Learning and Analytics in Intelligent Systems 42

Rangan Gupta
Francesco Bartolucci
Vasilios N. Katsikis
Srikanta Patnaik *Editors*

Recent Advancements in Computational Finance and Business Analytics

Proceedings of the 2nd International Conference on Computational Finance and Business Analytics – ICCFBA-2024

Springer

Learning and Analytics in Intelligent Systems 42

Series Editors

The main aim of the series is to make available a publication of books in hard copy form and soft copy form on all aspects of learning, analytics and advanced intelligent systems and related technologies. The mentioned disciplines are strongly related and complement one another significantly. Thus, the series encourages cross-fertilization highlighting research and knowledge of common interest. The series allows a unified/integrated approach to themes and topics in these scientific disciplines which will result in significant cross-fertilization and research dissemination. To maximize dissemination of research results and knowledge in these disciplines, the series publishes edited books, monographs, handbooks, textbooks and conference proceedings.

Indexed by EI Compendex.

Rangan Gupta · Francesco Bartolucci ·
Vasilios N. Katsikis · Srikanta Patnaik
Editors

Recent Advancements in Computational Finance and Business Analytics

Proceedings of the 2nd International Conference on Computational Finance and Business Analytics – ICCFBA-2024

Springer

Editors
Rangan Gupta
Department of Economics
University of Pretoria
Pretoria, South Africa

Francesco Bartolucci
Department of Economics
University of Perugia
Perugia, Italy

Vasilios N. Katsikis
Department of Economics
National and Kapodistrian University
of Athens
Athens, Greece

Srikanta Patnaik
Interscience Institute of Management
and Technology
Bhubaneswar, Odisha, India

ISSN 2662-3447 ISSN 2662-3455 (electronic)
Learning and Analytics in Intelligent Systems
ISBN 978-3-031-70597-7 ISBN 978-3-031-70598-4 (eBook)
https://doi.org/10.1007/978-3-031-70598-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

In the rapidly evolving landscape of data-driven decision-making, the significance of computational finance and business analytics has become increasingly paramount. Recognizing the need to drive forward research in this critical domain, the International Conference on Computational Finance and Business Analytics (CFBA-2024) convened a distinguished gathering of eminent researchers, seasoned academicians, domain experts, and innovative practitioners from around the globe.

This premier forum provided a platform for dynamic interactions and fruitful collaborations among scholars exploring the realms of financial analytics, human resource analytics, marketing analytics, and diverse facets of business analytics. The conference featured a rich tapestry of original and unpublished research contributions, delving into a wide array of topics, including computational intelligence for financial modeling, machine learning applications in finance, talent acquisition analytics, digital marketing strategies, fraud analytics, and much more.

The conference proceedings encapsulate a wealth of knowledge and insights, reflecting the latest developments and cutting-edge practical applications in the field of computational finance and business analytics. The curated papers included in this volume represent the forefront of research and innovation, showcasing the most advanced techniques and technologies employed in tackling complex business challenges.

We are honored to present these proceedings as a valuable resource for researchers, academicians, and practitioners who seek to deepen their understanding and contribute to the continued advancement of computational finance and business analytics. We extend our sincere gratitude to all the authors, presenters, and attendees who have made CFBA-2024 a resounding success, propelling the growth and evolution of this dynamic and crucial field.

Srikanta Patnaik

Editorial

The landscape of data-driven decision-making has undergone a remarkable transformation, and the 2nd International Conference on Computational Finance and Business Analytics (CFBA-2024) stood as a testament to the growing significance of these cutting-edge fields. Held at the Interscience Institute of Management & Technology in Bhubaneswar, India, from April 5–6, 2024, this premier forum brought together a diverse gathering of researchers, academicians, experts, and innovative practitioners from around the world. Computational finance, a dynamic and evolving discipline, has revolutionized the financial industry by harnessing the power of mathematical and statistical methods, combined with advanced computer algorithms and models. This field enables the analysis and prediction of market behavior, the management of risk, and the optimization of investment strategies. By processing and analyzing large volumes of financial data using sophisticated techniques, computational finance has empowered financial professionals to make informed decisions and gain unprecedented insights into market dynamics.

Complementing the advancements in computational finance, business analytics has emerged as a transformative force, leveraging data analysis and statistical methods to extract meaningful insights and drive strategic decision-making. Across various functional areas, from marketing and finance to operations and supply chain management, business analytics has proven instrumental in understanding customer behavior, optimizing processes, forecasting demand, and identifying strategic opportunities. The comprehensive program at CFBA-2024 provided a dynamic platform for the exchange of ideas, the sharing of cutting-edge research, and the forging of new collaborations. Keynote speeches, research paper presentations, and interactive sessions offered attendees the opportunity to gain invaluable insights, explore the latest developments, and contribute to the advancement of computational finance and business analytics.

The conference proceedings, a treasure trove of knowledge and innovation, capture the essence of this transformative gathering. The curated papers within these pages represent the leading edge of research, showcasing the most sophisticated techniques and technologies employed in tackling complex business challenges. These proceedings serve as a valuable resource for researchers, academicians, and practitioners seeking to deepen their understanding and drive the continued evolution of these dynamic and crucial fields. As the world becomes increasingly data-driven, the significance of computational finance and business analytics will only continue to grow. The CFBA-2024 conference has served as a catalyst, igniting collaborations, inspiring groundbreaking research, and paving the way for innovative solutions that will shape the future of finance, management, and beyond

Keeping this in view, the advancements of CFBA-2024 have focused on both core areas of finance and technological-enabled advancements in the area of finance. We did not limit ourselves to finance we also covered the technological advancements of human resource, marketing, and business analytics as well. The organizing committee has received around 100 papers for the conference which were considered for review

and editing. Out of these 100 papers, 62 papers were accepted for presentation and publication, whereas 54 papers were registered, which are covered in this proceeding. These papers have been categorized into three major sections which are as follows :(i) Financial analytics, (ii) HR and marketing analytics, and (iii) business analytics.

Financial Analytics

Embarking on the journey through financial analytics unveils a rich tapestry of 25 papers, each representing a meticulous exploration of the intricate web that binds traditional financial practices with the cutting-edge methodologies of data science. Beyond the mere crunching of numbers, these scholarly endeavors symbolize a collective quest to decode the nuanced language of modern finance within an increasingly digitized global economy. At the forefront, researchers endeavor to fortify financial systems against the ever-looming specter of fraudulent activities, wielding the formidable arsenal of deep learning algorithms and advanced data analysis techniques. Moreover, the optimization of audit processes through the seamless integration of cloud computing and data center technologies heralds a new era of transparency and accountability, promising to usher financial operations into a realm of unprecedented efficiency and trust. Digging deeper, the exploration of risk assessment methodologies emerges as a guiding beacon for portfolio managers and investors alike, navigating the treacherous waters of market volatility with newfound precision and foresight. Yet, it is the convergence of emerging technologies like blockchain and big data analytics that truly propels financial analytics into uncharted territories, offering invaluable insights into market trends, risk mitigation strategies, and investment opportunities. Through the lens of specific companies such as BYD and NIO Auto, researchers illuminate the practical applications of financial analytics, empowering decision-makers to navigate the complexities of industry-specific challenges and steer their organizations toward sustainable growth and prosperity. And as policymakers and investors alike pore over the intricate nuances of economic policies and taxation impacts, they glean invaluable insights from these papers, guiding their course toward informed decision-making and strategic planning in an ever-evolving economic landscape.

HR and Marketing Analytics

In the domain of HR and marketing analytics, a curated selection of five papers serves as a sign, illuminating the complex relationship between human behavior and organizational dynamics. These scholarly works transcend the realm of statistics; they represent a concerted effort to unravel the enigma of human decision-making within the context of marketing strategies and corporate culture. Armed with sophisticated data mining techniques and predictive analysis tools, researchers embark on a quest to decode the whims of consumers, optimize marketing campaigns, and encourage a culture of innovation and collaboration within organizations. Through empirical research and data-driven methodologies, they offer invaluable insights into the complex tapestry of human behavior, guiding organizations toward more effective marketing strategies and employee management practices. Moreover, the exploration of the correlation between organizational culture

and performance sheds light on the intangible factors that drive organizational success, emphasizing the importance of nurturing a positive workplace environment conducive to growth and productivity. By leveraging these insights, organizations can forge deeper connections with their customers, attract and retain top talent, and chart a course toward sustained success in an increasingly competitive marketplace.

Business Analytics

In the dynamic realm of business analytics, a collection of 24 papers serves as a compass, guiding organizations through the uncharted waters of the digital economy. These scholarly endeavors transcend the realm of data analysis; they represent a transformative journey toward business innovation and growth in an increasingly digitalized world. From the intricacies of cross-border e-commerce to the allocation of artificial intelligence resources and the optimization of real estate regulation strategies, these papers offer a panoramic view of the digital frontier. Leveraging the power of machine learning algorithms, statistical analysis, and optimization frameworks, researchers provide actionable insights for enhancing user experiences, optimizing resource allocation, and formulating agile business strategies. By harnessing the power of data analytics, these papers empower organizations to capitalize on emerging opportunities and stay ahead of the curve in a rapidly evolving business landscape. Furthermore, the analysis of CEO traits and their impact on company IPOs sheds light on the pivotal role of leadership in driving organizational success, emphasizing the importance of visionary decision-making in an era defined by technological disruption. Armed with these insights, organizations can navigate the digital frontier with confidence, unlocking new pathways for growth, innovation, and prosperity.

The CFBA-2024 conference proceedings included papers and presentations on cutting-edge research in the field of management. The conference provided a platform for researchers and practitioners to engage in meaningful discussions and to network with each other. Overall, the conference was a success and provided valuable insights into the latest trends and developments in the field of computational finance and business analytics.

Acknowledgement

We would like to express our sincere appreciation and gratitude to all those who have contributed to the success of the International Conference on Computational Finance and Business Analytics (CFBA-2024).

First of all, we would like to express our heartfelt thanks to all the contributors of CFBA-2024, who have contributed their research findings in this volume. Their active involvement, insightful discussions, and valuable contributions made the conference a truly enriching experience. We would also like to acknowledge EasyChair for their efficient conference management system. Their platform streamlined the submission and review process, making it easier for authors and reviewers to participate in the conference.

Our heartfelt appreciation goes out to our esteemed keynote speaker Shri. Jagadananda, and Prof. (Dr.) Francesco Bartolucci whose valuable insights and perspectives greatly enriched our conference. We are grateful for the time and effort they took to prepare and deliver their presentations. We would also like to thank the session chair members, Dr. Pramod Kumar Prusty and Mr. Subhajit Roul, for their invaluable contributions in ensuring that the conference ran smoothly. Their dedication and professionalism in managing the sessions were instrumental in making the conference a success.

Our deepest appreciation goes out to the program and technical committee for their diligent work in reviewing and selecting the papers that were presented at the conference. Their expertise and commitment to academic excellence helped to ensure that the papers presented were of the highest quality. Furthermore, we extend our deepest appreciation to the Organizing Committee for their hard work and dedication in planning and executing CFBA-2024. Their meticulous attention to detail, tireless coordination, and outstanding organizational skills contributed significantly to the smooth running of the conference.

Last but not least, we extend our sincere thanks to the Springer Series Editors of Learning and Analytics in Intelligent Systems Prof. George A. Tsihrintzis, Prof. Maria Virvou, and Prof. Lakhmi C. Jain for their support and guidance in publishing the conference proceedings "**Recent Advancements in Computational Finance and Business Analytics"** in their esteemed book series. We also extended our sincere thanks to Dr. Thomas Ditzinger, Editorial Director of Springer Interdisciplinary and Applied Sciences and Engineering division.

We are grateful to all the individuals and organizations that played a part in making CFBA-2024 a successful and memorable event. Without your support and dedication,

this conference would not have been possible. Thank you all for your contributions and commitment to advancing the field of computational finance and business analytics.

Rangan Gupta
Francesco Bartolucci
Vasilios N. Katsikis
Srikanta Patnaik

CFBA-2024 Conference Committee

Honorary Chair

Lakhmi C. Jain	KES International, UK

General Chair

Francesco Vigliarolo	Professor of Regional Economy at the Catholic University of La Plata UNESCO Chair in "Economic Systems and Human Rights" at the National University of La Plata Associate Researcher at the Center for the Study of Sociology of Work (CESOT), University of Buenos Aires

Publication and Program Chair

Srikanta Patnaik	I.I.M.T., Bhubaneswar, Odisha, India

Publication Chair

Vasilios N. Katsikis	National and Kapodistrian University of Athens, Greece

Program Chair

Rangan Gupta	Department of Economics, University of Pretoria Pretoria, South Africa

Program Co-chair

Deepika Gupta	Indian Institute of Management, Vishakhapatnam

Technical Chair

Ashok Mahapatra, B. Tech, MBA, Ph.D

Technical Co-chair

Sharada Prasad Sahoo	Berhampur University, India

Finance Chair

Dhananjaya Sarangi	IT Expert, World Skill Center, Bhubaneswar, India

Finance Co-chair

Bin Hu	Changsha Normal University

International Advisory Committee

Francesco Bartolucci	University of Perugia, Italy

Program Committee

David Boohene	West Africa University of Energy and Natural Resources, Ghana, West Africa
Georgios Mamanis	University of Applied Sciences of Thessaly, Greece
Grzegorz Wilk-Jakubowski	University of Social Sciences in Lodz
Swapnarag Swain	IIM Bodhgaya
Sankarshan Basu	IIM Bangalore
Mukesh Kumar Jha	IIM Amritsar
Aswin Alora	IIM Kozhikode
M.P. Sebastian	IIM Kozhikode
Pranith Kumar Roy	IIM Raipur
Dhananjay Bapat	IIM Raipur
Manojit Chattopadhyay	IIM Raipur
Sobhan Sarkar	IIM Ranchi

Satish Kumar	IIM Nagpur
Swamy Perumandla	IIM Nagpur
Dharani Munusamy	IIM Tiruchirappalli
M. S. Gajanand	IIM Tiruchirappalli
Priyesh V. P.	IIM Tiruchirappalli
Pradeep Kumar Dadabada	IIM Shillong
Neelam Rani	IIM Shillong
Shalini Agnihotri	IIM Vishakhapatnam
Mehul Raithatha	IIM Indore
Seshadev Sahoo	IIM Lucknow
Madhumita Chakraborty	IIM Lucknow
Sowmya Subramaniam	IIM Lucknow
Sanjay Dhir	IIT Delhi
M.P. Gupta	IIT Delhi
Manu Kumar Gupta	IIT Roorkee
Niladri Das	IIT Dhanbad
Indrajit Mukherjee	IIT Bombay
Usha Ananthakumar	IIT Bombay
Naresh Chandra Sahu	IIT Bhubaneswar
Madhusmita Dash	IIT Bhubaneswar
Tanmoy Som	IIT BHU
Bharati Mishra	IIIT Bhubaneswar
Shradhanjali Panda	Ravenshaw University, Cuttack
Sharada Prasad Sahoo	Berhampur University, Odisha
Manas Kumar Pal	Birla Global University, Odisha
Rabi Narayan Subudhi	KIIT University Bhubaneswar, Odisha
Sreekumar	RIMS, Rourkela, Odisha
Debendra Kumar Mahalik	Sambalpur University
Manoranjan Parhi	Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, Odisha
Padmalita Routray	Fakir Mohan University, Baleshwar, Odisha
Manoranjan Dash	Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, Odisha
Rudra Kalyan Nayak	VIT-Bhopal, India
Somen Dey	Motilal Nehru National Institute of Technology, Allahabad
Deepak Bansal	Sharda University
Deepti Patnaik	Kalinga University, Raipur

Organizing Committee

Organizing Chair

Taosheng Wang	Hunan International Economics University, China

Organizing Co-Chair

Pramod Kumar Prusty	IIMT, Bhubaneswar, India

Student Symposium Chair

Jagamohan Pattanayak	IIMT, Bhubaneswar, India

Organizing Committee

Sroojani Mohanty	IIMT, Bhubaneswar
Itishree Behera	IIMT, Bhubaneswar
Subhajit Roul	IIMT, Bhubaneswar
Alok Arun	IIMT, Bhubaneswar
Pradeep Kumar Baxla	IIMT, Bhubaneswar
Sankar Kumar Nayak	IIMT, Bhubaneswar
Amit Kumar Bariik	IIMT, Bhubaneswar
Siba Prasad Sahoo	IIMT, Bhubaneswar
Dipak Mishra	IIMT, Bhubaneswar
Dinesh Chandra Das	IIMT, Bhubaneswar
Rakesh Patnaik	IIMT, Bhubaneswar

Technical Program Committee

R. Ramachandran	Department of Commerce Annamalai University, Annamalai Nagar, Tamil Nadu India
Lakshmi Prasad Panda	Department of Humanities, GCE Kalahandi, Bhawanipatna
B.A. Prajapati	Veer Narmada South Gujarat University Surat, Gujarata
Sangya Shrivastava	RaiBussiness School, Raipur (C.G.)
P. Malyadri	Government Degree College, Osmania University, Andhra Pradesh, India

Ravi Sankar	School of Management Asian Institute of Technology, Thailand
Jayanthi Ranjan	Institute of Management Technology Ghaziabad, India
Vidya Sekhri	Institute of Management Studies, Ghaziabad, India
Poonam Garg	Institute of Management Technology Ghaziabad, India
Tabassum Sultana	Matrusri Institute of Post Graduate Studies (MIPGS), Hyderabad, India
Susanta Ku. Paikray	Department of Mathematics, Ravenshaw University, Cuttack, India
Kirti Rajhans	National Institute of Construction Management & Research, Pune, India
N. M. Leepsa	IIT-Kharagpur, West Bengal, India
Pragyan Nanda	IIMT, Bhubaneswar
Soma Mitra	IIMT, Bhubaneswar

Contents

Financial Analytics

About the Editors

Prof. (Dr.) Rangan Gupta is Professor at the Department of Economics, University of Pretoria. He holds a PhD in Economics from the University of Connecticut, USA. He conducts research in the fields of macroeconomics and financial economics. His research has appeared in high-impact factor international journals such as International Journal of Forecasting, International Review of Financial Analysis, International Review of Economics and Finance, Journal of Banking and Finance, Journal of International Money and Finance, Journal of International Financial Markets, Institutions and Money, Journal of Macroeconomics, Economic Modelling, Emerging Markets Review, Journal of Forecasting, Energy Economics, Energy Policy, Energy, Macroeconomic Dynamics, Quantitative Finance, Resources Policy.

Prof. (Dr.) Francesco Bartolucci is Full Professor of Statistics in the Department of Economics of the University of Perugia (IT), where he also coordinated the Doctorate program in "Mathematical and Statistical Methods for Economic and Social Sciences." Among his research interests, hidden Markov models and discrete latent variable models in general have a special role. On these topics he published papers appeared in prestigious journals of Statistics and Econometrics and he coauthored two books with an international publisher. He collaborated with many young researchers even within graduate programs and acted as Reviewer and Associate Editor for important field journals. He also acted as a reviewer within evaluation programs of the Italian University system. He participated in several research programs, even with the role of coordinator. He was Principal Investigator of the research project "Mixture and latent variable models for causal inference and analysis of socio-economic data," which was funded by the Italian Government. He is an Editor of Statistical Modelling: An International Journal.

Prof. (Dr.) Vasilios N. Katsikis is Associate Professor of Mathematics and Informatics at the Department of Economics, National, and Kapodistrian University of Athens, Greece. He received his Diploma of Mathematics from the National and Kapodistrian University of Athens, his MSc in Applied Mathematics as well as his PhD in Mathematics from the National Technical University of Athens. He has also worked as a post-doctoral researcher in Applied and Computational Mathematics under the financial support of the State Scholarship Foundation (IKY). His research interests lie in the fields of applied mathematics, computational finance, artificial intelligence, matrix analysis, and applied linear algebra. He has published more than 80 journal papers in high-quality journals, 4 research books, and 7 book chapters on the above-mentioned fields and has a variety of other research activities. He also serves as a reviewer for several journals and congresses.

Prof. (Dr.) Srikanta Patnaik is Director of IIMT, Bhubaneswar. He has received his Ph. D. (Engineering) on Computational Intelligence from Jadavpur University, India,

in 1999 and supervised 34 Ph. D. theses and more than 60 M. Tech theses in the area of machine intelligence, soft computing applications, and re-engineering. Dr. Patnaik has published more than 100 research papers in international journals and conference proceedings. He is the author of 2 textbooks and edited 82 books and few invited book chapters, published by leading international publishers like Springer-Verlag, Kluwer Academic, etc. He is Editor-in-Chief of International Journal of Information and Communication Technology and International Journal of Computational Vision and Robotics published from Inderscience Publishing House, England, and also Editor-in-Chief of Book Series on "Modeling and Optimization in Science and Technology" published from Springer, Germany.

Financial Analytics

Enterprise Financial Fraud Detection and Audit Optimization Based on Deep Learning

Peng Zhao[1(✉)] and Liquan Deng[2]

[1] Department of Financial Management, Gingko College of Hospitality Management, Chengdu, China
yihms2017@163.com

[2] School of Computer Science, Chengdu College of University of Electronic Science and Technology of China, Chengdu, China

Abstract. Aiming at the problem of enterprise financial fraud detection, this study aims to improve the accuracy and efficiency of detection through deep learning (DL) technology, so as to optimize the audit process and provide intelligent financial management guidance for enterprises. To achieve this goal, an advanced financial fraud detection model based on DL is designed and verified. This research is not only devoted to automatic extraction and in-depth analysis of complex features in enterprise financial data through self-defined DL model architecture, but also enables the model to accurately learn the key patterns of identifying fraud through large-scale and high-dimensional financial data training. Furthermore, the DL model is integrated into the traditional audit process, which realizes the automation and intelligent upgrade of the process. The results show that the proposed DL model has achieved excellent performance improvement in financial fraud detection, and all the evaluation indexes are over 90%, which is obviously superior to the traditional methods. This study not only provides a new technical means to ensure the financial security of enterprises, but also enhances the value of audit work.

Keywords: Deep Learning · Neural Network · Financial Fraud · Audit · Intelligence

1 Introduction

Financial fraud, as a long-standing problem in the field of enterprise management, refers to the purpose of misleading investors, creditors or other stakeholders by means of deliberately misreporting financial information and manipulating financial statements [1]. This behavior not only damages the economic interests of relevant stakeholders, but also seriously undermines the fairness and transparency of the market [2]. In recent years, with the global economic integration and the rapid development of financial markets, financial fraud has become increasingly complex and hidden, which has brought great challenges to traditional auditing and supervision methods [3].

As a branch of artificial intelligence, DL can process and analyze large-scale nonlinear data by simulating the working principle of neural network [4]. In the aspect

R. Gupta et al. (Eds.): CFBA 2024, LAIS 42, pp. 3–14, 2024.
https://doi.org/10.1007/978-3-031-70598-4_1

of financial fraud detection, DL technology shows great application potential. It can automatically extract effective features from massive financial data and build a high-precision fraud detection model, so as to realize rapid and accurate identification of financial fraud [5, 6]. This not only helps to improve the efficiency of internal audit of enterprises, but also provides strong technical support for regulatory authorities and maintains the stability of financial markets [7].

As a key link of enterprise financial management, audit is of great significance for preventing and discovering financial fraud. However, the traditional audit methods often rely on the experience and professional knowledge of auditors, and there are some problems such as strong subjectivity and low efficiency [8]. With the increase of business scale and complexity, traditional audit methods have been difficult to meet the actual needs. Therefore, audit optimization has become an urgent problem to be solved [9]. By introducing DL and other advanced technologies, we can realize the automation and intelligent transformation of audit process, improve audit efficiency and accuracy, and reduce audit cost, thus providing more reliable and efficient audit services for enterprises and regulatory authorities [10]. The purpose of this study is to build a DL-based enterprise financial fraud detection model and explore its application in audit optimization.

2 Theoretical Basis

2.1 Traditional Financial Fraud Detection Methods and the Application of DL in the Financial Field

Traditional financial fraud detection methods mainly rely on manual feature engineering and statistical analysis technology. These methods usually include ratio analysis, trend analysis, Benford's law and so on [11]. Ratio analysis identifies potential fraud by comparing various ratios in financial statements, such as current ratio and quick ratio. Trend analysis focuses on the historical trend of enterprise financial indicators to find abnormal fluctuations. This Ford's law is a statistical law based on the frequency of numbers, which is used to detect the abnormal distribution of numbers in financial statements, so as to identify possible fraud. However, traditional methods have many limitations [12]. First, they are highly dependent on the professional knowledge and experience of auditors and are subjective. Secondly, in the face of increasingly complex financial fraud, traditional methods are often difficult to effectively identify [13]. Finally, the traditional methods are inefficient in dealing with large-scale data and cannot meet the needs of modern enterprise financial management.

Traditional statistical methods usually need a lot of preprocessing work, including feature selection and feature engineering, which is time-consuming and may introduce human bias. In addition, these methods may not perform well in dealing with high-dimensional, nonlinear relationships and complex financial data. Although traditional machine learning techniques have achieved good results in some cases, they usually rely on a large number of labeled data for training, but in the real world, labeled financial data are often scarce. In addition, the generalization ability of these methods may also be limited, especially when faced with new data with different distribution from training data.

The implementation of automation, predictive modeling, and anomaly detection technology not only enhances the precision of detecting financial fraud but also brings significant efficiency gains and process optimization to the auditing process. Utilizing automated tools for processing vast amounts of financial data can substantially reduce the time and cost associated with manual auditing. By establishing precise forecasting models, auditors can more accurately assess the financial status of enterprises and promptly identify potential risks and issues. Furthermore, through the implementation of a well-defined threshold and algorithm, the anomaly detection system can automatically flag data points that deviate from the conventional pattern, thereby guiding auditors to conduct thorough analysis and investigation of these anomalies. In recent years, the utilization of Deep Learning (DL) in the financial sector has garnered increasing attention. Researchers are exploring the use of DL technology to automatically extract financial data characteristics and develop highly accurate fraud detection models.

2.2 DL Theoretical Basis

NN (Neural Network): NN is a computational model that mimics the interconnected structure of neurons in the human brain. It consists of numerous interconnected neurons, each receiving input signals from others and generating output signals based on specific rules. NN learns the mapping between input and output by adjusting the connection weights between neurons [14]. During training, NN employs the backpropagation algorithm to optimize weight parameters, gradually aligning the output signal with the target output.

CNN (Convolutional Neural Network): CNN is a type of NN designed specifically for processing data with grid structures. It extracts local features from input data by stacking convolutional and pooling layers alternately, progressively merging low-level features into high-level ones. CNN has achieved significant success in computer vision and is extensively used in image processing tasks within the financial sector.

RNN (Recurrent Neural Network): RNN is an NN tailored for handling sequential data. It captures temporal dependencies within sequences by incorporating cyclic connections. RNN excels in tasks such as speech recognition and natural language processing, and has been explored for analyzing time series data in finance.

LSTM (Long Short-Term Memory Network): LSTM is a variant of RNN designed to address the vanishing gradient problem encountered when processing lengthy sequences. By introducing gating mechanisms and memory units, LSTM can learn long-term dependencies within sequences, making it suitable for complex sequence modeling tasks.

Transformer: Transformer is an NN model based on the self-attention mechanism. It captures global dependencies within input sequences through multi-head self-attention and positional encoding. Transformer has revolutionized natural language processing and is gradually being applied to text analysis and modeling tasks within the financial domain.

Anomaly detection is an important application field of DL. In the financial field, anomaly detection is usually used to identify abnormal trading behavior, financial statement items, etc., to find potential fraud. DL models such as self-encoder and Generative adversarial network are widely used in anomaly detection tasks. These models build anomaly detection models by learning the distribution characteristics of normal data, and calculate the difference between the distribution of test data and normal data to identify abnormal samples. The successful application of DL in anomaly detection provides new ideas and methods for financial fraud detection.

3 Construction of a Financial Fraud Detection Model for Enterprises

3.1 Dataset Preparation and Processing

The careful selection of a dataset holds paramount importance in constructing an enterprise financial fraud detection model. This study endeavors to gather enterprise financial data from diverse channels, encompassing public financial statements, internal audit reports, market research data, and more. These data sources will encapsulate various facets of the enterprise, including financial status, operational outcomes, cash flow, among others, thereby furnishing comprehensive input characteristics for the model. The selection of the dataset entails primarily considering the following factors: Firstly, ensuring the authenticity and reliability of the data to uphold the credibility and accuracy of the data sources; Secondly, prioritizing the timeliness and representativeness of the data by selecting sources that reflect the recent financial condition of the enterprise; Finally, guaranteeing the integrity and diversity of the dataset to encompass ample samples and features requisite for model training and analysis.

Prior to utilizing the dataset for model training, a series of data preprocessing procedures are imperative. These operations aim at refining the data by purging redundancies, addressing outliers, standardizing feature scales, and extracting pertinent features. Key data preprocessing techniques include:

Data cleansing: Ensuring data integrity and consistency through actions such as eliminating duplicate records, filling in missing values, and handling anomalous data points.

Normalization: Standardizing features onto a uniform scale to prevent certain features from disproportionately influencing model training. Common normalization methods encompass min–max normalization and Z-score normalization.

Feature engineering: Extracting meaningful features conducive to model training through transformations and combinations of the original data. Examples include calculating financial ratios, constructing trend indicators, and so forth.

3.2 DL Model Selection and Design

When choosing DL model, this paper comprehensively considers the following factors: the applicability and performance of the model, and chooses the model that performs well in dealing with similar problems; The complexity and calculation cost of the model

ensure that the selected model can complete training and reasoning in an acceptable time; The interpretability and expansibility of the model, so as to explain and adjust the model in practical application. According to the characteristics and requirements of enterprise financial fraud detection, a custom DL model architecture is designed (as shown in Fig. 1). The architecture will combine the advantages of CNN and RNN to capture the local and global characteristics of financial data.

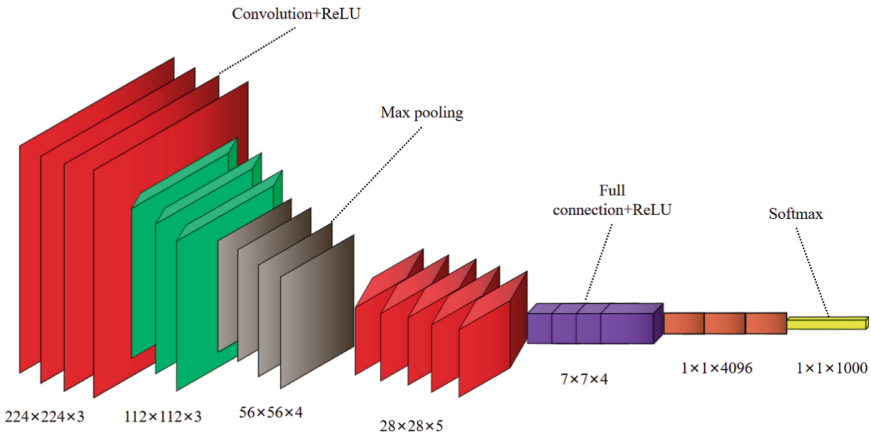

Fig. 1. Model architecture.

It is assumed that in a deep learning setup, the input layer contains c, the output layer contains d, and the hidden layer contains e. The weights connecting the hidden layer to the input layer are denoted as V_{iw}, and those connecting the hidden layer to the output layer are denoted as V_{jw}. Let the activation function for the hidden layer be f_1 and for the output layer be f_2. Specifically, the output of a node in the hidden layer is expressed as follows:

$$Z_k = f_1\left(\sum_{i=0}^{n} V_{ik}X_i\right), k = 1, 2, \cdots, q \tag{1}$$

For P sample size, the overall error value is:

$$E = \frac{1}{2}\sum_{P=1}^{P}\sum_{j=1}^{m}\left(t_j^p - y_j^p\right)^2 \tag{2}$$

The function of enterprise financial fraud detection model is:

$$y_j = f(x) = w \times \varphi(x_i) + b,\ \ i = 1, 2, \cdots; j = 1, 2, \cdots \tag{3}$$

Among them, $\varphi(x_i)$ is a nonlinear mapping from the input space to the high-dimensional feature space, y_j has two attributes, one is ST and the other is non-ST, and x_i represents the value of the i index.

Among all linear methods, the Mahalanobis distance measurement method is more obvious, which has the invariance of scale transformation. Let $y = Ax$, then the vector x_1, x_2, m_x, the transformed distance between them is y_1, y_2, m_y. Mahalanobis distance is:

$$\|x - m\|_M = (x - m)^T C^{-1}(x - m) \tag{4}$$

In the model, Gaussian radial basis function is used as the kernel function, and thus the enterprise financial fraud detection model based on radial basis function considering big data as input variables is obtained:

$$y = \sum_{i=1}^{n} \left(\overline{a}_i - \overline{a}_i^*\right) \exp\left[-\frac{|x_i - x|^2}{2\sigma^2}\right] + y_i \times \sum_{j=1}^{n} \left(\overline{a}_j - \overline{a}_j^*\right) \exp\left[-\frac{|x_i - x|^2}{2\sigma^2}\right] \pm \varepsilon \tag{5}$$

where ε represents the allowable error value. In this paper, the key to improve the algorithm is the setting of ε critical point. After preliminary exploration, ε can be set equal to the average value of $|g(x)|$ in all test samples minus twice the standard deviation, that is, the formula:

$$\varepsilon = \left|\overline{g(x)}\right| - 2\sigma \tag{6}$$

At the same time, attention mechanism is introduced to make the model pay attention to the features that are more important for fraud detection. Through the custom model architecture design, it is expected to improve the model's ability to identify complex fraud patterns.

3.3 Model Training and Optimization

During the model training phase, we employ effective strategies to enhance the model's performance. Initially, we carefully select a suitable loss function, such as the cross-entropy loss function, to quantify the discrepancy between the model's predictions and the actual labels. Subsequently, we utilize the RMSprop algorithm to iteratively update the model parameters and minimize the loss function. Moreover, we incorporate techniques like early stopping and learning rate decay to prevent overfitting and bolster the model's ability to generalize.

Hyperparameter tuning plays a pivotal role in the training of deep learning models. In this study, the optimal combination of hyperparameters is determined through experiments and grid search, encompassing variables such as learning rate, batch size, training epochs, among others. These parameters directly impact the model's training efficiency and performance. By fine-tuning the hyperparameters, our objective is to pinpoint a configuration that maximizes the model's performance on the validation dataset. Subsequently, a suite of appropriate evaluation metrics will be selected to assess the model's performance. Through a comparative analysis of different models based on these evaluation metrics, we identify the best-performing model for practical applications.