$:

Step 6.1. Draw $\hat{\hat{\epsilon}}_{ib}^t$ from $\mathcal{N}\left(0, \hat{\hat{\sigma}}_\epsilon^2\right)$ for each $i = 1, \dots, n^t$ with truncation on the left side at $\left(1 - \hat{\hat{\alpha}} - z_i^t \hat{\hat{\beta}}\right)$.

Step 6.2. Compute the $\hat{S}_{ib}^{t**} = \hat{\hat{\alpha}} + \hat{\hat{\beta}} z_i^t + \hat{\epsilon}_{ib}^t$ again for each $i = 1, \dots, n^t$.

Step 6.3. Apply the maximum likelihood method to estimate the truncated regression of S_{ib}^{t**} on z_i^t yielding estimators $\hat{\hat{\alpha}}^*, \hat{\hat{\beta}}^*$ and $\hat{\hat{\sigma}}_\epsilon^*$.

Step 7. Construct bootstrapped confidence intervals for α, β and σ_ϵ using the bootstrapping value $\hat{\hat{\alpha}}, \hat{\hat{\beta}}$ and $\widehat{\widehat{\sigma_\epsilon}}$.

At the end of this procedure, the set of biased-corrected efficiency estimates is obtained $\hat{\hat{S}}_i^t = E_{it}$, which will be used as in the truncated regression[18].

3.3.3 Final Double Bootstrapped Truncated Regression Specification

To investigate the impact of FinTech development on the efficiency of Chinese commercial banks, this chapter follows the previous literature (Chang et al., 2012; Du et al., 2018) and defines our model as follows:

$$E_{it} = \beta_0 + \beta_1 E_{i,t-1} + \beta_2 FI_{it-1} + \sum \delta_i Macrolevel_{it} + \sum \gamma_i Microlevel_{it} + \epsilon_{it} \qquad (11)$$

Where E_{it} represents the Malmquist efficiency score and the decomposition bootstrapped bias-corrected efficiency scores; FI_{it-1} indicates the lagged FinTech index measured as the logarithm of the aggregated development level of FinTech ; $Macrolevel_{it}$ denotes the macroeconomic variables while $Microlevel_{it}$ stands for the microeconomic variables; i identifies the bank; t is the time component; and ϵ is the random error. The analysis also controls the possibility of endogeneity (Taylor, 1999; Rossi et al., 2009) In particular, the lag value of efficiency is included ($E_{i,t-1}$). as the efficiency level of a bank at time *t* depends on the effect of efficiency from the previous period (*t-1*).

The study hypothesises that higher levels of FinTech development (FI_{it-1}) increase the efficiency levels of commercial banks' (hypothesis 1). Hence, it expects β_2 to exert a positive effect on banks' efficiency.

[18] The above procedure is implemented in R using the FEAR packaged for estimating the double bootstrapped truncated regression by Wilson (2008).

In addition to the main variable of interest, a set of financial variables are included in our model specification. Banks' size, measured as the total asset of banks, is used to control for the risk level and financing access preferences (Kishan and Opiela, 2000). Past literature refers that larger commercial banks should have lower efficiency levels as they are more difficult to manage due to they are less contributed to volatility connectedness (Wang et al., 2018; Fang et al., 2019). Hence, we expect a negative impact of size on the efficiency levels of banks.

The loan-to-deposit ratio (LDR) is also applied as an indicator of resource allocation for commercial banks. According to Van den End (2016), higher LDR reflects higher capital usage, therefore, increasing the efficiency of resource allocation. As such, The analysis expects the coefficient of LDR to be positive. The variable Return on Assets (ROA) is used as a proxy for firms' profitability performance (Bonin et al., 2005). Higher levels of ROA indicate higher profitability levels (Ozili and Uadiale, 2017). Hence, the profitability is expected to exert a positive effect on the efficiency levels of banks.

In addition, the Capital Adequacy Ratio (CAR) is used a proxy for risk tolerance level. It is generally accepted that a higher CAR reflects a bank's greater resilience to risk contributing to its efficiency (Pessarossi and Weill, 2015). This chapter expects a positive impact of CAR on the efficiency changes. Finally, IPO is a dummy variable equal to 1 if the banks are listed and 0, otherwise. Listing helps commercial banks to expand the capital source and strengthen the market discipline, thereby standardising internal management and improving the efficiency of banks (Alqahtani et al., 2017). As such, IPO should have a positive effect on efficiency levels.

At a macro level, the real Gross Domestic Product (GDP) growth is included to reflect the development of the economy. Better economic situations promote loans and deposits of businesses for commercial banks (Behr et al., 2017). As such, it expects a positive relationship with banks' efficiency. The Consumer Price Index (CPI) is employed to measure inflation (Lozano-Vivas and Humphrey, 2002; Lensink et al., 2008). The higher inflation level is considered to have a positive impact on the profitability of banks in China (Tan and Floros, 2012). As such, it expects a positive effect on efficiency levels. Finally, the proportion of assets of the five major banks in China is used as an indicator of industry concentration. According to Semih Yildirim and Philippatos (2007), industry concentration should negatively affect the efficiency of banks as the lack of competition in the highly concentrated structure of the banking system.

Next, this chapter investigates whether the relationship between banks' efficiency levels and FinTech differs across different bank groups (i.e., CCBs, SOCBs, and JSCBs). Corporate social responsibility activities and business survival vary among the bank groups and there are differences in their treatments of FinTech in terms of acceptance and policy guidelines. In particular, CCBs have smaller capitalisation but more targeted business compared to SOCBs and JSCBs, incentives under the local government and the imposition of FinTech have greatly expanded original business. As such, it is argue that the impact of FI on efficiency should be higher for CCBs. Equation (12) is augmented with a dummy variable representing each bank type. Specifically, $SOCB_{it}$ is a dummy variable equal to 1, and 0 otherwise while $JSCB_{it}(CCB_{it})$ is a dummy variable that takes the value of 1, and 0 otherwise. $SOCB_{it}$ is omitted in the following equation as the benchmark group (Luo et al., 2017).

$$\begin{aligned} E_{it} = \beta_0 + \beta_1 E_{i,t-1} + \beta_2 FI_{it-1} + \beta_3 JSCB_{it} + \beta_4 CCB_{it} \\ + \beta_5 JSCB_{it} * FI_{it-1} + \beta_6 CCB_{it} * FI_{it-1} \\ + \sum \delta_i Macrolevel_{it} + \sum \gamma_i Microlevel_{it} + \epsilon_{it} \end{aligned} \tag{12}$$

If the hypothesis stands, then the impact of FinTech development on efficiency is expected to be higher for CCBs than for the commercial banks counterparts.

Finally, this study aims to explore the impact of regional financial development on CCBs' absorption of FinTech. The motivation for this analysis is due to the importance of geographical impact on CCBs as the success of CCBs is systematically and positively correlated to their economic region (Cai et al., 2016; Li and Song, 2021).

To focus on the specific role of FinTech and financial development for CCBs we follow the past literature (Ariff and Luc, 2008) and remove data from other commercial banks. As a proxy for financial development in China, the ratio of regional total loans to GDP is applied (Yuxiang and Chen, 2011; Xu, 2012). This variable measures the depth of the financial sector (Ljungwall and Li, 2007).

Furthermore, the province-based sample in the study is also classified into more and less financially developed ones according to the average ratio of total loans to GDP. We argue that CCBs located in more financially developed than average levels have more potential to improve efficiency due to the FinTech application. Hence, we include a dummy variable FD_{it-1} which is equal to 1 if the Chinese provinces are above the average ratio of total loans to GDP, and 0 otherwise. The specific regression is shown as follow:

$$E_{it} = \beta_0 + \beta_1 E_{i,t-1} + \beta_2 * FI_{it-1} + \beta_3 * FD_{it-1} + \beta_4 * FD_{it-1} * FI_{it-1}$$

$$+\sum \delta_i Macrolevel_{it} + \sum \gamma_i Microlevel_{it} + \epsilon_{it} \quad (13)$$

This chapter expects the effect of financial development on efficiency to be stronger for less financially developed provinces than their more developed counterparts.

3.3.4 Data Description: DEA Inputs and Outputs, FinTech Index and Control Variables

To construct the database, this study uses annual reports from the Chinese Stock Market & Accounting Research Database (CSMAR), which is a leading source of the Chinese financial data, published by Shenzhen CSMAR Data Technology Co., Ltd. CSMAR provides us with consolidated financial statements on listed commercial banks. In China, there are also many unlisted commercial banks, these include smaller regional banks and some policy banks which are not publicly traded. The main difference between listed and unlisted commercial banks are the regulatory management (Chen et al., 2021). Listed commercial banks provide more transparency and prudential management than unlisted commercial banks (Xu et al., 2018). So only listed commercial banks are considered in this chapter because they are influenced by market forces, and their statement are reflective and objective, providing convincing variables for analysis. Further, information on GDP and CPI is obtained from the Chinese Statistic Yearbook[19]. This dataset covers the period between 2011 and 2020 starting from the era of Internet finance intergration stage, focusing on Chinese commercial banks across 31 provinces (i.e., only autonomseous regions and province-level municipal cities are included[20]).

[19] China Statistical Yearbook is an annual statistical publication which comprehensively reflects the economic and social development of China.
[20] There are 34 such divisions claimed by the People's Republic of China, classified as twenty-two provinces, five autonomous regions, four municipalities, two special administrative regions (Hong Kong and Macau) and one claimed region (Taiwan). Normally, 31 provinces directly under the central government are counted as the observations when doing relevant research (Hasan *et al*., 2009; Gao *et al*., 2020).

Observations with negative sales and assets are removed and firms that do not have complete records on the variables are droped in the regressions. Following the standard practice in the literature (Brown et al., 2009), we control for the potential influence of outliers by excluding observations in the 1% tails of each of the regression variables. Commercial banks with missing data are also removed. The final balance panel consists of 909 bank-year observations with 101 Chinese commercial banks (i.e., 5 large SOCBs, 10 JSCBs and 86 CCBs). To measure the Malmquist estimators, we consider that banks act as intermediaries for the flow of funds. As such, input and output indicators are selected based on banks' balance sheets (Casu et al., 2004; Tortosa-Ausina et al., 2008). As such, the efficiency of Chinese commercial banks whose profit mainly comes from the deposit and loan spread is estimated.[21] Thus, this expiriment selects the following as input variables: deposit, interest expense, non-interest expense (including other operating fees) and labour cost (measured by employee payable). The output variables are loans, interest income (including fee and commission income) and non-interest income.

In terms of the second stage double bootstrapped truncated regression, the DFII based on the past literature is conducted (Hua and Huang, 2020; Sheng, 2020). In the last decade, China has emerged as a leading role in FinTech adoption and digitalization, revolutionizing traditional financial services and paving the way for a cashless society. Chinese FinTech has experienced exponential growth, fuelled by a dynamic blend of technological innovation and a burgeoning middle class eager for convenient financial solutions. Companies such as Ant Group, Tencent, and JD Finance have played pivotal roles in pioneering novel fintech services, from mobile payments and wealth management to peer-to-peer lending. China's transition to a cashless society is epitomized by the widespread adoption of mobile payment platforms like Alipay and WeChat Pay. These digital wallets have seamlessly integrated into daily life, facilitating transactions for everything from groceries to transportation.

To characteristic the development level of FinTech adoption, we use the Peking University DFII. The experiment tests hypotheses using data from the Ant Financial Service, which belongs to Alibaba's affiliate and owns the largest customer groups of FinTech services in China. This is an annual indicator based on 31 provinces, compiled through a hierarchical analysis of three dimensions, i.e., the breadth of digital financial coverage, the depth of

[21] Previous literature (Xiaogang et al., 2005; Das and Ghosh, 2006) follows the financial intermediation approach of Sealey Jr and Lindley (1997), according to which bank deposits are used as an input.

digital financial usage, and the degree of digitalisation of inclusive financial services. Table 3.1 shows the index system of DFII in Guo et al. (2020).

Table 3.1 Index system of DFII

Level I	Level II		Specific Indicators
Breadth of Coverage	Account coverage rate		Number of Alipay accounts owned per 10,000 people
			Proportion of Alipay users who have bank cards bound to their Alipay accounts
			Average number of bank cards bound to each Alipay account
Depth of Usage	Payment		Number of payments per capita
			Amount of payments per capita
			Proportion of number of high-frequency active users (50 times or more each year) to number of users with a frequency at least once each year
	Money Funds		Number of Yu'ebao purchases per capita
			Amount of Yu'ebao purchases per capita
			Number of people who have purchased Yu'ebao per 10,000 Alipay users
	Credit	Individual User	Number of users with an internet loan for consumption per 10,000 adult Alipay users
			Number of loans per capita
			Total amount of loans per capita
		Small & Micro Businesses	Number of users with an Internet loan for small & micro businesses per 10,000 adult Alipay users
			Number of loans per small & micro business Average amount of loan among small & micro businesses
	Insurance		Number of insured users per 10,000 Alipay users
			Number of insurance policies per capita
			Average insurance amount per capita
	Investment		Number of people engaged in Internet investment and money management Per 10,000 Alipay users
			Number of investments per capita
			Average investment amount per capita
	Credit Investigation		Number of credit investigations by natural persons per capita
			Number of users with access to credit-based livelihood services (including finance, accommodations, mobility, social contacts, etc.) per 10,000 Alipay users
Level of Digitalization	Mobility		Proportion of number of mobile payments
			Proportion of total amount of mobile payments
	Affordability		Average loan interest rate for small & micro businesses

		Average loan interest rate for individuals
	Credit	Proportion of number of Ant Check Later payments
		Proportion of total amount Ant Check Later payments
		Proportion of number of “Zhima Credit as deposit” cases (to number of full deposit cases)
		Proportion of total amount of “Zhima Credit as deposit” (to amount of full deposit)
	Convenience	Proportion of number of QR code payments by users
		Proportion of “average amount” or “total amount” of QR code payments by users

Based on the platform of Alibaba owning over 400 million users, the DFII systematically displays the FinTech inclusion degree from different financials filed in China. Hence, the DFII is introduced as an indicator for FinTech development. In addition, and as a robustness check of main research results, we include the FinTech index from three different dimensions are also included, which are provided by the DFII, i.e., the coverage and usage breadth and the digitisation level. The construction of the whole index system is divided into two steps. Firstly, the specific indexes are used to construct the second-level index after being nondimensionalised according to the logarithmic efficacy function, and the weights are assigned according to the coefficient of variation. Second, the weights of the second-level indexes are assigned to the first-level indexes using hierarchical analysis (AHP), as well as the weights of the first-level indexes are assigned to the total digital inclusive finance index. Table 3.2 shows DFII in China across province in 2011 and 2020. In 2011, the average Digital Financial Inclusion Index (DFII) stood at 40.00. Over the subsequent nine years, there was a remarkable surge, with the index sharply escalating to 341.22 by 2020. This significant upswing serves as a clear indicator of the rapid and robust development of FinTech applications in China during that period. In addition, while eastern China leads in FinTech adoption, the middle and western regions are rapidly closing the gap (Yang and Zhang, 2022). This trend signifies a nationwide movement towards financial inclusion, extending FinTech benefits to regional commercial banks and consumers who were previously excluded.

Table 3.2 Province level DFII in China

Province	DFII		Province	DFII	
	2011	2020		2011	2020
Average	40.00	341.22	Henan	28.40	340.81
Beijing	79.41	417.88	Hubei	39.82	358.64
Tianjin	60.58	361.46	Hunan	32.68	332.03
Hebei	32.42	322.70	Guangdong	69.48	379.53
Shanxi	33.41	325.73	Guangxi	33.89	325.17
Inner Mongolia	28.89	309.39	Hainan	45.56	344.05
Liaoning	43.29	326.29	Chongqing	42.89	344.76
Jilin	24.51	308.26	Sichuan	40.16	334.82
Heilongjiang	33.58	306.08	Guizhou	18.47	307.94
Shanghai	80.19	431.93	Yunnan	24.91	318.48
Jiangsu	62.08	381.61	Tibet	16.22	310.53
Zhejiang	77.39	406.88	Shaanxi	40.96	342.04
Anhui	33.07	350.16	Gansu	18.84	305.50
Fujian	61.76	380.13	Qinghai	18.33	298.23
Jiangxi	29.74	340.61	Ningxia	31.31	310.02
Shandong	38.55	347.81	Xinjiang	20.34	308.35

Then, regarding the calculations of commercial banks' efficiency, the DEA-Malmquist method is used in the first stage. Table 3.3 provides the descriptive statistics of the input and output variables for DEA-Malmquist method.

Table 3.3 Summary of input & output

	Variables	Observations	Mean	Std. Dev.	Min	Max
Input	Total deposit	1010	1422.77	3652.62	7.22	25134.73
	Interest rate cost	1010	37.80	77.47	0.04	445.76
	Non-interest rate cost	1010	1.19	3.06	0.00	19.74
	Labor expense	1010	3.76	8.68	0.00	56.81
Output	Total loan	1010	1079.62	2718.29	3.68	18624.31
	Interest rate income	1010	80.11	179.27	0.15	1092.52
	Non-interest rate income	1010	12.11	29.11	0.00	171.64

Note: This table presents the descriptive statistics for input and output variables during the period 2011-2020 for DEA-Malmquist at first stage (all in billion RMB). All variables are collected from CSMAR.

All data are in CNY billions. As we can observe from Table 3.3, there is a large profitability gap among different banks, and the total loan dominates banks' operations. This is consistent with the view of Dong et al. (2016), according to which the degree of development and operation of each type of bank varies considerably.

Table 3.4 provides the definition and summary of all variables applied in the second stage. This result shows that CCBs are characterised by the highest average growth rate of total assets and average ROA when compared with other bank groups. CCBs also show an increasing expansion and profitability indicated by the highest average growth rate of size and ROA, which is consistent with Ferri (2009). However, CCBs show the lowest average LDR, while the JOCB has the largest LDR. These preliminary results suggest that JOCBs have even better resource allocation capabilities than their SOCBs counterparts. Moreover, the SOCBs' average capital adequacy ratio is also the highest. This is consistent with the idea that such bank groups have an advantage in terms of size and national policy support, providing them with a stronger risk tolerance (Li et al., 2001).

Table 3.4 Descriptive statistic

Banks	Variables	SIZE	LDR	ROA	CAR	IPO	GDPg	CPI	IP	FI	FCB	FUD
All banks	Obs	909	909	909	909	909	909	909	909	909	909	909
	Mean	0.168	0.430	0.205	8.602	0.286	0.077	102.465	0.396	208.039	186.271	202.086
	St. Dev	0.178	0.340	0.294	6.699	0.452	0.020	0.136	0.047	94.218	91.598	90.473
	Min	-0.789	-0.350	-0.470	-0.154	0.000	0.052	102.271	0.323	40.621	34.278	49.933
	Max	1.891	1.130	2.500	59.610	1.000	0.115	102.679	0.473	354.645	327.235	338.046
SOCB	Obs	45	45	45	45	45	45	45	45	45	45	45
	Mean	0.103	0.723	0.031	13.805	1.000	0.077	102.465	0.396	207.732	187.285	202.086
	St. Dev	0.050	0.072	0.005	1.631	0.000	0.021	0.138	0.048	94.706	93.882	90.473
	Min	0.046	0.585	0.024	10.830	1.000	0.052	102.271	0.323	40.621	34.278	49.933
	Max	0.350	0.916	0.043	17.520	1.000	0.115	102.679	0.473	341.220	327.235	338.046
JSCB	Obs	90	90	90	90	90	90	90	90	90	90	90
	Mean	0.125	0.821	0.035	12.040	1.000	0.077	102.465	0.396	207.732	187.285	202.086
	St. Dev	0.131	0.126	0.004	1.304	0.000	0.020	0.137	0.048	94.173	93.353	90.473
	Min	-0.789	0.553	0.026	9.000	1.000	0.052	102.271	0.323	40.621	34.278	49.933
	Max	0.511	1.130	0.048	15.680	1.000	0.115	102.679	0.473	341.220	327.235	338.046
CCB	Obs	774	774	774	774	774	774	774	774	774	774	774
	Mean	0.176	0.372	0.233	7.930	0.167	0.077	102.465	0.396	208.090	186.101	201.994
	St. Dev	0.186	0.326	0.308	6.983	0.373	0.020	0.136	0.047	94.313	91.389	90.369
	Min	-0.774	-0.350	-0.470	-0.154	0.000	0.052	102.271	0.323	40.621	34.278	46.933
	Max	1.891	1.000	2.500	59.610	1.000	0.115	102.679	0.473	354.645	327.235	338.046

Note: This table presents the descriptive statistic for independent variables from 2011-2020 for the second truncated regression. SIZE is calculated as the total asset growth rate. LDR is the loan-deposit ratio. ROA represent the profitability of commercial banks. CAR is the capital adequacy ratio which represent the operation conditions of commercial banks. IPO is the dummy variable of whether listed. GDPg is national annual GDP growth rate. CPI is the national customer purchase index of current period. IP represent the industry concentration which is calculate by the growth rate of total asset of five large commercial banks in industry. FI is the logarithm of lagged FinTech index which collected from the Pecking University Digital Financial Inclusion Index. FCB and FUD are the index of FinTech coverage breadth and the FinTech usage depth separately which are applied as the institution variables for robustness check. We only report 9 periods of the observation (e.g., 2011-2012 as one period) as one year is loss for the calculation of the original MPI score

3.4 Empirical Results

In this section, the main hypotheses is tested in Section 3.4. Specifically, the role of FinTech on the different decomposition of commercial banks is tested using the DEA-Malmquist efficiency scores (Section 3.4.1) and the Simar and Wilson (2007) truncated regression analysis is applied (Section 3.4.2). In Section 3.6, it explores the role of different dimensions of FinTech and potential endogeneity problems.

3.4.1 First Stage: Banking Efficiency Estimations

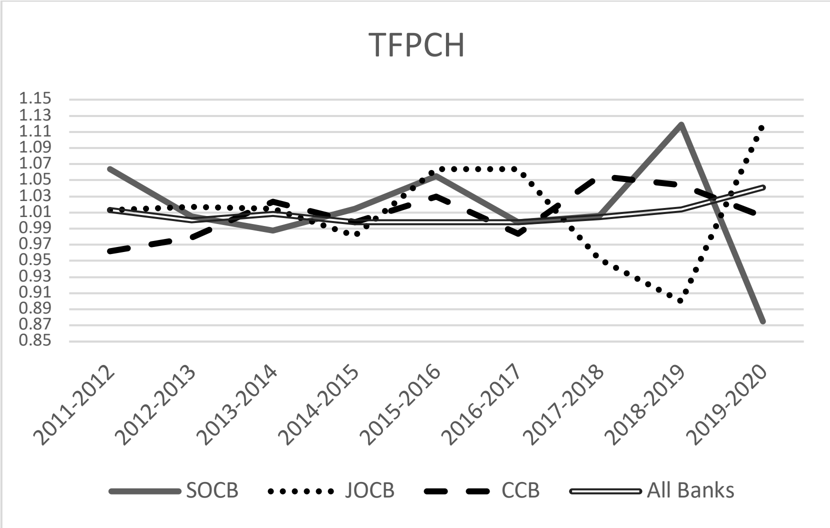
A basic premise of this study is that FinTech development has a positive effect on Chinese banks' efficiency levels. To assess this claim, we begin by estimating the DEA-Malmquist efficiency scores and respective TFPCH decomposition of commercial banks (i.e., EFFCH, TECHCH, PURE and SCALE). This decomposition is important for our analysis as it supplies essential information on internal technological changes (Isik and Hassan, 2003). Table 3.5 reports the Malmquist decomposition yearly scores, obtained using the (DEAP) 2.1 software. Overall, the average total factor productivity change is 1.008 during the sample period (2011-2020). This suggests that the TFPCH of commercial banks in China improves at an annual growth rate of 0.8%. When we consider the different TFPCH decomposition of commercial banks, the average productivity change of SOCBs (1.4%) is the same as that of JSCBs (1.4%) and CCBs (0.9%). From the aspect of decomposition, the TECHCH mainly contributes to productivity growth, which increases by 1.8% on average for all samples of commercial banks. This result is consistent with our expectations, as past studies have shown that Chinese commercial banks have increased their technological research, development efforts, and management practices, reflecting strong technological progress. For SOCBs, JSCBs and CCBs the average technology change is 2.3%, -0.2% and 1.5%, respectively. In terms of PURE and SCALE, we find that the average PURE remains unchanged during this period whilst SCALE also decreases by 0.3%. This result further illustrates that scale expansion does not bring further total factor productivity improvements to banks, and it even tends to saturate, leading to decreasing efficiency of scale (Duncan *et al.*, 2004). Figure 3.1 provides a visual account of the different Chinese commercial banks' efficiency scores, confirming the previous findings. Specifically, during the 2013-2016 period, the average TFPCH for all commercial banks presents an increasing trend with a slight drop afterwards (2016-2018). However, for the period between 2018 and 2020, the changes in the efficiency levels of the different types of banks are more volatile. To be specific, the TFPCH of SOCBs reaches a peak during 2018-2019, whereas the JSCBs has the lowest TFPCH. Interestingly, the figure indicates that during the 2019-2020 period, the TFPCH of SOCBs and JSCBs presents the opposite behaviour. In addition, the TFPCH of CCBs is relatively stable but also decreases during this period. These preliminary results may be explained by the macroeconomic environment during the 2019-2020 period, with a worldwide pandemic and a combination of the cyclical and structural slowdown of the Chinese economy (Bagchi et al., 2020). Finally, the TECHCH is found to has a similar pattern to the TFPCH, especially from 2017 onwards. This may be explained by the policy support for FinTech, with cooperation between banks and technological firms during the 2017-2020 period (Wang et al., 2021).

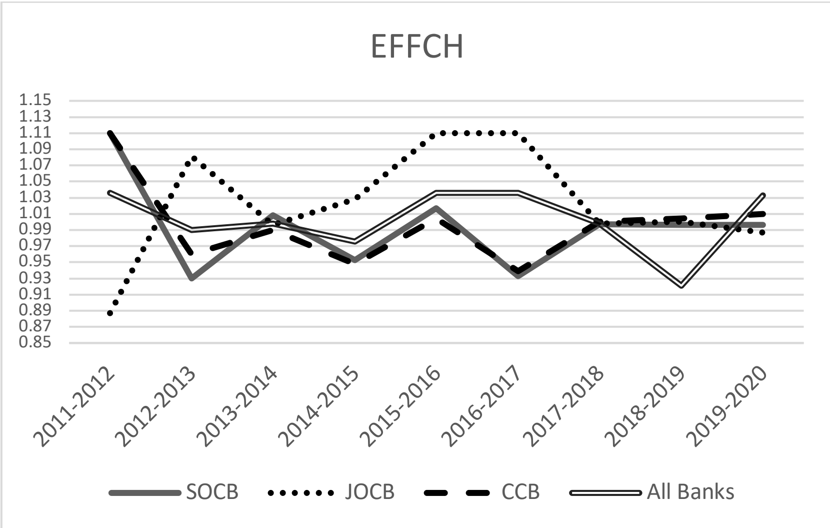
Table 3.5 First stage DEA-Malmquist efficiency scores

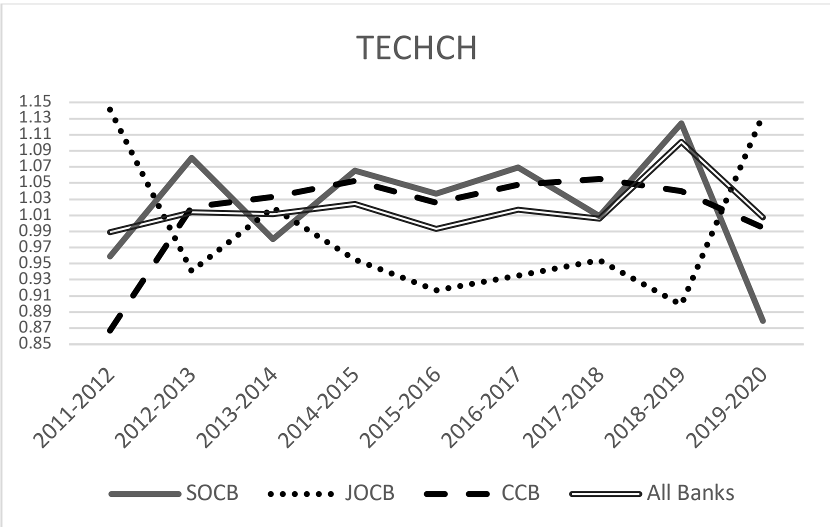
Banks	Year	TFPCH	EFFCH	TECHCH	PURE	SCALE
SOCB	2011-2012	1.064	1.110	0.959	1.000	1.110
	2012-2013	1.005	0.930	1.081	1.000	0.930
	2013-2014	0.988	1.008	0.980	1.000	1.008
	2014-2015	1.015	0.953	1.065	1.000	0.953
	2015-2016	1.055	1.017	1.037	0.999	1.018
	2016-2017	0.998	0.933	1.069	0.925	1.009
	2017-2018	1.006	0.997	1.009	1.000	0.997
	2018-2019	1.119	0.996	1.124	0.996	1.000
	2019-2020	0.875	0.996	0.879	1.000	0.996
	2011-2020	1.014	0.993	1.023	0.991	1.002
JSCB	2011-2012	1.013	0.887	1.141	0.936	0.948
	2012-2013	1.017	1.081	0.941	1.054	1.025
	2013-2014	1.015	0.996	1.019	0.999	0.997
	2014-2015	0.981	1.028	0.955	0.997	1.031
	2015-2016	1.064	1.110	0.917	1.050	0.982
	2016-2017	1.064	1.110	0.935	0.937	1.062
	2017-2018	0.952	0.998	0.954	1.009	0.989
	2018-2019	0.900	1.000	0.899	1.000	1.000
	2019-2020	1.119	0.987	1.134	0.999	0.989
	2011-2020	1.014	1.022	0.988	0.998	1.003
CCB	2011-2012	0.962	1.110	0.867	1.101	1.008
	2012-2013	0.979	0.960	1.020	0.928	1.034
	2013-2014	1.023	0.990	1.033	1.002	0.988
	2014-2015	0.998	0.948	1.053	0.974	0.973
	2015-2016	1.030	1.004	1.026	0.906	1.108
	2016-2017	0.984	0.939	1.048	1.017	0.923
	2017-2018	1.055	1.000	1.055	1.002	0.998
	2018-2019	1.044	1.004	1.040	1.012	0.992
	2019-2020	1.005	1.010	0.995	1.025	0.985
	2011-2020	1.009	0.996	1.015	0.996	1.001
All banks	2011-2012	1.013	1.036	0.989	1.012	1.022
	2012-2013	1.000	0.990	1.014	0.994	0.996
	2013-2014	1.008	0.998	1.011	1.000	0.998
	2014-2015	0.998	0.976	1.024	0.990	0.986
	2015-2016	0.998	1.036	0.993	0.985	1.022
	2016-2017	0.998	1.036	1.017	1.012	0.998
	2017-2018	1.004	0.998	1.006	1.004	0.995
	2018-2019	1.014	0.921	1.101	0.954	0.965
	2019-2020	1.041	1.033	1.007	1.046	0.987
	2011-2020	1.008	1.003	1.018	1.000	0.997

Note: This table presents the yearly efficiency result and the Malmquist decomposition of first stage by the DEA-Malmquist method. The results during the 2011-2020 are presented by the mean of each individual commercial bank during the period. The efficiency results based on different types commercial banks are reported separately.

Figure 3.1 Efficiency of different Chinese commercial banks







Note: The figure present the TFPCH, EFFCH and TECHCH evolution of different types of commercial banks for all years.

Overall, the above results show that the efficiency changes of SOCBs and JSCBs are more volatile, while the efficiency changes of CCBs are relatively more stable. These results are consistent with the work of Zha et al. (2016). Given the above, this study is interested in further exploring the relationship between FinTech development and efficiency changes in Chinese commercial banks. The results and analysis based on the hypotheses are presented in the next section.

3.4.2 The Role of FinTech Development

Table 3.6 presents the estimated results on the baseline model by using the double bootstrapped truncated regression technique (Equation 11). The results in Columns 1-5 reflect the decomposition of banks' efficiency measure (i.e., TFPCH, EFFCH, TECHCH, PURE and SCALE in Columns 1, 2, 3, 4 and 5, respectively). The findings reveal that the higher level of FinTech development (FI) does exert a significant and positive impact on each component of bank efficiency. This implies that FinTech can be instrumental to enhancing commercial banks' efficiency levels. This finding is significant, not only statistically but also economically. For instance, let us consider the effect of FI on TFPCH measure (Column 1). The coefficient of 0.061 on the FI variable implies that FinTech development increases the productivity of commercial banks by 6.1%. Our results show that, independent of the decomposition of bank efficiency used, the coefficient of FI is positive and statistically significant. These results consistent which show the important role of FinTech development on banks' efficiency. This is in line with past studies demonstrating the positive role of FinTech on banks' performance (Zhao et al., 2022; Leong et al., 2017; Fuster et al., 2019). These results are also consistent with past literature on bank efficiency which uses SFA studies as a proxy for Chinese banks' efficiency over the period 2003-2017 (Lee et al., 2021). Similarly, the finding of Ntwiga (2020) who use DEA methods for banks' efficiency in Kenya from 2009-2018 also supports these results. The results also indicate that the negative coefficient on the lag of the dependent variable (LEFF) suggests that the efficiency levels of Chinese banks in the current year is significantly and negatively affected by its previous year's efficiency levels. In other words, banks' past efficiency levels have a negative impact on their current efficiency levels, demonstrating the importance of banks' efficiency, consistent with the existing literature (Fiordelisi et al., 2011).

Table 3.6 Double bootstrapped truncated regression result (Baseline model)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.190*** (-4.65)	-0.316*** (-9.61)	-0.112** (-3.28)	-0.287*** (-8.38)	-0.102** (-2.97)
FI	0.061*** (5.26)	0.029*** (4.08)	0.049*** (6.86)	0.025** (3.24)	0.015* (2.11)
SIZE	-0.414*** (-3.82)	-0.251*** (-4.34)	-0.172** (-2.81)	-0.315*** (-5.29)	-0.064 (-1.21)
GDP	0.053*** (3.82)	0.007 (0.92)	0.047*** (5.98)	0.003 (0.39)	-0.009 (-1.32)
CPI	0.176*** (17.90)	0.100*** (18.57)	0.096*** (18.38)	0.096*** (18.76)	0.087*** (18.89)
IP	-0.614*** (-4.49)	-0.156*** (-2.09)	-0.438*** (-5.45)	-0.071 (-0.92)	0.027 (0.38)
LDR	-0.050 (-0.46)	-0.019 (-0.32)	-0.025 (-0.41)	-0.006 (-0.11)	0.021 (0.039)
ROA	0.077 (1.65)	0.052* (2.04)	0.051 (1.88)	0.027 (1.05)	-0.006 (-0.24)
CAR	0.042* (2.04)	0.063*** (5.23)	-0.068*** (-5.39)	0.047*** (3.79)	0.029** (2.74)
IPO	0.042 (1.19)	0.013 (0.67)	0.041* (2.03)	0.040* (2.05)	-0.015 (-0.84)
Observation	909	909	909	909	909
Constant	-1.780*** (-17.74)	-0.010*** (-18.35)	-9.654*** (-18.14)	-9.744*** (-18.56)	-8.764*** (-18.65)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
Wald $\chi 2$	32.91***	28.18***	47.73***	31.93***	7.20**

Notes: The table presents the truncated regressed result at the second stage for Hypothesis 1: Higher levels of FinTech development have a positive impact on CB's efficiency levels. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Regarding the control variables, bank size has a significantly negative effect on efficiency scores. Higher size levels indicate a faster total asset growth rate. This is consistent with Dong et al. (2017), who claim that faster growth of bank size can lead to inefficient performance of commercial banks. Banks' larger size would affect their efficiency which is consistent with our expectations.

The coefficient of the LDR variable indicates that this variable exerts a negative effect on the efficiency change (i.e., TFPCH, EFFCH, TECHCH and PURE) and a positive effect on the SCALE proxy. These results are consistent with the argument of Yamori et al. (2017) in which excessively aggressive lending may lead to a decline in marginal profit, affecting productivity and technological efficiency. Lending is also the main way for commercial banks to expand which can explain the positive influence on the scale efficiency change (Taiwo et al., 2017). In addition, ROA is positively correlated with the total factor productivity score. Higher ROA values reflect better profitability of banks (Zhao et al., 2022). CAR, which indicates the level of risk tolerance, shows a positive relationship with TFPCH, EFFCH, PURE and SCALE. A bank with a higher level of CAR is found to operate more effectively, which explains the positive influence on efficiency changes (Kwan and Eisenbeis, 1995). Finally, the listing coefficient has a positive effect on the productivity change, especially significantly on TECHCH and PURE due to the fact that listing strengthens the internal ownership structure and establishes external discipline, thereby improving operating incentives.

Overall, FinTech development improves commercial banks' efficiency levels. FinTech development in China generates more intensive competition in the financial industry but also brings technological improvement (Hua and Huang, 2020).

3.4.3 The Effect of FinTech on Different Banks' Ownership

This chapter also considers the role of FinTech on the heterogeneity of banks. Past literature has shown that banks are most likely to behave differently under different ownership structures (Jiang et al., 2013; Huang et al., 2017). According to Zhao et al. (2022), while SOCBs are required to finance certain government and state-owned projects, other commercial banks are not faced with these requirements. As such, FinTech should exert a differential effect on banks' efficiency changes. To further test the argument, a set of ownership dummy variables and their interaction terms with FinTech development are introduced (Equation 12). Estimation results are reported in Table 3.7. The coefficient on FI remains positive and statistically significant. The coefficients on all ownership variables (JSCB and CCB) are negative and statistically significant, indicating that these banks underperform SOCBs on their profit efficiency levels. In other words, these results show that regardless of the type of ownership, all banks have very low efficiency levels. This might be the result of past deeper reforms implemented in the banking sector. These reforms helped banks to rectify internal governance mechanisms and strengthen external monitoring.

Table 3.7 Double bootstrapped truncated regression result of different types of commercial banks

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.156*** (-4.01)	-0.160*** (-4.91)	-0.072* (-2.23)	-0.131*** (-3.61)	0.059 (1.81)
FI	0.047*** (8.02)	0.036*** (8.25)	0.018*** (4.97)	0.038*** (4.99)	0.046*** (11.34)
JSCB	-0.081*** (-8.73)	-0.042*** (-8.57)	-0.039*** (-7.48)	-0.038*** (-6.26)	-0.022*** (-5.17)
CCB	-0.079*** (-11.06)	-0.035*** (-9.00)	-0.043*** (-10.48)	-0.035*** (-6.32)	-0.022*** (-6.28)
FI*JSCB	0.016*** (11.35)	0.007*** (9.28)	0.009*** (10.87)	0.007*** (6.31)	0.005*** (6.62)
FI*CCB	0.017*** (9.11)	0.009*** (8.83)	0.008*** (7.73)	0.008*** (6.37)	0.005*** (5.34)
SIZE	-0.318** (-2.86)	-0.189** (3.16)	-0.134* (-2.11)	-0.247*** (-3.94)	-0.053 (-1.00)
GDP	0.203*** (9.02)	0.116*** (7.73)	0.101*** (7.36)	0.119*** (4.85)	0.141*** (10.32)
CPI	0.243*** (19.72)	0.144*** (14.93)	0.125*** (15.59)	0.141*** (7.09)	0.128*** (14.62)
IP	-2.806*** (-10.63)	-1.670*** (-9.23)	-1.261*** (-8.01)	-1.678*** (-5.26)	-1.875*** (-11.40)
LDR	-0.115 (-1.03)	-0.059 (-0.98)	-0.059 (-0.92)	-0.053 (-0.84)	-0.031 (-0.58)
ROA	0.049 (1.02)	0.032 (1.23)	0.036 (1.28)	0.007 (0.25)	-0.023 (-0.99)
CAR	0.033 (1.43)	0.049*** (4.01)	-0.072*** (-5.49)	0.033** (2.64)	0.018 (1.72)
IPO	0.029 (0.62)	-0.001 (-0.05)	0.012 (0.46)	0.046 (1.81)	-0.038 (-1.75)
Observation	909	909	909	909	909
Constant	-2.663*** (-19.25)	-1.587*** (-14.47)	-1.357*** (-15.00)	-1.548*** (-6.88)	-1.443*** (-14.52)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	101.75***	102.38***	61.34***	40.56***	129.87***

Note: The table presents the truncated regressed result at the second stage for Hypothesis 2: Higher levels of FinTech development have a stronger positive impact on the efficiency of CCBs than on their SOCBs and JSCBs counterparts. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. The SOCBs (omitted), JSCBs and CCBs are dummy variables. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

More importantly, when considering the role of FinTech on the relationship between their levels of efficiency, the findings suggest that FinTech development has a positive effect on all types of banks. This impact is statistically significant and higher for CCBs than JSCBs and SOCBs, with the latter presenting a relatively lower effect on efficiency changes. This is a new result which adds to the literature on efficiency and FinTech, demonstrating how effective FinTech is in improving the efficiency of banks, especially JSCBs and CCBs. There are 1.6% and 1.7% changes for the CCBs and JSCBs interaction terms respectively as in 5.2. Our findings are also in line with the argument of Wang et al. (2020) and Lee et al. (2021), according to which the FinTech spillover effect is higher for CCBs than JSCBs and SOCBs. The efficiency changes of CCBs are more sensitive to FinTech development and the technological promotions largely improve the efficiency of small-size commercial banks. As Liu et al. (2021) suggest, CCBs still have a gap with JSCB and SOCB at both technology and management levels. Due to the high cost of independent R&D, CCBs mainly rely on cooperation with FinTech companies to upgrade their technology. As a result, when FinTech develops to a higher level, CCBs can adjust their cooperation methods and structures more quickly, so their efficiency improvement is more significant.

These results show that the operation model, scale and ownership of banks affect the absorption of FinTech spillovers by commercial banks. FinTech decentralises the services that enable commercial banks to provide more commercialised services to users, simplifying banks' infrastructure and reducing costs (Gomber et al., 2018). Therefore, a rapid response capacity and good technological absorption can help commercial banks rethink their development models and paths. More importantly, local small banks with flexibility benefit most from FinTech development. The small scale and geographical restrictions of CCBs are the main limitations for their development. However, the aforementioned characteristics become advantages when absorbing the technology spillover (Zhao et al., 2022).

3.4.4 The Role of Regional Financial Development

In addition, the impact of regional financial development on the FinTech absorption of CCBs is tested. It is argued that their efficiency performance should be linked to their location, in line with previous studies which demonstrate that CCBs' performance is linked to local financial development (Chen et al., 2020). The FinTech variable (FI) is interacted with the financial development dummy to gauge the extent to which the impact of FinTech on banks' efficiency is stronger in a more financially developed setting. Results are provided in Table

3.8. The study observes that regional financial development exerts a positive effect on banks' efficiency, and this effect remains positive and statistically significant once FI is taken into consideration. In other words, the impact of FinTech on efficiency changes is higher for CCBs located in more financially developed provinces than their less financially developed counterparts. This is a new result which adds to the literature on financial development and banks (Sufian and Habibullah, 2009; Wu et al., 2007). Taking for example, the effect on TFPCH. 0.2% change on TFPCH for FI*FD as in 5.2. The coefficient of FI remains positive and statistically significant. The coefficient on the LEFF of the dependent variable remains negative and statistically significant.

Table 3.8 Double bootstrapped truncated regression result of the impact of financial development

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	0.354*** (9.46)	0.375*** (12.85)	0.532*** (17.63)	0.444*** (14.76)	0.438*** (15.31)
FI	0.008*** (3.63)	0.005*** (3.68)	0.005*** (4.14)	0.010*** (3.51)	0.008** (3.08)
FD	0.769*** (9.00)	0.482*** (9.68)	0.367*** (7.82)	0.489*** (9.37)	0.275*** (5.87)
FI*FD	0.002*** (4.64)	0.002*** (5.42)	0.001*** (5.73)	0.002*** (6.28)	0.001*** (4.97)
SIZE	-0.241 (-1.83)	-0.343*** (-4.34)	-0.095 (-1.29)	-0.419*** (-5.05)	-0.247*** (-3.28)
GDP	0.096*** (3.37)	0.056*** (3.34)	0.102*** (6.48)	-0.059** (-2.63)	-0.030 (1.52)
CPI	0.093*** (6.01)	0.039*** (4.10)	0.027** (3.13)	0.001 (0.12)	-0.009 (-1.04)
IP	-3.189*** (-4.31)	-2.109*** (-4.83)	-1.887*** (-4.72)	-1.654** (-3.01)	-1.217* (-2.49)
LDR	-0.065 (-0.51)	-0.013 (-0.18)	-0.067 (-0.95)	0.037 (0.47)	0.028 (0.39)
ROA	0.095 (1.66)	0.057 (1.66)	0.038 (1.16)	0.028 (0.77)	0.017 (0.52)
CAR	0.047 (1.73)	0.053** (3.26)	-0.034* (-2.24)	0.042* (2.44)	0.031* (2.03)
IPO	0.024 (0.44)	-0.006 (-0.17)	0.024 (0.77)	0.015 (0.46)	-0.032 (-1.04)
Observation	774	774	774	774	774
Constant	-9.043*** (-5.62)	-3.577*** (-3.65)	-2.487*** (-2.74)	0.787 (0.66)	1.716 (1.60)
Banks×Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	101.99***	117.77***	85.68***	93.06***	39.81***

Note: The table presents the truncated regressed result at the second stage for Hypothesis 3: FinTech exerts a positive and higher effect on CCBs in less financially developed areas than in their more financially developed counterparts. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. The FD is dummy variable for cities higher than the average financial development level. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

The coefficient of the interactive term between FI and FD is significantly positive at 0.002, indicating that for the same level of FinTech development, the TFPCH of CCBs in more financially developed cities would be 0.002 more than those in less financially developed cities. Similarly, for other Malmquist decompositions, the coefficients on the interactive terms are significantly positive.

Overall, these findings suggest that regional financial development promotes the innovation of firms, but specific government factors should be considered when analysing the local banking sectors, as CCBs are supported by the local government. Regional financial development is mainly related to government policy and economic growth. More financially developed cities are also more economically developed cities that have more emerging technologies and richer technical resources. In addition, customers in more financially developed regions are usually more willing to try new services of CCBs. These internal factors have affected the absorption of FinTech by CCBs because of different levels of regional financial development. Moreover, from an external perspective, the path and the level of FinTech development in different financially developed regions are also different. FinTech application in more financially developed regions tends to be wider and more quickly put into use.

3.5 Robustness Check

In this section, several robustness checks are applied. This chapter tests the role of FinTech in different dimensions, perform estimation techniques and address the possibility of endogeneity.

3.5.1 Different Dimensions of FinTech

Shifting the attention to the role of the different dimensions of FinTech on banks' efficiency changes, the measurements are worth noting when we consider more specific dimensions as Niu et al. (2022) shows.

Following the empirical results, the positive impact of FinTech on the efficiency of commercial banks is illstrauted. To check the robustness of the main result, this section aims to test whether the different dimensions of FinTech would generate the same positive impact on the efficiency change of commercial banks. As Table 3.9 shows, the coverage breadth

(FCB) and usage depth (FUD) of FinTech are used as instrumental variables for the FinTech development index. The coefficient of alternative FinTech indices stays consistent with the empirical result, which confirms the main hypothesis that FinTech development has a positive impact on the efficiency change of commercial banks. In addition, the breadth of FinTech coverage has a greater positive impact on the efficiency changes of commercial banks than the usage depth of FinTech. Regarding the decomposition of efficiency change, TFPCH is most affected by the development of FinTech, which increased by 2.2% and 2.7% respectively for coverage breath and usage depth.

Table 3.10 and Table 3.11 show the robustness results of Hypothesis 2, which uses FCB and FUD as instrumental variables, respectively. The results, consistent with Table 3.7, show that FinTech development has a relatively higher impact on the efficiency changes of CCBs, especially compared with SOCBs, but the difference of influence between CCBs and JSCBs is very slight. Table 3.12 and Table 3.13 present the robustness check of Hypothesis 3. Both results of using instrumental variables FCB and FUD are consistent with the result of Table 3.8: the CCBs in higher financially developed cities would increase more TFPCH, EFFCH, and TECHCH from FinTech development.

Table 3.9 Robustness check on different dimension FinTech index

Variables	TFPCH (1)	TFPCH (2)	EFFCH (1)	EFFCH (2)	TECHCH (1)	TECHCH (2)	PURE (1)	PURE (2)	SCALE (1)	SCALE (2)
LEFF	-0.322 (-0.95)	-0.206*** (-5.01)	-0.121*** (-4.51)	-0.327*** (-9.84)	0.062* (2.26)	-0.137*** (-3.94)	-0.107*** (-3.91)	-0.301*** (-8.67)	-0.014 (-0.54)	-0.109** (-3.14)
FCB	0.022*** (5.34)		0.018*** (8.12)		0.007** (3.16)		0.018*** (7.64)		0.022*** (10.96)	
FUD		0.027*** (5.95)		0.013*** (4.42)		0.003*** (7.95)		0.001*** (3.83)		-0.001* (-1.65)
SIZE	-0.206 (-1.91)	-0.421*** (-3.93)	-0.196*** (-3.36)	-0.250*** (-4.35)	-0.021 (-0.34)	-0.173** (-2.84)	-0.274*** (-4.58)	-0.312*** (-5.26)	-0.105* (-2.03)	-0.065 (-1.21)
GDP	0.123*** (5.07)	0.058*** (4.23)	0.083*** (6.21)	0.009 (1.21)	0.054*** (3.99)	0.052*** (6.56)	0.082** (5.87)	0.007 (0.81)	0.116*** (9.77)	-0.007 (-1.24)
CPI	0.238*** (14.31)	0.174*** (17.64)	0.135*** (17.84)	0.097*** (18.54)	0.108*** (13.93)	0.092*** (17.61)	0.128*** (15.19)	0.094*** (17.73)	0.118*** (18.19)	0.082*** (17.58)
IP	-2.765*** (-5.39)	-0.544*** (-4.48)	-2.15*** (-7.75)	-0.113 (-1.69)	-0.859** (-3.18)	-0.386*** (-5.45)	-2.099*** (-7.11)	-0.044 (-0.65)	-2.708*** (-11.02)	-0.019 (-0.31)
LDR	-0.054 (-0.46)	-0.050 (-0.46)	-0.017 (-0.27)	-0.019 (-0.33)	-0.024 (-0.37)	-0.026 (-0.41)	-0.004 (-0.06)	-0.009 (-0.15)	0.012 (0.22)	0.024 (0.44)
ROA	0.044 (0.88)	0.081 (1.74)	0.023 (0.86)	0.054* (2.07)	0.033 (1.15)	0.054* (2.01)	-0.003 (-0.10)	0.029 (1.09)	-0.035 (-1.51)	-0.007 (-0.30)
CAR	0.028 (1.19)	0.049* (2.12)	0.047*** (3.73)	0.064*** (5.21)	-0.065*** (-4.94)	-0.069*** (-5.42)	0.032* (2.51)	0.048*** (3.83)	0.019 (1.78)	0.029** (2.71)
IPO	0.037 (1.00)	0.042 (1.20)	0.012 (0.60)	0.013 (0.64)	0.036 (1.68)	0.041* (2.04)	0.038 (1.89)	0.041* (2.04)	-0.010 (-0.59)	-0.015 (-0.83)
Observation	909	909	909	909	909	909	909	909	909	909
Constant	-2.422*** (-14.33)	-1.774*** (-17.53)	-1.370*** (-19.82)	-0.997*** (-18.35)	-1.098*** (-13.91)	-0.929*** (-17.42)	-1.300*** (-15.12)	-0.094*** (-17.55)	-1.184*** (-18.02)	-0.082*** (-17.34)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	32.23***	40.66***	75.05***	30.60***	10.13***	63.70***	75.12***	35.08***	121.38***	5.09*

Note: The table presents the robustness check based on the truncated regressed result at the second stage for Hypothesis 1. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. The FCB and FUD are instrumental variables which represent the coverage breadth and the usage depth of FinTech separately. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Table 3.10 Robustness result for H2 (FCB: The coverage breadth)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.013 (-0.37)	-0.063* (-2.24)	0.085** (2.93)	-0.046 (-1.45)	0.084** (3.18)
FCB	0.028*** (5.21)	0.025*** (13.16)	0.012*** (5.63)	0.025*** (7.29)	0.031*** (17.61)
JSCB	-0.013*** (-4.01)	-0.008*** (-5.26)	-0.006*** (-3.75)	-0.008*** (-4.96)	-0.005*** (-3.64)
CCB	-0.014*** (-5.09)	-0.006*** (-4.96)	-0.008*** (-5.66)	-0.008*** (-5.48)	-0.004*** (-3.79)
FCB*JSCB	0.007*** (5.19)	0.004*** (6.53)	0.003*** (4.62)	0.004*** (5.80)	0.002*** (4.57)
FCB*CCB	0.009*** (6.09)	0.004*** (6.46)	0.005*** (7.16)	0.004*** (6.04)	0.003*** (5.61)
SIZE	-0.122 (-1.09)	-0.158** (-2.69)	-0.013 (-0.21)	-0.221*** (-3.61)	-0.101 (-1.95)
GDP	0.163*** (5.56)	0.110*** (10.11)	0.089*** (6.96)	0.113*** (6.43)	0.144*** (14.65)
CPI	0.272*** (10.08)	0.161*** (31.09)	0.138*** (16.80)	0.151*** (10.32)	0.146*** (30.84)
IP	-3.456*** (-6.25)	-2.379*** (-13.92)	-1.596*** (-7.81)	-0.240*** (-7.57)	-2.749*** (-17.66)
LDR	-0.065 (-0.55)	-0.019 (-0.31)	-0.037 (-0.57)	-0.010 (-0.16)	0.004 (0.08)
ROA	0.054 (1.06)	0.025 (0.93)	0.034 (1.17)	0.000 (0.00)	-0.034 (-1.46)
CAR	0.029 (1.21)	0.045*** (3.59)	-0.064*** (-4.79)	0.031* (2.39)	0.017 (1.56)
IPO	0.031 (0.64)	0.001 (0.02)	0.017 (0.63)	0.046 (1.80)	-0.037 (-1.71)
Observation	909	909	909	909	909
Constant	-2.990*** (-9.69)	-1.795*** (-30.10)	-1.500*** (-15.98)	-1.697*** (-10.03)	-1.676*** (-30.61)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	32.12***	193.63***	39.96***	60.14***	316.41***

Note: The table presents the robustness check based on the truncated regressed result at the second stage for Hypothesis 2. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. The FCB is the institution variable which represents the coverage breadth. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Table 3.11 Robustness result for H2 (FUD: The usage depth)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.295*** (-7.08)	-0.392*** (-12.49)	-0.231*** (-6.42)	-0.372*** (-11.43)	-0.230*** (-7.11)
FUD	0.012*** (4.44)	0.007*** (4.48)	0.012*** (6.98)	0.006*** (3.91)	0.001 (0.81)
JSCB	-0.006* (-2.36)	-0.005*** (-3.32)	-0.001 (-0.91)	-0.005** (-3.06)	-0.004** (-2.65)
CCB	-0.007** (-3.08)	-0.002* (-2.08)	-0.003** (-2.69)	-0.004*** (-3.45)	-0.002* (-2.08)
FUD*JSCB	0.004*** (3.87)	0.002*** (4.38)	0.001* (1.86)	0.002*** (3.70)	0.002*** (3.53)
FUD*CCB	0.004*** (4.60)	0.001** (3.17)	0.002*** (4.42)	0.001*** (3.79)	0.002*** (3.74)
SIZE	-0.307** (-2.86)	-0.199*** (3.44)	-0.094 (-1.59)	-0.245*** (-4.12)	-0.071 (-1.32)
GDP	-0.004 (-0.31)	-0.015* (-2.20)	0.006 (0.82)	-0.166* (-2.38)	-0.007 (-1.25)
CPI	0.206*** (16.05)	0.118*** (20.79)	0.114*** (18.12)	0.113*** (18.94)	0.098*** (16.67)
IP	-1.355*** (-11.04)	-0.532*** (-8.42)	-0.891*** (-13.07)	-0.513*** (-7.94)	-0.376*** (-6.18)
LDR	0.106 (2.67)	-0.044 (-0.74)	-0.071 (-1.18)	-0.045 (-0.76)	0.003 (0.05)
ROA	0.123** (2.67)	0.071** (2.80)	0.079** (3.03)	0.050 (1.93)	-0.005 (-0.22)
CAR	0.058** (2.69)	0.069*** (5.85)	-0.069*** (-5.69)	0.054*** (4.42)	0.033** (3.03)
IPO	0.033 (0.76)	0.005 (0.19)	0.017 (0.70)	0.060* (2.41)	-0.042 (-1.89)
Observation	909	909	909	909	909
Constant	-2.180*** (-16.46)	-1.232*** (-20.90)	-0.001*** (-18.73)	-1.174*** (-18.98)	-1.007*** (-16.43)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	48.62***	58.77***	75.02***	47.45***	12.28***

Note: The table presents the robustness check based on the truncated regressed result at the second stage for Hypothesis 2. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. The FUD is institution variable which represents the usage depth of FinTech. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Table 3.12 Robustness result for H3 (FCB: The coverage breadth)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.185*** (-5.35)	-0.223*** (-6.74)	-0.356*** (-7.18)	-0.175*** (-7.05)	-0.562*** (-17.01)
FCB	0.039*** (14.48)	0.018*** (9.72)	0.020*** (12.41)	0.022*** (12.65)	0.003*** (5.25)
FD	0.100*** (11.36)	0.049* (1.82)	0.056* (2.32)	0.012 (0.42)	0.020*** (3.54)
FCB*FD	0.002*** (5.48)	0.001*** (8.46)	0.001*** (2.32)	0.002*** (6.58)	0.001** (3.05)
SIZE	-0.335** (-2.92)	-0.301*** (-3.81)	-0.085 (-1.21)	-0.314*** (-3.85)	-0.268*** (-3.66)
GDP	0.259*** (12.02)	0.122*** (8.27)	0.171*** (13.23)	0.152*** (10.27)	0.034*** (4.00)
CPI	0.210*** (12.58)	0.087*** (7.38)	0.080*** (8.03)	0.094*** (8.37)	0.046*** (4.18)
IP	-1.183*** (-14.97)	-0.626*** (-11.43)	-0.656*** (-13.54)	-0.735*** (-14.13)	-0.454* (2.44)
LDR	0.032 (0.29)	0.057 (0.77)	0.016 (0.23)	0.057 (0.75)	0.048 (0.71)
ROA	0.183*** (3.64)	0.089* (2.51)	0.082** (2.59)	0.069 (1.91)	-0.000 (-0.01)
CAR	0.083*** (3.47)	0.059*** (3.58)	-0.037* (-2.49)	0.050** (2.96)	0.022 (1.45)
IPO	0.039 (0.82)	0.010 (0.29)	0.042 (1.41)	0.041 (1.18)	-0.026 (-0.88)
Observation	774	774	774	774	774
Constant	-1.940*** (-11.84)	-0.765*** (-6.67)	-0.703*** (-7.18)	-0.818*** (-7.42)	-0.471*** (4.24)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	268.18***	96.70***	156.62***	160.20***	64.47***

Note: The table presents the robustness check based on the truncated regressed result at the second stage for Hypothesis 3. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Table 3.13 Robustness result for H3 (FUD: The usage depth)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.272** (-7.05)	-0.428*** (-13.18)	-0.491*** (-14.69)	-0.298*** (-11.45)	-0.562*** (-17.01)
FUD	0.008*** (7.05)	0.001 (0.47)	0.003*** (5.17)	0.002** (3.10)	0.003*** (5.25)
FD	0.104*** (9.73)	0.586*** (9.45)	0.489*** (8.37)	0.557*** (8.31)	0.202*** (3.55)
FUD*FD	0.003*** (5.76)	0.002*** (5.78)	0.002*** (6.28)	0.002*** (5.79)	0.001** (3.05)
SIZE	-0.286* (-2.24)	-0.364*** (-4.61)	-0.108 (-1.48)	-0.318*** (-3.74)	-0.268*** (-3.66)
GDP	0.004 (0.25)	0.009 (1.03)	0.045*** (4.81)	0.007 (0.72)	0.034*** (4.01)
CPI	0.143*** (8.12)	0.025* (2.23)	0.042*** (4.11)	0.044*** (3.79)	0.046*** (4.18)
IP	-2.828*** (-9.34)	-0.601** (-3.10)	-1.185*** (-6.55)	-1.055*** (-5.29)	-0.455* (2.44)
LDR	-0.075 (-0.61)	-0.012 (-0.16)	-0.056 (-0.81)	-0.001 (-0.01)	0.048 (0.71)
ROA	0.136* (2.42)	0.053 (1.53)	0.053 (1.62)	0.034 (0.89)	-0.000 (-0.01)
CAR	0.063* (2.42)	0.051** (3.18)	-0.031* (-1.99)	0.044* (2.49)	0.022 (1.45)
IPO	0.029 (0.55)	-0.006 (-0.18)	0.024 (0.78)	0.032 (0.92)	-0.026 (-0.88)
Observation	774	774	774	774	774
Constant	-1.406*** (-7.92)	-2.431* (-2.11)	-0.410*** (-3.95)	-0.423*** (-3.61)	-0.471*** (4.25)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	108.04***	100.42***	74.08***	69.07***	64.47**

Note: The table presents the robustness check based on the truncated regressed result at the second stage for Hypothesis 3. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

3.5.2 Non-linear Effects

It could be argued that FinTech may have a negative effect up to a certain (turning) point and positive thereafter. Past studies argue that FinTech is characterised by high operation and research and development (R&D) cost, and, therefore, might not improve banks' efficiency at early stages due to intensive competition (Zhao et al., 2022). Hence, to capture potential non-linear influences, we also allow for non-linear transformation in our main models. We add a quadratic term on the FI variable in our three main specifications. These are as follows:

$$E_{it} = \beta_0 + \beta_1 E_{i,t-1} + \beta_2 FI_{it-1} + \beta_3 FI^2{}_{it-1} + \sum\delta_i Macrolevel_{it} + \sum\gamma_i Microlevel_{it} + \epsilon_{it} \quad (14)$$

$$\begin{gathered} E_{it} = \beta_0 + \beta_1 E_{i,t-1} + \beta_2 FI_{it-1} + \beta_3 FI^2{}_{it-1} + \beta_4 JSCB_{it} + \beta_5 CCB_{it} \\ +\beta_6 JSCB_{it} * FI_{it-1} + \beta_7 CCB_{it} * FI_{it-1} + \beta_8 JSCB_{it} * FI^2{}_{it-1} + \beta_9 CCB_{it} * FI^2{}_{it-1} \\ +\sum\delta_i Macrolevel_{it} + \sum\gamma_i Microlevel_{it} + \epsilon_{it} \end{gathered} \quad (15)$$

$$\begin{gathered} E_{it} = \beta_0 + \beta_1 E_{i,t-1} + \beta_2 * FI_{it-1} + \beta_3 FI^2{}_{it-1} + \beta_4 FD^2{}_{it-1} \\ \beta_5 FD_{it-1} * FI_{it-1} + \beta_6 FD^2{}_{it-1} * FI^2{}_{it-1} + \sum\delta_i Macrolevel_{it} + \sum\gamma_i Microlevel_{it} + \epsilon_{it} \end{gathered} \quad (16)$$

Table 3.14 shows that the models are linear and consistent with our main findings. The coefficient of FI remains positive and significant, as well as the coefficient of FI^2. This shows that FinTech development exerts a positive effect on banks' efficiency throughout the sample period. This study further investigates whether the linear nexus is related to banks' ownership structure (see Table 3.15). The interactive term of FI^2 and ownership dummy variables are included. The interactive terms are found positive and significant for JSCBs and CCBs. This finding shows that JSCBs and CCBs are more sensitive to the spillover effect brought by FinTech compared to their SOCBs counterparts. This is consistent with our hypothesis and the arguments of past studies (Wang et al., 2021; Cheng et al., 2022). Finally, the effect of both financial development and FinTech development is took into consideration. The interactive terms of FI^2 and FD^2 on Table 3.16 are positive and statistically significant. This denotes the positive and more significant impact of FinTech on banks in more financially developed areas. Therefore, these results are robust to the main findings.

Table 3.14 Double bootstrapped truncated regression result (non-linear model)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.228*** (-5.34)	-0.256*** (-7.61)	-0.151*** (-4.06)	-0.229*** (-6.58)	-0.013 (-0.39)
FI	0.028*** (6.94)	0.019*** (8.43)	0.011*** (4.83)	0.018*** (8.35)	0.023*** (11.94)
FI^2	0.015*** (7.88)	0.009*** (9.26)	0.006*** (6.10)	0.009*** (9.07)	0.010*** (11.78)
SIZE	-0.448*** (-4.07)	-0.267*** (-4.56)	-0.187** (-3.01)	-0.328*** (-5.46)	-0.085 (-1.64)
GDP	0.392*** (8.75)	0.225*** (9.15)	0.193*** (7.72)	0.206*** (8.69)	0.215*** (10.58)
CPI	0.238*** (14.69)	0.143*** (15.58)	0.121*** (14.56)	0.133*** (17.17)	0.127*** (19.37)
IP	-7.468*** (-8.49)	-4.544*** (-9.48)	-3.381*** (-6.92)	-4.149*** (-9.09)	-4.524*** (-11.52)
LDR	-0.123 (-1.09)	-0.067 (-1.12)	-0.055 (-0.87)	-0.051 (-0.83)	-0.028 (-0.54)
ROA	0.031 (0.65)	0.021 (0.80)	0.032 (1.15)	-0.002 (-0.08)	-0.032 (-1.40)
CAR	0.030 (1.35)	0.050*** (4.12)	-0.077*** (-5.96)	0.035** (2.82)	0.018 (1.69)
IPO	0.046 (1.29)	0.014 (0.75)	0.043* (2.13)	0.042* (2.10)	-0.012 (-0.67)
Observation	909	909	909	909	909
Constant	-2.308*** (-14.79)	-1.384*** (-15.60)	-1.180*** (-14.71)	-1.292*** (17.21)	-1.223*** (-19.24)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi2}$	90.25***	51.58***	61.05***	104.70***	70.87***

Notes: The table presents the truncated regressed result at the second stage for Hypothesis 1: Higher levels of FinTech development have a positive impact on CB's efficiency levels. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Table 3.15 Double bootstrapped truncated regression result of different types of commercial banks

(non-linear model)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.296*** (-6.07)	-0.270*** (-7.94)	-0.189*** (-5.08)	-0.234*** (-6.65)	-0.022 (-0.61)
FI	0.097*** (22.40)	0.037*** (20.00)	0.028*** (14.73)	0.038*** (19.45)	0.039*** (21.84)
FI^2	0.059*** (25.11)	0.022*** (24.41)	0.019*** (19.46)	0.023*** (23.97)	0.021*** (25.26)
JSCB	-0.199*** (-6.97)	-0.086*** (-7.77)	-0.055*** (-4.56)	-0.086*** (-7.51)	-0.057*** (-5.71)
CCB	-0.204*** (-7.66)	-0.069*** (-7.32)	-0.075*** (-7.17)	-0.078*** (-8.03)	-0.051*** (-5.98)
FI*JSCB	0.077*** (19.19)	0.028*** (17.14)	0.023*** (13.04)	0.031*** (17.24)	0.024*** (15.37)
FI*CCB	0.092*** (26.11)	0.031*** (25.06)	0.033*** (27.37)	0.034*** (24.71)	0.027*** (22.67)
FI^2*JSCB	0.017*** (16.78)	0.007*** (15.40)	0.005*** (11.81)	0.008*** (15.99)	0.006*** (14.61)
FI^2*CCB	0.019*** (22.56)	0.007*** (22.69)	0.007*** (23.32)	0.007*** (22.94)	0.006*** (21.12)
SIZE	-0.122 (-0.99)	-0.112 (-1.86)	-0.058 (-0.89)	-0.141* (-2.28)	0.038 (0.71)
GDP	0.906*** (19.75)	0.333*** (17.08)	0.304*** (14.16)	0.329*** (15.88)	0.316*** (16.89)
CPI	0.893*** (34.14)	0.339*** (43.15)	0.305*** (47.16)	0.343*** (40.09)	0.301*** (40.91)
IP	-2.488*** (-28.78)	-8.881*** (-29.87)	-8.191*** (-26.94)	-9.146*** (-29.68)	-8.147*** (-30.27)
LDR	-0.181 (-1.48)	-0.082 (-1.37)	-0.092 (-1.43)	-0.071 (-1.16)	-0.048 (-0.89)
ROA	0.058 (1.11)	0.036 (1.41)	0.041 (1.47)	0.015 (0.57)	-0.022 (-0.92)
CAR	0.033 (1.34)	0.054*** (4.38)	-0.079*** (-5.99)	0.038** (3.06)	-0.020 (1.82)
IPO	0.055 (1.10)	0.012 (0.48)	0.024 (0.92)	0.065* (2.56)	-0.031 (-1.38)
Observation	909	909	909	909	909
Constant	-1.041*** (-33.95)	-3.913*** (-42.65)	-3.533*** (-45.64)	-3.971*** (-40.10)	-3.498*** (-41.05)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
$Wald\ \chi2$	599.92***	581.89***	468.29***	620.08***	592.64***

Note: The table presents the truncated regressed result at the second stage for Hypothesis 2: Higher levels of FinTech development have a stronger positive impact on the efficiency of CCBs than on their SOCBs and JSCBs counterparts. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. The SOCBs (omitted), JSCBs and CCBs are dummy variables. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Table 3.16 Double bootstrapped truncated regression result of the impact of financial development

(non-linear model)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.114** (-2.79)	-0.165*** (-4.58)	-0.067 (-1.18)	-0.085* (-6.75)	0.078* (2.01)
FI	0.007*** (3.64)	0.009*** (8.54)	0.003*** (3.76)	0.013*** (10.71)	0.016*** (13.80)
FI^2	0.007*** (9.14)	0.005*** (13.89)	0.004*** (10.46)	0.007*** (14.63)	0.007*** (16.16)
FD^2	0.102*** (10.72)	0.404*** (8.69)	0.488*** (10.33)	0.326*** (6.70)	0.092* (2.07)
FI*FD	0.012*** (8.35)	0.005*** (5.98)	0.008*** (9.85)	0.005*** (6.11)	0.002*** (4.14)
FI^2*FD^2	0.001*** (7.61)	0.001*** (5.44)	0.001*** (9.02)	0.001*** (6.04)	0.000*** (4.69)
SIZE	-0.427*** (-3.78)	-0.251*** (-3.98)	-0.151* (-2.41)	-0.255*** (-3.89)	0.005 (0.08)
GDP	0.218*** (11.14)	0.126*** (12.22)	0.151*** (14.05)	0.142*** (12.07)	0.140*** (12.73)
CPI	0.054*** (4.05)	0.052*** (6.89)	0.013 (1.65)	0.045*** (5.65)	0.042*** (5.75)
IP	-3.767*** (-11.26)	-2.383*** (-14.75)	-2.291*** (-13.47)	2.004*** (14.73)	1.972*** (15.54)
LDR	-0.117 (-1.11)	-0.049 (-0.85)	-0.086 (-1.47)	-0.043 (-0.71)	-0.038 (-0.67)
ROA	0.093* (1.97)	0.048 (1.81)	0.060* (2.26)	0.022 (0.79)	-0.023 (-0.92)
CAR	0.062** (2.79)	0.062*** (4.92)	-0.063*** (-5.03)	0.042** (3.25)	0.016 (1.38)
IPO	0.013 (0.29)	-0.007 (-0.31)	0.017 (0.69)	0.039 (1.53)	-0.038 (-1.62)
Observation	774	774	774	774	774
Constant	-0.498*** (-3.65)	-0.491*** (-6.34)	-0.091 (-1.18)	-0.565*** (-6.75)	-0.532*** (-6.89)
Banks×Year FE	Yes	Yes	Yes	Yes	Yes
$Wald\ \chi2$	217.42***	284.83***	339.94***	263.37***	268.48***

Note: The table presents the truncated regressed result at the second stage for Hypothesis 3: FinTech exerts a positive and higher effect on CCBs in less financially developed areas than in their more financially developed counterparts. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. The FD is dummy variable for cities higher than the average financial development level. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

3.5.3 Endogeneity

To address the endogeneity, we introduced an exogenous variable. The People's Bank of China, the central bank of China, issued in July 2016 the Supervisory Guidance on the 13th Five-Year Development Plan for Informatisation in the Chinese Banking Sector (hereinafter referred to as "the Opinions") (Cheng et al., 2022). This Opinions document clearly states the mission to promote emerging high technology in the banking sector. We consider this as an exogenous policy shock to banks and define $shock_{it}$ to be equal to 1 in years after 2016 and 0 otherwise. The model for the three hypotheses is specified as follows.

$$E_{it} = \beta_0 + \beta_1 E_{i,t-1} + \beta_2 FI_{it-1} + \beta_3 shock_{it} + \beta_3 FI_{it-1} * shock_{it}$$
$$+\sum \delta_i Macrolevel_{it} + \sum \gamma_i Microlevel_{it} + \epsilon_{it} \quad (17)$$

$$E_{it} = \beta_0 + \beta_1 E_{i,t-1} + \beta_2 FI_{it-1} + \beta_3 shock_{it} + \beta_4 JSCB_{it} + \beta_5 CCB_{it} + \beta_6 FI_{it-1} * shock_{it}$$
$$+\beta_7 JSCB_{it} * shock_{it} + \beta_8 CCB_{it} * shock_{it}$$
$$+\beta_9 JSCB_{it} * FI_{it-1} * shock_{it} + \beta_{10} CCB_{it} * FI_{it-1} * shock_{it}$$
$$+\sum \delta_i Macrolevel_{it} + \sum \gamma_i Microlevel_{it} + \epsilon_{it} \quad (18)$$

$$E_{it} = \beta_0 + \beta_1 E_{i,t-1} + \beta_2 * FI_{it-1} + \beta_3 shock_{it} + \beta_4 * FD_{it-1}$$

$$+\beta_5 FI_{it-1} * shock_{it} + \beta_6 FD_{it-1} * shock_{it} + \beta_7 * FI_{it-1} * FD_{it-1} * shock_{it}$$
$$+\sum \delta_i Macrolevel_{it} + \sum \gamma_i Microlevel_{it} + \epsilon_{it} \quad (19)$$

As shown in Table 3.17, the coefficient on FI is significantly positive, which is consistent with the main finding. Interestingly, the coefficient on the shock is significantly negative, while the coefficient on FI*shock is significantly positive, suggesting that the Opinions aims to promote banks' R&D and system upgrades for innovation-driven development, which may have a negative impact on banks' current efficiency since the increase in R&D investment is not matched by a corresponding output in the short term. The Opinions does not work well unless there are high levels of FinTech infrastructure development.

Table 3.18 reports the results considering the banks’ ownership, where the coefficients on FI and FI*shock remain positive and the shock still has a negative effect on bank efficiency, whereas the coefficients on JCSB*shock and CCB*shock are both significantly negative, indicating that the Opinions has a less negative impact on JSCBs and CCBs efficiency than SOCBs. The coefficients of FI*JSCB*shock and FI*CCB*shock are both significantly positive, and the efficiencies of JSCBs and CCBs are affected more positively by the FinTech development than SOCBs under the Opinions.

Table 3.19 presents the results of Eq. 19, where the significantly negative coefficient on FD*shock indicates that the Opinions weakened the impact of the degree of financial development on bank efficiency, while the significantly positive coefficient on FI*FD*shock indicates that CCBs in areas with higher financial development were more affected by the development of FinTech after the Opinions was issued.

Table 3.17 Double bootstrapped truncated regression result (endogeneity test of policy shock)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.199*** (-5.17)	-0.366*** (-10.54)	-0.088** (-2.63)	-0.336*** (-9.67)	-0.141*** (-3.87)
FI	0.046*** (4.09)	0.028*** (3.59)	0.035*** (4.78)	0.021** (2.73)	0.044* (2.36)
shock	-0.971*** (-16.90)	-0.464*** (-9.18)	-0.705*** (-20.71)	-0.552*** (-16.01)	-0.384*** (-6.97)
FI*shock	0.082*** (16.83)	0.039*** (9.16)	0.059*** (20.64)	0.047*** (15.93)	0.033*** (6.95)
SIZE	-0.274** (-2.58)	-0.206*** (-3.32)	-0.069 (-1.12)	-0.257*** (-4.16)	-0.034 (-0.61)
GDP	0.043** (2.77)	0.009 (0.95)	0.038*** (4.16)	0.005 (0.53)	-0.002 (-0.21)
CPI	0.213*** (27.47)	0.126*** (17.56)	0.128*** (27.13)	0.124*** (26.38)	0.107*** (13.41)
IP	-0.404 (-1.61)	0.024 (0.16)	-0.296* (-1.98)	0.151 (1.03)	0.314* (2.36)
LDR	0.065 (0.63)	0.049 (0.80)	0.061 (0.98)	0.067 (1.08)	0.082 (1.47)
ROA	0.125** (2.76)	0.072** (2.69)	0.082** (3.10)	0.051 (1.91)	0.005 (0.19)
CAR	0.062** (2.92)	0.072*** (5.76)	-0.055*** (-4.40)	0.058*** (4.61)	0.036** (3.20)
IPO	0.036 (1.05)	0.011 (0.57)	0.036 (1.83)	0.039* (1.97)	-0.015 (-0.86)
Observation	909	909	909	909	909
Constant	-2.168*** (26.89)	-1.283*** (-17.34)	-1.301*** (-26.51)	-1.267*** (-25.91)	-1.093*** (-13.28)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	279.97***	98.35***	450.185***	260.56***	53.89***

Notes: The table presents the truncated regressed result at the second stage for Hypothesis 1: Higher levels of FinTech development have a positive impact on CB's efficiency levels. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Next, the role of this policy shock on banks' ownership type is tested. In Table 3.18, the coefficients on the two interaction terms are observed $JSCB_{it} * shock_{it}$ and $CCB_{it} * shock_{it}$ are negative and statistically significant, denoting that the policy shock affects banks' efficiency negatively. This effect is stronger for JSCBs than CCBs. This is in line with evidence presented by other studies (Cheng et al., 2022) Accordingly, due to the off-site surveillance nature of JSCBs (Sufian and Habibullah, 2012), they are more direct in responding to national calls and are more strongly influenced by them. On the contrary, the operations of CCBs are largely local in scope, which is different from JSCB, and their management is more influenced by local policies which often lag the newly released policy. However, once we consider FinTech, this negative effect is mitigated. Specifically, the coefficients of $JSCB_{it} * FI_{it-1} * shock_{it}$ and $CCB_{it} * FI_{it-1} * shock_{it}$ are positive and statistically significant. The positive impact of FinTech is higher for CCBs than JSCBs. This once more denotes the differential effect of bank ownership and the role of the government on CCBs. Overall, these findings show the clear and positive impact of FinTech on banks' efficiency levels.

Table 3.18 Double bootstrapped truncated regression result of different types of commercial banks
(Endogeneity test of policy shock)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	0.003 (0.09)	-0.124*** (-3.69)	-0.104** (3.28)	-0.291 (0.71)	-0.150*** (-4.68)
FI	0.016*** (6.81)	0.017*** (3.56)	0.021*** (5.01)	0.018*** (11.64)	0.018*** (14.89)
shock	-0.148*** (-3.34)	-0.237*** (-8.58)	-0.088*** (-4.09)	-0.120*** (-7.33)	-0.069*** (-4.83)
JSCB	0.176 (1.57)	0.169 (0.29)	0.089 (1.49)	0.019 (0.32)	-0.014 (-0.28)
CCB	0.060 (0.60)	0.183 (0.34)	0.051 (0.93)	-0.043 (-0.79)	0.033 (0.74)
FI*shock	0.042*** (3.34)	0.067*** (8.57)	0.025*** (4.09)	0.037*** (7.34)	0.065*** (4.85)
JSCB*shock	-0.079*** (-7.39)	-0.077*** (-11.87)	-0.049*** (-9.66)	-0.061*** (-8.19)	-0.040*** (-10.84)
CCB*shock	-0.011*** (-7.98)	-0.074*** (-8.30)	-0.014*** (-7.57)	-0.081** (-3.03)	-0.029*** (-6.37)
JSCB*FI*shock	0.014*** (7.35)	0.014*** (11.78)	0.009*** (9.63)	0.011*** (8.18)	0.007*** (10.85)
CCB*FI*shock	0.020*** (7.96)	0.014*** (8.29)	0.004*** (-7.57)	0.015** (3.03)	0.005*** (6.40)
SIZE	0.041 (0.34)	-0.069 (-1.09)	-0.079 (1.20)	-0.058 (-0.87)	0.051 (0.91)
GDP	0.159*** (6.02)	0.039* (2.47)	0.086*** (5.50)	0.144*** (9.27)	0.157*** (12.34)
CPI	0.333*** (14.46)	0.106*** (8.08)	0.169*** (13.81)	0.212*** (13.16)	0.174*** (20.12)
IP	-1.238*** (-5.29)	-0.396* (-2.50)	-0.527*** (-3.77)	-1.454*** (-9.46)	-1.639*** (-12.58)
LDR	0.083 (0.71)	0.033 (0.53)	0.068 (2.41)	0.082 (1.23)	0.075 (1.37)
ROA	0.104* (2.08)	0.055* (2.04)	0.068* (2.41)	0.034 (1.18)	0.006 (0.27)
CAR	0.047* (2.01)	0.057*** (4.53)	-0.050*** (-3.81)	0.042** (1.18)	0.027* (2.51)
IPO	0.042 (0.86)	0.006 (0.25)	0.023 (0.86)	0.055* (2.05)	-0.031 (-1.39)
Observation	909	909	909	909	909
Constant	-3.517*** (-14.09)	-1.113*** (-7.72)	-1.780*** (-13.26)	-2.294*** (-13.28)	-1.914*** (-20.09)
Banks × Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi2}$	127.63***	289.82***	109.52***	303.15***	501.17***

Note: The table presents the truncated regressed result at the second stage for Hypothesis 2: Higher levels of FinTech development have a stronger positive impact on the efficiency of CCBs than on their SOCBs and JSCBs counterparts. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. The SOCBs (omitted), JSCBs and CCBs are dummy variables. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Finally, in Table 3.19, the impact of the policy shock on how regional financial development relates to banks' efficiency. These results of using the Opinion as the exogenous variable remain robust to the main findings. The implemented policy exerts a negative and statistically significant effect on the relationship between regional financial development and CCBs' efficiency levels ($FD_{it-1} * shock_{it}$). However, this effect is mitigated once we take into consideration FinTech ($FI_{it-1} * FD_{it-1} * shock_{it}$). These findings confirm that FinTech exerts a positive impact on CCBs' efficiency levels, independent of any external negative policy shock, such as the new technology policy regulation promulgated by the Chinese government. Taken together, these findings suggest that the results are consistent with those of the baseline models, providing evidence of the effect of FinTech on banks' efficiency levels.

Table 3.19 Double bootstrapped truncated regression result of the impact of financial development

(endogeneity test of policy shock)

Variables	TFPCH	EFFCH	TECHCH	PURE	SCALE
LEFF	-0.113** (-2.89)	-0.140*** (-4.17)	-0.161*** (-4.30)	-0.133*** (-3.80)	-0.187*** (6.37)
FI	0.063*** (10.62)	0.022*** (5.91)	0.048*** (12.98)	0.022*** (5.45)	0.006*** (12.54)
shock	-0.201*** (-6.81)	-0.104*** (-6.56)	-0.101*** (-6.14)	-0.088*** (-5.79)	-0.024*** (-10.63)
FD	0.506*** (11.79)	0.281*** (11.88)	0.191*** (7.96)	0.213*** (8.61)	0.021 (0.91)
FI*shock	0.073*** (7.03)	0.037*** (6.78)	0.037*** (6.38)	0.032*** (5.99)	0.085*** (10.58)
FD*shock	-0.607* (-1.89)	-0.264 (-1.52)	-0.559** (-3.11)	-0.368* (-2.14)	-0.524** (-2.79)
FI*FD*shock	0.020* (1.83)	0.088 (1.46)	0.019** (3.07)	0.013* (2.09)	0.018** (2.82)
SIZE	-0.326** (-3.00)	-0.204*** (-3.36)	-0.107* (-1.73)	-0.232*** (-3.62)	0.061 (1.04)
GDP	0.025 (1.47)	-0.014 (-1.40)	0.033*** (3.39)	-0.013 (-1.27)	0.047*** (12.14)
CPI	0.052*** (3.98)	0.045*** (6.21)	0.012 (1.56)	0.030*** (3.87)	0.144*** (11.49)
IP	-2.843*** (-10.64)	-1.317*** (-9.32)	-1.698*** (-11.21)	-0.656*** (-7.08)	-1.127*** (12.74)
LDR	-0.051 (-0.52)	-0.004 (-0.07)	-0.045 (-0.80)	0.005 (0.08)	-0.049 (-0.91)
ROA	0.143** (3.21)	0.082** (3.26)	0.092*** (3.54)	0.053* (2.02)	-0.002 (-0.06)
CAR	0.064** (3.05)	0.067*** (5.68)	-0.068*** (-5.65)	0.049*** (4.02)	0.017 (1.47)
IPO	0.034 (0.83)	0.002 (0.09)	0.034 (1.43)	0.052* (2.09)	-0.025 (-1.14)
Observation	774	774	774	774	774
Constant	-4.697*** (-3.48)	-4.282*** (-5.65)	-0.811 (-1.04)	-3.278*** (-4.15)	-2.053*** (-12.20)
Banks×Year FE	Yes	Yes	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	136.08***	70.02***	180.11***	56.56***	163.26***

Note: The table presents the truncated regressed result at the second stage for Hypothesis 3: FinTech exerts a positive and higher effect on CCBs in less financially developed areas than in their more financially developed counterparts. The dependent variables are bias-corrected efficiency score derived from DEA-Malmquist method. The FD is dummy variable for cities higher than the average financial development level. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

3.6 Conclusions

This paper extends the existing literature on the FinTech development and efficiency changes of commercial banks in China. To quantify FinTech development, we introduce the DFII. This study employs a two-stage bootstrap-DEA-Malmquist truncated regression to test the hypotheses. The efficiency changes of 101 Chinese commercial banks from 2011-2020 is evaluated and the DEA-Malmquist method is applied in the first stage. This study also introduces the double bootstrapped truncated regression from Simar and Wilson (2007) to estimate the relationship between FinTech development and efficiency change of Chinese commercial banks. These results confirm that FinTech development exerts a positive impact on Chinese commercial banks' efficiency changes. Once accounting for the different types of Chinese commercial banks, these results show that CCBs are relatively more sensitive to the positive impact of FinTech development than other bank groups. In addition, the relationship between FinTech and the local financial development level is explored. This research finds that FinTech exerts a stronger and more positive effect on efficiency changes in CCBs located in more financially developed cities.

The results remain consistent under a battery of robustness checks. First, FCB and FUD are shown, two dimensions of DFII, a positive impact of FinTech on the efficiency change of commercial banks. Using those dimensions in our setup validates that FinTech development has a higher impact on the efficiency changes of CCBs, especially compared with SOCBs. The same factors also show that CCBs in higher financially developed cities would increase more TFPCH, EFFCH, and TECHCH from FinTech development. Second, the main results are replicated but allow for non-linear transformation in the main models. These findings suggest that the original findings are robust. Finally, to account for endogeneity we perform a DID approach using the "the Opinions" directive as an exogenous policy shock. This robustness check shows that FinTech is associated with higher banks' efficiency levels after the policy shock, even if the shock itself has a negative effect in banking efficiency. In the presence of higher FinTech development and regional development, the negative effect on CCBs is mitigated affirming the main results.

Overall, FinTech is still a forward-looking technological framework for the development of the financial system. These findings highlight the variation in efficiency changes among different commercial banks and the importance of their operations management. In that framework, this research underlines once more the impact of the Chinese government's guidance and supervision over the banking system. Regarding the common question of whether FinTech is a shock to the Chinese commercial banks, our current understanding is that FinTech helps Chinese commercial banks improve their productivity in the long run, but its integration should be guided by the different characteristics of commercial banks. This study's findings also highlight that less financially developed regions should integrate FinTech in conjunction with their levels of financial development. local governments in China should continue to collaborate with commercial banks, and especially CCBs, to maximise the positive impact of FinTech.

Chapter 4 Digital Supply Chain Finance: Feature Selection in Credit Risk Assessment

4.1 Introduction

In recent years, technological innovation and transformation of the new technology-led FinTech applications in the digital economy are gradually merging with traditional industries and generating new developments[22]. DSCF, a product of the digital technology surrounding SCF, is a complex web-like system formed due to the combination of big data, cloud computing, IT and blockchain technologies. The SCF platform provided by traditional financial institutions is infinitely extended by the participating entities in this engagement process (Du et al., 2020). The upstream and downstream operational structure of companies in the supply chain is not limited to the traditional chain organisation but has evolved into an organisational structure (Scuotto et al., 2017). Governments, financial institutions, logistics, and other proponents of SCF activities are all reflected in this intertwining of interests, guiding supply chain forecasting, planning, execution, and decision-making activities through DSCF platforms. The construction of a modern DSCF system is the integration of traditional process fragmentation, using new technological tools to keep companies closely connected while refining the division of labor and reducing the frictional costs between each link through information technology; DSCF is a deep integration of various industrial chains and finance (Korpela et al., 2017). Due to the application and penetration of digital technologies, SCF has undergone significant changes in the valuation of the soft power of companies, target credit assessment, and asset risk control (Banerjee et al., 2021). For companies financed based on DSCF, financial institutions are increasingly incorporating the digitalisation of companies into their credit assessment (Ivanov and Dolgui, 2021). Meanwhile, credit risk assessment models are being improved to accommodate the increasing complexity of the data. The introduction of machine learning methods has contributed significantly to the development of credit risk assessment, but the effectiveness of an extensive range of machine learning models in dealing with the credit risk assessment problem in DSCF remains to be investigated. The motivation of this paper is driven by three aspects: Firstly, the model for credit risk assessment is various and ambivalent. For instance, LR has defaulted to the most common method for credit risk

[22] see Deloitte (2021) at https://www2.deloitte. com/mt/en/pages/technology/articles/mt-what-is-digital-economy.html, accessed on 12 January 2021

assessment even if it shows less non-linear fitting ability in forecasting the credit risk (Denison et al., 2002). While SVM is believed to provide the highest accuracy in forecasting (Khemakhem and Boujelbene, 2017; Danenas and Garsva, 2012), the MLP is also argued to outperform other traditional approaches (Bahnsen and Gonzalez, 2011). The performance of modern machine learning models in empirical data remains to be tested. Secondly, most of the credit risk assessment variables in the existing literature are selected manually, and their selection is subjective and arbitrary, e.g., Wang et al. (2020) summarised the existing literature and came up with four first-level indicators, 11 s-level indicators, and 20 third-level indicators. However, the selection of feature variables for enterprises is diverse and advanced with the time that we cannot clarify the proper indicators for assessment. Thirdly, there are gaps in the research on DSCF, especially from the perspective of credit risk assessment, and most existing articles investigate DSCF from a theoretical perspective, not to mention the lack of a corresponding indicator system. Thus, this chapter uses 1357 observations from 85 Chinese-listed SMEs over the period 2016–2019 as the sample andselects the important feature automatically through XGBoost at the first stage, then compare the performance of MLP and other machine learning models in credit risk assessment. This study enriches the theory and practice of enterprise credit risk assessment in the DSCF environment. The effectiveness of the XGBoost-MLP approach for credit risk assessment in DSCF is investigated. Based on the traditional single credit risk assessment model, the feature selection is considered in the first stage by using XGBoost as the model, and then is compared to each traditional model including LR, KNN, NB, DT, RF, SVM and MLP, and its combination with XGBoost in the second stage. The hybrid method of XGBoost-MLP is observed to have optimal performance, which contributes to the enhancement and development of the theory of enterprise risk assessment models in the DSCF environment, and also provides new ideas to improve the accuracy of enterprise credit risk prediction. Further, the impact of feature selection on credit risk assessment under the XGBoost method is explored in depth by observing the effect of risk assessment models with different feature thresholds. Feature selection plays an important role in credit risk assessment, and selecting the most appropriate features as indicators for credit risk assessment analysis helps to improve the accuracy of the model. This extends the application of traditional credit risk assessment indicator systems and provides strong evidence for banks and other financial institutions to make sound financing decisions. Finally, the study on DSCF features is conducted by comparing the assessment results with and without DSCF features; we find that the credit risk assessment of firms is better when their DSCF features are considered. Based on feature screening, adding indicators of DSCF features further improves the modern credit risk assessment indicator system and enriches the relevant theory. The paper proceeds as follows. In Section 4.2, the background of DSCF and credit risk assessment with machine learning is presented as the literature review. Section 4.3 includes the theory and methodology. Section 4.4 exhibits experimental design. Section 4.5 reports the results and discussion of the experiment. Section 4.6 offers robustness check. Section 4.7 provides the conclusion.

4.2 Literature Review

4.2.1 Background of DSCF

Since the 1970s, driven by rising consumption levels, market demand and minimisation of production costs, there has been a gradual shift in the pattern of division of labour from within a single enterprise to between multiple enterprises. The role of inter-firm coordination and facilitation through new supply chain enterprises, leading to the derivation of a supply chain production model. Timme & Williams-Timme (2000) first introduced the concept of SCF and then Berger et al. (2004) defined SCF from the perspective of SME lending. They argued that SMEs have difficulty in obtaining loans due to a lack of good credit support and proposed a new financing model in which large enterprises or financial institutions control transactions to finance SMEs that are difficult to finance. Initially, supply chain management neglected the flow of capital until the late 20th century when the importance of capital flow to the entire supply chain came into focus and SCF was created. Hofmann (2005) argued that multiple firms and external service participants participate in the management and integration of financial resources to increase the value of all participants in the supply chain. He also innovatively incorporates corporate values by managing the stakeholders in the supply chain to strengthen the corporate culture of the core companies, which can effectively reduce the credit risk in SCF. The core of SCF is composed of financial institutions, core enterprises and information platforms, which focus on financing and cost settlement in the supply chain, thereby optimising and reducing the costs of enterprises in the supply chain (Supply chain Europe, 2007). Further, Camerinelli (2009) defines SCF as the provision of financial services by financial institutions to companies in the supply chain to help them manage logistics and information flows. Lyons et al. (2012) argue that supply chains contain a large number of enterprises with complex structures, and that they can be considered as a whole where countermeasures can be formulated by integrating information on all commodities and materials, information on transactions and financial transactions to ultimately improve the competitiveness of the supply chain.

Digitisation has been a popular trend in recent years, and its application does not happen overnight but is advanced in layers. With the advancement of technology, FinTech represented by AI, blockchain, cloud computing and big data is being deeply integrated with traditional SCF, forming a new generation of DSCF platforms. The root of DSCF is the supply chain. The essence of the supply chain is actually the supply and demand chain, which refers to the chain consisting of a series of supply and demand links from the supply chain to the customer. The supply chain includes physical flow, capital flow and information flow, in which the physical flow and capital flow forming a complete closed-loop, i.e. the use of funds to purchase raw materials, raw materials are converted into products, products are further converted into funds, and then part of the converted funds are used to purchase raw materials again, opening a new cycle[23]. SCF is an activity that brings in external capital when a company is not operating well or when it wants to expand its business. Scholars usually define the concept of SCF: from the supply chain perspective, e.g., Hofmann (2005), Guillén et al. (2007) thought that SCF integrates production and financing into the management framework of a firm's supply chain, and thus manages it in an integrated manner. Gomm (2010), Caniato et al. (2016) believed that SCF uses optimal strategies to plan, manage and control cash flows in the supply chain to help improve the operational efficiency of the supply chain, while Wuttke et al. (2013), Wandfluh et al. (2016) illustrated that SCF can strengthen the relationship between upstream and downstream firms and core firms and optimise the financing structure in the supply chain. From another type of financial perspective, such as Atkinson (2008) and Gobbi and Sette (2014) considered that SCF is a financing business conducted through a third-party trading platform, which can effectively reduce the financing cost of enterprises and improve the cash flow turnover of the supply chain. Jing and Seidmann (2014) and Caniato et al. (2016) argued that SCF is a process of optimising the financial management of the supply chain, focusing on core enterprises and financing institutions.

Compared to traditional SCF, an important feature of DSCF is the "enterprise data on the chain", i.e., the enterprises in the supply chain register and confirm their transaction information on the chain, which is a different way of digitising enterprises than the internet (Goldfarb and Tucker, 2019). For the realisation of this feature, the digitalisation of both financial institutions and enterprises is essential, with the digitalisation of enterprises also playing an important role in the risk control of SCF (see Figure 4.1). Firstly, the information recording, IoT technology plays a role in collecting and recording information, warehouse

[23] https://www.ft.com/content/8ca7b05d-f1a8-4ddd-8fda-3383f11e5143

management system (WMS), supplier relationship management (SRM), customer relationship management (CRM), etc. are all supply chain information collection and recording systems. Secondly, the dissemination of information, in digital form, makes it possible to share, collaborate and monitor information in real-time across locations. Enterprise resource planning (ERP), for example, is called the internal information internet of the enterprise. Thirdly, information processing, i.e., the fast and accurate processing of information, e.g., advanced planning and scheduling (APS) is an information processing system for supply chain management. The new generation platform features intelligent multi-party connection, mutual trust of chain enterprises, multi-level credit penetration and closed-loop ecological risk control, which is expected to drive the development of enterprise financing business in relatively risk-controlled batches by transferring core enterprise credit at multiple levels and closing the loop of funds in an operational manner.

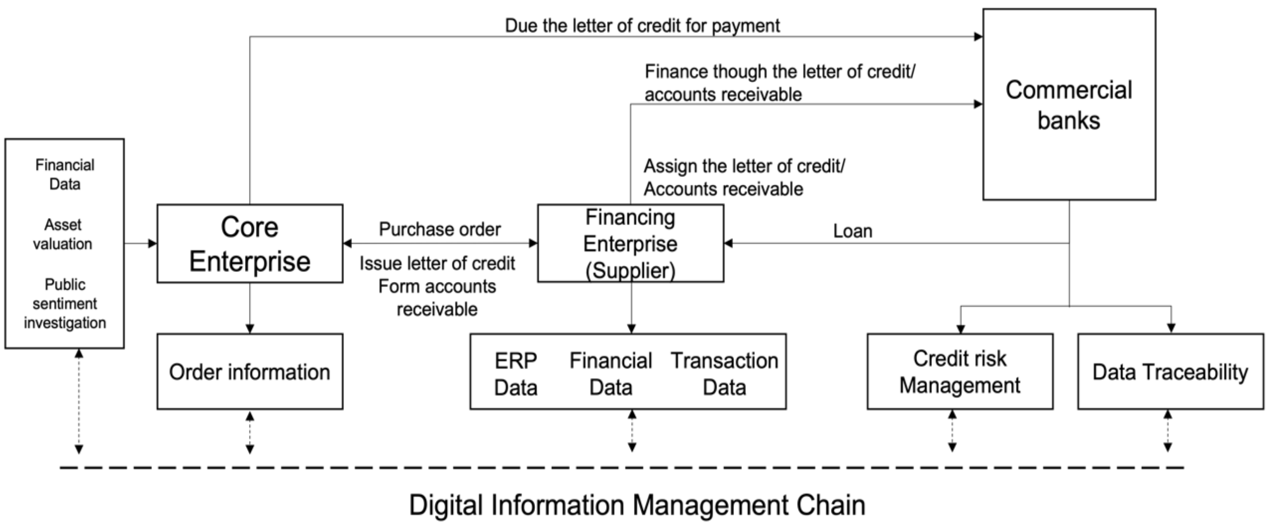


Figure 4.1 Framework of DSCF

Moreover, the role of machine learning method in DSCF is also significant (Olan et al., 2022). DSCF and machine learning have a symbiotic relationship that transforms the landscape of financial operations within the supply chain. The integration of machine learning into DSCF is motivated by the need for enhanced efficiency, risk mitigation, and strategic decision-making. Machine learning algorithms, equipped with the capability to analyze vast datasets, play a pivotal role in predicting market trends, optimizing inventory management, and assessing the creditworthiness of suppliers. By automating routine tasks, such as invoice processing and fraud detection, machine learning not only reduces errors but also frees up valuable human resources for more strategic endeavors. The real-time monitoring capabilities of machine learning ensure that financial decisions align with current market conditions, contributing to agile and adaptive supply chain finance strategies. Ultimately, the motivation behind applying machine learning in digital supply chain finance lies in its ability to drive innovation, improve accuracy, and foster a more resilient and responsive financial ecosystem within the dynamic landscape of global supply chains.

4.2.2 Machine Learning and Credit Risk Models

The issue of credit risk assessment in SCF has attracted the attention of scholars. Hallikas et al. (2002) used internal audit and computer cameras and analysed the causes of risk through interviews with two core enterprises and nine suppliers and classified the risks of SCF into four parts: demand, transaction, pricing and finance. Finch (2004) analysed the literature on the need for core firms to determine whether to use SMEs as a supply partner for critical operations and to establish appropriate information systems for review and found that improved information management of SMEs contributed to credit risk reduction. Yurdakul & İç (2004) developed a credit assessment and decision-making model for determining the credibility of manufacturing firms. Ghadge et al. (2013) developed a holistic, systematic and quantitative risk assessment process to measure overall risk behaviour. By capturing dynamic risks in case studies of manufacturing firms, the overall risk impact of SCF can be predicted and a whole picture of risk behaviour exhibited is constructed. With the gradual improvement of SCF applications, the credit risks they face are becoming increasingly complex. Subjective assessments based on experience and traditional linear models are no longer able to accurately predict risks, and assessment models based on machine learning techniques are now more popular. Many research results have been achieved in the assessment of enterprise credit risks in SCF. Zhu et al. (2017) used an integrated ensemble machine learning approach to assess SME credit risk in Chinese SCF. The RS-boosting method was found to outperform other methods in improving the accuracy of risk prediction. Zhu et al. (2019) further used a new hybrid ensemble machine learning method, RS-MultiBoosting, which improve the accuracy of credit risk assessment based on the SCF in China. Wang et al. (2020) then explored the mechanism of online SCF using Least Square Support Vector Machine (LS-SVM) method and found that LS-SVM method has higher accuracy in online SCF risk prediction.

Due to the complexity of credit risk, there are various models for credit risk assessment, which have undergone a series of improvements since their development. Prior to 1970, financial institutions such as commercial banks mainly carried out qualitative analysis of financing companies by professionals and credit assessment was more subjective. The methods used included expert scoring and profiling. After 1970, financial institutions used ZETA scoring models, Z-score models and other statistical distributions to assess the credit risk of financing companies. Orgler (1970) studied credit risk based on the characteristics of linear regression, and later linear regression methods also provided many references to credit risk assessment problems (Fitzpatrick, 1976; Lucas, 1992; Henley, 1995) However, in view of the shortcomings of linear discriminatory methods, non-linear statistical models such as LR and Probit have emerged as commonly used models for multivariate credit risk assessment. Wiginton (1980) assessed risk on the basis that LR can explain problems where the variable is a qualitative indicator, and Steenackers & Goovaens (1989) made a related follow-up application of personal loans. Cramer (2004) systematically investigated LR and showed that LR was more accurate in classification and that its low assumptions and high stability made it one of the most widely used methods for credit risk assessment. Profit regression was used by Grablowsky & Talley (1981) in their study of credit risk and the results showed that profit regression did not have as good an interpretation as LR.

Further, the classification tree method was first applied to credit risk assessment by Makowsik (1985), whose results were compared and confirmed its high accuracy in credit assessment applications, with the advantage of automatic variable selection and better handling of missing information (Carter & Catlett, 1987). While Cover (1968) proposed the K-Nearest Neighbour (KNN) discriminant method, and then Henley et al. (1996) applied the KNN analysis method to personal credit assessment and confirmed the feasibility of KNN in credit risk assessment. Hand (1981) used the KNN method and DT to identify loan risk and the results showed that the KNN method had better prediction accuracy. Subsequently, Bayesian algorithms were proposed by Pearl (1988) and have been used to good effect in the areas of representation of uncertain knowledge and inference. The research of Hsieh (2010) showed that Bayesian networks enable to intuitively represent the relationship between attributes and probabilities and have good explanatory power. As Bayesian classification models combine prior knowledge and sample information and use probability tables to quantify the dependencies between variables with better classification accuracy, they have attracted increasing attention from scholars. The Naïve Bayesian (NB) classification algorithm (Friedman et al., 1997), a milestone in Bayesian classification research, assumes that all feature variables are independent of each other where the class node is the parent of all attribute nodes in the structured graph, with no arcs between any other attribute nodes. A good classification with a simple structure can be obtained using an NB classifier when the correlation between feature variables is small, but its strict conditional independence is often not achieved under realistic conditions thereby greatly reducing its classification effectiveness (Langley et al., 1992).

As the application of machine learning methods in credit risk assessment continues to evolve, do Prado et al. (2016) used the Web Science database to analyse the journal literature on credit risk and bankruptcy research published between 1968 and 2014 using bibliometric methods. They found that LR has been a common approach though since Odom and Sharda (1990) first used Artificial Neural Network (ANN) for credit risk assessment, AI techniques represented by neural networks have been used more and more widely, and multiple or hybrid models with sophisticated AI techniques are a trend for further research. Since credit risk assessment models based on AI techniques do not require strict assumptions to be made and have advantages in dealing with non-linear problems (Denison et al., 2002), they have become more popular when facing increasingly complex credit risk. Davis et al. (1992) conducted a case study of neural networks in personal credit assessment and found that the neural network method was more accurate in classification, but the training time for the neural network data was longer. Desai et al. (1997) also used neural networks in personal credit assessment and showed that their performance was better. Piramuthu (1999) developed a neural network survival model using multi-layer perceptron (MLP) neural networks and fuzzy neural network-related principles. Lee & Chen (2005) used neural networks and the related theory of multivariate adaptive spline regression to investigate the feasibility of applying the related theory to credit assessment. Tsai (2008) applied the principles of MLP neural networks to corporate bankruptcy prediction and credit assessment. Marcano-Cedeño et al. (2011) developed a plasticity neural network model and then conducted an empirical study using relevant data. Coincidently, the theory of Support Vector Machine (SVM) was first proposed by Cortes and Vapnik in 1995, and SVM has quickly become a hot topic of research in machine learning in recent years. Stecking and Schebesch (2005) selected different kernel functions and then analysed the impact of these kernel functions on credit appraisal. Lai et al. (2006) modelled the problem of credit assessment and verified the feasibility of the theory in credit assessment by using the theory related to least squares support vector machines. Schebesch and Stecking (2008) developed a credit assessment model by combining these principles through a study of combined support vector machines and imbalanced data sets. Yu et al. (2010) developed a credit risk assessment model based on hybrid intelligent mining, in which rough set theory and the related theory of support vector machines were used.

Controversy surrounds the choice of credit risk assessment models. The advent of SVM has provided excellent algorithms for classification models, with a large number of kernel functions available for flexible solutions to a wide range of non-linear classification regression problems. However, model selection is also the main problem with SVMs, as the selection of kernels and the optimisation of kernels and regularisation parameters can often lead to severe overfitting if the model selection criteria are over-optimised, while the emergence of ANNs has effectively bridged the shortcomings of traditional methods. ANNs are widely used for the estimation and prognosis of complex processes due to their ability to classify research populations in complex environments using large amounts of uncertain information. The advantage of ANNs is that they do not require a strict distribution of the data, nor do they require a detailed representation of the function between the independent and dependent variables, and they are effective in solving non-normally distributed nonlinear credit assessment problems. However, neural networks also have their disadvantages, namely the long training time and the difficulty in identifying the relative importance of the input variables to obtain the optimal network. Among the ANNs, MLP neural networks have been used in risk assessment due to their outstanding performance.

4.3 Methodology

To accurately and effectively conduct a credit risk assessment, we construct the following model (See Figure 4.2). In the first stage, through feature selection, we extract the training sample set and select the features with higher scores based on the importance score of the calculated features. In the second stage, MLP is used for credit risk assessment based on the selected features. As credit risk assessment can essentially be seen as a classification problem, the MLP is used as a classification model in the credit assessment process. Further, the trained model is used to test the test set and ultimately, this research validates the proposed research question. Specifically, given the training set $\boldsymbol{X}$, $\boldsymbol{x_i}$ represents the original features as the input of credit risk assessment and $\boldsymbol{y_i}$ is the label of credit status ($\boldsymbol{Y = 0\ or\ 1}$, i.e. risky/non-risky). $\boldsymbol{X = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}}$. Based on the importance ranking $\boldsymbol{r_n}$ by XGBoost classifier in the first stage, we filter the features by thresholds $\boldsymbol{t_i}$ and remove features $\boldsymbol{x_n}$ where $\boldsymbol{r_n < t_i}$. Then the remaining features in subset are obtained $\boldsymbol{X'}$ as the input for retraining with MLP. Through the $\boldsymbol{ReLU}$-based MLP, $\boldsymbol{ReLU}$ as the activation function is more expressive for linear functions. For non-linear functions, $\boldsymbol{ReLU}$ does not have the vanishing gradient problem as the gradient of the non-negative interval is constant, allowing the convergence rate of the model to be maintained in a steady state. Thus, we obtain the output of $\boldsymbol{y = f(x) = \max(0, x)}$ and the performance of the model.

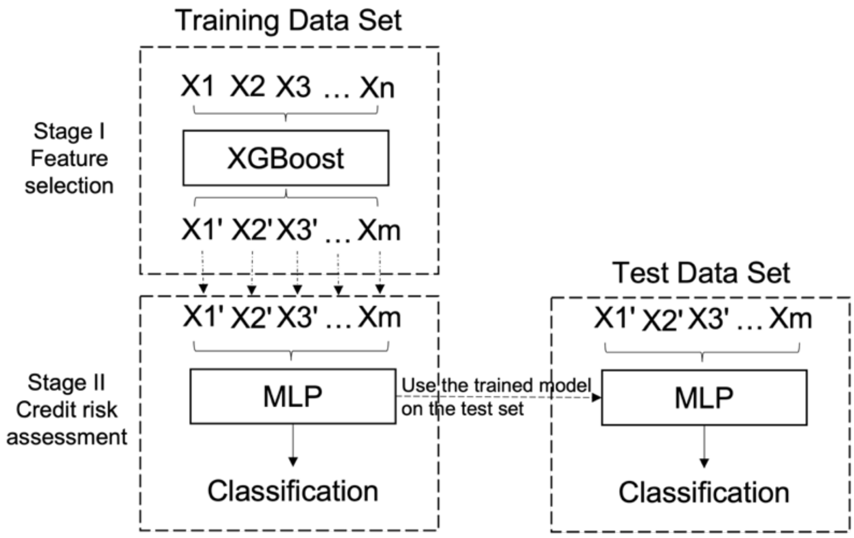


Figure 4.2 The flowchart of XGBoost-MLP.

4.3.1 Stage I: Feature Selection with XGBoost

XGBoost is an improved algorithm based on Gradient Boosting Decision Trees (GBDT) proposed by Chen and Guestrin (2016), which can efficiently build augmented trees and run in parallel. This is an ensemble learning method that the basic idea is to select some samples and features to generate a simple model (e.g. a DT) as the basic classifier and to learn the residuals of the previous model, minimise the target function and generate a new model, which is repeated to produce a combination of hundreds of linear or tree models with high accuracy. At its core, the new model is built in the direction of the corresponding gradient of the loss function, correcting for residuals while controlling complexity. Thus, the dataset in our paper containing n examples with e features is denoted as $X = \{(x_i, y_i): x_i \in R^e$, $y_i \in R, |X| = n\}$ and the set of all classification and regression trees (CART) (1984) is denoted as $F = \{f(x) = w_{q(x)},\ q: R^e \to T,\ w \in R^T\}$ where q is the rule structure for mapping the samples to the corresponding leaf nodes, T is the number of leaf nodes in a tree, and w is the weight of the leaf nodes. f represents the CART, including the structure of the tree q and the weight of the leaf nodes w. CART decision trees are divided into regression trees and classification trees, and CART regression trees, which assume that a DT is a binary tree. It constructs a DT by continuously splitting the features (into left and right halves). The predicted value of y_i based on the XGBoost algorithm can be expressed as:

$$\hat{y}_i = \theta(x_i) = \sum_{k=1}^{K} f_k(x_i) \tag{20}$$

Where $f_k \in F$ and K is the number of CART. $f_k(x)$ represents a DT, that function f can be interpreted as mapping the sample x into some leaf node of the tree, and each leaf node in the tree will correspond to a weight w.

A general objective function is firstly considered as follow:

$$obj(\theta) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{21}$$

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda\|w\|^2 \tag{22}$$

Among them, l is a derivable and convex loss function, which is used to measure the similarity between $\hat{y}$ and y. The second term Ω is a regular term, which contains two parts. The first one is γT, where T is leaf The number of nodes, γ is a hyperparameter that if γ is larger, the number of leaf nodes will be smaller. The other part is the L2 regularisation term, which penalises the weight of the leaf nodes so that there will be no leaf nodes with too large weights to prevent overfitting.

It is difficult to optimise and minimise the above objective function Eq. (22), so we transform it by greedily optimising the objective function by adding a base classifier f_t at each step so that each time it is added, the loss becomes smaller. In this way, an evaluation function is obtained that can be used to evaluate the performance of the current classifier f_t.

$$obj(t) = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{23}$$

Where y_i is the i target and $\hat{y}_i^t = \hat{y}_i^{(t-1)} + f_t(x_i)$ is the prediction for the tth iteration. Eq.(23) can also be called forward stepwise optimisation. To optimise this function more quickly, we do a second-order Taylor expansion at $f_t = 0$.

$$obj(t) \approx \sum_{i=1}^{n} \left[l\left(y_i, \hat{y}^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \gamma T + \Omega(f_t) \tag{24}$$

Where g_i denotes the first order partial derivative of l with respect to f and h_i denotes the second order partial derivative of l with respect to f.

$$g_i = \partial_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right) \tag{25}$$

$$h_i = \partial^2_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right) \tag{26}$$

Then the total number of samples is defined as n , each sample as i , the information of i is divided into some leaf node information and define the weight of each leaf belonging to i as j. $I_j = \{i|q(x_i) = j\}$ is the instance set of leaf i.

$$\begin{aligned}\widetilde{obj}(t) &= \sum_{i=1}^{n}\left[g_i f_t(x_i) + \frac{1}{2}h_i f_t^2(x_i)\right] + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2 \\ &= \sum_{j=1}^{n}\left[\left(\sum_{i\in I_j} g_i\right)w_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right)w_j^2\right] + \gamma T\end{aligned} \tag{27}$$

Define the $G_i = \sum_{i\in I_j} g_i$, $H_i = \sum_{i\in I_j} h_i$, then let the current function derivative of w be 0. At this point the objective function becomes quadratic with respect to w. The optimal weight for the fixed $q(\mathrm{x})$ is:

$$w_j^* = \frac{G_j}{H_j + \lambda} \tag{28}$$

Substituting Eq.(28) into the objective function gives:

$$\widetilde{obj}^* = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j + \lambda} + \gamma T \tag{29}$$

When selecting features for XGBoost-based classification, feature importance is integrated into the classification process. A new tree is created in each iteration, and the branch nodes in the tree are a feature variable, and the importance of these nodes is calculated. The importance of a feature is based on the squared improvement of the split nodes of the tree that a feature is selected for. Each time a feature is selected to be added to the tree as a splitting node, all possible splitting points are enumerated using a greedy algorithm, from which the splitting point with the best gain is selected. The best splitting point corresponds to the maximum gain and the gain is calculated by the formula:

$$G_{ain} = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \tag{30}$$

where the I_L and I_R are the instance sets of left and right nodes after splitting. Relevant features and split points improve the squared difference on a single tree, and the more improvement there is, the better the split point and the more important the feature is. When all trees are built, the calculated node importance is averaged over the forest. The more times a feature is selected as a split point, the more important it will be.

4.3.2 Stage II: Credit Risk Assessment Models

As part of the second stage, we utilise several models, namely a MLP, KNN, NB, DT, RF, and SVM. MLP, an ANN with forwarding agency that maps a set of input vectors to a set of output vectors, is thought of as a directed graph, consisting of multiple layers of nodes, each layer fully connected to the next (see Figure 4.3). In addition to the input nodes, each node is a neuron (or processing unit) with a non-linear activation function. A supervised learning method known as backpropagation is often used to train MLPs, which overcomes the weakness of the perceptron in their inability to recognise nonlinear data. MLP has been shown to be a general function approximation method that can be used to fit complex functions or to solve classification problems.

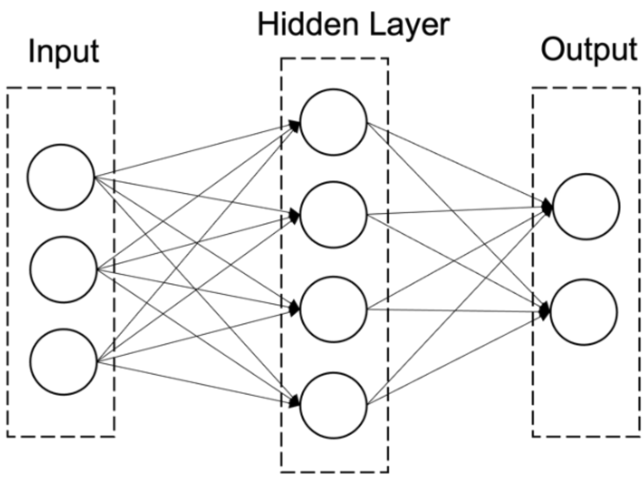


Figure 4.3 Structure of Multi-layer perceptron.

The LR model is widely used in corporate credit risk assessment research, and the LR model is used to calculate the relationship between the dependent variables and the independent variables as well as the strength of the relationship (Crook et al., 2007). In this paper, the subject of credit risk research, i.e. enterprise in the digital supply chain financial environment, is divided into two categories: one category is risky SMEs; the other category is non-risky enterprises, and the binary LR method is used to assess the credit risk based on the DSCF environment.

$$\ln\left(\frac{p}{p'}\right) = \beta_0 + \sum_{j=1}^{n} c_j \beta_j \tag{31}$$

Where $\ln\left(\frac{p}{p'}\right)$ is the dependent variable, i.e. the risky and non-risky enterprises assigned 0 and 1 respectively. p represents the probability of non-defaults and p' represents the probability of defaults. c_j (j=1, ... , n) is the independent variable which explains the value of variable j, i.e. the feature j of enterprise. β_j $(j = 1, \dots, n)$ is the coefficient of each independent variable.

KNN is a non-parametric estimation method in the field of pattern recognition (Cover and Hart, 1967). The algorithm is simple, fast and efficient, and the idea is to assume that a sample data x to be recognised, where most of the k nearest neighbour training sample representative points in the feature space belong to one of the categories, then $\boldsymbol{x}$ also belongs to this category. The sample vector $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$, the elements of the vector are the observed values of each sample feature. If the samples in the training set are divided into h classes, $C_1, C_2, \dots, C_h$, $C = (x_1, x_2, \dots, x_m)$, indicates that there are m samples in class i of the sample set. The similarity between the samples of $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$ and all the samples in the training sample set is first calculated, and the k samples that are most similar can be selected according to certain principles of similarity, and these K samples belong to class C_i. Assuming that there are attributes of the samples, the attribute indicators of the sample data can form a z-dimensional feature space, and all sample points have a unique point corresponding to it in this z-dimensional feature space, for any sample x to be identified can be put into this z-dimensional feature space, and by constructing a distance formula (generally using Euclidean distance), the k-nearest neighbours of the sample x can be found. Thus, for a sample training set given two samples $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$ and $x_j = (x_{j1}, x_{j2}, \dots, x_{jm})$, the Euclidean distance is as follow.

$$D(x_i, x_j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \cdots + (x_{im} - x_{jm})^2} \tag{32}$$

The NB classifier is the simplest Bayesian classifier with the advantage of high efficiency and good classification accuracy (Rish, 2001; Antonakis and Sfakianakis, 2009). In its structure, the class variables are treated as parents of the other attribute variables, and it is assumed that the attribute variables are independent of each other, provided that the class variables are known. Figure 4.4 gives a graphical depiction of a simple Bayesian classifier.

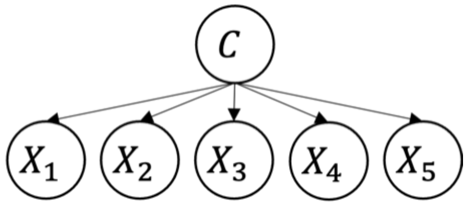


Figure 4.4 Naïve Bayesian classifier.

Assume $A_1, A_2, \dots, A_n$ are the n attribute variables of dataset D, $C_1, C_2, \dots, C_n$ are the m classes of dataset D, $X = \{X_1, X_2, \dots, X_n\}$ is a specific object, then the probability that this object belongs to class C_i can be calculated using the Bayesian formula to calculate:

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \tag{33}$$

Bayesian classification considers a given object $X = \{X_1, X_2, \dots, X_n\}$ belongs to the class with the highest posterior probability, under the premise of the Naïve Bayesian assumption, the description $(X_1, X_2, \dots, X_n|C)$ is simplified as follow:

$$P(C_i|X) = P(x_1, x_2, \dots, x_n|C_i)$$

$$= P(x_1|C_i)P(x_2|C_i, x_1)P(x_3|C_i, x_1, x_2) \dots P(x_n|C_i, x_1, x_2, \dots, x_{n-1})$$

$$= \prod_{j=1}^{n} P(x_j|C_i) \tag{34}$$

The DT is a binary tree decision method similar to that used in risk management theory and a conditional branching structure in discrete mathematical flowchart theory, where probability calculations are used to classify the categories. The DT model makes an inductive classification algorithm that learns from a sample of training data and then suitable decision rules are then used to analyse the test data samples. The structure of a DT consists of a root node, i.e., the decision point, internal nodes, and leaf nodes. Each internal node is the location where the root data attribute selection metric is performed and the attribute values are used to calculate the attributes of the classified sample, representing the conditions for conducting the test. The leaf nodes are the category identifiers obtained after the splitting is completed. In addition, each DT has a number of branches, with more branches representing the complexity of the classification process and each internal node of the DT is tested using only one input dimension. If a variable in the input dimension used is discrete, the internal node calculates the attribute value of that variable and then selects the corresponding branch. The DT algorithm divides the data into subsets depending on whether the selected attribute is discrete or numerical. The corresponding subsets are then divided recursively until the division is no longer required and a leaf node is placed to identify it. There are many classification algorithms for DTs, including the Iterative Dichotomiser 3 (ID3) algorithm and the C 4.5 algorithm proposed by Quinlan (1996).

The RF method is a classification model based on DT theory, but which differs from DT in that the RF does not generate only unique trees, and randomly uses variables and data in the process of generating DTs (Breiman, 1999). It is also known as a random DT because it uses variables and data randomly in the process of generating a DT that contains multiple DTs. RF contains the idea of integrated learning, which means that weak classifiers are learned and trained to combine into strong. In the RF model, this integrated learning theory is based on the Bagging algorithm (Boostrap aggregating). The difference is that the RF model creates a DT by splitting the set of attributes for random selection.

SVM is the linear classifier first proposed by Cortes & Vapnik (1995). The SVM has advantages in solving small-sample, non-linear and high-dimensional pattern recognition (Cusano et al., 2003). For non-linear problems, a non-linear transformation $\boldsymbol{f(x)}$ is used to map the input data into a high-dimensional feature space, and then go for linear classification in the high-dimensional feature space, which solves the low-dimensional space (Bao et al., 2019). In a linear classifier, the classifier is a hyperplane $f(x) = wx - b = 0$, If $f(x) > 0$, then the point belongs to class 1, and if $f(x) < 0$, then the point belongs to class -1. The optimal partitioned hyperplane constructed by the SVM is the one that maximises the maximum of the shortest distance from a point in class 1 to the hyperplane and the shortest distance from a point in class - 1 to the hyperplane, the solution showed as the following function, resulting in the weight vector w and offset b.

$$\min\Theta(w) = \frac{1}{2}\ \|w\|^2 \tag{35}$$

$$s.t\ \ y_i(w * x_i + b) \geq 1, i = 1, 2, \ldots, n \tag{36}$$

The linearly indivisible problem in low-dimensional space could be transformed into a linearly divisible problem in high-dimensional feature space by Kernel function, basically including linear, polynemoid, radial bias function and sigmoid.

4.4 Experimental Setup

To compare the performance of the XGBoost-MLP model with other traditional models for credit risk assessment of DSCF, we selected listed SMEs in China as the data sample. SMEs in China are a major demand-side of SCF which is certainly representative. Nevertheless, a controversy arises when examining the actual financial constraints faced by listed versus unlisted SMEs. Listed and unlisted SMEs in China differ primarily in terms of their status on the stock exchange. Listed SMEs are companies that have undergone an initial public offering (IPO) and have their shares traded on a stock exchange, while unlisted SMEs are privately held companies that do not have their shares publicly traded. This study primarily concentrates on listed SMEs for two key reasons. For one thing, the studies on SCF in China are relatively limited, and it is difficult to collect relevant data. The information disclosures are also lacking transparency. SMEs listed on the SME board offer greater public information. A second aspect, concerning financial constraints, listed firms face heightened scrutiny from investors, analysts, and regulatory bodies. This examination can generate pressure to maintain consistent performance and meet market expectations, thereby constraining their financial flexibility. Thus, listed SMEs are considered as a more representative sample for this experiment.

This chapter firstly select listed SMEs as the main subject of the credit risk assessment, which represents the main target of supply chain financial services. Secondly, large enterprises listed on the Main Board are selected as the core enterprises, which have the strong financial strength and enable them to act as important guarantors in the supply chain. The requirements for listing on the Main Board are the highest, with the listing criteria requiring the company to be established and in operation for at least three years, and to be profitable for three years, with an aggregate of more than RMB 30 million, and the company’s net cash flow from operations for three years to exceed an aggregate of RMB 50 million. The company is also required to have a cumulative total of more than RMB 300 million over three years, plus a total pre-issue share capital of not less than RMB 30 million. Companies that can successfully list on the Main Board are in a leading position in a certain industry [24]. Thirdly, the selected SMEs have real trading relationships with the core enterprises, and they are suppliers or customers of the core enterprises. Based on the above selection criteria, this chapter selectes 85 listed SMEs from 31 March 2016–31 December 2019 from the Small and Medium Enterprise Board of the Shenzhen and Shanghai Stock Exchange including a quarterly 1357-observations dataset of risky and non-risky enterprises.

All companies selected are private manufacturing companies that have been listed for more than 10 years. Although this method of data collection is commonly used in the existing literature (Zhang et al., 2015; Zhu et al., 2017; Zhu et al., 2019), it has certain limitations that make the results susceptible to error. Hence, certain improvements have been made on this basis. Firstly, most of the relevant data samples are collected through questionnaires on non-financial data related to the supply chains which is somewhat subjective and arbitrary and can bias the experimental results. Thus, we use publicly available financial data for the SCF part of the feature data to be measured. Secondly, the existing literature mostly takes SCF or online SCF as the research object, and there are gaps in research on the characteristics of the DSCF. In this paper, through the analysis and investigation of DSCF, digital features are added to the credit risk assessment. Thirdly, there are few data treatments in the existing literature that focus on feature selection. Zhu et al. (2017) use the DT to evaluate data samples and derive important rankings before conducting classification assessment. Although the algorithm of DT is simple and interpretable, the risk of overfitting is great, and the application scenario is limited. This research uses XGBoost as the first stage feature selection method, which improves based on GBDT by adding a regular term to the objective function of each iteration to further reduce the risk of overfitting, thus improving the performance for feature selection.

Further, to assess the likelihood of default among SMEs, this study employs ST and ST* stocks as indicators of the default sample. These stocks represent enterprises under special treatment with a risk alert. In the context of China's financial landscape, "special treatment" refers to the differentiated regulatory measures or interventions applied to certain companies that are facing financial distress or other issues. Companies under special treatment face

[24] https://www.szse.cn/English/products/equity/mainboards/index.html, accessed on 5 February 2021

restrictions on accessing new loans or credit facilities (Jiang and Jones, 2018). The regulatory intervention often implies that the company's financial health is compromised, making lenders hesitant to extend credit. Limited access to funding exacerbates liquidity issues and heightens credit risk.

The 85 listed SMEs comprise 11 enterprises under special treatment under risk alert, i.e., ST and *ST stocks, which are regard as risky SME with negative credit status, and 74 enterprises with normal financial status. Thus, the dependent variables are classified into two groups on the basis of the credit status, the dependent variables are assigned the value of 0 or 1 which indicates the risky and non- risky enterprises. This experiment selects 30% of the data set as the test set, i.e., 408 observations, with 49 negative examples and 359 positive examples, and 949 observations in the training set, with 127 negative examples and 822 positive examples.

In addition, the confusion matrix and its derived assessment metrics are used to evaluate the results of the sample data. In this paper, positive samples are creditworthy, i.e., risk-free firms, and negative samples are bad creditworthy, i.e., risky firms. The parameters mentioned below are calculated based on a confusion matrix shown in Table 4.1. True positive (TP) refers the number of defaults that are correctly predicted as defaults; false positive (FP) refers the number of non-defaults that are mistakenly predicted as defaults; true negative (TN) refers the number of non- defaults that are correctly predicted as non-default; false negative (FN) refers the number of defaults that are mistakenly predicted as non-defaults. The parameters used in this work are calculated with the following equations (Eqs. (37)–(43)).

Table 4.1 Confusion matrix.

		Actual condition	
		Positive (non-risky)	Negative (risky)
Test result	Positive (non-risky)	Ture positive (TP)	False positive (FP)
	Negative (risky)	False negative (FN)	Ture negative (TN)

The accuracy rate represents the proportion of correct samples to the total sample：

$$Average\ Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{37}$$

Precision indicates the number of samples that are predicted to be positive that are truly positive, and Recall indicates the number of positive cases in the sample that was correctly predicted. Precision is specific to the predicted output and recall is specific to the original sample. Type I error is defined as the number of true negative samples incorrectly predicted to be positive as a proportion of the number of all true negative samples. While Type II error is defined as the number of true positive samples incorrectly predicted to be negative as a proportion of the number of all true positive samples.

$$Precision = \frac{FP}{TP + FP} \tag{38}$$

$$Recall = \frac{TP}{TP + FN} \tag{39}$$

$$Type\ I\ Error = \frac{FP}{TP + FN} \tag{40}$$

$$Type\ II\ Error = \frac{FN}{TN + FP} \tag{41}$$

The F-measure is the composite index based on the accuracy and recall, the closer the F-measure is to 1, the better the classification model is.

$$F - Measure = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{42}$$

The Matthew correlation coefficient (MCC) considers true and false positives and false negatives and is often seen as an unbalanced measure that can be used even if these categories are of varied sizes.

MCC is the correlation coefficient between the observed category and the predicted binary category; it returns a value between -1 and +1. A coefficient of +1 indicates a perfect prediction, 0 indicates no better than a random prediction, and -1 indicates a complete inconsistency between prediction and observation.

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{43}$$

4.5 Experimental Results

Following the existing literature (Zhu et al., 2019; Wang and Ma, 2012; Wang et al., 2020), 17 independent variables are selected. Table 4.2 defines the variables for enterprise credit risk analysis based on the DSCF and Table 4.3 presents the descriptive statistics of all data.

Table 4.2 Variables for enterprise credit risk analysis.

Groups	Independent Variables
Status of financing company	Current ratio of SMEs
	Quick ratio of SMEs
	Working capital turnover of SMEs
	Accounts Receivable Turnover Ratio of SMEs
	Rate of return on total assets of SMEs
	Total assets growth rate of SMEs
	Credit rating of SME (The evaluation of SMEs creditworthiness is divided into 10 grade)
Status of core enterprise	Quick ratio of the CE
	Total assets growth rate of the CE
	Rate of return on total assets of the CE
	Credit rating of CE (The evaluation of CEs creditworthiness is divided into 10 grade)
Status of Supply chain	Transaction amount / SME sales or cost of sales (sales when the SME is upstream, cost of sales when the SME is downstream)
	Transaction amount/cost of sales of the core enterprise (sales when the core enterprise is an upstream supplier, cost of sales when the core enterprise is a downstream purchaser)
	Average rate of return on total assets in the industry
Status of digitalisation	Age of online platform construction
	Enterprise Resource Planning (ERP) system application (1/0)
	Age of ERP system application

Table 4.3 Descriptive statistics.

Code	Observations	Mean	Std. Dev	Minimum	Maximum
SME_CurrentRatio	1357	2.327	2.327	0.162	45.316
SME_QuickRatio	1357	1.802	1.972	0.161	45.191
SME_WorkingCapitalTurnover	1355	0.502	5.859	-3.101	189.143
SME_AccountReceivableTurnover	1319	12.710	92.873	0.000	1736.194
SME_ROA	1357	0.029	0.059	-0.909	0.248
SME_TotalAssetGrowthRate	1357	0.091	0.327	-0.579	5.779
SME_CreditRating	1357	8.757	1.365	2.000	10.000
CE_QuickRatio	1348	1.570	1.587	0.000	19.821
CE_TotalAssetGrowthRate	1348	0.163	0.304	-0.708	2.587
CE_ROA	1348	533.465	285.922	1.000	946.000
CE_CreditRating	1357	4.850	3.511	1	10
TransactionAmount/SME	1357	96.944	57.707	1.000	197.000
TransactionAmount/CE	1357	57.027	60.218	1.000	185.000
AverageIndustryROA	1357	3.248	1.481	1.000	5.000
ERP_Age	1324	4.546	5.348	0.000	19.000
ERP_Usage	1325	0.649	0.487	0.000	1.000
PlatformAge	1325	5.629	5.323	0.000	19.000

Note: This table presents the descriptive statistics of original features for credit risk assessment based on the existing works of literatures. The raw data is collected from CSMAR, Wind and annual reports manually.

4.5.1 Model Performance Evaluations

To compare the performance of the proposed XGBoost- MLP model with other machine learning models, LR, DT, SVM, RF and MLP were chosen as the single model for comparison, as well the hybrid model of XGBoost with DF, SVM, and RF. The results of XGBoost-MLP and other machine learning results using out-of-sample tests are shown in Table 4.4, the accuracy of XGBoost-MLP is the highest of the full sample (0.983). Compared to the average accuracies of LR (0.909), DT (0.936), SVM (0.961) and RF (0.966), the single machine learning model is overall lower than the hybrid XGBoost model although the MLP has a better classification evaluation among them. The comparison of the hybrid models shows that XGBoost - MLP has the best results, which validates our first research question. Further, the XGBoost-MLP model achieves reliable results for both recall and precision, and the XGBoost-MLP model has the highest F-measure score of 0.994 compared to other models, which indicates a well-balanced precision and recall. Type I error indicates the weight of this false-positive case, i.e., enterprises that are expected to be risky are judged to be risk-free, which is unfavorable for credit risk assessment. The Type I error of XGboost-MLP is 0.014, which is the lowest among the models measured, which is beneficial for credit risk assessment. In addition, MCC shows that the XGBoost- MLP has the best performance, i.e., 0.922.

Table 4.4 Performance of XGBoost-MLP and other machine learning methods. (Out-of-sample)

	Average accuracy	Recall	Precision	Type I error	Type II error	F-measure	MCC
LR	0.909	0.994	0.910	0.098	0.038	0.950	0.508
KNN	0.946	0.983	0.956	0.045	0.115	0.969	0.741
NB	0.897	0.938	0.943	0.056	0.423	0.941	0.545
DT	0.936	0.978	0.951	0.051	0.154	0.964	0.693
SVM	0.961	1.000	0.957	0.044	0.000	0.978	0.814
RF	0.966	1.000	0.962	0.039	0.000	0.981	0.839
MLP	0.973	0.986	0.983	0.017	0.009	0.985	0.922
XGBoost-KNN	0.953	0.986	0.961	0.039	0.096	0.974	0.776
XGBoost-NB	0.912	0.952	0.947	0.053	0.327	0.949	0.607
XGBoost-DT	0.963	0.986	0.972	0.028	0.096	0.979	0.921
XGBoost-SVM	0.963	1.000	0.960	0.042	0.000	0.979	0.826
XGBoost-RF	0.973	1.000	0.970	0.031	0.000	0.985	0.875
XGBoost-MLP	0.983	0.994	0.986	0.014	0.038	0.994	0.922

Notes: All models presented in Table 4.4 are estimated based on the out-of-samples test. Results are estimated based on the training set of 949-observations and the test set of 408-observations from 31 March 2016- 31 December 2019. LR is the logistic regression model; KNN is the k-nearest-neighbor model; NB is the naïve Bayes model; DT is the decision tree model; SVM is the support vector machine model with radial bias function as kernel function; RF is the random forest model; MLP is the multi-layer perceptron model.

In addition, to assess the effectiveness of the algorithm, this chapter also presents the performance of models by in-sample test (See Table 4.5). The average accuracy score of models is all higher than the results of the out-of-sample test and the results of XGBoost based models are close to 1, which indicates that the models are well trained.

Table 4.5 Performance of XGBoost-MLP and other machine learning methods. (In-sample)

	Average accuracy	Recall	Precision	Type I error	Type II error	F-measure	MCC
LR	0.917	0.991	0.919	0.087	0.056	0.954	0.573
KNN	0.978	0.994	0.981	0.019	0.040	0.987	0.900
NB	0.902	0.948	0.939	0.061	0.347	0.944	0.558
DT	1.000	1.000	1.000	0.000	0.000	1.000	1.000
SVM	0.967	1.000	0.964	0.037	0.000	0.982	0.850
RF	1.000	1.000	1.000	0.000	0.000	1.000	1.000
MLP	1.000	1.000	1.000	0.000	0.000	1.000	0.995
XGBoost-KNN	0.978	0.996	0.979	0.022	0.024	0.987	0.899
XGBoost-NB	0.906	0.958	0.935	0.067	0.274	0.947	0.558
XGBoost-DT	1.000	1.000	1.000	0.000	0.000	1.000	1.000
XGBoost-SVM	0.969	1.000	0.966	0.035	0.000	0.983	0.870
XGBoost-RF	1.000	1.000	1.000	0.000	0.000	1.000	1.000
XGBoost-MLP	0.999	1.000	0.999	0.001	0.000	0.999	0.995

Notes: All models presented in Table 4.5 are estimated based on the in-samples test. Results are estimated based on the training set of 949-observations from 31 March 2016- 31 December 2019. LR is the logistic regression model; KNN is the k-nearest-neighbor model; NB is the naïve Bayes model; DT is the decision tree model; SVM is the support vector machine model with radial bias function as kernel function; RF is the random forest model; MLP is the multi-layer perceptron model.

4.5.2 The Impact of Feature Selection

To further validate the impact of XGBoost feature selection on credit risk assessment, the importance of all the features is first to ranked, and Figure 4.5 shows the XGBoost feature importance ranking, with the horizontal axis showing the threshold of the selected features.

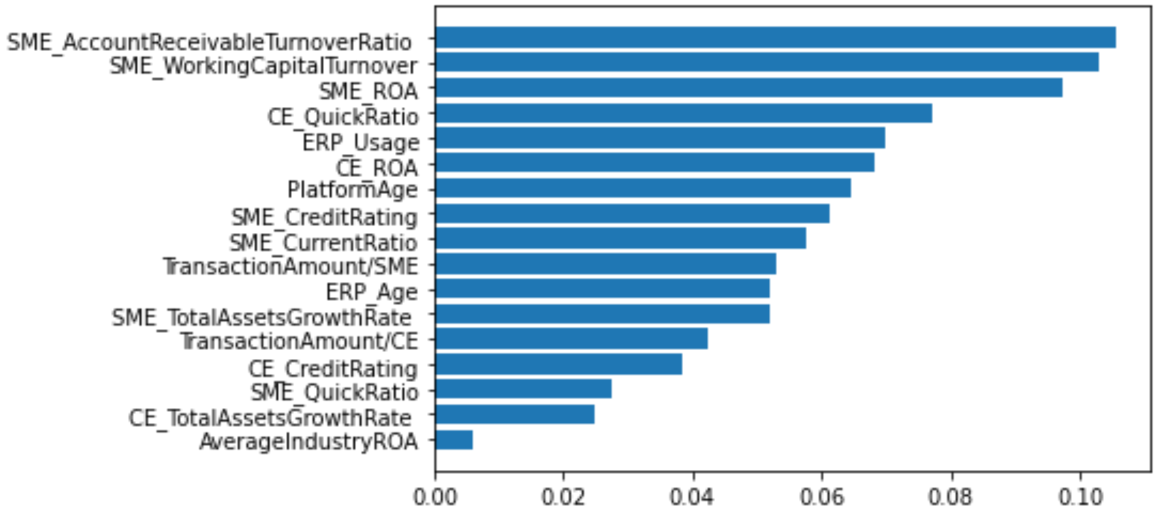


Figure 4.5 XGBoost feature importance ranking.

Then the accuracy of the assessment at different thresholds is examined and the change is plotted as Figure 4.6. An increasing threshold means that more useless features are removed, and the accuracy of the model increases as the threshold increases until the best accuracy of the model is assessed at a threshold of 0.03 (average accuracy is 0.983), when the quick ratio of SMEs, the growth rate of total assets of the core enterprise and the average industry ROA are removed. This indicates that these three indicators are detrimental to credit risk assessment and that removing these three characteristics will result in a more accurate model. Then, as the threshold continues to increase (above 0.03), the correctness of the model starts to decline, especially when the threshold is between 0.05 and 0.08, the correctness tends to drop sharply which shows when notable features are removed from the model the correctness rate deteriorates. This indicates that feature selection has a significant impact on the effectiveness of credit risk assessment models, and that reasonable feature selection can improve model effectiveness.

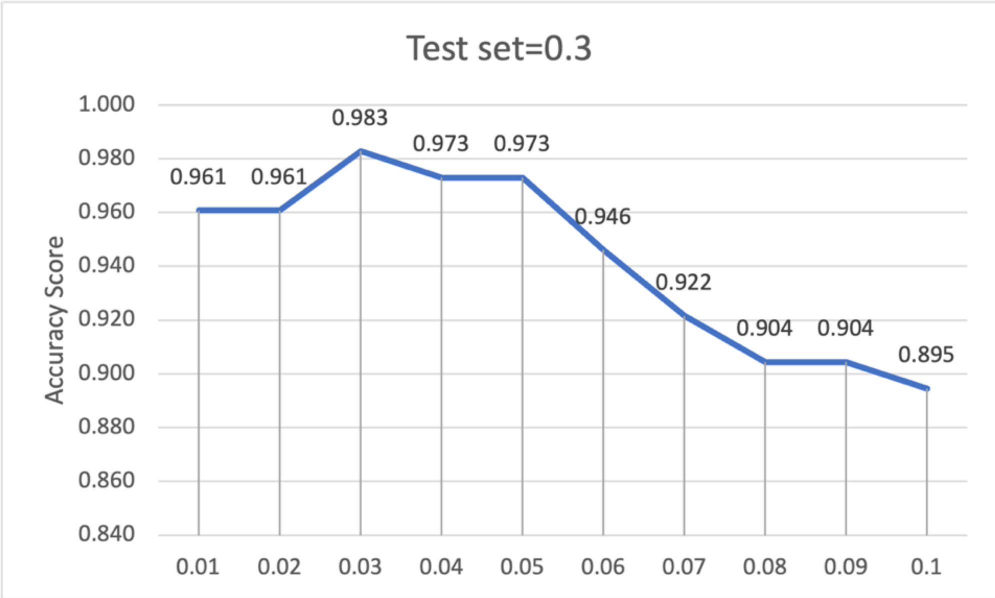


Figure 4.6 The model accuracy in different threshold level.

4.5.3 The Impact of DSCF Feature

Further, for the extent to which DSCF features affect the effectiveness of credit risk assessment, it finds that DSCF features occupy certain importance from the prominent features chart. i.e., whether the enterprise has an ERP system or not, the importance accounts for 0.07 in credit risk assessment, which is an important credit risk assessment factor. The age of enterprise's electronic information technology platform construction, with an important share of 0.065, and the year in which the ERP was used with an important share of 0.055, are the more important features. To further confirm that the inclusion of digital supply chain financial features has an impact on the effectiveness of the credit risk assessment model, we compared the results of the XGBoost-MLP model with/without digital supply chain financial features. As shown in Table 4.6, the average accuracy of XGBoost-MLP without DSCF feature is 0.946 which is lower than the result of XGBoost-MLP with DSCF features and the MCC also shows that the performance of XGBoost-MLP with DSCF features is better than it without DSCF features.

Table 4.6. Comparison of XGBoost-MLP with/without digital SCF features.

	Average accuracy	Recall	Precision	Type I error	Type II error	F-measure	MCC
XGBoost-MLP (Threshold=0.03) with DSCF features	0.983	0.994	0.986	0.014	0.038	0.994	0.922
XGBoost-MLP (Threshold=0.03) without DSCF features	0.946	0.986	0.954	0.048	0.096	0.970	0.739

Moreover, this paper uses partial dependence plots (PDP) to analyse the impact of each explanatory variable in the XGBoost-MLP model on credit risk assessment (Scikit-learn in Python is used for the PDP experiment.). PDP was introduced by Friedman (2001) which can be used to indicate how one of the features affects the model prediction if all other features are maintained constant. Figure 4.7 shows the PDP of traditional financing features of SMEs including the current ratio, working capital turnover ratio, account receivable turnover ratio, ROA, total asset growth rate and the credit rating score. The vertical axis of PDP represents the probability that an SME is judged non-risky, and the horizontal axis represents the change in features. Accordingly, Figure 4.7 indicates that the higher the current ratio, the higher probability of non-risky SMEs which is consistent with the result of Zhu et al. (2019). Similarly, the impact of accounting receivable turnover ratio and ROA also have a similar trend that the higher ratio the higher probability of non-risky SMEs. The change of total asset growth rate has a slight impact on the probability though the overall impact of the total asset growth rate remains between 0.85 to 0.9. And the working capital turnover ratio has the contrary trend of probability changes that the higher the working capital turnover ratio the lower possibility of non-risky SMEs. The highest probability of non-risky SMEs happens when the working capital ratio is below 1. Generally, the higher the accounts receivable turnover rate, the shorter the period of accounts receivable, which means that the return of funds is guaranteed and the risk of repayment is correspondingly lower. However, for working capital turnover, a high working capital turnover indicates that the company is under-capitalised and has a debt crisis. Based on the sample of the SMEs in the paper, the working capital turnover ratio of SMEs in China is generally low, and although the risk of loan repayment is low, it also indicates low capital utilisation and insufficient sales. Additionally, a higher credit rating has a higher probability of non-risky SMEs and the probability increases sharply when the credit rating of SMEs is improved.

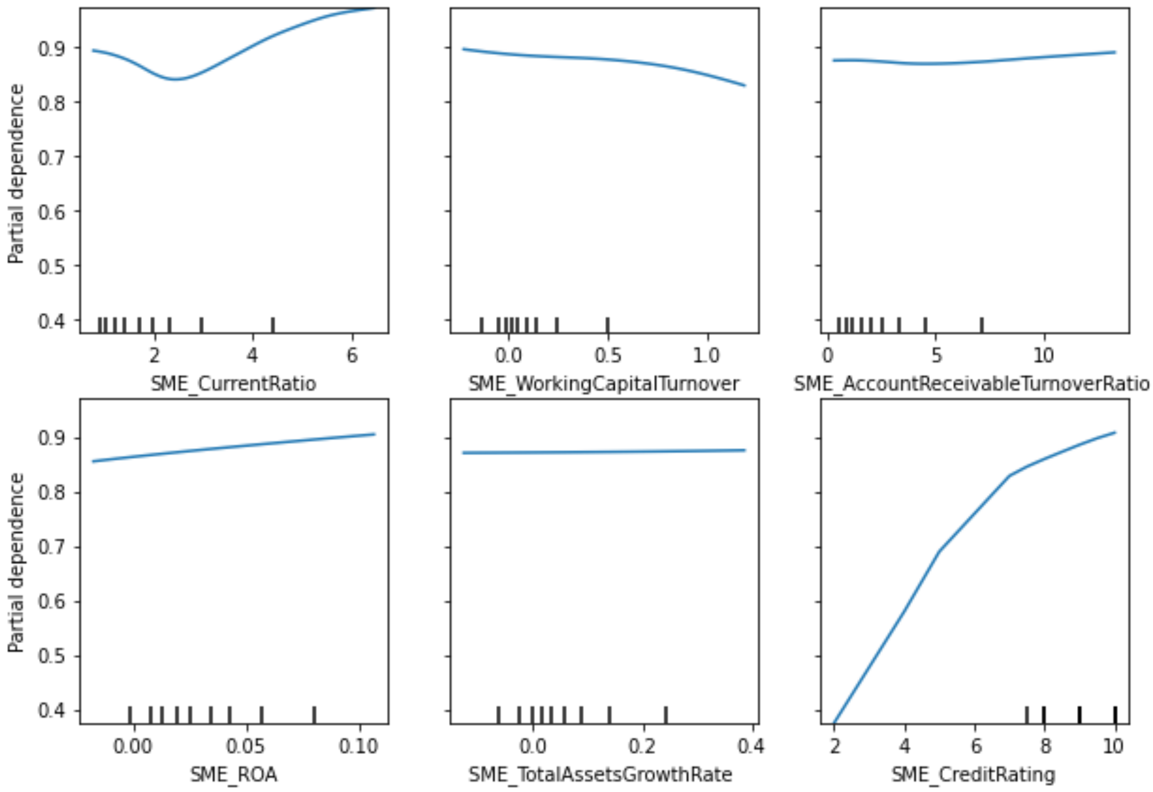


Figure 4.7 The PDP of SMEs' traditional financing features based on the XGBoost-MLP.

Figure 4.8 indicates the PDP of DSCF features of SMEs including the features of core enterprise, the digitalisation feature of SMEs and the trading features in the supply chain. There is a decreasing trend of the probability of non-risky SMEs following the increase of CE's Quick ratio, which is consistent with the result of Zhu et al. (2019). The high quick ratio of a core enterprise leads to excessive capital occupation in its quick assets, which are mostly accounts receivable in the supply chain, and this has an impact on its solvency, as there is a certain degree of uncertainty regarding the collection of accounts receivable. Therefore, for SMEs in the supply chain, a core enterprise with a high quick ratio does not enhance its own risk tolerance. It is also interesting to note that a higher return on assets (ROA) of the core firm does not improve the risk-free probability of the SME. Although a higher ROA indicates a better utilisation of the assets of the core enterprise, for SMEs in the supply chain, their own repayment ability is more important. Meanwhile, we find that core enterprise with a higher credit rating does not have the higher risk-free probability of SMEs but has the opposite effect. Combined with the fact that the weight of the credit rating of core enterprises is not prominent in the feature importance ranking in Figure 4.5, we believe that the current source of funds for SMEs in China is complex, and SCF is not the main source of funds for SMEs, which leads to the core enterprises' own advantage does not effectively enhance the risk-free probability of SMEs.

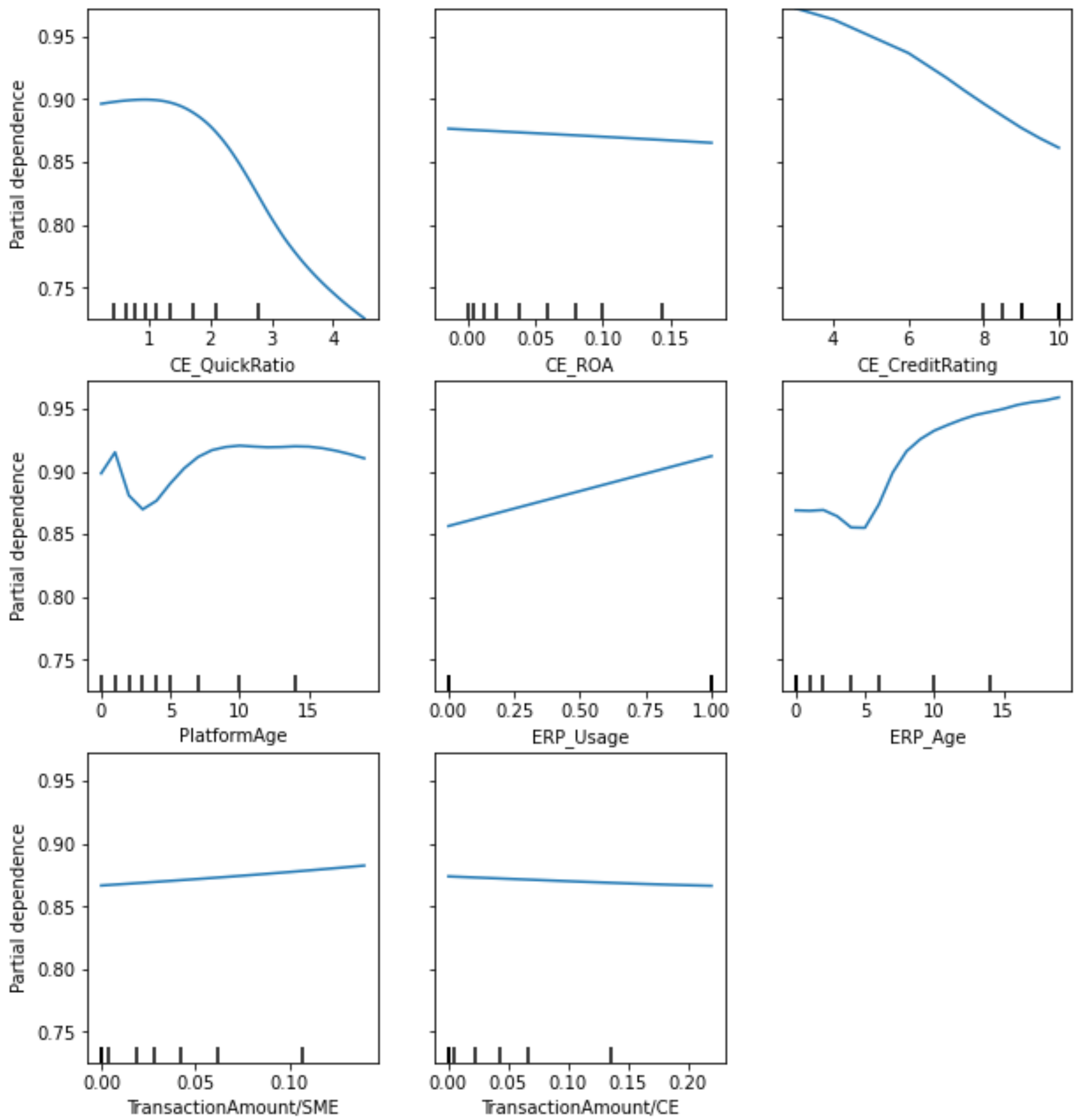


Figure 4.8 The PDP of DSCF features based on the XGBoost-MLP.

Nevertheless, the DSCF features of SMEs have a more positive impact. The change in the age of platform usage is non-linear, with the lowest probability of a firm being non-risky when the age of information platform usage is around 3 years. Whereas, when the age of platform use is in the range of 3 to 10 years, the probability of a firm being non-risky is positively affected. Furthermore, the longer the platform is used does not increase the risk-free probability of the firm, which starts to decrease after 15 years of usage. This study further uses the dummy variable to describe the usage of ERP by firms and the trend in Figure 4.8 shows that SMEs using ERP systems have a higher probability of being risk-free. The feature of ERP usage age is also non-linear, as the change in risk-free probability is not significant for firms with ERP usage of fewer than 5 years, but when firms have ERP usage of more than 5 years, the longer the usage time, the higher the risk-free probability.

Regarding the basis of supply chain financial cooperation, the variable of the transactions between the core firm and the SME divided by SME's sales or costs also shows a positive change, with a subsequent increase in the probability of risk-free for the firm. This indicates that in the supply chain when the main business of SMEs and core enterprises has a certain regularity and a large proportion, the solvency of SMEs has certain stability and security.

4.6 Robustness Check

The XGBoost -MLP model achieves better credit risk assessment results than the comparison models in the designed experimental environment. To assess the robustness of the XGBoost-MLP model in credit risk assessment, we attempt to vary the experimental setting of the model and investigate whether changing the test set proportion in the dataset has an impact on the performance of the models. The following Tables show the evaluation results for each model when the test set percentage is adjusted from 0.3 to 0.1 with the rest of the data set remaining unchanged.

Table 4.7 shows the performance of each model when the test set is 0.1, the average accuracy of XGBoost-MLP is the highest, we further focus on the F-measure which represents the harmonised average score of recall and precision. The F-measure score of XGBoost-MLP is also the highest among the tested models. In addition, the type I error of XGBoost-MLP is the lowest among the models, which indicates that XGBoost-MLP works best in screening for risky firms.

Table 4.7 Performance of XGBoost-MLP and other machine learning methods.

(Test set =0.1)

	Average accuracy	Recall	Precision	Type I error	Type II error	F-measure	MCC
LR	0.926	0.992	0.929	0.076	0.056	0.959	0.638
KNN	0.971	0.992	0.975	0.025	0.056	0.983	0.868
NB	0.860	0.889	0.946	0.051	0.722	0.917	0.487
DT	0.956	0.966	0.983	0.017	0.222	0.974	0.818
SVM	0.971	1.000	0.967	0.034	0.000	0.983	0.867
RF	0.963	1.000	0.959	0.042	0.000	0.979	0.832
MLP	0.978	1.000	0.975	0.025	0.000	0.987	0.901
XGBoost-KNN	0.978	0.992	0.983	0.017	0.056	0.987	0.902
XGBoost-NB	0.882	0.924	0.939	0.059	0.500	0.932	0.512
XGBoost-DT	0.934	0.983	0.943	0.059	0.111	0.963	0.695
XGBoost-SVM	0.977	1.000	0.975	0.025	0.000	0.987	0.901
XGBoost-RF	0.971	1.000	0.967	0.034	0.000	0.983	0.867
XGBoost-MLP	0.978	0.989	0.986	0.014	0.077	0.987	0.901

Notes: All models presented in Table 4.7 are estimated based on the test set=0.1. Results are estimated based on the training set of 949-observations and the test set of 408-observations from 31 March 2016- 31 December 2019. LR is the logistic regression model; KNN is the k-nearest-neighbor model; NB is the naïve Bayes mode; DT is the decision tree model; SVM is the support vector machine model with radial bias function as kernel function; RF is the random forest model; MLP is the multi-layer perceptron model.

Figure 4.9 shows the ranking of the feature importance of XGBoost-MLP at a test set of 0.1, with the ROA of SME being the most important feature. More specifically, we find by plotting the variation in model accuracy for different thresholds (See Figure 4.10) that the model has the highest accuracy of 0.978 when the threshold is at 0.04 or 0.05, i.e. removing the quick ratio of SMEs, the credit rating of CE, the proportion of trading transaction on CE sales or cost, the growth rate of total assets of the core enterprise and the average industry ROA. It is noteworthy that when the threshold rises to 0.06, the growth rate of total assets of the core enterprise and the average industry ROA is also removed, and then the accuracy of the model decreases significantly and the features removed include ERP age, usage status of ERP, the growth rate of total assets of the SME and the credit rating of the SME.

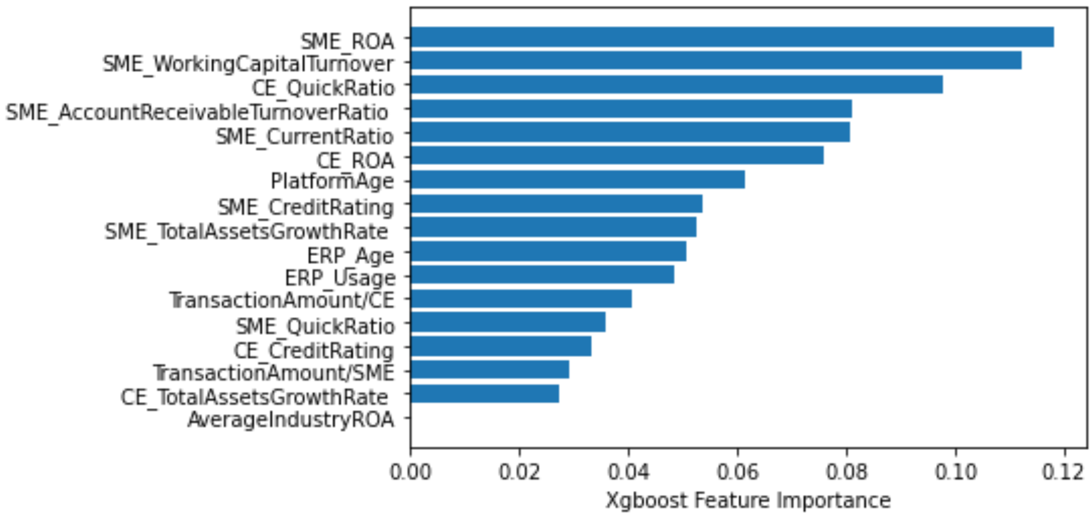


Figure 4.9 XGBoost feature importance ranking. (Test set=0.1)

Figure 4.10 The model accuracy in different threshold level. (Test set=0.1)

Table 4.8 shows the performance of each model when the test set is 0.2, with XGBoost-MLP having the highest average accuracy, and F-measure and type I error also having the best performance. In this case the working capital turnover of the SME is the most important feature in the feature importance ranking (see Figure 4.11). Figure 4.12 indicates that the average accuracy of the model is highest at a threshold of 0.03 at 0.978, i.e., removing the quick ratio of SMEs, the growth rate of total assets of the core enterprise and the average industry ROA.

Table 4.8 Performance of XGBoost-MLP and other machine learning methods. (Test set =0.2)

	Average accuracy	Recall	Precision	Type I error	Type II error	F-measure	MCC
LR	0.909	0.994	0.910	0.098	0.038	0.950	0.909
KNN	0.934	0.978	0.946	0.056	0.125	0.962	0.934
NB	0.875	0.918	0.934	0.065	0.475	0.926	0.875
DT	0.956	0.983	0.967	0.034	0.115	0.975	0.956
SVM	0.922	0.997	0.919	0.087	0.019	0.957	0.922
RF	0.963	1.000	0.959	0.042	0.000	0.979	0.963
MLP	0.946	0.986	0.953	0.048	0.096	0.969	0.946
XGBoost-KNN	0.934	0.983	0.942	0.060	0.100	0.962	0.934
XGBoost-NB	0.901	0.948	0.936	0.065	0.300	0.942	0.901
XGBoost-DT	0.945	0.966	0.960	0.300	0.200	0.968	0.945
XGBoost-SVM	0.959	1.000	0.955	0.047	0.000	0.977	0.959
XGBoost-RF	0.974	0.996	0.975	0.026	0.025	0.985	0.974
XGBoost-MLP	0.978	0.991	0.979	0.022	0.050	0.985	0.978

Notes: All models present in Table 4.8 are estimated based on the test set=0.2. Results are estimated based on the training set of 949-observations and test set of 408-observations from 31 March 2016- 31 December 2019. LR is logistic regression model; KNN is k-nearest-neighbor model; NB is naïve Bayes mode; DT is decision tree model; SVM is support vector machine model with radial bias function as kernel function; RF is random forest model; MLP is multi-layer perceptron model.

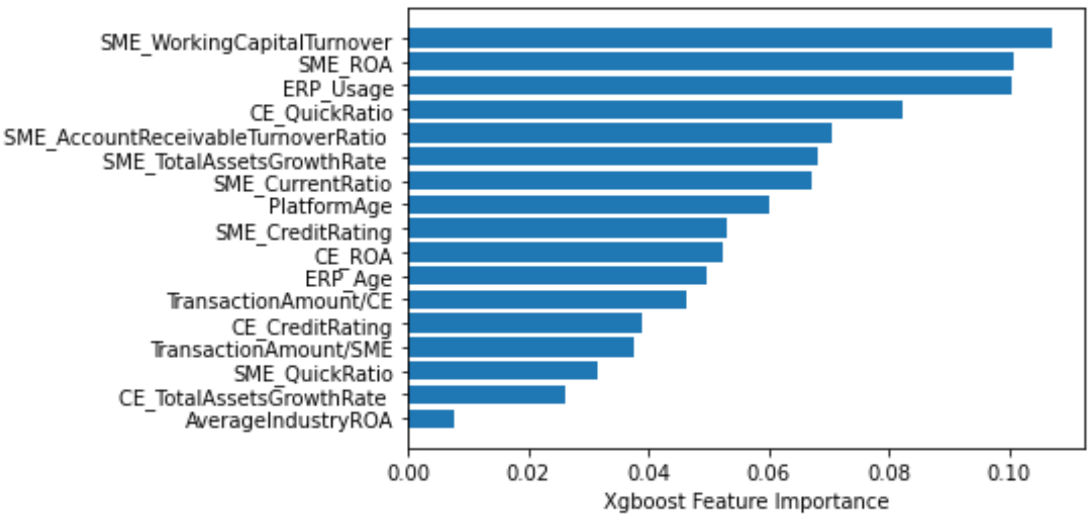


Figure 4.11 XGBoost feature importance ranking. (Test set=0.2)

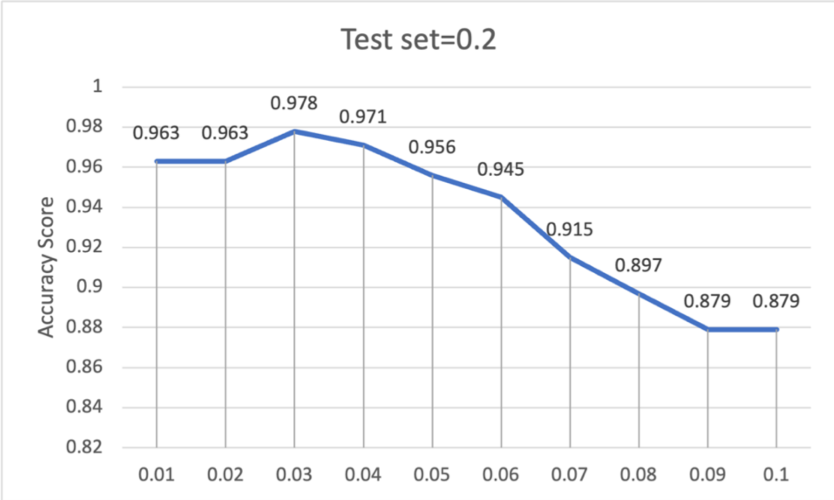


Figure 4.12 The model accuracy in different threshold level. (Test set=0.2)

Table 4.9 shows the performance of each model when the test set is 0.4. The average accuracy, F-measure and type I error of XGBoost-MLP are consistent with the previous tests. Figure 4.13 shows the ranking of feature importance, with the SME's accounts receivable turnover rate being the most important feature to assess at a test set of 0.4. Figure 4.14 indicates that the average accuracy of the model is highest at 0.974 when threshold is 0.03 or 0.04, that the quick ratio of SMEs, the credit rating of CE, the growth rate of total assets of the core enterprise, the average industry ROA and status of ERP usage are removed.

Table 4.9 Performance of XGBoost-MLP and other machine learning methods. (Test set =0.4)

	Average accuracy	Recall	Precision	Type I error	Type II error	F-measure	MCC
LR	0.906	0.991	0.908	0.100	0.054	0.948	0.906
KNN	0.937	0.985	0.945	0.058	0.095	0.965	0.937
NB	0.901	0.951	0.935	0.066	0.311	0.943	0.901
DT	0.937	0.977	0.952	0.049	0.149	0.964	0.937
SVM	0.939	1.000	0.934	0.070	0.000	0.966	0.939
RF	0.958	0.991	0.961	0.041	0.054	0.976	0.958
MLP	0.959	0.989	0.964	0.036	0.068	0.977	0.959
XGBoost-KNN	0.932	0.985	0.939	0.064	0.095	0.961	0.932
XGBoost-NB	0.915	0.966	0.938	0.064	0.216	0.952	0.915
XGBoost-DT	0.934	0.959	0.964	0.036	0.257	0.962	0.934
XGBoost-SVM	0.932	1.000	0.927	0.079	0.000	0.962	0.932
XGBoost-RF	0.961	0.991	0.965	0.036	0.054	0.978	0.961
XGBoost-MLP	0.974	0.991	0.979	0.022	0.050	0.985	0.974

Notes: All models present in Table 4.9 are estimated based on the test set=0.4. Results are estimated based on the training set of 949-observations and test set of 408-observations from 31 March 2016- 31 December 2019. LR is logistic regression model; KNN is k-nearest-neighbor model; NB is naïve Bayes mode; DT is decision tree model; SVM is support vector machine model with radial bias function as kernel function; RF is random forest model; MLP is multi-layer perceptron model.

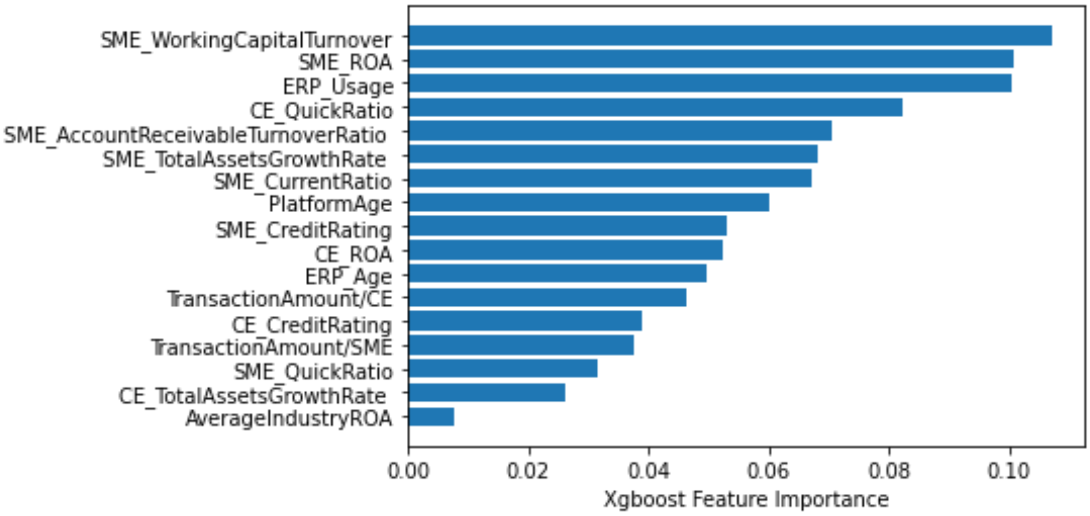


Figure 4.13 XGBoost feature importance ranking. (Test set=0.4)

Figure 4.14 The model accuracy in different threshold level. (Test set=0.4)

In summary, combining the different test set settings, it finds that the overall model evaluation results change to some extent as the test set changes, but the average accuracy of XGBoost -MLP is still the highest, indicating that robustness of XGBoost-MLP model. And the most optimal test setting is when the test set is 0.3.

4.7 Conclusions

With the development of the industrial IoT and the digital economy, various industries and sectors will form different industrial chains and supply chains on various digital platforms in the future. DSCF is breaking the shackles of the current inertia of building digital platforms centered on finance or banks, embedding DSCF into various industrial Internet of

Things and various digital economic platforms, and becoming an organic part of these digital economic platforms. In order to take DSCF as the research background, specifically from the perspective of credit risk assessment, this paper conducted research on the credit risk assessment methods of enterprises in the DSCF environment and its empirical evidence.

Firstly, the existing credit risk assessment methods in terms of their subjective and arbitrary feature selection and the poor effectiveness of linear assessment methods are analysed in this paper. Secondly, feature importance and the role of feature selection on credit risk assessment models through XGBoost feature selection are evaluated. Then, the role of digital features for credit risk assessment in SCF is validated. This chapter selected 1357 observations from 85 private Chinese-listed manufacturing SMEs over the period 2016–2019 to empirically test and compare the credit risk assessment models. After the feature selection by XGBoost, the five most important features were selected as accounts receivable turnover of SME, working capital turnover of SME, ROA of SME, quick ratio of CE, and ERP usage situations of SME, which improved the accuracy of risk identification by 98.3% compared to the traditional credit risk assessment models without the feature selection. The importance of the DSCF features was also verified through the XGboost feature selection. It is further found that the feature selection is essential to the performance of credit risk assessment results by varying the threshold value of XGBoost feature importance ranking. The effectiveness of the risk assessment model varies depending on the threshold value set by the lending decision-maker for the feature selection process in the risk assessment, and that reasonable feature selection will improve the model effectiveness. Considering the various threshold values for feature selection, the accounts receivable turnover ratio of SMEs is the most important risk assessment indicator. Finally, by comparing the inclusion and removal of digital features of enterprises, we found that digital features are important for the credit risk assessment effect of DSCF, and the model with the inclusion of digital features as an assessment indicator has a higher accuracy rate, with an increase of 3.7%. This further validates that the inclusion of DSCF features in credit risk assessment is beneficial in terms of the accuracy of its risk identification.

On this aforementioned basis, this paper provides the following recommendations for the mitigation of credit risks based on DSCF. For SCF platforms, including commercial banks and other financial institutions, as one of the main actors in supply chain financial services, they should be well prepared for their own risk management, credit assessment, and credit limits. Traditional credit risk assessment features such as accounts receivable turnover of

SMEs, working capital turnover of SMEs, ROA of SMEs and CEs, quick ratio of CE, and credit rating of SMEs are still key characteristics for lending decision-makers. Furthermore, in the case of companies with a high degree of digitalisation, such as those that actively use ERP systems or have a well-developed information technology network, the corresponding DSCF features such as the degree of ERP usage or the construction of an information technology platform should also be taken into account in the credit risk assessment. For core enterprises and SMEs, the enterprise's accounts receivable turnover and working capital turnover are two important indicators for credit risk assessment, so enterprises are expected to be flexible in working capital and to digitise assets such as pledges to improve the speed and efficiency of circulation of the pledge. Whereas ROA, as one of the most important traditional evaluation indicators, also points out that enterprises are supposed to improve their own financial system and management system to improve their production and operation capacity. Meanwhile, the construction of digital information platforms and the usage of ERP as new indicators also provide important reference bases for credit risk assessment, and enterprises are advised to strengthen their digital development process to achieve open and transparent business data and reduce their own credit risks.

Overall, SCF is a very promising business for commercial banks, and with the continuous innovation of technology, the application of DSCF is becoming more and more widespread, and its connotations are becoming more and more enriched. Although DSCF is a future development trend and has high research value, DSCF is still in its infancy and research on it is very limited. Thus, there are some limitations in our paper. Firstly, due to the availability of data, 85 Chinese enterprises are selected as the sample for this paper. Although they are representative of the empirical samples in the context of DSCF in China, the experimental results may be biased due to the small sample. Secondly, only by comparing traditional commonly used machine learning models as a comparative analysis in this paper, this research does not perform a comprehensive experimental analysis, so the inter-model study would be increased in future research.

Chapter 5 P2P Lending: Imbalanced Issue Analysis and Feature Selection with Machine Learning Method

5.1 Introduction

With the continuous development of technologies such as the Internet, big data and cloud computing, P2P lending is an innovative financial industry derived from a high degree of merger between finance and the Internet (Bachmann et al., 2011). P2P lending, as an important part of the Internet financial market, has been developing rapidly in recent years, realising direct lending between peer-to-peer without the involvement of financial intermediaries. In March 2005, the UK's Zopa website was launched, marking the birth of p2p network lending, and by the end of 2013, the transaction volume of the Zopa website was growing at an annual average rate of over 50%[25]. 2006 saw the launch of Prosper, the first p2p lending platform in the US; in 2007, Lending Club, currently the world's largest p2p lending platform, was established. By the end of 2013, the US p2p online lending market was entirely dominated by Lending Club and Prosper. The annual transaction volume in 2013 was approximately US$2.4 billion, of which Lending Club's transaction volume was US$2.065 billion, an increase of 200% compared to 2012's transaction volume of US$718 million; Prosper's transaction volume in 2013 was 3.56 billion. The volume of transactions in 2013 was $356 million, a growth rate of over 100% compared to 2012's volume of $152 million[26]. However, high returns often go hand in hand with risk. When P2P lending platforms provide users with convenient and reliable credit services, they need to predict the credit risk of users based on their basic information and transaction data (Lenz, 2016). For example, a credit-risk-free and compliant customer will repay the loan on time and in full after completing the transaction, while a credit-risk-missing customer will default due to his or her lack of repayment ability or willingness to repay, which will result in the platform incurring substantial financial losses. Therefore, for P2P platforms, establishing a personal credit risk prediction model to accurately identify users who are likely to default and reject their transaction requests can effectively reduce the platform's economic losses and safeguard its sound development.

The rise of P2P lending stems from the maturity of the credit system and the advancement of Internet technology (Klafft, 2008). The multi-dimensional data recording customers'

[25] https://www.ft.com/content/65795036-ddd6-3838-a100-d0ae4f401c7c
[26] https://www.ft.com/content/029c2ddc-d3df-11e3-b0be-00144feabdc0

various behaviours has been increasing and accumulating, and the thinking of big data credit has gradually stepped into people's view and is sought after by Internet finance and capital markets (Yan et al., 2015). The cross-recurrence of multi-faceted and multi-level behavioural information of P2P borrowers and the continuous presentation of information associated with customers' activities can map their credit status in terms of both willingness and ability to repay (Zhang et al., 2016). However, although multi-dimensional information data can cross-replicate the credit status of P2P borrowers, the computation of credit evaluation models tends to become more complex as the dimensionality of the data increases. At the same time, the data generated by credit transactions is usually highly unbalanced, i.e., the vast majority of people who make credit transactions can keep their contracts, and only a small number of people will incur defaults. Meanwhile, the problem of classifying imbalance data is often a dilemma: if a model predicts a creditworthy customer to be a non-creditworthy user, the platform will reject the user's transaction request, thereby reducing revenue; if it predicts a non-creditworthy user to be a creditworthy user, the platform will approve the user's transaction request and lend to the user, resulting in significant financial losses. Machine learning methods are crucial in credit risk assessment due to their ability to analyze large volumes of diverse data and identify complex patterns that may elude traditional rule-based systems. However, the imbalance data issue is a prevalent challenge in machine learning and data analysis where the distribution of classes within a dataset is significantly skewed, with one or more classes being underrepresented compared to others. This imbalance compromise the performance of machine learning models, as they tend to be biased towards the majority class, often leading to poor generalization and predictive accuracy for minority classes (Abd Elrahman and Abraham, 2013). Therefore, many scholars working on the improvement of evaluation models for the P2P lending imbalance data problem, such as (Niu et al., 2020; Song et al., 2020; Zhou et al., 2019).re

Lending Club, once the most popular online lending platform, provides the data set that most academics use to conduct their analysis (Emekter et al., 2015). Many scholars have used Lending club data for credit risk assessment analysis and model improvement studies, but there are inconsistencies in their treatment of the data. Firstly, many scholars have ignored the imbalance in the credit data (Bastani et al., 2019; Ma et al., 2021; Xia et al., 2017), e.g., Malekipirbazari and Aksakalli (2015) selected only 15 features for risk assessment by Random Forest model without considering the imbalance status of the data, and the results show that the model is ineffective in identifying the risky samples. Similarly, Teply and Polena (2020) selected 23 features for the optimal risk assessment model but did not consider the impact of imbalance in the data sample and the incomplete selection of features on the classification model. Arora and Kaur (2020) used 143 features for assessment, and although this study does not mention and address the imbalanced nature of the data, all models of the experiment, including the baseline model, yielded very good assessment results, with good identification of risky samples. Yu and Zhang (2021), in their study of poor evaluation due to missing samples, ignore the fact that imbalance in the sample and the selection of features can be a real cause of poor results. Secondly, it is found that some scholars have studied the imbalance issue in Lending Club data, but only selected partial specific features for credit evaluation and claimed that the low effectiveness of their results was due to the imbalance data problem. For example, Namvar et al. (2018) attempted to propose a new model to improve credit risk evaluation in an imbalance data environment, but only 25 features were selected for evaluation and the results were unsatisfactory, with poor overall prediction accuracy and no significant improvement in the ability to identify risky samples. COŞER et al. (2019)considered the imbalance data problem in their risk assessment and compared the balanced and imbalance datasets, but the feature selection level was unclear and only 14 numerical and 3 categorical variables were retained in the analysis. Finally, although some scholars have selected important features through a feature filtering mechanism and noted the imbalance data problem, e.g., Chen et al. (2021) and Moscato et al. (2021) attempted to address the imbalance data problem arising from P2P credit evaluation, the results are not significant. Owusu et al. (2022) selected only five features for evaluation after feature filtering and the results showed that the accuracy was still high, which may be a problem of information leakage rather than valid proof that the imbalance in the data was resolved. Similarly, Chang et al. (2022) filtered 16 Lending Club's features and under-sampling the imbalance data, but the overall results were not satisfactory.

Thus, the motivation for this chapter aims to address the following research questions:

RQ1: Whether an imbalanced data issue matters in the Lending Club dataset with complete features?

RQ2: Whether the feature incompleteness in the existing literature result in poor performances of the existing models?

RQ3: How to mitigate the imbalance data issue with limited features?

To examine the above questions, the data is collected from the Lending Club 2007-2018 Q4, which contains 151 original features and 1345310 observations. The first set of data was constructed from the features selected in the previous article and contains 17 features. The second set of data is cleaned by removing the missing data and information leakage features from the original features, resulting in a dataset of 87 features. By comparing the results obtained from these two datasets on the classification problem to explain the RQ1 that whether an imbalanced data issue matters in the Lending Club dataset with complete features. The imbalanced data issue won't affect the prediction performance if the evaluation metric of the dataset with 87 features performs better than that of the dataset with 17 features. Meanwhile, the poor performance of existing models with 17 features will be identified that could explain the RQ2. In addition, further investigation focuses on the relationship between feature selection and imbalance data which try to figure out the solution of RQ3. Thus, comprehensive baseline models (LR, DT, XGBoost), as well as learning models for imbalanced data (Bagging, Easy Ensemble, Instance Hardness Threshold (IHT), SMOTE, Tomek Links, Cost-sensitive), are built on the basis of which a plot of feature selection versus imbalance proportions are constructed. Figure 5.1 shows the flow chart of our research.

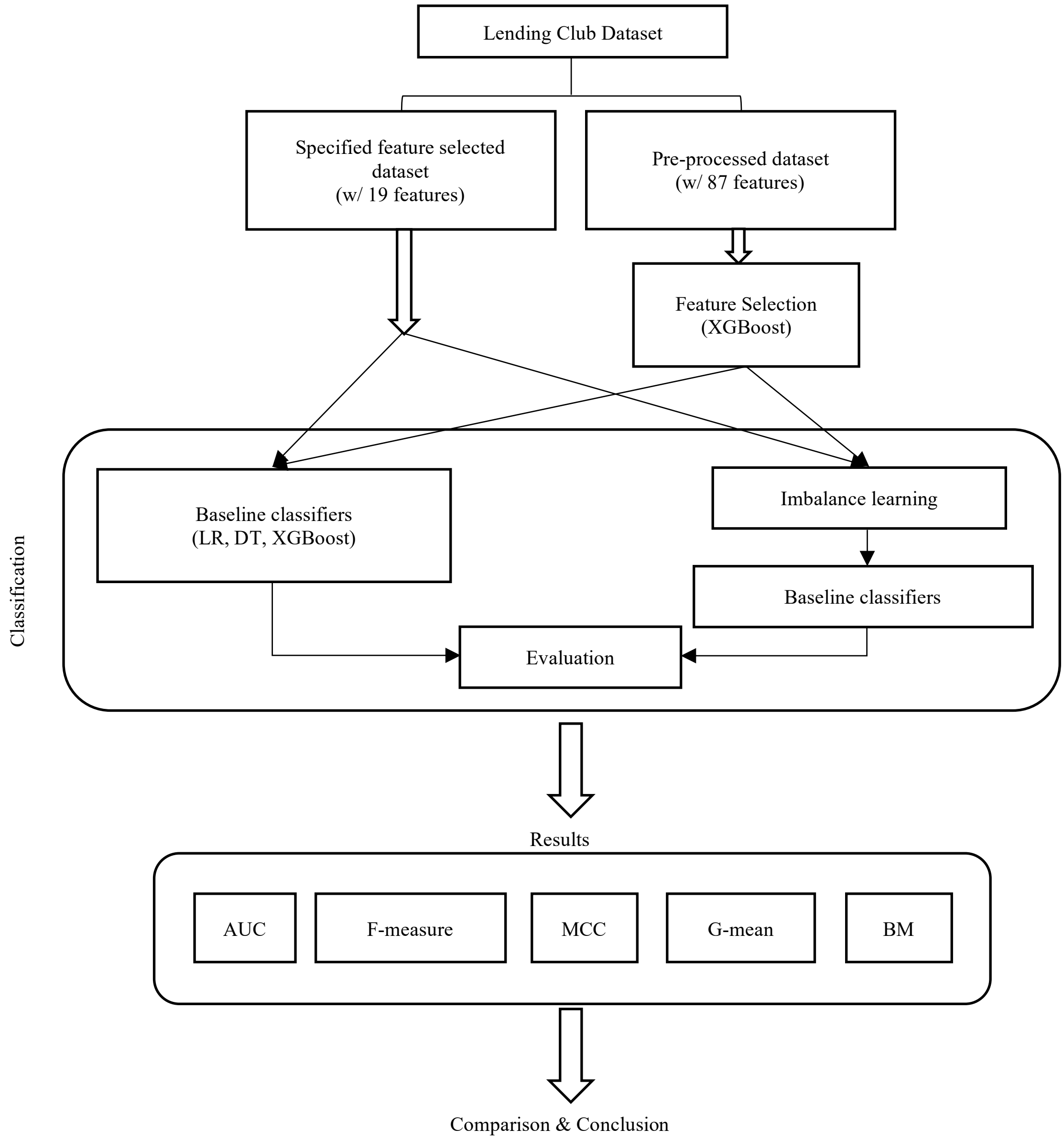


Figure 5.1 Research flow chart.

This research is of great interest to the enrichment of the theory and practice of credit risk assessment under the P2P lending model, and the main contributions are as follows. Firstly, there is a gap in the existing literature on the relationship between imbalanced data and feature selection due to there being no standard for feature selection and a lack of systematic theoretical support. The feature selection in previous studies is replicated and compared. The findings illustrate that too limited features reduce the effectiveness of credit assessment. While the more complete and richer dataset makes the credit risk assessment model achieve satisfactory results. This indicates that the selection of a dataset is important for credit risk assessment. The findings expand the consideration for model improvement in risk assessment from the actual P2P lending business thereby meeting the actual business needs. Meanwhile, a systematic study of learning methods for the imbalanced data problem is also conducted. A comprehensive range of over-sampling, under-sampling, ensemble, and cost-sensitive methods are used to fully train the data. The results enrich the study of the classification of imbalanced data. Further, the theoretical support for imbalanced datasets and feature selection is extended by exploring the dynamics of the imbalanced ratio and feature selection in depth. The analysis of different data imbalance ratios and feature selection results helps future research to select credit risk assessment features and deal effectively and rationally with imbalance problems. Finally, an optimal set of features is obtained through feature selection, which further improves the assessment system of P2P lending and enriches the related theory.

The paper is presented as follows. Section 5.2 contains the background of P2P lending and a literature review of the imbalanced data issue. Section 5.3 covers the methodology of the baseline classifier and the imbalance learning methods. Section 5.4 reports the design, acquisition, and processing of the experiments. Section 5.5 presents the results and analysis. Chapter 5.6 provides the conclusions.

5.2 Literature Review

P2P lending is direct lending between individuals and individuals through Internet platforms, i.e., individuals with funds, lending to other individuals with borrowing needs through electronic trading platforms (Emekter et al., 2015). As platforms offering peer-to-peer lending usually operate online, the cost of service is relatively low and therefore generates more revenue for investors and lower costs for borrowers. With the launch of the first P2P platform, Zopa, in the UK in 2005, the P2P lending business has grown rapidly around the world[27]. The UK's P2P industry is well regulated, with three layers of regulatory defence. The first layer is self-regulation by industry associations, with the Peer-to-Peer Finance Association (P2P Association) established on 15 August 2011, and its proposed "Principles of the Association" filling the gap in regulatory law to a certain extent[28]. The second level is government regulation. The UK government has clear requirements on minimum capital requirements, customer funds and information disclosure for P2P platforms, and the corresponding regulatory laws and regulations mainly consist of national macro financial laws, industry laws and regulations. The third level is the professional institutions regulation, since April 2014, the Financial Conduct Authority (FCA) has started to exercise regulatory functions on P2P lending business, requiring lending platforms to provide regular financial

[27] https://www.theguardian.com/money/2021/dec/11/zopa-peer-to-peer-lending-p2p-money
[28] https://www.p2pfinancenews.co.uk/2020/01/13/p2pfa-disbands-as-platforms-establish-new-group/

reports, customer funds reports, etc.[29] Further, Prosper was founded in February 2006, marking the beginning of the P2P lending industry in the US, and by 2017 the total amount of P2P lending in the US exceeded US$26 billion[30]. The US Securities and Exchange Commission (SEC) is the regulatory body for the business, and since 2008 the SEC has required all P2P lending platforms to be registered and their products regulated as securities. The SEC has required all P2P lending platforms to register and regulate their products as securities since 2008 (Lo, 2015). The US P2P business has an 'oligarchic' pattern. The two main P2P lending platforms, Lending Club and Prosper, largely dominate the market, and the size of P2P lending is growing rapidly, with Nash and Beardsley (2015) and Zhang et al. (2015) noting that the total P2P lending business in the US and Europe will continue to grow at a rate of double the annual rate.

Although P2P lending is growing at a phenomenal rate, there is still a huge difference between the size of the business and the credit business of traditional banks. The International Monetary Fund (IMF, 2015) reports that the total asset size of global P2P lending represents only 0.05% of total global banking assets[31]. Havrylchyk et al. (2017) also argue that the total size of the P2P lending business in the US is equivalent to only 0.7% of the total national retail sales in the U.S. P2P lending platforms try to act as intermediaries by leveraging the power of internet technology to connect individuals or businesses (borrowers) who wish to obtain loans with individuals and institutions (investors) who wish to invest. Berger and Gleisner (2009) found that P2P platforms act as financial intermediaries such as banks between borrowers and investors, reducing information asymmetry between borrowers and lenders and improving borrowers' credit ratings and that the costs incurred in this process are compensated by lower borrowing rates. Wang et al. (2015) compared the P2P lending business model with that of traditional banks and found that the information flow in the P2P lending process is more frequent and transparent, and thus more information-dependent. As P2P lending platforms operate online automated transaction systems outside of the bank regulatory system, none of the loans are included in their liabilities and operating costs are very low relative to banks (Cummins et al., 2019). P2P lending platforms, on the

[29] https://news.ateb-group.co.uk/advising-on-p2p-agreements
[30] https://techcrunch.com/2009/04/29/p2p-lending-marketplace-prosper-gets-off-the-bench-debuts-open-market-initiative/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAEwIVXcVsdcefd-Ai53koEwSg26HSQrXuACc10jaFFIesgqK9w-dd1chJ9Qgqo5J3X2LYlqf39vMA8c1Tdtebh_e-UAnOVJu1Ye9DKsLPuaE0kmGMvf_ifMxWFSMp2agcy7eqxBqEc-7IE0m-KieTOJyDpZ6ZhWFoZZeaAjGniBu
[31] https://www.imf.org/external/pubs/ft/wp/2015/wp1519.pdf

other hand, are mainly profitable from transaction aggregation and loan-related service fees for both borrowers and investors, with borrowers typically paying a percentage of the loan closing fee at loan inception, as well as penalties in case of late repayment, while investors pay a service fee based on the amount of their investment.

The imbalanced data problem arising from P2P lending platforms refers to the phenomenon of an imbalance in the ratio of the number of samples with different labels in the classification of risk assessment, i.e. a situation where in reality the default sample is much smaller than the compliance sample. Due to the complex nature of imbalanced datasets, different processing methods and techniques are classified into two categories. The first category is data-level methods that balance the data set at the data pre-processing stage by resampling to reduce the effect of skewed class distribution on the classification model during learning. Kubat and Matwin (1997) proposed a selective majority of sample extraction method, the One-sided Selection method, considering that it would be detrimental to the classification of small samples if the majority of samples taken during the under-sampling process were noisy or unreliable samples. Yen and Lee (2009) introduced clustering in under-sampling methods to select large classes of samples. Similarly, Padmaja et al. (2008) introduced clustering in the under-sampling process and proposed a new under-sampling method, majority filter-based minority prediction (MFMP). Laurikkala (2001) proposed an under-sampling method based on the neighborhood cleaning rule. The oversampling method balances the large and small samples in the dataset by increasing the number of small samples, so that the influence of the small samples in the classification learning process is increased. Nickerson et al. (2001) aimed to improve the classification performance of small class samples by introducing clustering methods in oversampling. Nowadays, the most widely used oversampling method is the Synthetic Minority Oversampling Technique (SMOTE) algorithm proposed by Chawla et al. (2002). The basic idea of this algorithm is to analyse the minor class samples and synthesise new minor class samples by k-nearest neighbours. However, there are two main problems with the SMOTE algorithm: firstly, there is a certain degree of blindness in the selection of nearest neighbours; secondly, it cannot overcome the problem of data distribution in imbalanced datasets and is prone to distribution marginalisation. Many scholars have improved the SMOTE algorithm in response to its problems. Borderline-SMOTE (Han et al., 2005), MSMOTE (Hu et al., 2009) and other algorithms have been proposed, which improve the SMOTE algorithm to some extent. Nowadays, hybrid sampling methods, which are a combination of under-sampling and over-sampling methods, are emerging in a growing number of research studies

(Marqués et al., 2013; Sun et al., 2018). The aim is to achieve a balanced distribution of categories in an imbalanced data set.

The second category is algorithm-level approaches, where the recognition rate of classification models for small classes of samples is improved by changing the algorithm. Integrated learning algorithms and cost-sensitive learning algorithms are commonly used to deal with classification problems on imbalanced datasets. Krawczyk et al. (2015) introduced cost-sensitive learning algorithms to deal with imbalanced classification problems based on the use of DT as base classifiers. Using hybrid methods, a combination of sampling methods, ensemble learning algorithms and cost-sensitive learning algorithms have been used for imbalance classification problems. Barandela et al. (2003) proposed the under-bagging algorithm by combining the under-sampling method and the bagging ensemble algorithm, which can effectively handle imbalance classification problems. Subsequently, Seiffert et al. (2010) proposed a random under-sampling boosting (RUSBoost) algorithm to solve the imbalance classification problem by combining a random under-sampling method with a boosting ensemble learning algorithm. Based on the RUSBoost algorithm, Galar et al. (2013) proposed the evolutionary under-sampling boosting (EUSBoost) algorithm by improving the under-sampling method. SMOTEBoost was proposed by Chawla et al. (2003), which combined SMOTE oversampling method and boosting ensemble learning algorithm. The algorithm first oversampled the imbalanced dataset to make the data categories balanced and then used the balanced dataset to train the classifier. Afterwards, Wang and Yao (2009) combined SMOTE algorithm and the bagging algorithm to propose a SMOTEBagging algorithm to solve the imbalance classification problem. Hu et al. (2009) proposed the MSMOTEBoost algorithm after improving the SMOTEBoost algorithm, which offers some performance improvements over the SMOTEBoost algorithm. An increasing number of scholars have now started to apply a combination of hybrid sampling methods and ensemble learning algorithms to the imbalance classification problem (Qian et al., 2014; Zhang and Chi, 2021).

Feature selection is a critical step in data processing, and it also plays an important role in imbalance data problems, that the aim is to select a subset of j features based on some rule that allows the classifier to achieve optimal performance, where j is a user-defined parameter (Wasikowski and Chen, 2009). Feature selection is extremely important as sampling techniques and algorithm-level approaches are not sufficient to solve the class imbalance problem for high-dimensional data, where the class imbalance problem usually occurs (Chawla et al., 2004). Van Der Putten and Van Someren (2004) found that feature selection was more important than the choice of a classification algorithm in improving performance when analysing the CoLL Challenge 2000 dataset. In a study of highly imbalanced textual classification problems, Forman (2003) found that the use of feature selection alone in a high-level dataset could largely resolve the imbalance in the dataset. Feature selection lies in the fact that a small number of features represent the majority of information in the original data, but in cases where the imbalance ratio is large, feature selection may eliminate key features that serve to identify the minority samples. Elkan (2001) found that feature selection did not bring enough benefit to his work. Guyon and Elisseeff (2003) also illustrate the limitations of feature selection through theoretical analysis.

5.3 Methodology

5.3.1 Baseline Models

For the binary classification assessment of credit risk, this study uses LR, DT and XGBoost as baseline classifiers. LR is a traditional classification and prediction model, which is widely used in data mining, economic forecasting and other fields, and is essentially a probabilistic prediction model that can be categorised by different attributes of the dependent variable, such as binary classification problems (Komarek, 2004). In this paper, for the binary classification problem of credit risk assessment, given a set of customer samples $T = \{(x_i, y_i)\}_{i=1}^{n}$, $x_i \in R^p$ is a customer feature variable, $y_i \in \{0, 1\}$ is a binary attribute variable that y_i denotes the i-th customer is a risky customer, $y_i = 0$ means the i-th customer is a non-risky customer. The purpose of a credit scoring model using LR is to assess the probability of a given customer being a "good customer" or a "bad customer", with the expression of the LR equation as $p(y = 1|X) = \frac{\exp(\beta_0+\beta^T X)}{1+\exp(\beta_0+\beta^T X)}$, where $p(y = 1|X)$ denotes the probability of the “bad customer”, X is the m-dimensional vector, β is the m-dimensional parameters to be estimated.

DT is a non-linear discriminant analysis method, which is a classification function approximation method developed in the field of machine learning (Curram and Mingers, 1994). It is essentially a process of classifying samples by building a series of tree rules, i.e., generating a series of tree classifier rules based on the attributes and classification results of known samples, and using these rules to classify and predict unknown data, which is a typical supervised single classifier. The DT has three types of nodes: a root node, a leaf node and an intermediate node, each node is a specific attribute of an attribute, and from each node, a new path can be generated by splitting one level down, with as many attribute sets as the node has. Thus, from the top root node, a path is followed through a number of intermediate nodes to a leaf node, which can be seen as a specific rule for classification, and the whole DT is a collection of classification rules made up of a series of paths. From the root node, there is one and only one path to the final leaf node, ensuring that the output of the DT is unique and that the results of the tree can be used to classify and predict data. The specific application process of DTs consists of three steps: Firstly, analyse the training sample set and generate an inverted tree topology through recursive calls. Secondly, analyse each path of the generated inverted tree from the root to the leaf nodes and generate specific classification rules. Thirdly, use the generated classification rules to classify and predict the new data to get the result. There are many algorithms for DT generation. The Concept Learning System (CLS) proposed by Hunt et al. (1966) is the earliest algorithm for DT generation, after which scholars have successively proposed ID3, C4.5, C5.0, CART and other algorithms (Quinlan, 1996), all of which can be regarded as improvements or derivatives of the CLS algorithm.

XGBoost (Extreme Gradient Boosting) is an integrated tree-based learning algorithm that represents an advanced gradient boosting system, proposed by Chen et al. (2015). For a given credit dataset with n samples and m features, the XGBoost model draws on the ideas of Gradient Boosting (GB), using feature sampling techniques to prevent overfitting and controlling the complexity of the model through regular terms. Similar to GB, both use additive functions and control the complexity of the model by means of a self-defined Data Matrix. The base classifier is then trained from the initial training set to improve the efficiency of each iteration, where the base classifiers are weak classifiers. The distribution of training samples is then adjusted according to the performance of the base classifier, memory consumption is reduced by using an approximation algorithm that finds split nodes. The misclassified training samples receive continuous attention in subsequent training and the weights are adjusted. The next base classifier is then trained based on the adjusted sample distribution and iterated until the number of base classifiers reaches a pre-specified number, which is eventually weighted and integrated. A weak classifier here means that the model performs only marginally better than a random guess, and the Boosting algorithm is suitable for addable base classifiers to minimise the loss function provided. The boosting tree uses an additive function algorithm with a forward distribution algorithm to implement the optimisation process of learning. The loss function measures how well the model fits the data at hand, and each step of the optimisation is easy to implement when the loss function is a squared loss and exponential loss function. In the XGBoost algorithm, the loss function is extended to the objective function by adding a regularisation term (Chen et al., 2015):

$$L_k\big(F(x_i)\big) = \sum\nolimits_{i=1}^{n} \psi\big(y_i, F_k(x_i)\big) + \sum\nolimits_{k=1}^{K} \Omega(f_k) \tag{44}$$

$F_k(x_i)$ denotes the prediction about the ith sample at the kth lift, and $\Omega(f_k) = \gamma T + 0.5 * \lambda\|\omega\|^2$. In regularisation, γ is the complexity parameter and λ is a fixed coefficient. $\Omega(\cdot)$ is the regularisation term that penalises the complexity of the model. The regularised objective function is inspired by the regularised greedy forest algorithm and tends to smooth the contribution of the base classifier to avoid overfitting from occurring; with the regular term removed, the objective function is reduced to the loss function of GB. However, it is difficult to optimise directly in the function space for the objective function represented. Similarly, XGBoost is trained in the additive form at step k as:

$$L_k\big(F(x_i)\big) = \sum\nolimits_{i=1}^{n} \psi\big(y_i, F_{k-1}(x_i) + f_k(x_i)\big) + \Omega(f_k) \tag{45}$$

Unlike traditional GBDT algorithms which only use information from the first-order derivatives, the XGBoost algorithm performs a second-order Taylor expansion of the loss function and adds a regular term to the objective function to find the optimal solution, in order to weigh the decline of the objective function against the complexity of the model and avoid overfitting. The Taylor expansion of the objective function and the introduction of the regularisation term.

$$Obj^{(t)} = \sum\nolimits_{i=1}^{n} L\left(y_i, y_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) + constant$$

$$\approx \sum_{i=1}^{n} \left[L\Big(y_i, y_i^{(t-1)}\Big) + \delta_{y_i^{(t-1)}} L\Big(y_i, y_i^{(t-1)}\Big) f_t(x_i) + \frac{1}{2}\delta^2_{y_i^{(t-1)}} y_i^{(t-1)} f_t^2(x_i)\right] + \Omega(f_t) + constant \tag{46}$$

Divide the constant term and find the first-order derivative g_i and second-order derivative h_i for each sample, grouping the objective functions by the leaf node statute.

$$Obj^{(t)} = \sum\nolimits_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \Omega(f_t)$$

$$= \sum_{j=1}^{T} \left[G_j \omega_j + \frac{1}{2} \left(H_j + \lambda \right) \omega_j^2 \right] + \gamma T \tag{47}$$

The leaf node weight ω_j in the above equation has a closed-form solution, and the solution and corresponding objective function values are as follows.

$$\omega_j^* = \frac{G_j}{H_j + \lambda} \tag{48}$$

$$Obj = \frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma T \tag{49}$$

Finally, the XGBoost algorithm corrects the estimation of the data in each training session in order to modify the weights and penalise the misclassified samples by increasing their weights, so that the misclassified points are continuously focused on, and after several iterations, a number of base classifiers are obtained, which are then integrated and fused and weighted (i.e. the larger the error rate the smaller the weight of the base classifier, the smaller the weight of the base classifier) or allowed to vote to obtain a final integrated learning model.

5.3.2 Feature Selection

For the pre-processed dataset, the data dimensionality is still high, and in order to select the best feature set, we further perform feature filtering by XGBoost. Ranking the importance of feature variables is an important application of XGBoost (Chen et al., 2019). The importance of a feature is the sum of the number of times it appears in all trees, meaning that the more an attribute is used to build a DT in the model, the more important it is in comparison. XGBoost is also implemented with a number of optimisation improvements: Firstly, the traditional greedy algorithm of enumerating all possible splitting points for each feature is inefficient to find the best segmentation point, the XGBoost algorithm enumerates several possible candidates for the segmentation point based on the percentile method and then finds the best segmentation point from the candidates. Secondly, the XGBoost algorithm takes into account the case where the training data is sparse, and can specify the default direction of branching for missing values or specified values. This greatly improves the efficiency of the algorithm.

5.3.3 Imbalance Learning Models

Tomek Links

Tomek improved the Condensed Nearest Neighbor (CNN) in 1976 by proposing a framework for detecting noisy data in the majority class by under-sampling samples from the bounded majority class without destroying the underlying information in the data space. That is, if two samples belong to different classes and are close together, they can be linked into a Tomek link pair. Once two samples form a Tomek link pair, it means that either one of the samples is noisy or both samples are on the boundary. The idea is that given a pair of samples (S_i, S_j), S_i belonging to the majority class and S_j belonging to the minority class, the distance between the two points is denoted by $d(S_i, S_j)$, and S_i and S_j form a Tomek link pair if there is no arbitrary sample S_k such that $d(S_i, S_k) < d(S_i, S_j)$ or $d(S_j, S_k) < d(S_i, S_j)$ holds. In order to avoid the effect of a small number of minority samples on classification, the majority of samples in each Tomek link pair are removed and the majority of samples after under-sampling are denoted as S_{maj}. This operation makes the boundary between the two classes obvious and reduces the effect of class overlap on classification performance.

Instance Hardness Threshold (IHT)

Smith et al. (2014) proposed the concept of instance hardness (IH) for the problem of imbalanced data for binary classification. This approach expresses the probability of a data point in the training set being misclassified in terms of the property of IH, i.e. data samples that lie on the boundary between two classes or exist in noisy form have a higher IH value due to the fact that the learning algorithm forces them to correctly overfit a given training sample $\langle x_i, y_i \rangle$, $p(y_i|x_i, h)$ is the conditional probability of the label y_i given by the weak learner h for the input feature vector x_i. The smaller the value of $p(y_i|x_i, h)$, the more incorrect h will be (Le et al., 2018). The IH of the training sample $\langle x_i, y_i \rangle$, denoted by i, with respect to h, is as follows.

$$IH_h(\langle x_i, y_i \rangle) = 1 - p(y_i|x_i, h) \tag{50}$$

In practice, h is induced by the learning algorithm g trained on t using the hyperparameter α, i.e., $h = g(t, \alpha)$. Then, $IH_h(\langle x_i, y_i \rangle) = 1 - p(y_i|x_i, t, h)$, but since y_i is conditionally independent of t given h that we can use $p(y_i|x_i, h)$. Thus, the hardness of the instance is dependent on the instances in the training data and the algorithm used to generate h. There are a number of methods that can be used to calculate the hardness of an instance, such as analysing the distribution of instances in t based on their category. To further investigate what causes IH in general, the dependence of IH on particular assumptions can be reduced by adding instance hardness to the set of hypothesis H and weighting each $h \in H$ by $p(h|t)$.

$$IH(\langle x_i, y_i \rangle) = \sum_H (1 - p(y_i|x_i, h))\, p(h|t)$$

$$= \sum_H p(h|t) - \sum_H p(y_i|x_i, h)\, p(h|t)$$

$$= 1 - \sum_H p(y_i|x_i, h) p(h|t) \quad (51)$$

SMOTE

In terms of oversampling, the Synthetic Minority Oversampling Technique (SMOTE) sampling algorithm is used, which differs from traditional oversampling methods in that it directly replicates minority samples to achieve data equalisation by creating new minority sample points from the original minority sample points (Chawla et al., 2002). For a given data set $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, where $x_i \in R^m, y_1 \in \{+1, -1\}, i = 1,2, \dots, n$, for a minority class sample x_i, firstly find its k neighbours among the sample points of the same minority class, i.e. the k nearest samples in the sample space to x_i, and then construct its attribute value $r_{ij}, j = 1, 2, \dots, m$ for each attribute of the sample. For each attribute j a sample x_l^j is randomly selected among the k neighbouring samples, and the difference between the original minority class sample x_l and the sample x_l^j on attribute j is multiplied by a random number within [0,1] plus the value of the original minority class sample x_l on attribute j. The specific formula is as follows:

$$r_{ij} = x_{l,j} + \left(x_{l,j}^j - x_{l,j}\right) * rand[0,1] \quad (52)$$

where $x_{l,j}$ is the value of the original minority class sample x_l on attribute j and $x_{l,j}^j$ is the value of $x_{l,j}$ on attribute j, $rand[0,1]$ is the random number within $[0,1]$. Then, the generated new minority samples are $[r_{l1}, r_{l2}, \dots, r_{lm}]$. Finally, the above operation is repeated according to the number of samples that need to be synthesised manually. The new minority class samples are added to the original dataset, thereby balancing the dataset.

Bagging

Bagging, also known as bootstrap, is an integrated learning method based on the idea of sampling with put-back (Breiman, 1996). The main idea is to partition the original dataset into multiple training subsets, i.e., before each iteration, a random sample is taken from the training set to be trained, based on this approach, the weak classifier generated in each iteration does not depend on the previously trained In this way, the weak classifiers generated in each iteration do not depend on the previously trained classifiers, meaning that each of the generated weak classifiers is not directly related to each other and can be trained in parallel. After all the sub-classifiers are trained, the test set is predicted separately, and the classification result with the most votes is selected as the final result of the whole classification system by voting (see Figure 5.2).

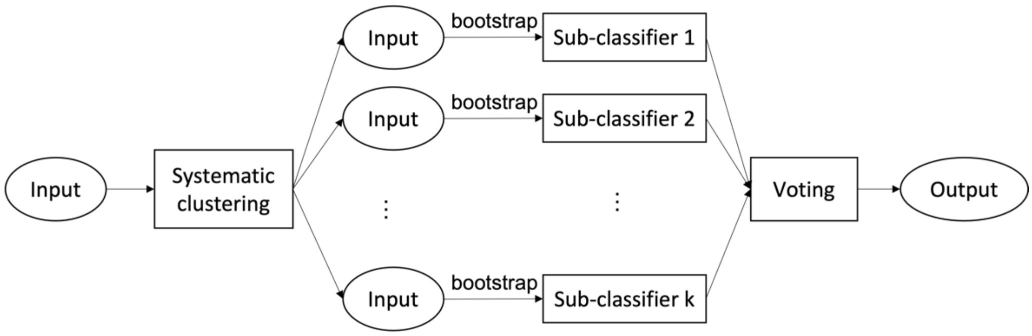


Figure 5.2 Working process of the bagging.

Easy Ensemble

Easy Ensemble is an efficient data augmentation algorithm for extremely imbalanced data. It combines under-sampling methods with ensemble learning by randomly dividing the majority class samples into the same number of subsets as the minority class, and then merging each of the removed majority class subsets with the minority class sample set to obtain a new training subset with a balanced proportion of sample classes (Liu et al., 2009). The Adaboost classifier is trained on the new subset, and finally, a strong classifier is obtained by integrating all the base classifiers using simple voting. Specifically, the number of minority class samples is assumed to be P and the number of majority class samples is N. P number of samples are randomly sampled from the majority class and combined with the minority class samples into the base classifier for training. Repeated sampling trains T base classifiers, the predicted probabilities of the T base classifiers are summed, and then the classification is determined by the $sign$ function.

$$H_i(x) = sign\left(\sum_{j=1}^{S_i} \alpha_{i,j} h_{i,j}(x) - \theta_i\right) \tag{53}$$

$$H(x) = sign\left(\sum_{i=1}^{T}\sum_{j=1}^{Si} \alpha_{i,j} h_{i,j}(x) - \sum_{i=1}^{T} \theta_i\right) \tag{54}$$

where $h_{i,j}$ is the base classifier, $\alpha_{i,j}$ is the weight of corresponding base classifier. S_i is the number of iterations for each base model, H_i represents the ensemble of each base model, and θ_i is the threshold of each ensemble model.

Cost-Sensitive

Unlike sampling methods that use different sampling strategies by changing the distribution of numbers, cost-sensitive learning methods address the problem of data imbalance by setting the cost of misclassification for different samples (Elkan, 2001). The basic idea is to create a misclassification cost matrix, which can be regarded as the cost of misclassifying a sample from one class into another class, where the cost of the correct class is zero. Assuming that the cost matrix C is known, cost-sensitive learning classifies a sample x into class j according to the loss minimisation criterion.

$$\emptyset^*(x, j) = \sum\nolimits_i p(i|x) C_{ij} \tag{55}$$

where $\sum_i p(i|x) C_{ij}$ denotes the expected loss of classifying sample x into class j, $p(i|x)$ denotes the posterior probability that the sample belongs to class i, and C_{ij} denotes the misclassification cost of classifying sample i into class j. It is worth noting that when all the elements of the cost matrix C have a value of 1, i.e., $C_{ij} = 1$ for any i and j, the cost-sensitive learning degenerates to a traditional classification learning algorithm that seeks the lowest classification error rate. The cost due to misclassification can be described by a misclassification cost matrix, Table 5.1 shows the cost matrix for the most common two-class classification.

Table 5.1 Confusion and cost matrix of Cost Sensitive method.

		Actual classes	
		Minority class	Majority class
Predicted classes	Minority class	$C(+,+)$	$C(+,-)$
	Majority class	$C(-,+)$	$C(-,-)$

Where $C(+,+)$ denotes the cost of partitioning the minority class samples to the majority class, $C(-,+)$ denotes the cost of partitioning majority class samples to minority class. When dealing with imbalanced data, the identification of minority class samples is more important than the identification of majority class samples. Thus, the cost of misclassifying minority classes is higher than the cost of misclassifying majority classes, which is expected to balance the difference in numbers between samples, i.e., $C(+,-) > C(-,+)$. If the classification is correct, it usually corresponds to a 0 penalty, i.e., $C(+,+) = C(-,-) = 0$.

5.4 Experiment Design

This research selects data from Lending Club 2007-2018 Q4, with 151 original features and 1345310 observations. Based on the loan status, Fully Paid observations are marked as the compliance sample (i.e. negative sample) and Charged off as the default sample (positive sample), with 1303607 negative sample records and 261655 positive sample records in the dataset, the imbalance ratio is 0.20 (positive samples : negative samples = 1:5).

5.4.1 Data Processing

The dataset with 87 features is obtained by only processing the basic data cleaning to remove noise. (1) Regarding the pre-processing of the data, features with more than 40% of missing values are removed and the missing values which attribute less than or equal to 40% are filled with mean. The removed features are as Table 5.2 shows below. (2) We remove features which lack useful information (e.g., ‘address’, ‘zip_code’). (3) The redundant information is manually removed. Due to the similarity of some features, for example, ‘loan_amnt’, ‘funded_amnt_inv’ and ‘funded_amnt’ are considered as the same record for loan amount (Song et al., 2020). In this case, only one feature is retained. (4) Post-loan features, such as ‘total_rec_int’, and ‘last_pymnt_amnt’, which leak the situation of monthly payback status are removed (Namvar et al., 2018). Further, the categorical features are segmented with one-hot encoding[32].

Table 5.2 Pre-processing of features exclude from the dataset due to the higher missing percentage above 40%

Feature name	% of missing value	Feature name	% of missing value
member_id	100	verification_status_joint	98.1
next_pymntd	100	dti_joint	98.1
orig_projected_additional_accrued_interest	99.7	annual_inc_joint	98.1
hardship_type	99.6	debt_settlement flag_date	97.5
hardship_reason	99.6	settlement_status	97.5
hardship_status	99.6	settlement_date	97.5
deferral_term	99.6	settlement_amount	97.5
hardship_amount	99.6	settlement_percentage	97.5
hardship_start_date	99.6	settlement_term	97.5
hardship_end_date	99.6	desc	90.8
payment_plan_start_date	99.6	mths_since_last_record	83
hardship_length	99.6	mths_since_recent_bc_dlq	76.3
hardship_dpd	99.6	mths_since_last_major_derog	73.7
hardship_loan_status	99.6	mths_since_recent_revol_delinq	66.6
hardship_payoff_balance_amount	99.6	il_util	65.4
hardship_last_payment_amount	99.6	mths_since_rcnt_il	61.1
sec_app_mths_since_last_major_derog	99.5	all_util	60
sec_app_revol_util	98.6	open_acc_6m	60
revol_bal_joint	98.6	inq_last_12m	60
sec_app_chargeoff_within_12_mths	98.6	total_cu_t	60
sec_app_open_act_il	98.6	open_rv_12m	60
sec_app_open_acc	98.6	open_il_12m	60
sec_app_mort_acc	98.6	open_rv_24m	60
sec_app_inq_last_6mths	98.6	max_bal_bc	60
sec_app_earliest_cr_line	98.6	total_bal_il	60
sec_app_fico_range_high	98.6	inq_fi	60
sec_app_num_rev_accts	98.6	open_il_24m	60
sec_app_fico_range_low	98.6	open_act_il	60
sec_app_collections_12_mths_ex_med	98.6	mths_since_last_delinq	50.5

The dataset with 17 features which represents the feature selection of some scholars with respect to credit risk assessment on Lending Club's dataset[33] is manually selected based on existing articles (Babaei and Bamdad, 2020; Bastani et al., 2019). The specific features we select are shown in Table 5.3.

[32] The N states are encoded using N-bit status registers, each with its own independent register bit and only one of which is valid at any given time. For each feature, if it has N possible values, then after one-hot encoding it becomes N binary features. These features are mutually exclusive, with only one activation at a time. As a result, this solves the problem that the classifier does not handle attribute data well. It also serves to expand the features to a certain extent.

[33] Description source: http://rstudio-pubs-static.s3.amazonaws.com/290261_676d9bb194ae4c9882f599e7c0a808f2.html

Table 5.3 The dataset with 17 features.

Feature name	Description
Annual_Income	The self-reported annual income provided by the borrower during registration.
Delinqency_2yrs	The number of 30+ days past-due incidences of delinquency in the borrower's credit file for the past 2 years.
Employment_Length	Employment length in years. Possible values are between 0 and 10 where 0 means less than one year and 10 means ten or more years.
Home_Ownership	Own, rent, mortgage.
Inquery_Last_6mths	The number of inquiries in the past 6 months.
Loan_Amount	The listed amount of the loan applied for by the borrower.
Purpose	14 loan purposes: wedding, credit card, car loan, major purchase, home improvement, debt consolidation, house, vacation, medical, moving, renewable energy, educational, small business, and other.
Open_Account	The number of open credit lines in the borrower's credit life.
Fico	A measure of credit risk, based on credit reports that range from 300 to 850. FICO is a registered trademark of Fair Isaac Corporation
Grade	LC assigned loan grade. (7 loan grades totally for borrowers from A to G, A-grade being the best grade.)
Sub_Grade	LC assigned loan subgrade. (35 loan subgrades totally for borrowers from A1 to G5, A1 denotes to the best subgrade.)
Dti	A ratio calculated using the borrower's total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower's self-reported monthly income.
Revoling_Utilisation	Revolving line utilisation rate, or the amount of credit the borrower is using relative to all available revolving credit.
Interest_Rate	The interest rate on the loan paid by the borrower.
Installment	The monthly payment owed by the borrower if the loan originates.
Public_Record	Number of derogatory public records.
Months_Since_Last_Delinqency	The number of months since the borrower's last delinquency.

5.4.2 Experimental Setup

DT and XGBoost were chosen as the base classifier as the best rank baseline model, and the parameters were selected by the Grid search mechanism [34] . Use max_depth=20, min_samples_leaf=5, and min_samples_split=2 for DT, and for XGBoost, max_depth=6, n_estimators=100 are used as parameters. To further avoid overfitting, the early stopping mechanism[35] is used to ensure that the XGBoost loss function stops training when it reaches

[34] Among all the alternative parameter choices, look for the best-performing parameter by iterating through the loop and testing every possibility. For k parameters, there are m_k values to take. The total number of configurations: $N = m_1 \times m_2 \times \ldots \times m_k$. If the hyperparameters are continuous, some empirical values can be chosen empirically, such as learning rate: $\alpha \in \{0.01, 0.1, 0.5, 1.0\}$, for different combinations of these hyperparameters, train and test the performance on the validation set separately, and finally select the best performing parameter configuration from them.

[35] The expectation of the ideal model is that as the model's error in the training set decreases, its error performance in the validation set does not deteriorate. Conversely, when the model performs well on the training set and poorly on the validation set, we assume that the model is overfitting. By calculating the model's performance on the validation set during training and stopping training when the model's performance on the validation set starts to decline, the problem of overfitting by continuing training could be avoided.

its minimum point (see Figure 5.3), avoiding too long and ineffective learning. This expiriment selects 20% of the dataset as the test set, i.e., 269062 observations with 53454 default samples (denotes as 1 in classification) and 215608 fully paid samples (denotes as 0 in classification). And 80% of the dataset as the train set, i.e., 1076248 observations with 215105 default samples and 861143 fully paid samples. The further adjustment in the imbalance ratio is processed in the train set.

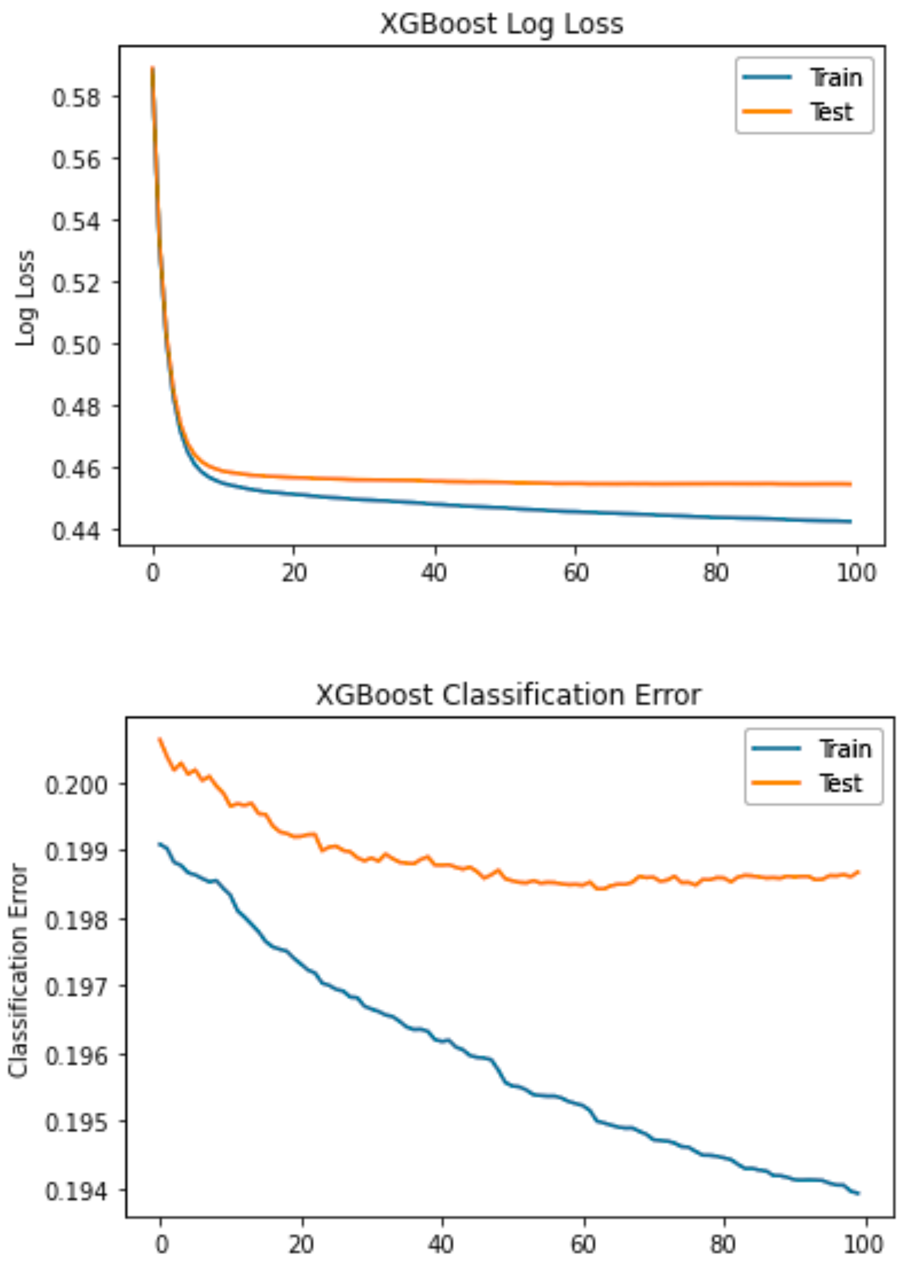


Figure 5.3 Log loss curve and classification error changes of XGBoost.

5.4.3 Evaluation Metrics

For the evaluation of binary imbalance data, traditional evaluation metrics such as accuracy cannot effectively assess the performance of the model for imbalance data problems due it cannot reflect the identification situation of true positive and false positive. Thus, F-measure, AUC, G-mean, Matthews correlation coefficient (MCC) and Bookmaker Informedness (BM) are chosen as evaluation metrics for imbalance data credit risk assessment, emphasising the significance of true positive and true negative identification for the model. F-Measure is a composite evaluation indicator that measures Precision and Recall comprehensively and is calculated as follows.

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \tag{56}$$

AUC is used to represent the area enclosed by the ROC curve and the coordinate axis, so the value of AUC is not greater than 1. Since the ROC curve always lies above $y = x$, the value of AUC ranges between 0.5 and 1. The closer the AUC is to 1, the better the reliability of the model, and equal to 0.5, the model is equivalent to random guessing and has no application value.

G-mean is a comprehensive evaluation metric measuring the accuracy of positive and negative classes, which contains $Sensitivity$ and $Specificity$, where $Sensitive = TP/(TP + FN)$, which measures the classifier's ability to identify positive class samples, and $Specificity = TN/(TN + FP)$, which measures the classifier's ability to identify negative classes ability. The greater the true-positive rate and true-negative rates of the classification model, the better the performance of the model; correspondingly, the greater the false-positive and false-negative rates of the classification model, the worse the performance of the model. The G-mean is therefore a good indicator of the performance of the classification model and is not affected by the imbalance of the dataset (Bekkar et al., 2013).

$$G - mean = \sqrt{Sensitive * Specificity} \tag{57}$$

The MCC also takes into account true-positives, true-negatives, false-positives and false-negatives, and is generally considered to be a more balanced indicator that can be applied even when the sample content of the two categories is extremely different (Boughorbel et al., 2017; Zhu, 2020). The MCC is essentially a correlation coefficient describing the relationship between the actual and predicted classifications and takes values in the range $[-1,1]$, where a value of 1 indicates a perfect prediction of the subject, a value of 0 indicates that the prediction is not as good as the random prediction, and -1 means that the predicted classification and the actual classification do not agree at all.

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{58}$$

BM is a comprehensive assessment metric that is often applied to measure the results of imbalance data (Zhang & Chi, 2021; Chicco et al., 2021).

$$BM = TPR + TNR \tag{59}$$

5.5 Results

Both datasets were evaluated with the same model. Thus, the following results are obainted. Table 5.4, 5.5 and 5.6 shows the classification results with different base estimators of the models when specific features are selected, while Table 5.7, 5.8 and 5.9 shows the classification results with a different base estimator of our dataset with more complete features by pre-processing the data.

Tables 5.4, 5.5 and 5.6 show the results of each model based on 18 feature-specific datasets with XGBoost, DT and LR as base classifiers, respectively. The AUC shows that the effectiveness of this dataset is not good. In Table 5.4, the results for XGBoost as a base classifier, for example, show that the AUCs of basically all models, including the imbalanced learning model, are below 0.6, except for ES-XGBoost, which has the highest score of 0.64737 but is still poor. In addition, the F-measure can be used to evaluate the results of the combined precision and recall trade-offs, and the low F-measure for the feature-specific dataset indicates that the models are not effective in identifying positive samples in the data. The F-measure for the feature-specific dataset evaluation results are all below 0.5, which indicates that the overall effectiveness of identifying positive samples in this data is poor, possibly due to the imbalanced data problem as in the previous study. The G-mean, however, is not affected by the imbalanced data and thus better reflects the training effect of the model, which to some extent reflects the true effect of the model training, but the overall result is still low, indicating that the model is not able to discriminate between positive and negative samples. Similarly, the MCC index is close to 0, which further indicates that the model results can barely discriminate between positive and negative samples. This may be due to the low classification effect caused by the imbalance of the data which makes it difficult to identify a few samples. However, we added a variety of imbalance learning methods and we found that the overall assessment improved again with the addition of further imbalance learning methods, with better results than a single classifier for each metric, but the overall results were still poor. The ES-XGBoost has the highest rank among the overall metrics. This result suggests to some extent that although the unbalanced learning methods improve the evaluation of this data sample, they still do not effectively address the problem of misclassification of positive and negative samples, so the problem may lie in the data set itself. Similar results could be observed from Table 5.5 and Table 5.6 that DT and LR as base classifiers. The results suggest that the conclusions drawn in previous studies, i.e. the inefficiency of selecting a small number of specific features and thus attributing their results to unbalanced data, still merit further observation.

Table 5.4 Classification results with XGBoost as the base estimator (specific dataset).

Model	AUC	Rank	F-Measure	Rank	MCC	Rank	G-Mean	Rank	BM	Rank	Ave rank
XGBoost	0.54778	5	0.28290	5	0.09290	6	0.4855	5	0.09555	5	5.2
Bagging-XGBoost	0.53561	7	0.16159	7	0.14597	2	0.3071	7	0.07122	7	6
Easy Ensemble-XGBoost	**0.64737**	1	**0.42442**	1	**0.23780**	1	**0.64638**	1	**0.29474**	1	1
IHT-XGBoost	0.56479	2	0.31906	2	0.11740	3	0.53055	2	0.12957	2	2.2
Tomek Links-XGBoost	0.55353	3	0.29685	3	0.10090	4	0.50344	3	0.10705	3	3.2
SMOTE-XGBoost	0.55126	4	0.28916	4	0.09914	5	0.49221	4	0.10253	4	4.2
CS-XGBoost	0.54611	6	0.27590	6	0.09198	7	0.47514	6	0.09223	6	6.2

Table 5.5 Classification results with DT as the base estimator (specific dataset).

Model	AUC	Rank	F-Measure	Rank	MCC	Rank	G-Mean	Rank	BM	Rank	Ave rank
DT	0.52959	5	0.13031	5	0.14799	7	0.26949	5	0.05918	5	5.4
Bagging-DT	0.52944	7	0.12910	7	0.14906	6	0.26794	7	0.05888	7	6.8
Easy Ensemble-DT	**0.65624**	1	**0.43407**	1	**0.25336**	1	**0.65584**	1	**0.31247**	1	1
IHT-DT	0.61269	2	0.39409	2	0.20222	2	0.53940	2	0.22538	2	2
Tomek Links-DT	0.55203	3	0.21639	3	0.18446	3	0.36427	3	0.10405	3	3
SMOTE-DT	0.53145	4	0.13886	4	0.14966	5	0.27963	4	0.06290	4	4.2
CS-DT	0.53072	6	0.13520	6	0.14972	4	0.27523	6	0.06144	6	5.6

Table 5.6 Classification results with LR as the base estimator (specific dataset).

Model	AUC	Rank	F-Measure	Rank	MCC	Rank	G-Mean	Rank	BM	Rank	Ave rank
LR	0.51529	5	0.09086	5	0.07556	5	0.22463	5	0.03058	5	5
Bagging-LR	0.50771	6	0.04696	6	0.05541	6	0.15721	6	0.01542	6	6
Easy Ensemble-LR	0.62498	3	0.40228	3	0.20087	3	0.62285	3	0.24997	3	3
IHT-LR	0.61216	4	0.39267	4	0.18704	4	0.57783	4	0.22432	4	4
Tomek Links-LR	0.50040	7	0.00434	7	0.00819	7	0.04672	7	0.00080	7	7
SMOTE-LR	**0.62736**	1	**0.40457**	1	**0.20529**	1	**0.62658**	1	**0.25472**	1	1
CS-LR	0.62504	2	0.40234	2	0.20096	2	0.62292	2	0.25008	2	2

Table 5.7 Classification results with XGBoost as the base estimator (our dataset)

Model	AUC	Rank	F-Measure	Rank	MCC	Rank	G-Mean	Rank	BM	Rank	ave rank
XGBoost	0.99009	2	**0.98771**	1	**0.98471**	1	0.99006	2	0.98018	2	1.6
Bagging-XGBoost	0.98964	4	0.98754	2	0.98451	2	0.98960	5	0.97927	5	3.6
Easy Ensemble-XGBoost	**0.9926**	1	0.98622	5	0.9828	5	**0.99260**	1	**0.98521**	1	2.6
IHT-XGBoost	0.8460	7	0.61716	7	0.55565	7	0.83204	7	0.69201	7	7
Tomek Links-XGBoost	0.98992	3	0.98740	3	0.98433	3	0.98988	4	0.97983	4	3.4
SMOTE-XGBoost	0.98938	5	0.98663	4	0.98338	4	0.98934	6	0.97875	6	5
CS-XGBoost	0.9809	6	0.93413	6	0.96403	6	0.99001	3	0.98003	3	4.8

Table 5.8 Classification results with DT as the base estimator (our dataset)

Model	AUC	Rank	F-Measure	Rank	MCC	Rank	G-Mean	Rank	BM	Rank	Ave rank
DT	0.98648	6	0.98032	2	0.97547	2	0.98643	6	0.97295	6	4.4
Bagging-DT	0.98793	2	**0.98422**	1	**0.98036**	1	0.98788	2	0.97586	2	1.6
Easy Ensemble-DT	**0.98979**	1	0.97917	4	0.9740	4	**0.98979**	1	**0.97959**	1	2.2
IHT-DT	0.98767	3	0.97708	6	0.97138	6	0.98765	3	0.97533	3	4.2
Tomek Links-DT	0.98652	5	0.98019	3	0.97531	3	0.98647	5	0.97303	5	4.2
SMOTE-DT	0.98721	4	0.97896	5	0.97374	5	0.98718	4	0.97442	4	4.4
CS-DT	0.98566	7	0.96922	7	0.96160	7	0.98566	7	0.97132	7	7

Table 5.9 Classification results with LR as the base estimator (our dataset)

Model	AUC	Rank	F-Measure	Rank	MCC	Rank	G-Mean	Rank	BM	Rank	Ave rank
LR	0.97341	6	0.96495	3	0.**95654**	3	0.97318	6	0.94682	6	4.8
Bagging-LR	0.97350	5	**0.96575**	1	0.95757	1	0.97326	5	0.9470	5	2.6
Easy Ensemble-LR	0.97645	2	0.94742	6	0.93446	6	0.97645	2	0.95291	2	3.8
IHT-LR	0.75452	7	0.50272	7	0.41301	7	0.71409	7	0.50905	7	7
Tomek Links-LR	0.97410	4	0.96517	2	0.95677	2	0.97389	4	0.94821	4	3.4
SMOTE-LR	**0.97716**	1	0.95763	4	0.94707	4	**0.97711**	1	**0.95432**	1	2.4
CS-LR	0.97632	3	0.94794	5	0.93509	5	0.97631	3	0.95265	3	4

Tables 5.7, 5.8 and 5.9 show the results for each model with XGBoost, DT and LR as base classifiers for our dataset after pre-processing, respectively. Using Table 5.7 as an example, we find that the ES model is still the best choice for the XGBoost-based classifier based on the validity metric AUC. In this sample evaluation, the models were generally able to obtain very good classification results, with AUCs above 0.8. Even a single XGBoost model could achieve an AUC of 0.99009. This result shows that although there are data imbalances in the lending club dataset, the evaluation effect is not affected by the data imbalance when the features are comprehensive. Even the excessive use of unbalanced learning methods can lead to a decrease in effectiveness.

Furthermore, the relationship between feature selection and imbalanced data in depth is explored. For the dataset with 87 features, each base classifier and its ES model are used to measure the importance ranking of features and set different thresholds for selecting relevant features. Randomly removing a proportion of positive samples to form a more severe proportion of imbalance, allows the test to examine in detail the variation in evaluation effects between different levels of data imbalance and different feature selections. Thus, the full feature dataset with the XGBoost evaluation model is evaluated and the results are presented in Figure 5.4.

From Figure 5.4, it is observed that under the single XGBoost model evaluation, when the imbalance of the dataset is 1:5, the feature selection can get the optimal solution by removing the features with importance ranking below 0.01. As the imbalance becomes more serious, the evaluation of the model does not decrease much when the features are more complete, even if the imbalance ratio deteriorates from 1:5 to 1:500, at the feature selection threshold of 0.01 still had an AUC of 94.64%, F-measure of 94.93%. In addition, for feature selection, the model works best with a threshold of 0.01, however, compared to the results in Table 5.5, the more complete the features are, the better the model evaluation results are found. As the threshold increases, i.e., the more features we remove, the evaluation of the model becomes progressively worse, and the performance of the model decreases significantly as the number of unbalanced data problems increases. Even in the case of an imbalance greater than 1:100, too limited data can lead to the model not working properly and the metrics being 0.

An interaction plot between sample selection and feature selection with the ES-XGBoost model is further shown in Figure 5.5. As the imbalance learning model with the best average performance for both dataset, the results of ES-XGBoost shows a similar trend to the results of the single XGBoost model. For sample selection, as the ratio of imbalance increases, the effectiveness of the model evaluation decreases. And as the threshold for feature selection increases, i.e. more features are removed, the model's effectiveness also decreases. Moreover, it is worth noting that when the imbalance ratio and feature selection increase simultaneously, unlike the results of the single XGBoost model, ES-XGBoost is able to obtain reasonable evaluation results despite a high degree of imbalance and a very limited number of features, e.g. when the imbalance ratio is 1:500 and the threshold of feature selection is 0.1, the model AUC still reaches 88.742%, the F-measure of 74.154% and G-mean of 88.714%. This demonstrates that ES-XGBoost is very effective for imbalanced data processing, and is able to solve the problem of imbalanced data with large differences in binary classes. It also illustrates that for datasets with limited features, the ES-XGboost method also alleviates the poor performance to a certain extent.

To demonstrate the robustness of the results, the experimental results with LR and with DT as the base classifier are shown in Figure 5.6 which presents the Interaction heat plot of the imbalance ratio & feature selection threshold with DT as the base classifier. And Figure 5.7 shows the ES-DT heat plot. Likewise, Figure 5.8 and Figure 5.9 show the heat plot with LR and ES-LR as the method respectively. It is observed that as the imbalance indicator becomes more severe, the evaluation of each model becomes less effective. Similarly, as the threshold increases, i.e. as fewer features are used for evaluation, the models become less effective. When both the degree of imbalance and the degree of feature reduction increase, the models become sharply and severely worse, which is in line with our expectations.

Based on the evaluation results of various models with different base classifiers, we conclude that ES-XGBoost has the best evaluation results. ES addresses the issue of extremely imbalanced datasets, ensuring that the model pays more attention to the minority class. This helps in situations where one class is significantly underrepresented. While XGBoost is also known for its powerful performance, efficient computation, and ability to capture complex patterns in the data. It provides a strong base classifier for the easy ensemble to aggregate and create a robust final prediction. So, by comparing Figure 5.4 and Figure 5.5 we also conclude that ES-XGBoost is effective in alleviating the imbalance of the data with limited features. Combining the strengths of ES and XGBoost, the resulting ensemble model has the potential to provide better generalisation and accuracy, especially when dealing with imbalanced datasets and complex relationships in the data. Then, a further feature importance analysis was performed on the data by XGBoost and derived the minimum feature set. The variation of the AUC results of the model with ES-XGBoost as the classifier and modulated feature selection by threshold is shown in Figure 5.10. Since the selected feature set has 87 initial features, each feature has a relatively small importance share, and the result is selected with a feature importance rank of 0.001. It is found that the AUC reaches its highest value when the threshold is at 0.007, from which this study selects the feature set with feature importance above the threshold of 0.007 as the minimum feature set (see Table 5.10). Compared with the dataset with 17 features, the minimum feature set includes the term, verification status, number of mortgage accounts, number of trades opened in the past 24 months, number of currently active revolving trades, number of accounts currently 120 days past due (updated in past 2 months), the average current balance of all accounts, total bankcard high credit/credit limit and application type of users. This result identifies the attributions of this information in credit risk assessment.

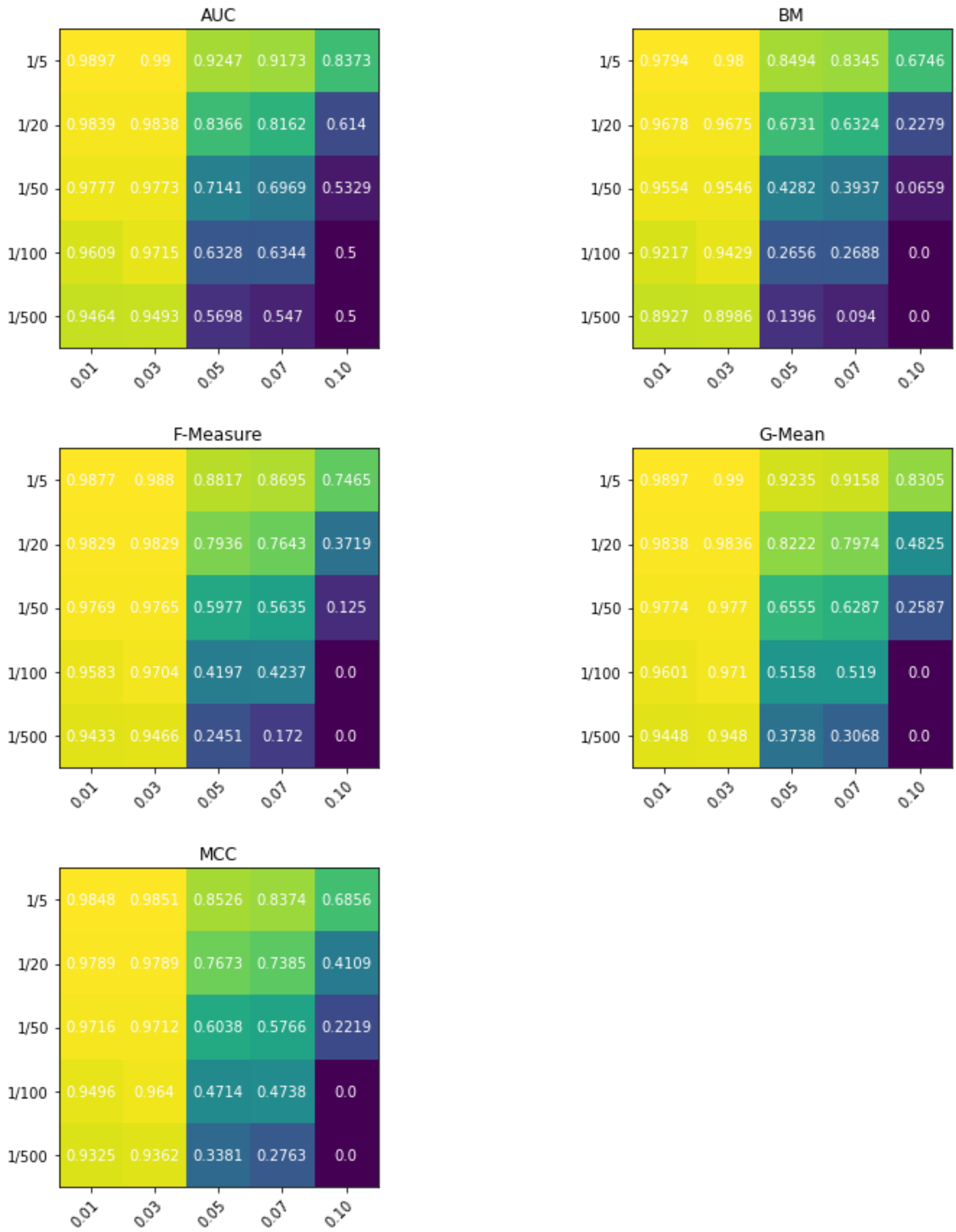


Figure 5.4 Interaction heat plot of imbalance ratio & feature selection threshold with XGBoost in selected dataset.

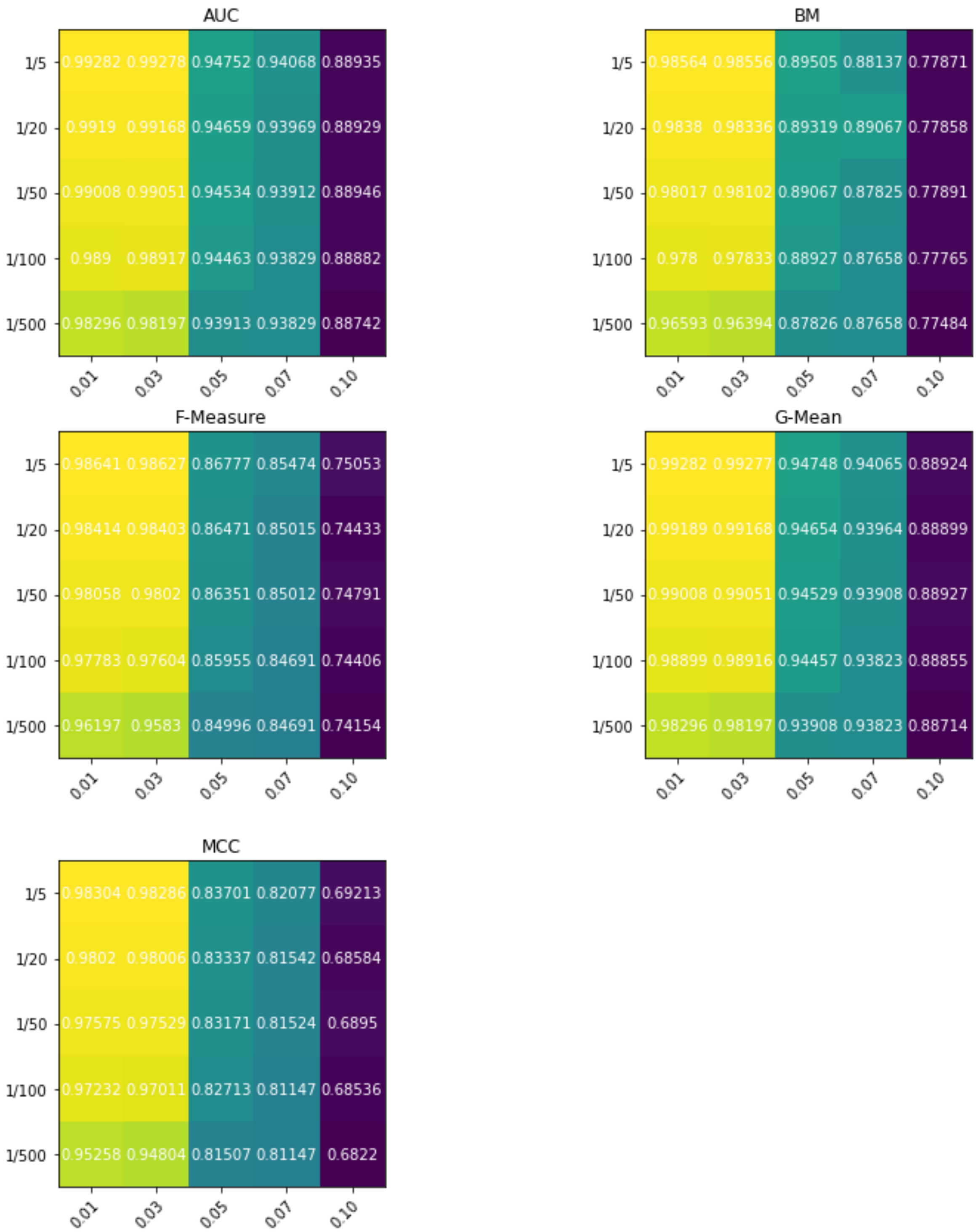


Figure 5.5 Interaction heat plot of imbalance ratio & feature selection threshold with ES-XGBoost in complete dataset.

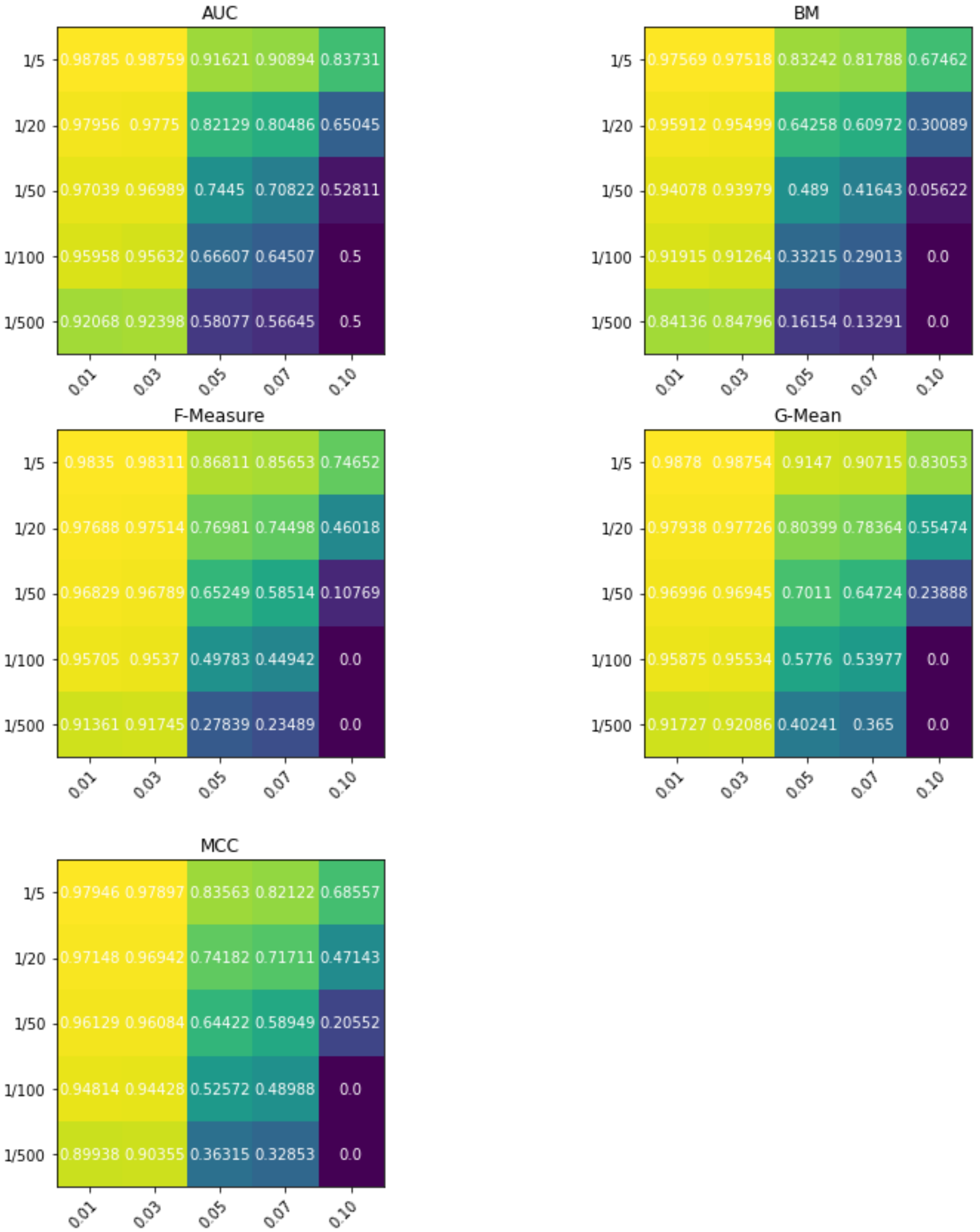


Figure 5.6 Interaction heat plot of imbalance ratio & feature selection threshold with DT in selected dataset.

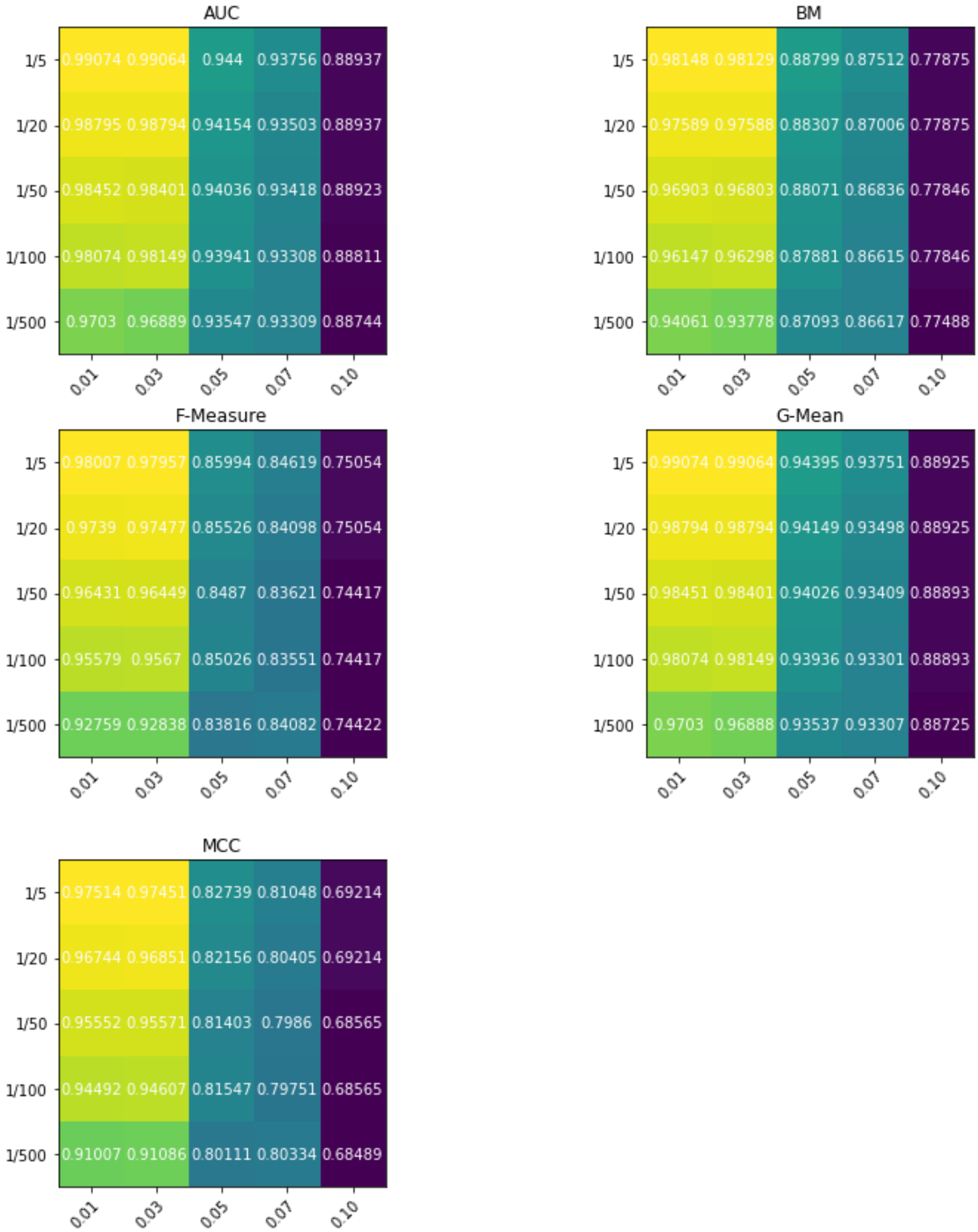


Figure 5.7 Interaction heat plot of imbalance ratio & feature selection threshold with ES-DT in complete dataset.

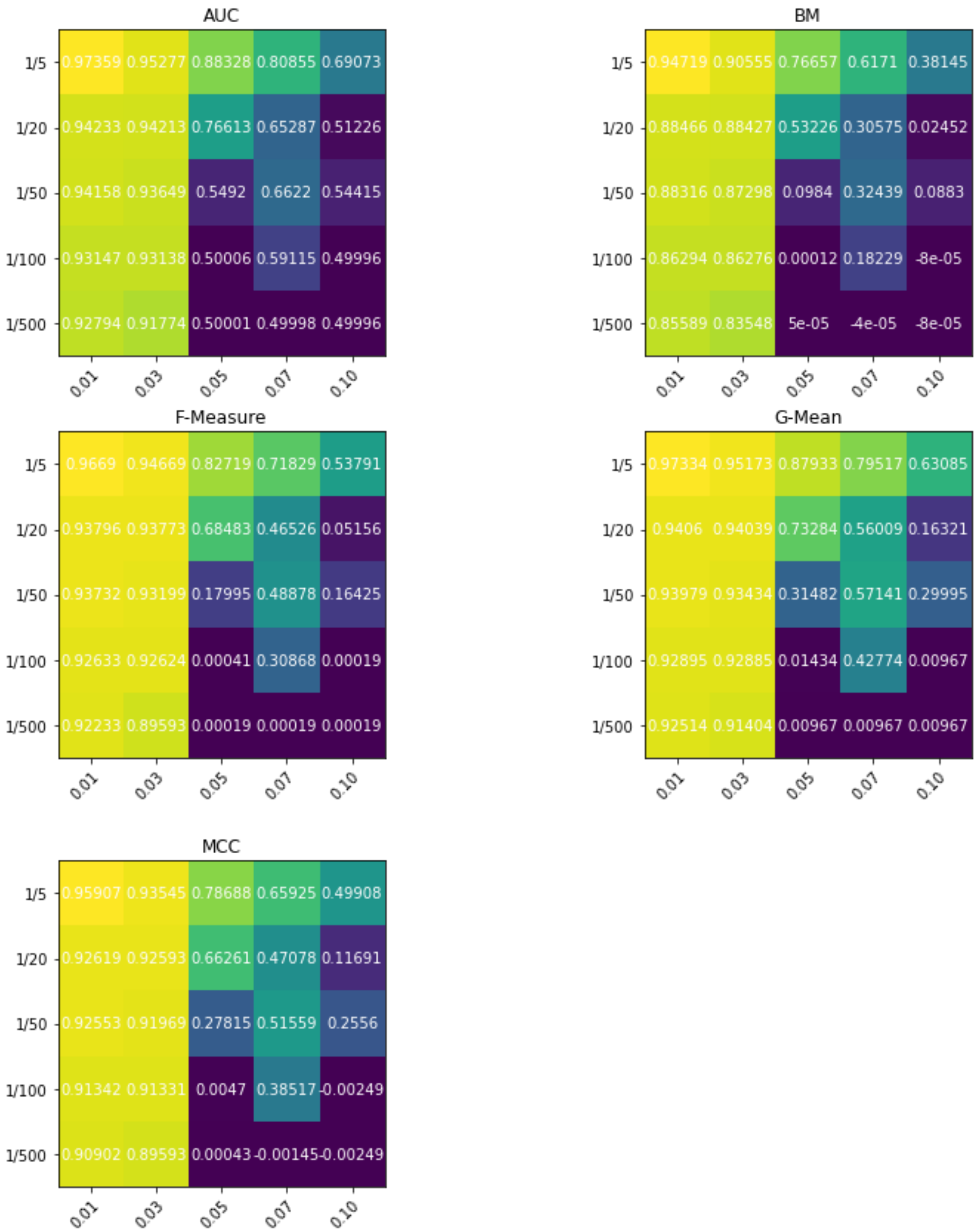


Figure 5.8 Interaction heat plot of imbalance ratio & feature selection threshold with LR in selected dataset.

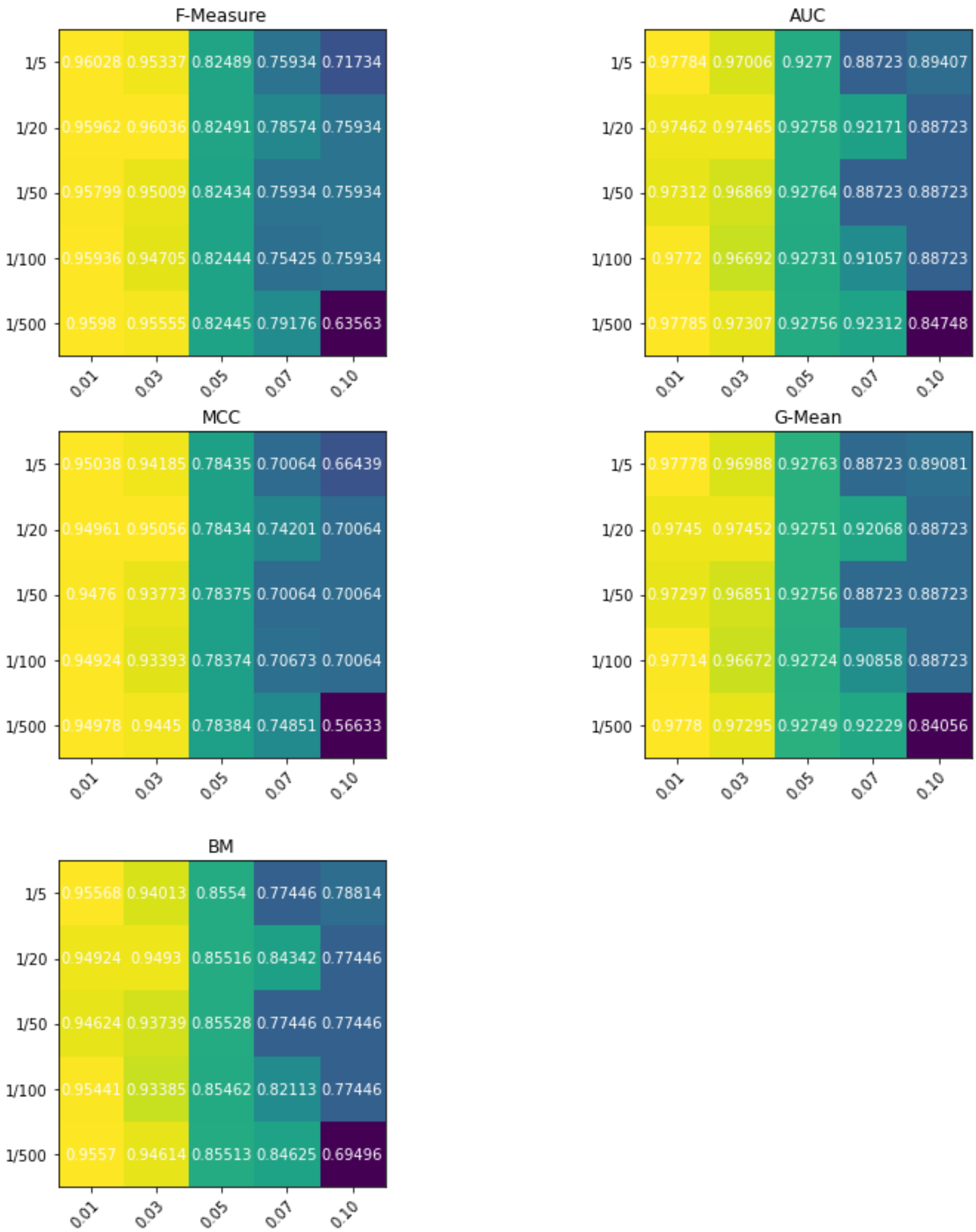


Figure 5.9 Interaction heat plot of imbalance ratio & feature selection threshold with ES-LR in complete dataset.

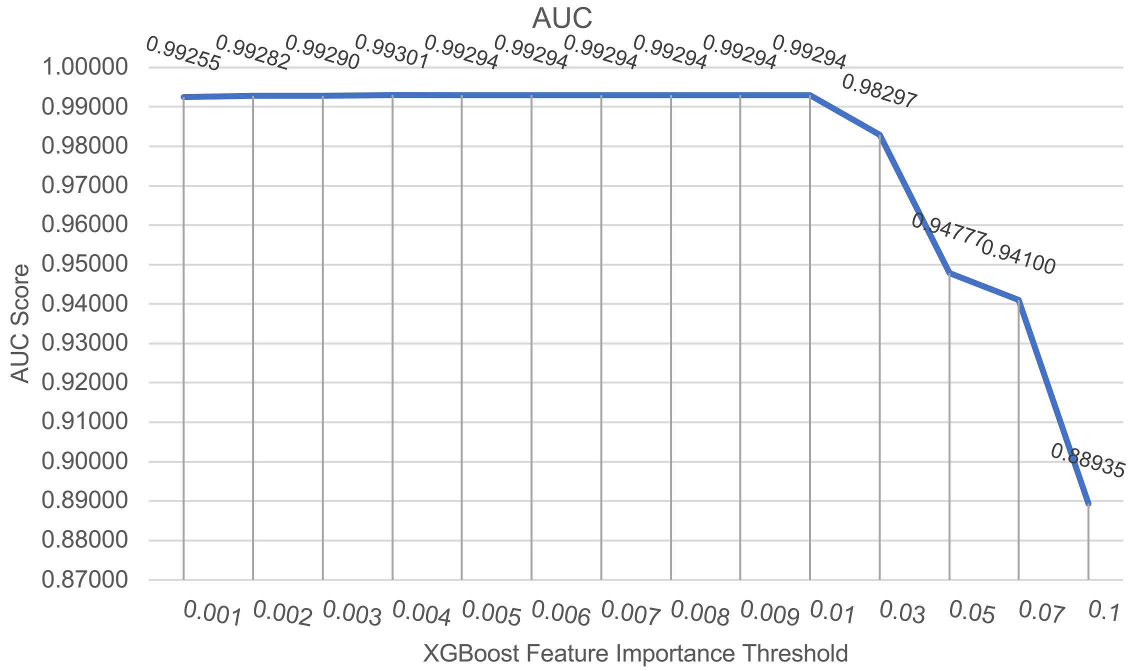


Figure 5.10 The AUC of ES-XGBoost model in different threshold.

Table 5.10 Minimum feature set based on the feature selection.

Importance rank	Feature name	Importance	Description
1	grade	0.21411	LC assigned loan grade
2	sub_grade	0.20161	LC assigned loan subgrade
3	term	0.06025	The number of payments on the loan. Values are in months and can be either 36 or 60.
4	home_ownership_MORTGAGE	0.03236	The home ownership status provided by the borrower during registration or obtained from the credit report. Our values are: RENT, OWN, MORTGAGE, OTHER
5	home_ownership_RENT	0.02474	The home ownership status provided by the borrower during registration or obtained from the credit report. Our values are: RENT, OWN, MORTGAGE, OTHER
6	verification_status_Not Verified	0.02270	Indicates if income was verified by LC, not verified, or if the income source was verified.
7	mort_acc	0.02202	Number of mortgage accounts.
8	acc_open_past_24mths	0.01662	Number of trades opened in past 24 months.
9	num_actv_rev_tl	0.01518	Number of currently active revolving trades.
10	num_tl_120dpd_2m	0.01364	Number of accounts currently 120 days past due (updated in past 2 months)
11	avg_cur_bal	0.01340	Average current balance of all accounts.
12	purpose_small_business	0.01290	A category provided by the borrower for the loan request.
13	emp_length	0.01236	Employment length in years. Possible values are between 0 and 10 where 0 means less than one year and 10 means ten or more years.
14	dti	0.01135	A ratio calculated using the borrower’s total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower’s self-reported monthly income.
15	fico	0.01053	A measure of creditworthiness, based on credit reports that range from 300 to 850. FICO is developed by Fair Isaac Corporation
16	loan_amnt	0.01033	The listed amount of the loan applied for by the borrower. If at some point in time, the credit department reduces the loan amount, then it will be reflected in this value.
17	purpose_medical	0.00828	A category provided by the borrower for the loan request.
18	delinq_2yrs	0.00821	The number of 30+ days past-due incidences of delinquency in the borrower's credit file for the past 2 years
19	installment	0.00762	The monthly payment owed by the borrower if the loan originates.
20	total_bc_limit	0.00758	Total bankcard high credit/credit limit
21	annual_inc	0.00745	The self-reported annual income provided by the borrower during registration.
22	mths_since_recent_inq	0.00744	Months since most recent inquiry.
23	application_type_Individual	0.00730	Indicates whether the loan is an individual application or a joint application with two co-borrowers.

3

5.6 Conclusions

P2P lending is becoming an important way forward in the financial sector as an innovation in the consumer finance business model. As the number of users increases and the amount of lending grows, credit risk as the core risk that needs attention is also gradually rising. When using FinTech tools for risk management, training on historical data to obtain an effective credit risk assessment model is the basis of risk control modelling, but the problem of data imbalance leads to poor performance of traditional classification and prediction models on the test set. Also due to the data-driven nature of P2P lending, how to select suitable variables for training for high-dimensional data is also the key to effective risk assessment. This chapter takes a practical approach to the problem by using the initial dataset of Lending Club 2007-2018 Q4 to investigate credit risk assessment (containing 151 features). By arranging two datasets to compare the assessment results, this chapter focuses on three research questions. Firstly, aiming to figure out whether an imbalanced data issue matters in the Lending Club dataset with complete features, this study obtained a dataset with 87 features by removing missing values and noise from the initial dataset based on the basic data pre-processing only. The good performance of this dataset illustrates that a more complete feature set would cause a less imbalanced effect on the credit risk assessment model.

Secondly, to further identify whether the feature incompleteness in the existing literature result in poor performances of the existing models, the dataset with 17 features are selected based on the dataset already used by previous papers for the imbalance data study. By using LR, DT and XGBoost as base classifiers, the performance of two datasets is obtained and compared. The results revealed that the evaluation was poor based on the dataset chosen by the previous scholars, while the imbalanced learning approach did not effectively improve the evaluation. The result illustrates that the possible problems with the dataset are not just a matter of imbalanced data. In contrast, the evaluation of the dataset with 87 features is very good and the model is still very effective in identifying defaulting users even when trained using only the base classifier. And the results show that the Easy Ensemble approach is very effective in alleviating the problem of imbalanced data. This comparison illustrates the importance of data features and that too limited feature selection can lead to serious risk assessment problems.

Finally, to investigate the solution of mitigating the imbalanced data issue with limited features, the dynamics between the imbalanced data problem and feature selection are tested based on the dataset with 87 features. Using the XGBoost algorithm for feature selection and adjusting the threshold to reduce the features, while the imbalance ratio of the training set is adjusted to make it progressively more imbalanced. Thus, a heat map is shown to represent the dynamics between the imbalance ratio and feature selection. For the traditional single-base classifier evaluation model, the evaluation model is found that becomes less effective as the imbalance increases in the absence of feature reduction. Similarly, a gradual reduction in features also makes the evaluation model less effective when the data imbalance ratio is constant. When both the imbalance ratio and the feature reduction increase, the traditional classification model deteriorates dramatically, rendering the assessment unworkable. As a representative of the unbalanced learning approach, the same heat map demonstration was also performed using Easy Ensemble. The results are consistent in that too few features and too much imbalance will both cause the assessment model to deteriorate, but with the fewest features and the highest level of unbalance, the training model maintains relatively good assessment results and does not deteriorate dramatically. This suggests that Easy Ensemble 's imbalance learning method is effective in dealing with both imbalanced data issues and problems with limited features. It also shows that too few data features can lead to poorer model evaluation when the data has some data imbalance problem. Ultimately, the minimum feature set which includes 23 features according to the XGBoost feature importance ranking and considering the model evaluation metrics and comprehensive operation is selected.

On this basis, this chapter makes the following recommendations based on the risk assessment of P2P lending. Firstly, appropriate risk management by P2P enterprises is always inseparable from the development of the credit industry, which is due to the fact that credit and finance are always complementary to each other. However, at present, there are various problems such as fragmentation of data collection scenarios, prominent data silos, serious data homogenisation and difficulty in guaranteeing data quality, thus making it difficult to match the development of data supply and risk assessment, and failing to form a useful complement to P2P lending risk management, thereby restricting the healthy and sustainable development of the P2P lending industry. Therefore, full consideration should be given to incorporating valid data into the construction of P2P lending default risk prediction models. Secondly, in the future practical application environment, the sources of data will be rich and diverse, and the processing of multi-source data will be based on machine learning. The processing of data from multiple sources will be one of the difficulties in modelling financial risk management based on machine learning algorithms. In the future era of explosive data growth, credit data will exist not only in the form of data presentation but also in the form of text and even including voice and video. Thus, it is attempted to incorporate these multimodal data into the default risk prediction modelling to increase the diversity of the data structure of the prediction model. Finally, the forecasting models constructed in this paper are all based on machine learning algorithms, which are themselves a form of black-box forecasting. Financial institutions need not only prediction models that perform well, but also prediction models that can be interpreted. When prediction models can be interpreted, the decision process of the model can be informed, and thus they can be improved to obtain a more accurate model. In particular, when financial institutions use machine learning to model risk problems, they need to be able to understand and trust the models, and only then can they be widely used. Therefore, combining machine learning algorithms with empirical analysis is recommended for future research in order to improve the interpretability of predictive models.

Chapter 6 General Conclusion

6.1 Summary

This chapter summarises the main findings on the subject of the study, distils the connotations and implications of the research findings, and reviews the contributions of the relevant literature. It also shows the limitations of current research and how to improve the prospect for future research. Based on the rapid and widespread emergence of FinTech, this thesis provides an investigation and review of the state of FinTech development and its implications for the development of different financial service cooperation concepts, business models, and financial needs, in the context of the new functions that FinTech has injected into traditional financial institutions and the application of new technologies to financial services. The thesis begins with a thorough review of previous academic research, on the basis of which it defines the connotation of FinTech and argues that FinTech is a holistic concept that can be understood in three dimensions: industry, technology, and integration. As a result, through empirical analysis, this study firstly clarifies the positive impact of FinTech on the efficiency of traditional commercial banks. Secondly, this thesis boasts a structured collation and analysis of systems for credit risk management through advanced machine learning combined with specific financial services. The specific processes of combining each service with FinTech are analysed and summarised in terms of two specific business models, i.e. SCF and P2P lending. Finally, this study presents a comprehensive analysis of machine learning models and data analysis methods for credit risk assessment-related classification problems and validates the effectiveness of the hybrid models. The followings summarise the contributions as well as the conclusions of each of the three independent empirical studies.

Chapter 3 employs a two-stage double bootstrapped truncated regression model proposed by Simar and Wilson (2007) to explore the impact of FinTech development on the efficiency of Chinese commercial banks. This study provides a systematic review of FinTech development in the Chinese banking industry and identifies the positive effect of emerging FinTech development on the efficiency of Chinese commercial banks. Using financial data of Chinese commercial banks for the period 2011-2020, the first stage employs the DEA-Malmquist approach to efficiency estimation and analyses the dynamics changes of Chinese commercial banks' efficiency over the decade. In the second stage, the Digital FinTech Inclusion is applied to the truncated regression model as a representative index of FinTech development. The positive impact of FinTech development on the efficiency of commercial banks is expanded from the perspective of bank ownership and regional financial development. The impact of FinTech development varies for different ownership structures of Chinese commercial banks. Smaller commercial banks, such as city commercial banks, which rely on the local economy and business development, are more positively impacted by FinTech development than state-owned banks and joint-stock banks. Regional financial development was included as a consideration for bank efficiency in response to the nature of the business structure of city commercial banks linked to the local economy, and the positive impact of FinTech development on the efficiency of urban commercial banks was greater in regions with higher FinTech development than in those with less developed financial development. From the perspective of political effect, this study identifies the extent of the influence of FinTech on the efficiency of Chinese commercial banks and provides additional policy guidance for the development of local financial institutions. In addition, this study of local financial development as a macro-factor for the development of FinTech in Chinese city commercial banks' commercial banks deepens the significance of financial inclusion development and the growth of local financial institutions. From the perspective of research improvement, the application of a two-stage bootstrap-Malmquist truncated regression method provides further relevant research with a more robust and complete reference.

The objective of Chapter 4 is to conduct a credit risk assessment for DSCF through a hybrid XGBoost-MLP model. This study explores in detail the content and risk management elements of DSCF at a model level. Based on the analysis of the business processes of DSCF and the review of previous studies, the incorporation of digital features into the credit risk assessment model increases the effectiveness of credit assessment in this area. The essay validates the importance of digital features in the risk assessment of DSCF by comparing the credit risk assessment results of various base classifier models. The findings also verify the advantages of the MLP model and the superiority of the hybrid model with feature selection. This study enriches credit risk assessment at both theoretical and practical levels and clarifies the issue of feature selection for the DSCF field. Meanwhile, the extended credit risk assessment system promotes data-driven decision-making within DSCF institutions. Moreover, by considering a broader range of features, such as ERP system construction and usage and SCF platform status, the system offers a more accurate and in-depth risk assessment. This heightened risk evaluation helps banks and financial institutions identify potential risks and make well-calibrated financing decisions.

Chapter 5 attempts to investigate the imbalanced data issue based on P2P lending credit risk assessment. The review of studies on the imbalanced data issue in P2P lending reveals that there is confusion and subjective assumptions in the existing literature between the imbalanced data problem and incomplete sample characteristics. The inefficiency of the model on this basis is not because of an imbalance in the data but may be due to a lack of features. In the absence of a systematic analysis of the relationship between imbalanced data and feature selection in P2P lending credit risk assessment, this study compares the feature selection problem of Lending Club, a P2P lending platform, with the existing literature on the subject. By constructing two comparative feature sets, the feature set selected based on the existing literature is used as the baseline feature set and the dataset obtained through data cleaning is used as the complete dataset. Meanwhile, this study collates the model of imbalance learning as a benchmark for the research method on the imbalance problem. The trade-off relationship between the degree of data imbalance and the degree of feature selection in the analysis of credit risk assessment is investigated. when the data features are certain, the higher the degree of data imbalance will make the assessment model less effective, and in the case of relatively complete data features, the data imbalance problem has less impact on the credit risk assessment model. However, for a given level of data imbalance, the more missing features severely worsen the effectiveness of the model. At the same time, with a low level of data imbalance, missing features have a significant impact on the credit risk assessment model. Based on this finding, the minimum effective feature set is generated for Lending Club's dataset. Further, by using and comparing the imbalance learning methods, the Easy-Ensemble method is the most effective and it can successfully mitigate the imbalance problem and feature missing problem of the data. This finding provides fundamental theoretical support for the subsequent research on P2P lending risk assessment and the exploration of the imbalanced data problem. Moreover, the identified machine learning techniques help in balancing the data distribution, ensuring that the credit risk model can accurately capture the risks associated with both low and high-risk borrowers. Finally, by implementing an interactive trade-off that incorporates imbalanced data issues and feature selection, this study improves can fine-tune the credit risk models regularly in P2P lending.

6.2 Contributions and Importance of the Thesis

At present, the rapid development of FinTech has a significant impact on the business and operation of traditional financial institutions, which are facing rapid innovation in the economic and financial environment, market competition and customer demand. The transformation of their operations urgently requires the empowerment of FinTech to optimise their business models, improve innovation efficiency and management effectiveness, and ultimately enhance the efficiency and functionality of financial services. In order to accelerate the transformation of their business development philosophy, traditional financial institutions should strategically recognise the importance of FinTech and integrate FinTech into their overall business strategies, as well as their top-level design and planning, and actively promote the transformation of their business operations. Chapter 3 provides the theoretical support and impetus for the transformation and development.

Simultaneously, this thesis emphasises the concrete hands-on analysis of FinTech applications for various detailed operations. Emerging information technologies provide effective means to implement FinTech strategies. Firstly, using cloud computing, big data, and other technologies to deepen the transformation of the data system architecture to provide strong support for rapid response to market and customer expectations. Secondly, it establishes an enterprise-level business architecture to realise the coordination and integration of business demands and enhance the efficiency of business innovation. Thirdly, the goals of focusing on the application of cutting-edge technologies in the financial sector, building application models based on AI and other new technologies, enhancing perception, analysis, decision-making, and prediction capabilities, improving operational efficiency and output capacity, and boosting financial business development are theoretically supported.

6.3 Limitations and Future Research Dimensions

Due to the availability of data and the design of the model, there are still some limitations and shortcomings in the research process of this paper, which can be further developed and improved in the future.

From the perspective of the research design, the innovation of this thesis includes the prediction of credit risk assessment through machine learning methods and the analysis of data characteristics. However, due to the poor interpretability of the machine learning approach, the empirical analysis and derivation of the model results are lacking. As Kruse et al. (2022) suggested, the application of AI technology and machine learning algorithm models in the financial industry is becoming more and more common and in-depth, promoting the intelligent development of the industry, while the algorithm black box problem, security issues, bias issues, etc. are also exposed and become potential risks affecting the healthy development of the financial industry. Especially for machine learning applications such as credit rating, the development and systematic analysis of explainability models and algorithms are very important. Moreover, the limitations of machine learning models in this thesis are intertwined with the challenges of running them efficiently in terms of both size and time. As models become more complex and sophisticated, their size tends to increase, posing challenges for storage, memory, and bandwidth requirements. Larger models demand more computational resources, which can strain hardware capabilities and hinder deployment on devices with limited processing power. Additionally, the time required to train and run these models can be a significant bottleneck, especially when dealing with massive datasets or real-time applications. Longer training times not only impede rapid experimentation and development but also increase the overall cost of computation. Balancing the trade-off between model complexity and practical deployment considerations remains a crucial challenge in the field of machine learning, requiring continuous efforts to optimize size and runtime efficiency without compromising performance.

From the perspective of empirical analysis, the data availability is the main limitation. The study of SCF in Chapter 4 is limited by the lack of profound information on the amount and mode of financing of SCF. Due to the non-mandatory nature of SCF data disclosure by listed companies, this article only compiles superficial data on whether enterprises have developed DSCF, the number of years of development of relevant systems and supply chain financial capability through manual collection. However, the detailed information on the amount of SCF generated by enterprises, the mode of SCF and specific supply chain financial products cannot be obtained. In addition, the use of different SCF models, such as accounts receivable financing model, prepayment financing model and inventory financing model, and the exploration of the heterogeneity of risk assessment of DSCF based on various models are more conducive to revealing its operation process and internal logic. Thus, the limitation of data availability stands as a formidable challenge in the realm of machine learning. Machine learning algorithms heavily rely on large and diverse datasets to generalize patterns effectively. Nevertheless, obtaining and curating such datasets can be a daunting task, especially in niche domains or emerging fields where relevant data may be scarce. Data availability issues in this thesis is worth noting and could be improved in the future work.

This thesis emphasises the specific impact and application of FinTech development on financial institutions and services. It is interesting to further examine and sort the potential insights in the following aspects:

Firstly, from the perspective of the empirical research objective, as shown in Chapter 3, the impact of FinTech development on the efficiency of commercial banks is investigated. For commercial banks, a traditional financial institution, efficiency, risk and profitability are significant factors in operations (Bitar et al., 2018). Thus, except for the efficiency of commercial banks, it is also worthy of further exploration to address the issues related to the development of FinTech on commercial banks' risk-taking and profitability. Combined with the existing literature (Carletti et al., 2020; Banna and Alam, 2021), the contradictions of traditional financial institutions in the post-epidemic era are prominent. The development and cooperation of FinTech are even more foremost. The comprehensive analysis of the combination of multiple factors is worthy of further analysis and research.

Secondly, from the perspective of the research method, the operational research of combining FinTech and traditional businesses from the perspective of risk management is conducted using machine learning methods for credit risk assessment on DSCF and P2P lending. There is more extensive research on the application of machine learning methods to credit risk assessment issues, such as multiple classification approaches in risk rating issues and classification issues for time series data. Due to the immerging technology and financial businesses, the method of risk assessment is also constantly updated (Zheng et al., 2019). Based on different financial business, the usage of advanced machine learning methods is one of the innovation points of future research.

Thirdly, from the perspective of the theoretical research, focusing on the service of lending, the trend of the combination of FinTech and financial services is upgrading. With the continuous elaboration of technology and business models, P2P lending gradually develops from the reception of lending in traditional commercial banks. Moreover, the development of decentralised finance today has made DeFi lending a reality. What changes have taken place in the connotation of the lending business during the development process, and what problems have emerged in the lending process and risk assessment are all worthy of future research. It is an insightful expansion of existing research to sort out its process and measure the risk assessment of specific services.

Finally, from the perspective of the thorough FinTech development environment, future research should explore the ethical implications of FinTech innovations, such as algorithmic bias, fair lending practices, and responsible AI usage. This research would lead to guidelines and principles that prioritize ethical decision-making in FinTech development and deployment. Meanwhile, researching strategies to foster collaboration between FinTech startups and traditional financial institutions would be instrumental in harnessing the strengths of both sectors. Identifying barriers to collaboration and proposing mechanisms for knowledge exchange could accelerate the adoption of innovative FinTech solutions.

In summary, future research in FinTech should encompass a broad spectrum of specific and generic themes, from specialized advancements in credit risk assessment and blockchain applications to more overarching considerations like ethics, collaboration, and the impact of emerging technologies. By addressing these diverse areas of research, FinTech can continue to drive innovation and shape the future of the financial industry in a responsible and sustainable manner.

Appendix

Appendix A (Chapter 3)

A. 1 Robustness Tests with SFA Cost Efficiency

For the measurement of bank efficiency, the frontier analysis method has been used frequently and is divided into two types of methods, parametric and non-parametric, based on the principle of calculation. Aigner & Schmidt (1977) proposed a stochastic frontier approach to calculate the boundaries of the production function and analysed the theory and estimation methods of the SFA method in detail, laying the foundation for subsequent empirical studies. Subsequently, many scholars have applied various parametric methods to the efficiency evaluation of banking institutions, such as Berger et al. (1993), Meste (1996), Altunbaş et al (2001). SFA assumes that the presence of the stochastic error term and inefficiency leads to the deviation of the bank to be examined from the stochastic frontier bank. Greene (1990) discusses the validity of using truncated versus half-normal distributions to define the type of distribution for inefficiency and concludes that different distribution types can sometimes affect the average efficiency of financial institutions. The effect of the type of distribution on the calculation of efficiency has been explored by many scholars, and the truncated distribution has been found to have statistically significant results compared to the normal distribution (Berger and DeYoung, 1997; Greene, 1993; Mester, 1996; Yuengert, 1993). However, the validity of SFA is reduced by the arbitrary nature of the assumptions of the distribution of the inefficiency term on the stochastic frontier and the difficulty of testing the assumptions (Bauer et al., 1998; Coelli, 1996). Furthermore, once the actual distribution of the inefficiency term deviates from the set distribution form, it is not possible to distinguish between the inefficiency term and the stochastic error term in bank efficiency using SFA.

Many debates exist on the use of parametric or non-parametric approaches to bank efficiency measurement, and synthesising the previously mentioned literature, there is no optimal choice in the comparison of DEA and SFA. In order to improve the study of the impact of FinTech on commercial banks and to make the results more robust, this paper will measure efficiency through the parametric method and analyse whether its efficiency is affected in the same manner as in the main text.

The SFA model consists of a stochastic frontier cost analysis and a stochastic frontier output function analysis. Frontier cost refers to the minimum cost that can be achieved at a given level of output; frontier output refers to the maximum output that can be achieved at a given level of input. In this study, a translog (transcendental logarithmic) stochastic cost frontier which is the most widely used functional form in the bank efficiency literature is used to verify the robustness of the experimental results. The input and output sectors consist with the main paper which involve four inputs - Total deposit (w1), Interest rate cost (w2), Non-interest rate cost (w3) and Labor expense (w4) and three outputs - Total loan (y1), Interest rate income (y2) and Non-interest rate income (y3). We also use time tendency (t) to denote technical change (Nguyen et al., 2016). The translog stochastic cost frontier model is as follow:

$$\ln TC_{it} = \alpha_0 + \sum_{m=1}^{3} \alpha_m \ln y_{itm} + \frac{1}{2}\sum_{m=1}^{3}\sum_{k=1}^{3} \alpha_{mk} \ln y_{1itm} \ln y_{itk}$$

$$+ \sum_{n=1}^{4} \beta_n \ln w_{itn} + \frac{1}{2}\sum_{n=1}^{4}\sum_{l=1}^{4} \beta_{nl} \ln w_{itn} \ln w_{itl}$$

$$+ \frac{1}{2}\sum_{m=1}^{3}\sum_{n=1}^{4} \beta_{mn} \ln y_{itm} \ln w_{itn} + \gamma_1 t + \frac{1}{2}\gamma_2 t^2$$

$$+ \sum_{m=1}^{3} \gamma_3 t \ln y_{itm} + \sum_{n=1}^{4} \gamma_4 t \ln w_{itn} + v_{it} + \mu_{it} \qquad (A.1)$$

where the subscript i denotes the cross-sectional dimension across banks, subscript t denotes the time. TC is the total observed cost, consisting of interest expenses, other operating expenses and personnel expenses. The combined error term includes random noise v_{it}, which is assumed to follow a normal distribution, and cost inefficiency μ_{it}, which is assumed to follow a truncated distribution. Further, using the linear homogeneity condition, equation (1) can be transformed into a cost function by normalising the dependent variable and all input prices by the input 4 (w4), as follows.

$$\ln(TC_{it}/w_{it4}) = \alpha_0 + \sum_{m=1}^{3} \alpha_m \ln y_{itm} + \frac{1}{2}\sum_{m=1}^{3}\sum_{k=1}^{3} \alpha_{mk} \ln y_{1itm} \ln y_{itk}$$

$$+\sum_{n=1}^{4}\beta_n \ln(w_{itn}/w_{it4}) + \frac{1}{2}\sum_{n=1}^{4}\sum_{l=1}^{4}\beta_{nl}\ln(w_{itn}/w_{it4})\ln(w_{itl}/w_{it4})$$

$$+\frac{1}{2}\sum_{m=1}^{3}\sum_{n=1}^{4}\beta_{mn}\ln y_{itm}\ln(w_{itn}/w_{it4}) + \gamma_1 t + \frac{1}{2}\gamma_2 t^2$$

$$+\sum_{m=1}^{3}\gamma_3 t\ln y_{itm} + \sum_{n=1}^{4}\gamma_4 t\ln(w_{itn}/w_{it4}) + v_{it} + \mu_{it} \qquad (2)$$

Thus, the individual cost efficiency score is calculated as $CE_{it} = \exp(-\mu_{it})$ which range from 0 to1.

The empirical analysis is carried out using SFA and the specific quantitative model is that of Battese and Coelli (1995). The parameter estimation part was done using Frontier 4.1 software, which is mainly used to implement the estimation and statistical tests for the translog stochastic parameters. The results of the model are presented in Table A.1. Since $\gamma = {\sigma_u}^2/({\sigma_u}^2 + {\sigma_v}^2)$, if $\gamma = 0$, the technically inefficient term does not exist and all errors are due to random disturbances; if $\gamma = 1$, the random disturbances do not exist and all errors are due to technical inefficiency. As the Table A.6.1 shows, $\gamma = 0.532$, the technical inefficiency exist and the SFA method is suitable for efficiency estimation of commercial banks.

Table A.6.1 SFA results for cost efficiency

Variables	**Coef.**
Constant	9.726*** (3.19)
$\ln y_{1it}$	0.179 (1.52)
$\ln y_{2it}$	0.486 (0.86)
$\ln y_{3it}$	-0.588 (-0.99)
$0.5 \cdot \ln^2 y_{1it}$	-0.003 (-0.67)
$0.5 \cdot \ln^2 y_{2it}$	-0.015 (-0.80)
$0.5 \cdot \ln^2 y_{3it}$	0.054*** (2.77)
$\ln(w_{1it}/w_{4it})$	1.198*** (2.54)
$\ln(w_{2it}/w_{4it})$	-0.558*** (-2.80)
$\ln(w_{3it}/w_{4it})$	-0.399 (-1.00)
$0.5 \cdot \ln^2(w_{1it}/w_{4it})$	-0.320***

	(-7.09)
$0.5 \cdot \ln^2(w_{2it}/w_{4it})$	-0.014*** (-2.97)
$0.5 \cdot \ln^2(w_{3it}/w_{4it})$	-0.229*** (-8.30)
$0.5 \cdot \ln(w_{1it}/w_{4it}) \cdot \ln(w_{2it}/w_{4it})$	0.086*** (3.92)
$0.5 \cdot \ln(w_{3it}/w_{4it}) \cdot \ln(w_{2it}/w_{4it})$	-0.322*** (-1.99)
$0.5 \cdot \ln(w_{1it}/w_{4it}) \cdot \ln(w_{3it}/w_{4it})$	0.491*** (7.54)
$0.5 \cdot \ln y_{1it} \cdot \ln(w_{1it}/w_{4it})$	-0.015 (-0.50)
$0.5 \cdot \ln y_{1it} \cdot \ln(w_{2it}/w_{4it})$	-0.018 (-1.34)
$0.5 \cdot \ln y_{1it} \cdot \ln(w_{3it}/w_{4it})$	0.022 (0.89)
$0.5 \cdot \ln y_{2it} \cdot \ln(w_{1it}/w_{4it})$	0.279*** (4.60)
$0.5 \cdot \ln y_{2it} \cdot \ln(w_{2it}/w_{4it})$	-0.039 (-1.95)
$0.5 \cdot \ln y_{2it} \cdot \ln(w_{3it}/w_{4it})$	-0.401*** (-10.82)
$0.5 \cdot \ln y_{3it} \cdot \ln(w_{1it}/w_{4it})$	-0.243*** (-2.90)
$0.5 \cdot \ln y_{3it} \cdot \ln(w_{2it}/w_{4it})$	0.083*** (2.44)
$0.5 \cdot \ln y_{3it} \cdot \ln(w_{3it}/w_{4it})$	0.349*** (6.84)
t	0.043 (0.78)
$0.5 \cdot t^2$	-0.002*** (-1.97)
$t \cdot \ln y_{1it}$	-0.011*** (-6.28)
$t \cdot \ln y_{2it}$	0.013** (1.93)
$t \cdot \ln y_{3it}$	-0.004 (-0.55)
$t \cdot \ln(w_{1it}/w_{4it})$	-0.014*** (-2.23)
$t \cdot \ln(w_{2it}/w_{4it})$	0.002 (1.07)
$t \cdot \ln(w_{3it}/w_{4it})$	0.013*** (2.44)
$\sigma^2 = {\sigma_u}^2 + {\sigma_v}^2$	0.011
$\gamma = {\sigma_u}^2/({\sigma_u}^2 + {\sigma_v}^2)$	0.532
${\sigma_u}^2$	0.006
${\sigma_v}^2$	0.005

Note: The table presents the results based on SFA cost frontier.
* Significant at 10%, ** significant at 5% and *** significant at 1%.
See notes on Table 1 for the definition of the variables.

Thus, we calculated the cost efficiency of commercial banks and the specific results are shown in the Table A.2. Accordingly, the cost efficiency of all commercial banks shows an upward trend, but the cost efficiency of SOCBs shows a more moderate trend of change, while the cost efficiency of JSCBs shows a stronger change, and its change tends to be more in line with the development of the overall national economy.

Table A.6.2 Cost efficiency score based on the SFA method

	SOCB	JSCB	CCB	All Banks
2011	0.802	0.804	0.784	0.797
2012	0.783	0.789	0.785	0.786
2013	0.790	0.817	0.799	0.802
2014	0.798	0.824	0.806	0.809
2015	0.786	0.808	0.800	0.798
2016	0.790	0.812	0.819	0.807
2017	0.802	0.842	0.825	0.823
2018	0.806	0.848	0.831	0.828
2019	0.808	0.812	0.813	0.811
2020	0.799	0.823	0.824	0.815
2011-2020	0.796	0.818	0.809	0.808

We further investigate the influence of FinTech development. Based on Hypothesis 1, we estimated the impact of FinTech indexes on the cost efficiency of commercial banks through truncated regression. The results are collated in Table A.3. The coefficients of FI in columns (a), (b) and (c) are significantly positive, which is consistent with the results of the double bootstrapped DEA-Malmquist truncated regression in the main text, indicating that the overall development of FinTech have an positive effect on the cost efficiency of commercial banks. Meanwhile, the expanding coverage of FinTech breadth and the improvement of usage depth of FinTech would increase the cost efficiency of Chinese commercial banks.

Table A.6.3 Truncated regression results based on the SFA efficiency score

	(a)	(b)	(c)
LTE	-0.197*** (-9.05)	-0.001*** (-7.09)	-0.004*** (1.46)
FI	0.069*** (12.96)	0.057*** (13.06)	0.033*** (11.10)
SIZE	-0.009 (-1.18)	-0.010 (-1.33)	-0.002 (-0.28)
GDP	0.573*** (4.53)	0.951*** (6.58)	0.177*** (0.28)
CPI	0.105*** (6.67)	0.293** (2.51)	0.570** (3.97)
IP	-0.888*** (-4.42)	-0.654*** (-3.59)	-3.437*** (-2.71)
LDR	-0.002 (-0.18)	-0.004 (-0.45)	-0.005 (-0.52)
ROA	0.001 (0.02)	-0.026 (-0.49)	-0.001 (-0.02)
CAR	0.000 (1.14)	0.000 (0.86)	-0.000 (-0.75)
IPO	0.005** (2.05)	0.005** (2.09)	0.007* (2.23)
Observation	1010	1010	1010

Constant	0.987*** (5.60)	0.995*** (7.68)	0.994*** (6.51)
Banks×Year FE	Yes	Yes	Yes
$Wald\ \chi2$	168.57***	171.41***	123.99***

Notes: The table presents the truncated regressed result for Hypothesis 1. The dependent variables are SFA Cost efficiency, the independent variables of (a), (b) and (c) are aggregate FinTech index (FI), coverage breadth (FCB) and the usage depth (FUD) of FinTech at t-1 separately. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Further, we analysed the different impacts of FinTech on SOCBs, JSCBs and CCBs according to Hypothesis 2. The results are presented in the Table A.4, where we find that the cross-sectional coefficient of FinTech with CCBs is the highest among all commercial banks in the sample, while the cross-sectional coefficient for SOCBs is the lowest. This result is consistent with the results in the main text.

Table A.6.4 Truncated regression result of different types of commercial banks based on the SFA efficiency score

	(a)	(b)	(c)
LTE	0.000*** (7.36)	0.002*** (6.11)	0.000*** (9.06)
FI	0.095*** (3.53)	0.080*** (4.65)	0.047*** (6.40)
JSCB	-0.013** (-2.71)	-0.013** (-2.73)	-0.010* (-1.97)
CCB	-0.014*** (-4.09)	-0.015*** (-3.79)	-0.013** (-2.84)
FI*JSCB	0.059** (2.69)	0.056** (2.71)	0.044 (1.97)
FI*CCB	0.071*** (4.01)	0.073*** (3.72)	0.055** (2.77)
SIZE	-0.000 (-0.65)	-0.005 (-0.64)	-0.000 (-1.23)
GDP	0.447*** (3.41)	0.669*** (4.58)	0.097 (0.78)
CPI	0.120*** (5.05)	0.070*** (5.37)	0.100*** (7.19)
IP	-3.16*** (-6.38)	-2.892*** (-8.88)	-2.382*** (-16.74)
LDR	-0.016 (-1.36)	-0.013 (-1.09)	-0.013 (-1.11)
ROA	0.047 (0.07)	0.008 (0.13)	0.024 (0.04)
CAR	0.001 (0.35)	0.002 (0.53)	0.002 (0.41)
IPO	0.047 (1.29)	0.047 (1.30)	0.048 (1.32)
Observation	1010	1010	1010
Constant	1.013*** (3.81)	1.008*** (7.04)	1.011*** (6.90)
Banks×Year FE	Yes	Yes	Yes
$Wald\ \chi2$	19.509***	31.579***	40.994***

Notes: The table presents the truncated regressed result for Hypothesis 2. The dependent variables are SFA Cost efficiency, the independent variables of (a), (b) and (c) are aggregate FinTech index (FI), coverage breadth (FCB) and the usage depth (FUD) of FinTech at t-1 separately. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

Finally, we present the impact of FinTech on commercial banks in regions with different levels of financial development based on Hypothesis 3. According to Table A.5, we find that the cost efficiency of commercial banks in regions with higher levels of financial development are affected by FinTech more than those in regions with lower levels of FinTech development (see column (a) and (b)), the result that is consistent with the results in main text, both contrary to the original hypothesis.

Table A.6.5 Truncated regression result of the impact of financial development based on the SFA efficiency score

	(a)	(b)	(c)
LTE	-0.002*** (-9.08)	-0.001*** (-7.12)	0.000*** (5.13)
FI	0.070*** (12.92)	0.057*** (13.00)	0.016 (0.39)
FD	0.007** (2.44)	0.008 (0.50)	0.007** (2.94)
FI*FD	0.027* (1.69)	0.026* (1.77)	0.032 (0.82)
SIZE	-0.008 (-1.13)	-0.009 (-1.25)	-0.005 (-0.63)
GDP	0.578*** (4.67)	0.957*** (6.62)	0.314* (2.55)
CPI	0.105*** (6.68)	0.029* (2.51)	0.061*** (3.71)
IP	-0.890*** (-4.43)	-0.649*** (-3.57)	-1.794*** (-10.71)
LDR	-0.001 (-0.10)	-0.003 (-0.37)	-0.017 (-1.62)
ROA	0.003 (0.25)	0.004 (0.32)	0.004 (0.30)
CAR	0.005 (1.15)	0.003 (0.83)	-0.000 (-0.70)
IPO	0.006* (2.26)	0.006* (2.23)	0.006* (2.12)
Observation	860	860	860
Constant	0.987*** (5.62)	0.995*** (7.70)	1.007*** (5.58)
Banks×Year FE	Yes	Yes	Yes
$\boldsymbol{Wald\ \chi 2}$	169.17***	171.79***	9.50***

Notes: The table presents the truncated regressed result for Hypothesis 3. The dependent variables are SFA Cost efficiency, the independent variables of (a), (b) and (c) are aggregate FinTech index (FI), coverage breadth (FCB) and the usage depth (FUD) of FinTech at t-1 separately. * Significant at 10%, ** significant at 5% and *** significant at 1%. See notes on Table 1 for the definition of the variables.

In summary, the performance of SFA on the efficiency estimation are consistent with the results of the DEA method in the main text, which demonstrates the robustness of the results in this paper.

Bibliography

Abd Elrahman, S.M. and Abraham, A., 2013. A review of class imbalance problem. *Journal of Network and Innovative Computing, 1*(2013), pp.332-340.

Aigner, D., Lovell, C.K. and Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. *Journal of econometrics, 6*(1), pp.21-37.

Allen, F., McAndrews, J. and Strahan, P., 2002. E-finance: an introduction. *Journal of financial services research*, *22*(1), pp.5-27.

Alqahtani, F., Mayes, D.G. and Brown, K., 2017. Islamic bank efficiency compared to conventional banks during the global crisis in the GCC region. *Journal of International Financial Markets, Institutions and Money*, *51*, pp.58-74.

Alt, R. and Puschmann, T., 2012. The rise of customer-oriented banking-electronic markets are paving the way for change in the financial industry. *Electronic Markets*, *22*, pp.203-215.

Alt, R., Beck, R. and Smits, M.T., 2018. FinTech and the transformation of the financial industry. *Electronic markets*, *28*, pp.235-243.

Altunbaş, Y., Gardener, E.P., Molyneux, P. and Moore, B., 2001. Efficiency in European banking. *European Economic Review*, *45*(10), pp.1931-1955.

Anshari, M., Almunawar, M.N. and Masri, M., 2020. Financial technology and disruptive innovation in business: Concept and application. *International Journal of Asian Business and Information Management (IJABIM), 11*(4), pp.29-43.

Ariff, M. and Luc, C., 2008. Cost and profit efficiency of Chinese banks: A non-parametric analysis. *China economic review*, *19*(2), pp.260-273.

Arner, D.W., Barberis, J. and Buckley, R.P., 2015. The evolution of FinTech: A new post-crisis paradigm. *Geo. J. Int'l L.*, *47*, p.1271.

Arora, N. and Kaur, P.D., 2020. A Bolasso based consistent feature selection enabled random forest classification algorithm: An application to credit risk assessment. *Applied Soft Computing*, *86*, p.105936.

Arslanian, H., Fischer, F., Arslanian, H. and Fischer, F., 2019. FinTech and the Future of the Financial Ecosystem. *The Future of Finance: The Impact of FinTech, AI, and Crypto on Financial Services*, pp.201-216.

Ataullah, A. and Le, H., 2006. Economic reforms and bank efficiency in developing countries: the case of the Indian banking industry. *Applied Financial Economics*, *16*(9), pp.653-663.

Athey, S., 2018. The impact of machine learning on economics. In *The economics of artificial intelligence: An agenda* (pp. 507-547). University of Chicago Press.

Atkinson, W., 2008. Supply chain finance: the next big opportunity. *Supply Chain Management Review*, *12*(3), pp.57-60.

Austin, R.E. and Dunham, L.M., 2022. Do FinTech Acquisitions Improve the Operating Performance or Risk Profiles of Acquiring Firms?. *Journal of Economics and Business*, p.106078.

Babaei, G. and Bamdad, S., 2021. A new hybrid instance-based learning model for decision-making in the P2P lending market. *Computational Economics*, *57*, pp.419-432.

Bachmann, A., Becker, A., Buerckner, D., Hilker, M., Kock, F., Lehmann, M., Tiburtius, P. and Funk, B., 2011. Online peer-to-peer lending-a literature review. *Journal of Internet Banking and Commerce*, *16*(2), p.1.

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J. and Vanthienen, J., 2003. Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the operational research society*, *54*, pp.627-635.

Bagchi, B., Chatterjee, S., Ghosh, R., Dandapat, D., Bagchi, B., Chatterjee, S., Ghosh, R. and Dandapat, D., 2020. Impact of COVID-19 on global economy. *Coronavirus Outbreak and the Great Lockdown: Impact on Oil Prices and Major Stock Markets Across the Globe*, pp.15-26.

Bahnsen, A.C. and Gonzalez, A.M., 2011, December. Evolutionary algorithms for selecting the architecture of a MLP neural network: A credit scoring case. In *2011 IEEE 11th International Conference on Data Mining Workshops* (pp. 725-732). IEEE.

Balyuk, T., Berger, A.N. and Hackney, J., 2020. What is fueling FinTech lending? The role of banking market structure. *The Role of Banking Market Structure* (June 23, 2020).

Banerjee, A., Lücker, F. and Ries, J.M., 2021. An empirical analysis of suppliers' trade-off behaviour in adopting digital supply chain financing solutions. *International Journal of Operations & Production Management*, *41*(4), pp.313-335.

Banna, H. and Alam, M.R., 2021. Impact of digital financial inclusion on ASEAN banking stability: implications for the post-Covid-19 era. *Studies in Economics and Finance*, *38*(2), pp.504-523.

Bao, W., Lianju, N. and Yue, K., 2019. Integration of unsupervised and supervised machine learning algorithms for credit risk assessment. *Expert Systems with Applications*, *128*, pp.301-315.

Barandela, R., Valdovinos, R.M. and Sánchez, J.S., 2003. New applications of ensembles of classifiers. *Pattern Analysis & Applications*, *6*, pp.245-256.

Bastani, K., Asgari, E. and Namavari, H., 2019. Wide and deep learning for peer-to-peer lending. *Expert Systems with Applications*, *134*, pp.209-224.

Battese, G.E. and Coelli, T.J., 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical economics*, *20*, pp.325-332.

Bauer, P.W., Berger, A.N., Ferrier, G.D. and Humphrey, D.B., 1998. Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods. *Journal of Economics and business*, *50*(2), pp.85-114.

Bauer, P.W., Berger, A.N., Ferrier, G.D. and Humphrey, D.B., 1998. Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods. *Journal of Economics and business*, *50*(2), pp.85-114.

Bavoso, V., 2020. The promise and perils of alternative market-based finance: the case of P2P lending in the UK. *Journal of Banking Regulation, 21*, pp.395-409.

Begenau, J., Farboodi, M. and Veldkamp, L., 2018. Big data in finance and the growth of large firms. *Journal of Monetary Economics*, *97*, pp.71-87.

Behr, P., Foos, D. and Norden, L., 2017. Cyclicality of SME lending and government involvement in banks. *Journal of Banking & Finance*, *77*, pp.64-77.

Bekkar, M., Djemaa, H.K. and Alitouche, T.A., 2013. Evaluation measures for models assessment over imbalanced data sets. *J Inf Eng Appl, 3*(10).

Bell, J., 2022. What is machine learning?. *Machine Learning and the City: Applications in Architecture and Urban Design*, pp.207-216.

Berg, T., Fuster, A. and Puri, M., 2022. FinTech lending. *Annual Review of Financial Economics*, *14*, pp.187-207.

Berger, A.N. and DeYoung, R., 1997. Problem loans and cost efficiency in commercial banks. *Journal of banking & finance*, *21*(6), pp.849-870.

Berger, A.N. and Humphrey, D.B., 1991. The dominance of inefficiencies over scale and product mix economies in banking. *Journal of Monetary Economics*, *28*(1), pp.117-148.

Berger, A.N., 1993. "Distribution-free" estimates of efficiency in the US banking industry and tests of the standard distributional assumptions. *Journal of productivity Analysis*, *4*(3), pp.261-292.

Berger, A.N., 2003. The economic effects of technological progress: Evidence from the banking industry. *Journal of Money, credit and Banking*, pp.141-176.

Berger, A.N., Demirgüç-Kunt, A., Levine, R. and Haubrich, J.G., 2004. Bank concentration and competition: An evolution in the making. *Journal of Money, credit and Banking*, pp.433-451.

Berger, A.N., Hasan, I. and Zhou, M., 2009. Bank ownership and efficiency in China: What will happen in the world's largest nation?. *Journal of Banking & Finance*, *33*(1), pp.113-130.

Berger, S.C. and Gleisner, F., 2009. Emergence of financial intermediaries in electronic markets: The case of online P2P lending. *Business Research*, *2*, pp.39-65.

Bhatore, S., Mohan, L. and Reddy, Y.R., 2020. Machine learning techniques for credit risk evaluation: a systematic literature review. *Journal of Banking and Financial Technology*, *4*, pp.111-138.

Bitar, M., Pukthuanthong, K. and Walker, T., 2018. The effect of capital ratios on the risk, efficiency and profitability of banks: Evidence from OECD countries. *Journal of international financial Markets, Institutions and Money*, *53*, pp.227-262.

Bonin, J.P., Hasan, I. and Wachtel, P., 2005. Bank performance, efficiency and ownership in transition countries. *Journal of banking & finance*, *29*(1), pp.31-53.

Bons, R.W., Alt, R., Lee, H.G. and Weber, B., 2012. Banking in the Internet and mobile era. *Electronic markets*, *22*(4), pp.197-202.

Boughorbel, S., Jarray, F. and El-Anbari, M., 2017. Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. *PloS one, 12*(6), p.e0177678.

Breiman, L., 1984. Friedman. *JH and Olsen, RA and Stone, CJ*.

Breiman, L., 1996. Bagging predictors. *Machine learning*, *24*, pp.123-140.

Breiman, L., 2001. Random forests. *Machine learning*, *45*, pp.5-32.

Brown, J.R., Fazzari, S.M. and Petersen, B.C., 2009. Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *The Journal of Finance*, *64*(1), pp.151-185.

Buchak, G., Matvos, G., Piskorski, T. and Seru, A., 2018. FinTech, regulatory arbitrage, and the rise of shadow banks. *Journal of financial economics*, *130*(3), pp.453-483.

Byanjankar, A., Heikkilä, M. and Mezei, J., 2015, December. Predicting credit risk in peer-to-peer lending: A neural network approach. *In 2015 IEEE symposium series on computational intelligence* (pp. 719-725). IEEE.

Cai, W., Xu, F. and Zeng, C., 2016. Geographical diversification and bank performance: Evidence from China. *Economics Letters*, *147*, pp.96-98.

Camanho, A.S. and Dyson, R.G., 2006. Data envelopment analysis and Malmquist indices for measuring group performance. *Journal of productivity Analysis*, *26*(1), pp.35-49.

Camerinelli, E., 2009. Supply chain finance. *Journal of Payments Strategy & Systems*, *3*(2), pp.114-128.

Caniato, F., Gelsomino, L.M., Perego, A. and Ronchi, S., 2016. Does finance solve the supply chain financing problem?. *Supply chain management: an international journal*.

Caniato, F., Gelsomino, L.M., Perego, A. and Ronchi, S., 2016. Does finance solve the supply chain financing problem?. *Supply chain management: an international journal*.

Carletti, E., Claessens, S., Fatás, A. and Vives, X., 2020. *Post-Covid-19 World*. Centre for Economic Policy Research.

Carter, C. and Catlett, J., 1987. Assessing credit card applications using machine learning. *IEEE Intelligent Systems*, *2*(03), pp.71-79.

Casu, B., Girardone, C. and Molyneux, P., 2004. Productivity change in European banking: A comparison of parametric and non-parametric approaches. *Journal of Banking & Finance*, *28*(10), pp.2521-2540.

Caves, D.W., Christensen, L.R. and Diewert, W.E., 1982. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica: Journal of the Econometric Society*, pp.1393-1414.

Chang, A.H., Yang, L.K., Tsaih, R.H. and Lin, S.K., 2022. Machine learning and artificial neural networks to construct P2P lending credit-scoring model: A case using Lending Club data. *Quantitative Finance and Economics*, *6*(2), pp.303-325.

Chang, T.P., Hu, J.L., Chou, R.Y. and Sun, L., 2012. The sources of bank productivity growth in China during 2002–2009: A disaggregation view. *Journal of Banking & Finance*, *36*(7), pp.1997-2006.

Chawla, B., 2002. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. *Smote: synthetic minority over-sampling technique, J. Artif. Intell. Res*, *16*, pp.321-357.

Chawla, N.V., Japkowicz, N. and Kotcz, A., 2004. Special issue on learning from imbalanced data sets. *ACM SIGKDD explorations newsletter*, *6*(1), pp.1-6.

Chawla, N.V., Lazarevic, A., Hall, L.O. and Bowyer, K.W., 2003. SMOTEBoost: Improving prediction of the minority class in boosting. In *Knowledge Discovery in Databases: PKDD 2003: 7th European Conference on Principles and Practice of Knowledge Discovery in Databases, Cavtat-Dubrovnik, Croatia, September 22-26, 2003. Proceedings 7* (pp. 107-119). Springer Berlin Heidelberg.

Chen, L., Emanuel, D., Li, L.Z. and Yang, M., 2021. Regulatory changes and loan loss provisions management by Chinese banks. *Journal of Accounting in Emerging Economies, 11*(4), pp.651-675.

Chen, M.A., Wu, Q. and Yang, B., 2019. How valuable is FinTech innovation?. *The Review of Financial Studies*, *32*(5), pp.2062-2106.

Chen, N., Ribeiro, B. and Chen, A., 2016. Financial credit risk assessment: a recent review. *Artificial Intelligence Review, 45*, pp.1-23.

Chen, S., Xu, H., Liu, D., Hu, B. and Wang, H., 2014. A vision of IoT: Applications, challenges, and opportunities with china perspective. *IEEE Internet of Things journal, 1*(4), pp.349-359.

Chen, T. and Guestrin, C., 2016, August. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

Chen, T., He, T., Benesty, M. and Khotilovich, V., 2019. Package 'xgboost'. *R version*, *90*, pp.1-66.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I. and Zhou, T., 2015. Xgboost: extreme gradient boosting. *R package version 0.4-2*, *1*(4), pp.1-4.

Chen, X., Liu, C. and Li, S., 2019. The role of supply chain finance in improving the competitive advantage of online retailing enterprises. *Electronic Commerce Research and Applications, 33*, p.100821.

Chen, Y.R., Leu, J.S., Huang, S.A., Wang, J.T. and Takada, J.I., 2021. Predicting default risk on peer-to-peer lending imbalanced datasets. *IEEE Access*, *9*, pp.73103-73109.

Chen, Z., Li, K. and He, L.Y., 2020. Has internet finance decreased the profitability of commercial banks?: evidence from China. *Emerging Markets Finance and Trade, 56*(13), pp.3015-3032.

Chen, Z., Li, Y., Wu, Y. and Luo, J., 2017. The transition from traditional banking to mobile internet finance: an organizational innovation perspective-a comparative study of Citibank and ICBC. *Financial Innovation*, *3*(1), pp.1-16.

Chen, Z., Poncet, S. and Xiong, R., 2020. Local financial development and constraints on domestic private-firm exports: Evidence from city commercial banks in China. *Journal of Comparative Economics*, *48*(1), pp.56-75.

Cheng, M., Qu, Y., Jiang, C. and Zhao, C., 2022. Is cloud computing the digital solution to the future of banking? *Journal of Financial Stability, 63*, p.101073.

Cheng, X. and Degryse, H., 2010. The impact of bank and non-bank financial institutions on local economic growth in China. *Journal of financial services research, 37*, pp.179-199.

Chortareas, G.E., Girardone, C. and Ventouri, A., 2013. Financial freedom and bank efficiency: Evidence from the European Union. *Journal of Banking & Finance*, *37*(4), pp.1223-1231.

Christensen, C.M., 2013. *The innovator's dilemma: when new technologies cause great firms to fail*. Harvard Business Review Press.

Chueca Vergara, C. and Ferruz Agudo, L., 2021. FinTech and Sustainability: Do They Affect Each Other?. *Sustainability*, *13*(13), p.7012.

Chuen, D.L.K. and Teo, E.G., 2015. Emergence of FinTech and the LASIC principles. *The Journal of Financial Perspectives: FinTech, 3*(3), pp.24-37.

Claessens, S. and Laeven, L., 2003. Financial development, property rights, and growth. *the Journal of Finance*, *58*(6), pp.2401-2436.

Clark, J.A. and Siems, T.F., 2002. X-efficiency in banking: Looking beyond the balance sheet. *Journal of Money, Credit and Banking*, pp.987-1013.

Coelli, T.J., 1996. *A guide to FRONTIER version 4.1: a computer program for stochastic frontier production and cost function estimation* (Vol. 7, pp. 1-33). CEPA Working papers.

Cole, R.A., Cumming, D.J. and Taylor, J., 2019. Does FinTech compete with or complement bank finance?. *Available at SSRN 3302975*.

Cortes, C. and Vapnik, V., 1995. Support-Vector Networks. Mach. Leaming 20, 273–297.

Coşer, A., Maer-matei, M.M. and Albu, C., 2019. PREDICTIVE MODELS FOR LOAN DEFAULT RISK ASSESSMENT. *Economic Computation & Economic Cybernetics Studies & Research*, *53*(2).

Cousin, V., 2011. *Banking in China*. Springer.

Cover, T., 1968. Estimation by the nearest neighbor rule. *IEEE Transactions on Information Theory*, *14*(1), pp.50-55.

Craig, V.V., 2005. China's opening to the world: What does it mean for US banks. *FDIC Banking Rev., 17*, p.1.

Cramer, J.S., 2004. Scoring bank loans that may go wrong: a case study. *Statistica Neerlandica*, *58*(3), pp.365-380.

Cummins, M., Lynn, T., Mac an Bhaird, C. and Rosati, P., 2019. Addressing information asymmetries in online peer-to-peer lending. *Disrupting Finance: FinTech and Strategy in the 21st Century*, pp.15-31.

Curram, S.P. and Mingers, J., 1994. Neural networks, decision tree induction and discriminant analysis: An empirical comparison. *Journal of the Operational Research Society*, *45*(4), pp.440-450.

Cusano, C., Ciocca, G. and Schettini, R., 2003, December. Image annotation using SVM. In *Internet imaging V* (Vol. 5304, pp. 330-338). SPIE.

Danenas, P. and Garsva, G., 2012. Credit risk evaluation modeling using evolutionary linear SVM classifiers and sliding window approach. *Procedia Computer Science*, *9*, pp.1324-1333.

Das, A. and Ghosh, S., 2006. Financial deregulation and efficiency: An empirical analysis of Indian banks during the post reform period. *Review of Financial Economics*, *15*(3), pp.193-221.

DAVIS, R.H., Edelman, D.B. and Gammerman, A.J., 1992. Machine-learning algorithms for credit-card applications. *IMA Journal of Management Mathematics*, *4*(1), pp.43-51.

De Roure, C., Pelizzon, L. and Thakor, A., 2022. P2P lenders versus banks: Cream skimming or bottom fishing?. *The Review of Corporate Finance Studies*, *11*(2), pp.213-262.

Denison, D. G.; Holmes, C. C.; Mallick, B. K.; Smith, A. F. *Bayesian Methods for Nonlinear Classification and Regression*; John Wiley & Sons, 2002; Vol. 386.

Desai, V.S., Conway, D.G., Crook, J.N. and Overstreet Jr, G.A., 1997. Credit-scoring models in the credit-union environment using neural networks and genetic algorithms. *IMA Journal of Management Mathematics*, *8*(4), pp.323-346.

Dong, Y., Firth, M., Hou, W. and Yang, W., 2016. Evaluating the performance of Chinese commercial banks: A comparative analysis of different types of banks. *European Journal of Operational Research*, *252*(1), pp.280-295.

Dong, Y., Girardone, C. and Kuo, J.M., 2017. Governance, efficiency and risk taking in Chinese banking. *The British Accounting Review*, *49*(2), pp.211-229.

Donnelly, G., Iyer, R. and Howell, R.T., 2012. The Big Five personality traits, material values, and financial well-being of self-described money managers. *Journal of economic psychology*, *33*(6), pp.1129-1142.

Du, K., Worthington, A.C. and Zelenyuk, V., 2018. Data envelopment analysis, truncated regression and double-bootstrap for panel data with application to Chinese banking. *European Journal of Operational Research*, *265*(2), pp.748-764.

Du, M., Chen, Q., Xiao, J., Yang, H. and Ma, X., 2020. Supply chain finance innovation using blockchain. *IEEE Transactions on Engineering Management, 67*(4), pp.1045-1058.

Du, M., Chen, Q., Xiao, J., Yang, H. and Ma, X., 2020. Supply chain finance innovation using blockchain. *IEEE Transactions on Engineering Management*, *67*(4), pp.1045-1058.

Duarte, J., Siegel, S. and Young, L., 2012. Trust and credit: The role of appearance in peer-to-peer lending. *The Review of Financial Studies*, *25*(8), pp.2455-2484.

Duncan, E. and Elliott, G., 2004. Efficiency, customer service and financial performance among Australian financial institutions. *International Journal of bank marketing*, *22*(5), pp.319-342.

Elkan, C., 2001, August. The foundations of cost-sensitive learning. In *International joint conference on artificial intelligence* (Vol. 17, No. 1, pp. 973-978). Lawrence Erlbaum Associates Ltd.

Emekter, R., Tu, Y., Jirasakuldech, B. and Lu, M., 2015. Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. *Applied Economics, 47*(1), pp.54-70.

Emekter, R., Tu, Y., Jirasakuldech, B. and Lu, M., 2015. Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. *Applied Economics*, *47*(1), pp.54-70.

Epetimehin, F.M., 2013. Impact of risk appetite on the value of a firm. *European Scientific Journal, 9*(22).

Erel, I. and Liebersohn, J., 2022. Can FinTech reduce disparities in access to finance? Evidence from the Paycheck Protection Program. *Journal of Financial Economics, 146(1)*, pp.90-118.

Escribano, A., Fosfuri, A. and Tribó, J.A., 2009. Managing external knowledge flows: The moderating role of absorptive capacity. *Research policy*, *38*(1), pp.96-105.

Fang, J., Lau, C.K.M., Lu, Z., Tan, Y. and Zhang, H., 2019. Bank performance in China: A Perspective from Bank efficiency, risk-taking and market competition. *Pacific-Basin Finance Journal, 56,* pp.290-309.

Farboodi, M. and Veldkamp, L., 2017. Financial technology, unpredictability and illiquidity in the long run.

Farboodi, M. and Veldkamp, L., 2020. Long-run growth of financial data technology. *American Economic Review*, *110*(8), pp.2485-2523.

Farboodi, M., Mihet, R., Philippon, T. and Veldkamp, L., 2019, May. Big data and firm dynamics. In *AEA papers and proceedings* (Vol. 109, pp. 38-42).

Färe, R., Grosskopf, S., Norris, M. and Zhang, Z., 1994. Productivity growth, technical progress, and efficiency change in industrialized countries. *The American economic review*, pp.66-83.

Farrell, M.J., 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, *120*(3), pp.253-281.

Feng, G. and Zhang, M., 2021. A literature review on digital finance, consumption upgrading and high-quality economic development. *Journal of Risk Analysis and Crisis Response, 11*(4).

Fernandes, F.D.S., Stasinakis, C. and Bardarova, V., 2018. Two-stage DEA-Truncated Regression: Application in banking efficiency and financial development. *Expert Systems with Applications*, *96*, pp.284-301.

Ferri, G., 2009. Are new tigers supplanting old mammoths in China's banking system? Evidence from a sample of city commercial banks. *Journal of Banking & Finance*, *33*(1), pp.131-140.

Ferrier, G.D. and Lovell, C.K., 1990. Measuring cost efficiency in banking: Econometric and linear programming evidence. *Journal of econometrics*, *46*(1-2), pp.229-245.

Fethi, M.D. and Pasiouras, F., 2010. Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey. *European journal of operational research*, *204*(2), pp.189-198.

Finch, P., 2004. Supply chain risk management. *Supply chain management: an International Journal*, *9*(2), pp.183-196.

Fiordelisi, F., Marques-Ibanez, D. and Molyneux, P., 2011. Efficiency and risk in European banking. *Journal of banking & finance, 35*(5), pp.1315-1326.

Firth, M., Lin, C. and Wong, S.M., 2008. Leverage and investment under a state-owned bank lending environment: Evidence from China. *Journal of corporate finance*, *14*(5), pp.642-653.

Fitzpatrick, D.B., 1976. An analysis of bank credit card profit. *Journal of Bank Research*, *7*, pp.199-205.

Florez-Lopez, R. and Ramon-Jeronimo, J.M., 2015. Enhancing accuracy and interpretability of ensemble strategies in credit risk assessment. A correlated-adjusted decision forest proposal. *Expert Systems with Applications, 42*(13), pp.5737-5753.

Forman, G., 2003. An extensive empirical study of feature selection metrics for text classification. *J. Mach. Learn. Res.*, *3*(Mar), pp.1289-1305.

Freedman, S. and Jin, G.Z., 2008. Do social networks solve information problems for peer-to-peer lending? Evidence from Prosper. com.

Freedman, S. and Jin, G.Z., 2017. The information value of online social networks: Lessons from peer-to-peer lending. *International Journal of Industrial Organization*, *51*, pp.185-222.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pp.1189-1232.

Friedman, N., Geiger, D. and Goldszmidt, M., 1997. Bayesian network classifiers. *Machine learning*, *29*, pp.131-163.

Fu, X.M. and Heffernan, S., 2009. The effects of reform on China's bank structure and performance. *Journal of Banking & Finance*, *33*(1), pp.39-52.

Galar, M., Fernández, A., Barrenechea, E. and Herrera, F., 2013. EUSBoost: Enhancing ensembles for highly imbalanced data-sets by evolutionary undersampling. *Pattern recognition*, *46*(12), pp.3460-3471.

Galindo, J. and Tamayo, P., 2000. Credit risk assessment using statistical and machine learning: basic methodology and risk modeling applications. *Computational economics*, *15*, pp.107-143.

Gao, P., Zhang, H., Wu, Z. and Wang, J., 2020. Visualising the expansion and spread of coronavirus disease 2019 by cartograms. *Environment and Planning A: Economy and Space*, *52*(4), pp.698-701.

García-Herrero, A., Gavilá, S. and Santabárbara, D., 2009. What explains the low profitability of Chinese banks?. *Journal of Banking & Finance*, *33*(11), pp.2080-2092.

Gazel, M. and Schwienbacher, A., 2021. Entrepreneurial FinTech clusters. *Small Business Economics, 57*, pp.883-903.

Ghadge, A., Dani, S., Chester, M. and Kalawsky, R., 2013. A systems approach for modelling supply chain risks. *Supply chain management: an international journal*, *18*(5), pp.523-538.

Gianiodis, P.T., Ettlie, J.E. and Urbina, J.J., 2014. Open service innovation in the global banking industry: Inside-out versus outside-in strategies. *Academy of Management Perspectives*, *28*(1), pp.76-91.

Gobbi, G. and Sette, E., 2014. Do firms benefit from concentrating their borrowing? Evidence from the great recession. *Review of Finance*, *18*(2), pp.527-560.

Goldfarb, A. and Tucker, C., 2019. Digital economics. *Journal of Economic Literature*, *57*(1), pp.3-43.

Goldfarb, A. and Tucker, C., 2019. Digital economics. *Journal of Economic Literature*, *57*(1), pp.3-43.

Goldsmith, R.W., 1969. *Financial structure and development* (No. HG174 G57).

Gomber, P., Kauffman, R.J., Parker, C. and Weber, B.W., 2018. On the fintech revolution: Interpreting the forces of innovation, disruption, and transformation in financial services. *Journal of management information systems, 35*(1), pp.220-265.

Gomber, P., Kauffman, R.J., Parker, C. and Weber, B.W., 2018. On the FinTech revolution: Interpreting the forces of innovation, disruption, and transformation in financial services. *Journal of management information systems*, *35*(1), pp.220-265.

Gomm, M.L., 2010. Supply chain finance: applying finance theory to supply chain management to enhance finance in supply chains. *International Journal of Logistics: Research and Applications*, *13*(2), pp.133-142.

González, L.O., Gil, L.I.R., Cunill, O.M. and Lindahl, J.M.M., 2016. The effect of financial innovation on European banks' risk. *Journal of Business Research*, *69*(11), pp.4781-4786.

Goodell, J.W., Kumar, S., Lim, W.M. and Pattnaik, D., 2021. Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance, 32*, p.100577.

Gorjón Rivas, S., 2018. The growth of the FinTech industry in China: a singular case. *Economic bulletin/Banco de España, n. 4, 2018, 13 p*.

Grablowsky, B.J. and Talley, W.K., 1981. Probit and discriminant functions for classifying credit applicants-a comparison. *Journal of Economics and Business*, *33*(3), pp.254-261.

Greene, W.H., 1990. A gamma-distributed stochastic frontier model. *Journal of econometrics*, *46*(1-2), pp.141-163.

Greene, W.H., 1999. Frontier production functions. *Handbook of applied econometrics volume 2: Microeconomics*, pp.75-153.

Guillaumont Jeanneney, S., Hua, P. and Liang, Z., 2006. Financial development, economic efficiency, and productivity growth: Evidence from China. *The Developing Economies, 44*(1), pp.27-52.

Guillen, G., Badell, M. and Puigjaner, L., 2007. A holistic framework for short-term supply chain management integrating production and corporate financial planning. *International Journal of Production Economics*, *106*(1), pp.288-306.

Guo, F., Wang, J., Wang, F., Kong, T., Zhang, X. and Cheng, Z., 2020. Measuring China's digital financial inclusion: Index compilation and spatial characteristics. *China Economic Quarterly*, *19*(4), pp.1401-1418.

Guyon, I. and Elisseeff, A., 2003. An introduction to variable and feature selection. *Journal of machine learning research*, *3*(Mar), pp.1157-1182.

Hallikas, J., Virolainen, V.M. and Tuominen, M., 2002. Risk analysis and assessment in network environments: A dyadic case study. *International journal of production economics*, *78*(1), pp.45-55.

Han, H., Wang, W.Y. and Mao, B.H., 2005. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In *Advances in Intelligent Computing: International Conference on Intelligent Computing, ICIC 2005, Hefei, China, August 23-26, 2005, Proceedings, Part I 1* (pp. 878-887). Springer Berlin Heidelberg.

Hand, D. J. Discrimination and Classification. *Wiley Series in Probability and Mathematical Statistics* **1981**.

Hartman, T.E., Storbeck, J.E. and Byrnes, P., 2001. Allocative efficiency in branch banking. *European Journal of Operational Research*, *134*(2), pp.232-242.

Hasan, I., Wachtel, P. and Zhou, M., 2009. Institutional development, financial deepening and economic growth: Evidence from China. Journal of Banking & Finance, 33(1), pp.157-170.

Hassanniakalager, A., Sermpinis, G. and Stasinakis, C., 2021. Trading the foreign exchange market with technical analysis and Bayesian Statistics. *Journal of Empirical Finance*, *63*, pp.230-251.

Hassanniakalager, A., Sermpinis, G., Stasinakis, C. and Verousis, T., 2020. A conditional fuzzy inference approach in forecasting. *European Journal of Operational Research*, *283*(1), pp.196-216.

Havrylchyk, O. and Verdier, M., 2018. The financial intermediation role of the P2P lending platforms. *Comparative Economic Studies*, *60*, pp.115-130.

Havrylchyk, O., Mariotto, C., Rahim, T. and Verdier, M., 2017. What drives the expansion of the peer-to-peer lending?. *URL http://dx. doi. org/10.2139/ssrn*, *2841316*.

He, L., Xue, M. and Gu, B., 2020. Internet-of-things enabled supply chain planning and coordination with big data services: Certain theoretic implications. *Journal of Management Science and Engineering, 5*(1), pp.1-22.

Heffernan, S., 2005. *Modern banking*. John Wiley & Sons.

Henley, W. and Hand, D.J., 1996. AK-Nearest-Neighbour Classifier for Assessing Consumer Credit Risk. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *45*(1), pp.77-95.

Henley, W. and Hand, D.J., 1996. AK-Nearest-Neighbour Classifier for Assessing Consumer Credit Risk. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *45*(1), pp.77-95.

Henley, W.E., 1995. *Statistical aspects of credit scoring*. Open University (United Kingdom).

Hoehle, H., Scornavacca, E. and Huff, S., 2012. Three decades of research on consumer adoption and utilization of electronic banking channels: A literature analysis. *Decision Support Systems*, *54*(1), pp.122-132.

Hofmann, E., 2005. Supply chain finance: some conceptual insights. *Beiträge Zu Beschaffung Und Logistik*, *16*, pp.203-214.

Hsieh, N.C. and Hung, L.P., 2010. A data driven ensemble classifier for credit scoring analysis. *Expert systems with Applications*, *37*(1), pp.534-545.

Hu, S., Liang, Y., Ma, L. and He, Y., 2009, October. MSMOTE: Improving classification performance when training data is imbalanced. In *2009 second international workshop on computer science and engineering* (Vol. 2, pp. 13-17). IEEE.

Hua, X. and Huang, Y., 2021. Understanding China's FinTech sector: development, impacts and risks. *The European Journal of Finance*, *27*(4-5), pp.321-333.

Hua, X. and Huang, Y., 2021. Understanding China's FinTech sector: development, impacts and risks. *The European Journal of Finance*, *27*(4-5), pp.321-333.

Hunt, E.B., Marin, J. and Stone, P.J., 1966. Experiments in induction.

Iman, N., 2019. Traditional banks against FinTech startups: A field investigation of a regional bank in Indonesia. *Banks and Bank systems*, *14*(3), pp.20-33.

Isik, I. and Hassan, M.K., 2003. Financial deregulation and total factor productivity change: An empirical study of Turkish commercial banks. *Journal of Banking & Finance*, *27*(8), pp.1455-1485.

Ivanov, D. and Dolgui, A., 2021. A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. *Production Planning & Control*, *32*(9), pp.775-788.

Jia, Q., Guo, Y., Li, R., Li, Y. and Chen, Y., 2018. A conceptual artificial intelligence application framework in human resource management.

Jiang, C., Yao, S. and Feng, G., 2013. Bank ownership, privatization, and performance: Evidence from a transition country. *Journal of banking & finance*, *37*(9), pp.3364-3372.

Jiang, C., Yao, S. and Zhang, Z., 2009. The effects of governance changes on bank efficiency in China: A stochastic distance function approach. *China Economic Review*, *20*(4), pp.717-731.

Jiang, Y. and Jones, S., 2018. Corporate distress prediction in China: A machine learning approach. *Accounting & Finance, 58*(4), pp.1063-1109.

Jing, B. and Seidmann, A., 2014. Finance sourcing in a supply chain. *Decision support systems*, *58*, pp.15-20.

Jordan, M.I. and Mitchell, T.M., 2015. Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), pp.255-260.

Kerle, P., 2007. Steady Supply: The Growing Role of Supply Chain Finance in Europe. *Supply Chain. Eur*, *16*, p.18.

Khemakhem, S. and Boujelbene, Y., 2017. Artificial intelligence for credit risk assessment: artificial neural network and support vector machines. *ACRN Oxford Journal of Finance and Risk Perspectives*, *6*(2), pp.1-17.

Kishan, R.P. and Opiela, T.P., 2000. Bank size, bank capital, and the bank lending channel. *Journal of Money, credit and banking*, pp.121-141.

Klafft, M., 2008, March. Peer to peer lending: auctioning microcredits over the internet. In *Proceedings of the International Conference on Information Systems, Technology and Management, A. Agarwal, R. Khurana, eds., IMT, Dubai*.

Komarek, P., 2004. *Logistic regression for data mining and high-dimensional classification*. Carnegie Mellon University.

Kong, S.T. and Loubere, N., 2021. Digitally down to the countryside: FinTech and rural development in China. *The Journal of Development Studies*, *57*(10), pp.1739-1754.

Korpela, K., Hallikas, J. and Dahlberg, T., 2017. Digital supply chain transformation toward blockchain integration.

Krawczyk, B., Schaefer, G. and Woźniak, M., 2015. A hybrid cost-sensitive ensemble for imbalanced breast thermogram classification. *Artificial intelligence in medicine*, *65*(3), pp.219-227.

Kruse, L., Wunderlich, N. and Beck, R., 2019. Artificial intelligence for the financial services industry: What challenges organizations to succeed.

Kubat, M. and Matwin, S., 1997, July. Addressing the curse of imbalanced training sets: one-sided selection. In *Icml* (Vol. 97, No. 1, p. 179).

Kwan, S.H. and Eisenbeis, R.A., 1995. An analysis of inefficiency in banking: A stochastic cost frontier approach. In *Working paper, Working Papers in Applied Economics*. Federal Reserve Bank of San Francisco.

Lai, K.K., Yu, L., Huang, W. and Wang, S., 2006, January. A novel support vector machine metamodel for business risk identification. In *PRICAI* (pp. 980-984).

Laidroo, L. and Avarmaa, M., 2020. The role of location in FinTech formation. *Entrepreneurship & Regional Development*, *32*(7-8), pp.555-572.

Langley, P., Iba, W. and Thompson, K., 1992, July. An analysis of Bayesian classifiers. In *Aaai* (Vol. 90, pp. 223-228).

Laurikkala, J., 2001. Improving identification of difficult small classes by balancing class distribution. In *Artificial Intelligence in Medicine: 8th Conference on Artificial Intelligence in Medicine in Europe, AIME 2001 Cascais, Portugal, July 1–4, 2001, Proceedings 8* (pp. 63-66). Springer Berlin Heidelberg.

Law, S.H., Azman-Saini, W.N.W. and Ibrahim, M.H., 2013. Institutional quality thresholds and the finance–growth nexus. *Journal of Banking & Finance*, *37*(12), pp.5373-5381.

Le, T., Hoang Son, L., Vo, M.T., Lee, M.Y. and Baik, S.W., 2018. A cluster-based boosting algorithm for bankruptcy prediction in a highly imbalanced dataset. *Symmetry*, *10*(7), p.250.

Le, T.D., Ho, T.H., Nguyen, D.T. and Ngo, T., 2021. FinTech credit and bank efficiency: International evidence. *International Journal of Financial Studies*, *9*(3), p.44.

Lee, C.C., Li, X., Yu, C.H. and Zhao, J., 2021. Does FinTech innovation improve bank efficiency? Evidence from China’s banking industry. *International Review of Economics & Finance, 74*, pp.468-483.

Lee, T.S. and Chen, I.F., 2005. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. *Expert Systems with applications*, *28*(4), pp.743-752.

Lensink, R., Meesters, A. and Naaborg, I., 2008. Bank efficiency and foreign ownership: Do good institutions matter?. *Journal of Banking & Finance*, *32*(5), pp.834-844.

Lenz, R., 2016. Peer-to-peer lending: Opportunities and risks. *European Journal of Risk Regulation*, *7*(4), pp.688-700.

Leo, M., Sharma, S. and Maddulety, K., 2019. Machine learning in banking risk management: A literature review. *Risks, 7*(1), p.29.

Leong, C., Tan, B., Xiao, X., Tan, F.T.C. and Sun, Y., 2017. Nurturing a FinTech ecosystem: The case of a youth microloan startup in China. *International Journal of Information Management, 37*(2), pp.92-97.

Leong, K. and Sung, A., 2018. FinTech (Financial Technology): what is it and how to use technologies to create business value in FinTech way?. *International Journal of Innovation, Management and Technology*, *9*(2), pp.74-78.

Levine, R., 1997. Financial development and economic growth: views and agenda. *Journal of economic literature*, *35*(2), pp.688-726.

Li, S., Liu, F., Liu, S. and Whitmore, G.A., 2001. Comparative performance of Chinese commercial banks: Analysis, findings and policy implications. *Review of Quantitative Finance and Accounting*, *16*, pp.149-170.

Li, X. and Song, C., 2021. Does the target market affect bank performance? Evidence from the geographic diversification of city commercial banks in China. *Frontiers of Business Research in China, 15(1),* pp.1-25.

Li, Y., 2020. Analyzing efficiencies of city commercial banks in China: An application of the bootstrapped DEA approach. *Pacific-Basin Finance Journal*, *62*, p.101372.

Li, Y., Stasinakis, C. and Yeo, W.M., 2022. A hybrid XGBoost-MLP model for credit risk assessment on digital supply chain finance. *Forecasting*, *4*(1), pp.184-207.

Lin, H.L., Tsao, C.C. and Yang, C.H., 2009. Bank reforms, competition and efficiency in China's banking system: Are small city bank entrants more efficient?. *China & World Economy*, *17*(5), pp.69-87.

Lin, X. and Zhang, Y., 2009. Bank ownership reform and bank performance in China. *Journal of Banking & Finance*, *33*(1), pp.20-29.

Liu, X.Y., Wu, J. and Zhou, Z.H., 2008. Exploratory undersampling for class-imbalance learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *39*(2), pp.539-550.

Liu, Y., Luan, L., Wu, W., Zhang, Z. and Hsu, Y., 2021. Can digital financial inclusion promote China's economic growth?. *International Review of Financial Analysis*, *78*, p.101889.

Liu, Z., 2020. Literature review of supply chain finance based on blockchain perspective. Open *Journal of Business and Management, 9*(1), pp.419-429.

Ljungwall, C. and Li, J., 2007. Financial sector development, FDI and economic growth in China. *China Center for Economic Research.*

Lo, B., 2015. It Ain't Broke: The Case for Continued SEC Regulation of P2P Lending. *Harv. Bus. L. Rev. Online*, *6*, p.87.

Lozano-Vivas, A. and Humphrey, D.B., 2002. Bias in Malmquist index and cost function productivity measurement in banking. *International Journal of Production Economics, 76*(2), pp.177-188.

Lucas, A., 2004. Updating scorecards: removing the mystique. *Readings in Credit Scoring: Foundations, Developments, and Aims. Oxford University Press: New York*, pp.93-109.

Luo, D., Dong, Y., Armitage, S. and Hou, W., 2017. The impact of foreign bank penetration on the domestic banking sector: new evidence from China. *The European Journal of Finance, 23*(7-9), pp.752-780.

Lyons, A.C., Coronado Mondragon, A.E., Piller, F., Poler, R., Lyons, A.C., Mondragon, A.E.C., Piller, F. and Poler, R., 2012. Supply chain performance measurement. *Customer-Driven Supply Chains: From Glass Pipelines to Open Innovation Networks*, pp.133-148.

Ma, C., Zhang, H.H. and Wang, X., 2014. Machine learning for big data analytics in plants. *Trends in plant science, 19*(12), pp.798-808.

Ma, Z., Hou, W. and Zhang, D., 2021. A credit risk assessment model of borrowers in P2P lending based on BP neural network. *PloS one*, *16*(8), p.e0255216.

Mahalle, A., Yong, J. and Tao, X., 2021, May. Regulatory challenges and mitigation for account services offered by FinTech. *In 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)* (pp. 280-287). IEEE.

Makowski, P., 1985. Credit scoring branches out. *Credit World*, *75*(1), pp.30-37.

Malekipirbazari, M. and Aksakalli, V., 2015. Risk assessment in social lending via random forests. *Expert Systems with Applications*, *42*(10), pp.4621-4631.

Marcano-Cedeño, A., Marin-De-La-Barcena, A., Jiménez-Trillo, J., Pinuela, J.A. and Andina, D., 2011. Artificial metaplasticity neural network applied to credit scoring. *International journal of neural systems*, *21*(04), pp.311-317.

Marcelin, I. and Mathur, I., 2014. Financial development, institutions and banks. *International Review of Financial Analysis*, *31*, pp.25-33.

Marqués, A.I., García, V. and Sánchez, J.S., 2013. On the suitability of resampling techniques for the class imbalance problem in credit scoring. *Journal of the Operational Research Society*, *64*(7), pp.1060-1070.

Matthews, K., 2013. Risk management and managerial efficiency in Chinese banks: A network DEA framework. *Omega*, *41*(2), pp.207-215.

Matthews, K., Guo, J. and Zhang, N., 2007. Rational Inefficiency and non-performing loans in Chinese Banking: A non-parametric Bootstrapping Approach.

McKinnon, R.I., 1973. Money and capital in economic development (Washington, DC: Brookings Institution, 1973). *McKinnonMoney and Capital in Economic Development1973*.

Mester, L.J., 1996. A study of bank efficiency taking into account risk-preferences. *Journal of banking & finance*, *20*(6), pp.1025-1045.

Michels, J., 2012. Do unverifiable disclosures matter? Evidence from peer-to-peer lending. *The Accounting Review*, *87*(4), pp.1385-1413.

Moretto, A., Grassi, L., Caniato, F., Giorgino, M. and Ronchi, S., 2019. Supply chain finance: From traditional to supply chain credit rating. *Journal of Purchasing and Supply Management, 25*(2), pp.197-217.

Moretto, A., Grassi, L., Caniato, F., Giorgino, M. and Ronchi, S., 2019. Supply chain finance: From traditional to supply chain credit rating. *Journal of Purchasing and Supply Management, 25*(2), pp.197-217.

Mosavi, A., Ozturk, P. and Chau, K.W., 2018. Flood prediction using machine learning models: Literature review. *Water, 10*(11), p.1536.

Moscato, V., Picariello, A. and Sperlí, G., 2021. A benchmark of machine learning approaches for credit score prediction. *Expert Systems with Applications*, *165*, p.113986.

Mullainathan, S. and Spiess, J., 2017. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, *31*(2), pp.87-106.

Namvar, A., Siami, M., Rabhi, F. and Naderpour, M., 2018. Credit risk prediction in an imbalanced social lending environment. *arXiv preprint arXiv:1805.00801*.

Nash, R.M. and Beardsley, E., 2015. The future of finance. *The rise of the new Shadow Bank, Goldman Sachs*.

Nasteski, V., 2017. An overview of the supervised machine learning methods. *Horizons. b*, *4*, pp.51-62.

Navaretti, G.B., Calzolari, G., Mansilla-Fernandez, J.M. and Pozzolo, A.F., 2018. FinTech and banking. Friends or foes?. Friends or Foes.

Nguyen, D.K., Sermpinis, G. and Stasinakis, C., 2022. Big data, artificial intelligence and machine learning: A transformative symbiosis in favour of financial technology. *European Financial Management*.

Nguyen, T.P.T., Nghiem, S.H., Roca, E. and Sharma, P., 2016. Bank reforms and efficiency in Vietnamese banks: evidence based on SFA and DEA. *Applied Economics*, *48*(30), pp.2822-2835.

Nickerson, A., Japkowicz, N. and Milios, E.E., 2001, January. Using unsupervised learning to guide resampling in imbalanced data sets. In *International workshop on artificial intelligence and statistics* (pp. 224-228). PMLR.

Niu, G., Jin, X., Wang, Q. and Zhou, Y., 2022. Broadband infrastructure and digital financial inclusion in rural China. *China Economic Review, 76*, p.101853.

Niu, K., Zhang, Z., Liu, Y. and Li, R., 2020. Resampling ensemble model based on data distribution for imbalanced credit risk evaluation in P2P lending. *Information Sciences*, *536*, pp.120-134.

Ntwiga, D.B., 2020. Technical efficiency in the Kenyan banking sector: Influence of FinTech and banks collaboration. *Journal of Finance and Economics, 8*(1), pp.13-22.

Odeck, J., 2009. Statistical precision of DEA and Malmquist indices: A bootstrap application to Norwegian grain producers. *Omega*, *37*(5), pp.1007-1017.

Odom, M.D. and Sharda, R., 1990, June. A neural network model for bankruptcy prediction. In *1990 IJCNN International Joint Conference on neural networks* (pp. 163-168). IEEE.

Olan, F., Arakpogun, E.O., Jayawickrama, U., Suklan, J. and Liu, S., 2022. Sustainable supply chain finance and supply networks: The role of artificial intelligence. *IEEE Transactions on Engineering Management.*

Omarini, A., 2020. FinTech: A new hedge for a financial re-intermediation. Strategy and risk perspectives. *Frontiers in Artificial Intelligence*, *3*, p.63.

Oreski, S. and Oreski, G., 2014. Genetic algorithm-based heuristic for feature selection in credit risk assessment. *Expert systems with applications, 41*(4), pp.2052-2064.

Orgler, Y.E., 1970. A credit scoring model for commercial loans. *Journal of money, Credit and Banking*, *2*(4), pp.435-445.

Owusu, E., Quainoo, R., Appati, J.K. and Mensah, S., 2022, June. Loan Default Predictive Analytics. In *2022 IEEE World Conference on Applied Intelligence and Computing (AIC)* (pp. 617-622). IEEE.

Ozili, P.K. and Uadiale, O., 2017. Ownership concentration and bank profitability. *Future Business Journal*, *3*(2), pp.159-171.

Padmaja, T.M., Krishna, P.R. and Bapi, R.S., 2008, November. Majority filter-based minority prediction (MFMP): An approach for unbalanced datasets. In *TENCON 2008-2008 IEEE Region 10 Conference* (pp. 1-6). IEEE.

Paradi, J.C. and Schaffnit, C., 2004. Commercial branch performance evaluation and results communication in a Canadian bank—a DEA application. *European Journal of Operational Research*, *156*(3), pp.719-735.

Pearl, J., 1988. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan kaufmann.

Pessarossi, P. and Weill, L., 2015. Do capital requirements affect cost efficiency? Evidence from China*. Journal of Financial Stability, 19*, pp.119-127.

Phan, D.H.B., Narayan, P.K., Rahman, R.E. and Hutabarat, A.R., 2020. Do financial technology firms influence bank performance?. *Pacific-Basin finance journal*, *62*, p.101210.

Philippon, T., 2016. *The FinTech opportunity* (No. w22476). National Bureau of Economic Research.

Piramuthu, S., 1999. Financial credit-risk evaluation with neural and neurofuzzy systems. *European Journal of Operational Research*, *112*(2), pp.310-321.

Pope, D.G. and Sydnor, J.R., 2011. What's in a Picture? Evidence of Discrimination from Prosper. com. *Journal of Human resources*, *46*(1), pp.53-92.

Prado, J.W., Castro Alcântara, V., Melo Carvalho, F., Vieira, K.C., Machado, L.K. and Tonelli, D.F., 2016. Multivariate analysis of credit risk and bankruptcy research data: a bibliometric study involving different knowledge fields (1968---2014). *Scientometrics*, *106*(3), pp.1007-1029.

Puschmann, T., 2017. FinTech. *Business & Information Systems Engineering*, *59*, pp.69-76.

Qian, Y., 2000. The process of China's market transition (1978-1998): The evolutionary, historical, and comparative perspectives. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, pp.151-171.

Qian, Y., Liang, Y., Li, M., Feng, G. and Shi, X., 2014. A resampling ensemble algorithm for classification of imbalance problems. *Neurocomputing*, *143*, pp.57-67.

Quinlan, J.R., 1996. Improved use of continuous attributes in C4. 5. *Journal of artificial intelligence research*, *4*, pp.77-90.

Quinlan, J.R., 1996. Improved use of continuous attributes in C4. 5. *Journal of artificial intelligence research*, *4*, pp.77-90.

Rai, A., 2020. Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science, 48,* pp.137-141.

Rasiwala, F.S. and Kohli, B., 2021. Artificial intelligence in FinTech: Understanding stakeholders perception on innovation, disruption, and transformation in finance. *International Journal of Business Intelligence Research (IJBIR)*, *12*(1), pp.48-65.

Rossi, S.P., Schwaiger, M.S. and Winkler, G., 2009. How loan portfolio diversification affects risk, efficiency and capitalization: A managerial behavior model for Austrian banks. *Journal of banking & finance*, *33*(12), pp.2218-2226.

Sagi, O. and Rokach, L., 2018. Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(4), p.e1249.

Sarker, I.H., 2021. Machine learning: Algorithms, real-world applications and research directions. *SN computer science, 2*(3), p.160.

Schebesch, K.B. and Stecking, R., 2008. Using multiple SVM models for unbalanced credit scoring data sets. In *Data Analysis, Machine Learning and Applications: Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation eV, Albert-Ludwigs-Universität Freiburg, March 7–9, 2007* (pp. 515-522). Springer Berlin Heidelberg.

Schumpeter, J.A., 2000. Entrepreneurship as innovation. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Scuotto, V., Caputo, F., Villasalero, M. and Del Giudice, M., 2017. A multiple buyer–supplier relationship in the context of SMEs' digital supply chain management. *Production Planning & Control*, *28*(16), pp.1378-1388.

Sealey Jr, C.W. and Lindley, J.T., 1977. Inputs, outputs, and a theory of production and cost at depository financial institutions. *The journal of finance*, *32*(4), pp.1251-1266.

See, K.F. and He, Y., 2015. Determinants of technical efficiency in Chinese banking: A double bootstrap data envelopment analysis approach. *Global Economic Review*, *44*(3), pp.286-307.

Seiffert, C., Khoshgoftaar, T.M., Van Hulse, J. and Napolitano, A., 2009. RUSBoost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *40*(1), pp.185-197.

Semih Yildirim, H. and Philippatos, G.C., 2007. Efficiency of banks: Recent evidence from the transition economies of Europe, 1993–2000. *European Journal of Finance*, *13*(2), pp.123-143.

Sermpinis, G., Dunis, C., Laws, J. and Stasinakis, C., 2012. Forecasting and trading the EUR/USD exchange rate with stochastic Neural Network combination and time-varying leverage. *Decision Support Systems*, *54*(1), pp.316-329.

Sermpinis, G., Stasinakis, C. and Hassanniakalager, A., 2017. Reverse adaptive krill herd locally weighted support vector regression for forecasting and trading exchange traded funds. *European Journal of Operational Research*, *263*(2), pp.540-558.

Sheng, T., 2021. The effect of FinTech on banks' credit provision to SMEs: Evidence from China. *Finance Research Letters*, *39*, p.10

Shinde, P.P. and Shah, S., 2018, August. A review of machine learning and deep learning applications. In *2018 Fourth international conference on computing communication control and automation (ICCUBEA)* (pp. 1-6). IEEE.

Shy, O., 2001. *The economics of network industries*. Cambridge university press.

Simar, L. and Wilson, P.W., 1999. Estimating and bootstrapping Malmquist indices. *European journal of operational research*, *115*(3), pp.459-471.

Simar, L. and Wilson, P.W., 2000. A general methodology for bootstrapping in non-parametric frontier models. *Journal of applied statistics*, *27*(6), pp.779-802.

Simar, L. and Wilson, P.W., 2000. Statistical inference in nonparametric frontier models: The state of the art. *Journal of productivity analysis, 13(1)*, pp.49-78.

Simar, L. and Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, *136*(1), pp.31-64.

Smith, M.R., Martinez, T. and Giraud-Carrier, C., 2014. An instance level analysis of data complexity. *Machine learning*, *95*, pp.225-256.

Song, H., Han, S., Liu, W. and Ganguly, A., 2023. What role do FinTech companies play in supply chain finance? A signaling intermediary perspective. *Journal of Business & Industrial Marketing, 38*(6), pp.1279-1294.

Song, Y., Wang, Y., Ye, X., Wang, D., Yin, Y. and Wang, Y., 2020. Multi-view ensemble learning based on distance-to-model and adaptive clustering for imbalanced credit risk assessment in P2P lending. *Information Sciences*, *525*, pp.182-204.

Soni, G., Kumar, S., Mahto, R.V., Mangla, S.K., Mittal, M.L. and Lim, W.M., 2022. A decision-making framework for Industry 4.0 technology implementation: The case of FinTech and sustainable supply chain finance for SMEs. *Technological Forecasting and Social Change*, *180*, p.121686.

Stasinakis, C., Sermpinis, G., Psaradellis, I. and Verousis, T., 2016. Krill-Herd Support Vector Regression and heterogeneous autoregressive leverage: evidence from forecasting and trading commodities. *Quantitative Finance*, *16*(12), pp.1901-1915.

Staub, R.B., e Souza, G.D.S. and Tabak, B.M., 2010. Evolution of bank efficiency in Brazil: A DEA approach. *European journal of operational research*, *202*(1), pp.204-213.

Stecking, R. and Schebesch, K.B., 2006. Variable subset selection for credit scoring with support vector machines. In *Operations Research Proceedings 2005: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), Bremen, September 7–9, 2005* (pp. 251-256). Springer Berlin Heidelberg.

Steenackers, A. and Goovaerts, M., 1989. A credit scoring model for personal loans. *Insurance: Mathematics & Economics*, *8*(1), pp.31-34.

Stiglitz, J.E. and Weiss, A., 1981. Credit rationing in markets with imperfect information. *The American economic review*, *71*(3), pp.393-410.

Sufian, F. and Habibullah, M.S., 2009. Bank specific and macroeconomic determinants of bank profitability: Empirical evidence from the China banking sector. *Frontiers of Economics in China, 4*(2), pp.274-291.

Sufian, F. and Habibullah, M.S., 2012. Globalizations and bank performance in China. *Research in International Business and Finance, 26*(2), pp.221-239.

Sun, J., Harimaya, K. and Yamori, N., 2013. Regional economic development, strategic investors, and efficiency of Chinese city commercial banks. *Journal of Banking & Finance*, *37*(5), pp.1602-1611.

Sun, J., Lang, J., Fujita, H. and Li, H., 2018. Imbalanced enterprise credit evaluation with DTE-SBD: Decision tree ensemble based on SMOTE and bagging with differentiated sampling rates. *Information Sciences*, *425*, pp.76-91.

Taiwo, J.N., Ucheaga, E.G., Achugamonu, B.U., Adetiloye, K. and Okoye, O., 2017. Credit risk management: Implications on bank performance and lending growth. *Saudi Journal of Business and Management Studies*, *2*, pp.584-590.

Tan, Y. and Floros, C., 2012. Bank profitability and inflation: the case of China. *Journal of Economic Studies.*

Tang, H., 2019. Peer-to-peer lenders versus banks: substitutes or complements?. *The Review of Financial Studies*, *32*(5), pp.1900-1938.

Taylor, J.B., 1999. The robustness and efficiency of monetary policy rules as guidelines for interest rate setting by the European Central Bank. *Journal of Monetary Economics*, *43*(3), pp.655-679.

Teply, P. and Polena, M., 2020. Best classification algorithms in peer-to-peer lending. *The North American Journal of Economics and Finance*, *51*, p.100904.

Thakor, A.V., 2020. FinTech and banking: What do we know?. *Journal of Financial Intermediation*, *41*, p.100833.

Timme, S.G. and Williams-Timme, C., 2000. The financial-SCM connection. *SUPPLY CHAIN MANAGEMENT REVIEW, V. 4, NO. 2 (MAY/JUNE 2000), P. 33-40: ILL.*

Tomek, I., 1976. AN EXPERIMENT WITH THE EDITED NEAREST-NIEGHBOR RULE.

Tortosa-Ausina, E., Grifell-Tatjé, E., Armero, C. and Conesa, D., 2008. Sensitivity analysis of efficiency and Malmquist productivity indices: An application to Spanish savings banks. *European Journal of Operational Research*, *184*(3), pp.1062-1084.

Tsai, C.F., 2008. Financial decision support using neural networks and support vector machines. *Expert Systems*, *25*(4), pp.380-393.

Van den End, J.W., 2016. A macroprudential approach to address liquidity risk with the loan-to-deposit ratio. *The European Journal of Finance*, *22*(3), pp.237-253.

Van Der Putten, P. and Van Someren, M., 2004. A bias-variance analysis of a real world learning problem: The CoIL challenge 2000. *Machine learning*, *57*, pp.177-195.

Van Gestel, T., Baesens, B., Van Dijcke, P., Suykens, J., Garcia, J. and Alderweireld, T., 2005. Linear and nonlinear credit scoring by combining logistic regression and support vector machines. *Journal of credit Risk, 1*(4).

Varian, H.R., 2014. Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *28*(2), pp.3-28.

Vives, X., 2017. The impact of FinTech on banking. *European Economy*, (2), pp.97-105.

Wandfluh, M., Hofmann, E. and Schoensleben, P., 2016. Financing buyer–supplier dyads: an empirical analysis on financial collaboration in the supply chain. *International Journal of Logistics Research and Applications*, *19*(3), pp.200-217.

Wang, F., Ding, L., Yu, H. and Zhao, Y., 2020. Big data analytics on enterprise credit risk evaluation of e-Business platform. *Information Systems and e-Business Management*, *18*, pp.311-350.

Wang, G. and Ma, J., 2012. A hybrid ensemble approach for enterprise credit risk assessment based on Support Vector Machine. *Expert Systems with Applications*, *39*(5), pp.5325-5331.

Wang, G.J., Xie, C., Zhao, L. and Jiang, Z.Q., 2018. Volatility connectedness in the Chinese banking system: Do state-owned commercial banks contribute more?. *Journal of International Financial Markets, Institutions and Money*, *57*, pp.205-230.

Wang, J., Shen, Y. and Huang, Y., 2016. Evaluating the regulatory scheme for internet finance in China: the case of peer-to-peer lending. *China Economic Journal, 9*(3), pp.272-287.

Wang, R. and Wu, Y., 2021. Application of blockchain technology in supply chain finance of beibu gulf region. *Mathematical Problems in Engineering*, 2021, pp.1-10.

Wang, R., Liu, J. and Luo, H., 2021. FinTech development and bank risk taking in China. *The European Journal of Finance*, *27*(4-5), pp.397-418.

Wang, S. and Yao, X., 2009, March. Diversity analysis on imbalanced data sets by using ensemble models. In *2009 IEEE symposium on computational intelligence and data mining* (pp. 324-331). IEEE.

Wang, X., Wang, X., Ren, X. and Wen, F., 2022. Can digital financial inclusion affect CO2 emissions of China at the prefecture level?

Wang, Y., Li, X., Abdou, H. and Ntim, C.G., 2015. Financial development and economic growth in China. *Investment Management and Financial Innovations*, *13*(3), pp.8-18.

Wang, Y., Xiuping, S. and Zhang, Q., 2021. Can FinTech improve the efficiency of commercial banks?—An analysis based on big data. *Research in international business and finance*, *55*, p.101338.

Wasikowski, M. and Chen, X.W., 2009. Combating the small sample class imbalance problem using feature selection. *IEEE Transactions on knowledge and data engineering*, *22*(10), pp.1388-1400.

Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S. and Wrobel, S., 2019. A review of machine learning for the optimization of production processes. *The International Journal of Advanced Manufacturing Technology, 104*(5-8), pp.1889-1902.

Wiginton, J.C., 1980. A note on the comparison of logit and discriminant models of consumer credit behavior. *Journal of Financial and Quantitative Analysis*, *15*(3), pp.757-770.

Wonglimpiyarat, J., 2017. FinTech banking industry: a systemic approach. *foresight*, *19*(6), pp.590-603.

Wu, H.L., Chen, C.H. and Shiu, F.Y., 2007. The impact of financial development and bank characteristics on the operational performance of commercial banks in the Chinese transitional economy. *Journal of Economic Studies, 34*(5), pp.401-414.

Wuttke, D.A., Blome, C. and Henke, M., 2013. Focusing the financial flow of supply chains: An empirical investigation of financial supply chain management. *International journal of production economics*, *145*(2), pp.773-789.

Xia, Y., Liu, C., Li, Y. and Liu, N., 2017. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. *Expert systems with applications*, *78*, pp.225-241.

Xiang, X., Lina, Z., Yun, W. and Chengxuan, H., 2017. China's path to FinTech development. *European Economy*, (2), pp.143-159.

Xiaogang, C., Skully, M. and Brown, K., 2005. Banking efficiency in China: Application of DEA to pre-and post-deregulation eras: 1993–2000. *China Economic Review*, *16*(3), pp.229-245.

Xu, J.X., Li, N. and Ahmad, M.I., 2018. Banking performance of China and Pakistan. *Entrepreneurship and Sustainability Issues, 5*(4), pp.929-942.

Xu, S.J., 2012. The impact of financial development on energy consumption in China: based on SYS-GMM estimation. In *Advanced materials research* (Vol. 524, pp. 2977-2981). Trans Tech Publications Ltd.

Yamori, N., Harimaya, K. and Tomimura, K., 2017. Corporate governance structure and efficiencies of cooperative banks. *International Journal of Finance & Economics*, *22*(4), pp.368-378.

Yan, J., Yu, W. and Zhao, J.L., 2015. How signaling and search costs affect information asymmetry in P2P lending: the economics of big data. *Financial Innovation*, *1*, pp.1-11.

Yang, T. and Zhang, X., 2022. FinTech adoption and financial inclusion: Evidence from household consumption in China. *Journal of Banking & Finance, 145*, p.106668.

Yao, S., Han, Z. and Feng, G., 2008. Ownership reform, foreign competition and efficiency of Chinese commercial banks: A non-parametric approach. World Economy, 31(10), pp.1310-1326.

Yen, S.J. and Lee, Y.S., 2009. Cluster-based under-sampling approaches for imbalanced data distributions. *Expert Systems with Applications*, *36*(3), pp.5718-5727.

Yu, L. and Zhang, X., 2021. Can small sample dataset be used for efficient internet loan credit risk assessment? Evidence from online peer to peer lending. *Finance Research Letters*, *38*, p.101521.

Yu, L., Wang, S. and Lai, K.K., 2010. Developing an SVM-based ensemble learning system for customer risk identification collaborating with customer relationship management. *Frontiers of Computer Science in China*, *4*, pp.196-203.

Yuengert, A.M., 1993. The measurement of efficiency in life insurance: estimates of a mixed normal-gamma error model. *Journal of Banking & Finance*, *17*(2-3), pp.483-496.

Yurdakul, M. and İç, Y.T., 2004. AHP approach in the credit evaluation of the manufacturing firms in Turkey. *International journal of production economics*, *88*(3), pp.269-289.

Yuxiang, K. and Chen, Z., 2011. Financial development and environmental performance: evidence from China. *Environment and Development Economics*, *16*(1), pp.93-111.

Zetsche, D.A., Buckley, R.P., Arner, D.W. and Barberis, J.N., 2017. From FinTech to TechFin: the regulatory challenges of data-driven finance. *NYUJL & Bus.*, *14*, p.393.

Zha, Y., Liang, N., Wu, M. and Bian, Y., 2016. Efficiency evaluation of banks in China: A dynamic two-stage slacks-based measure approach. *Omega*, *60*, pp.60-72.

Zhang, J., Wang, P. and Qu, B., 2012. Bank risk taking, efficiency, and law enforcement: Evidence from Chinese city commercial banks. *China Economic Review*, *23*(2), pp.284-295.

Zhang, L., Hu, H. and Zhang, D., 2015. A credit risk assessment model based on SVM for small and medium enterprises in supply chain finance. *Financial Innovation*, *1*(1), pp.1-21.

Zhang, L., Hu, H. and Zhang, D., 2015. A credit risk assessment model based on SVM for small and medium enterprises in supply chain finance. *Financial Innovation*, *1*(1), pp.1-21.

Zhang, T. and Chi, G., 2021. A heterogeneous ensemble credit scoring model based on adaptive classifier selection: An application on imbalanced data. *International Journal of Finance & Economics*, *26*(3), pp.4372-4385.

Zhang, W., 1998. China's SOE reform: A corporate governance perspective. Institute of Business Research Working Paper, (1998E04).

Zhang, Y., Jia, H., Diao, Y., Hai, M. and Li, H., 2016. Research on credit scoring by fusing social media information in online peer-to-peer lending. *Procedia Computer Science*, *91*, pp.168-174.

Zhao, J., Li, X., Yu, C.H., Chen, S. and Lee, C.C., 2022. Riding the FinTech innovation wave: FinTech, patents and bank performance. *Journal of International Money and Finance*, *122*, p.102552.

Zhao, Q., Tsai, P.H. and Wang, J.L., 2019. Improving financial service innovation strategies for enhancing china's banking industry competitive advantage during the FinTech revolution: A Hybrid MCDM model. *Sustainability*, *11*(5), p.1419.

Zheng, J., Liu, X. and Bigsten, A., 2003. Efficiency, technical progress, and best practice in Chinese state enterprises (1980–1994). *Journal of Comparative Economics*, *31*(1), pp.134-152.

Zheng, X.L., Zhu, M.Y., Li, Q.B., Chen, C.C. and Tan, Y.C., 2019. FinBrain: when finance meets AI 2.0. *Frontiers of Information Technology & Electronic Engineering*, *20*(7), pp.914-924.

Zhou, J., Li, W., Wang, J., Ding, S. and Xia, C., 2019. Default prediction in P2P lending from high-dimensional data based on machine learning. *Physica A: Statistical Mechanics and its Applications*, *534*, p.122370.

Zhu, Q., 2020. On the performance of Matthews correlation coefficient (MCC) for imbalanced dataset. *Pattern Recognition Letters, 136,* pp.71-80.

Zhu, Y., Xie, C., Wang, G.J. and Yan, X.G., 2017. Comparison of individual, ensemble and integrated ensemble machine learning methods to predict China's SME credit risk in supply chain finance. *Neural Computing and Applications, 28*, pp.41-50.

Zhu, Y., Xie, C., Wang, G.J. and Yan, X.G., 2017. Comparison of individual, ensemble and integrated ensemble machine learning methods to predict China's SME credit risk in supply chain finance. *Neural Computing and Applications*, *28*, pp.41-50.

Zhu, Y., Zhou, L., Xie, C., Wang, G.J. and Nguyen, T.V., 2019. Forecasting SMEs' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach. *International Journal of Production Economics*, *211*, pp.22-33.