Research on the Application of Artificial Intelligence and Big Data Technology in Financial Fraud Detection

Ziyue Wang

Independent Researcher, New York, NY 10012, United States

Abstract: *Financial fraud is a hidden criminal activity that has a serious impact on the stability of financial markets and investor confidence. Traditional fraud detection methods are often inefficient and unable to meet the rapidly changing fraudulent means. The development of artificial intelligence and big data technology has provided new solutions for financial fraud detection. This article aims to explore the current application status and methods of artificial intelligence and big data technology in financial fraud detection, and analyze the challenges and future development directions.*

Keywords: Financial fraud detection; AI Plus Big data; application.

1. INTRODUCTION

With the continuous development of financial markets and technological advancements, financial fraud has shown diverse and covert characteristics. Traditional manual detection methods can no longer meet the needs of real-time, efficient, and accurate fraud detection[1-3]. Traditional financial fraud detection methods mainly rely on rule engines and manual auditing. The limitations of this method are that it cannot handle massive amounts of financial data, cannot detect complex fraud patterns, and is susceptible to fraudulent behavior. Therefore, financial institutions urgently need a more intelligent and efficient fraud detection solution to improve the accuracy and efficiency of fraud detection, and reduce the risks and losses caused by fraud.

2. The Principles of Artificial Intelligence and Big Data Technology in Detecting Financial Fraud

On the one hand, big data technology provides a data foundation for artificial intelligence. The financial field has generated a massive amount of data, including transaction data, user behavior data, market data, etc. Big data technology can efficiently store, manage, and process this data, providing abundant data resources for artificial intelligence[4].

On the other hand, artificial intelligence technology learns patterns of fraudulent behavior from massive financial data through methods such as machine learning and deep learning. In machine learning, algorithms can construct models and train them based on the data characteristics of known fraud cases and non fraud cases, thereby learning the characteristics and patterns of fraudulent behavior[5]. In deep learning, by constructing a multi-layer neural network model, the system can automatically extract and learn advanced features of data, thereby achieving more complex and refined fraud detection. In addition, artificial intelligence technology utilizes real-time data flow processing technology, allowing the system to monitor financial transactions and user behavior in real time, and provide timely warnings for abnormal behavior. By combining the real-time data processing capabilities of big data technology and the intelligent analysis capabilities of artificial intelligence technology, the efficiency and accuracy of fraud detection can be effectively improved. At the same time, financial fraud often involves multiple data sources and dimensions of information. Based on the correlation analysis ability of artificial intelligence technology, systems can mine the correlation and correlation between different data sources and discover potential fraud networks and patterns[6].

3. Application of Artificial Intelligence Technology in Financial Fraud Detection

3.1 Requirements Analysis

The application of artificial intelligence technology in financial fraud detection begins with the analysis of various needs. This includes an understanding of the complexity of financial markets and the constantly changing methods of fraud, as well as an awareness of the low efficiency of traditional fraud detection methods. Meanwhile, for financial institutions and users, timely detection and prevention of fraudulent behavior is crucial to maintain the stability of the financial market and the rights and interests of users. Therefore, requirements analysis focuses on high efficiency, accuracy, and real-time performance, and requires the system to be able to process large-scale data and identify diverse patterns of fraudulent behavior. In this situation, the application of artificial intelligence technology has become an indispensable choice. It can learn and discover potential fraud patterns from massive data through methods such as machine learning and deep learning, thus achieving an intelligent and efficient financial fraud detection system.

3.2 Feature Engineering

In financial fraud detection, feature engineering involves in-depth analysis of financial transaction data and user behavior data, extracting meaningful features for fraud detection. The collected relevant features can include various information such as transaction amount, transaction frequency, transaction location, transaction time, and user historical behavior patterns. Through feature engineering, raw data can be transformed into feature representations that can be understood and utilized by the model, thereby improving the performance and effectiveness of fraud detection models. The goal of feature engineering is to select representative, relevant, and discriminative features, while considering the correlation and interaction effects between features, in order to maximize the predictive ability of the model. In practical applications, feature engineering needs to comprehensively consider business requirements, data characteristics, and model algorithms. Through in-depth understanding and mining of data, an efficient feature set suitable for financial fraud detection is constructed, providing a reliable foundation for subsequent modeling and prediction.

3.3 Model construction

In financial fraud detection, model construction is the core link of artificial intelligence technology application. Through model construction, machine learning and deep learning technologies are used to learn and discover potential fraud patterns from massive financial data, in order to achieve an intelligent and efficient fraud detection system. In the process of building a model, it is necessary to select appropriate algorithms and model structures, such as logistic regression, decision tree, support vector machine, neural network, etc., and train and optimize the model to improve its generalization ability and accuracy. At the same time, it is necessary to consider the results of feature engineering and input the extracted meaningful features into the model for training. During the model construction process, it is necessary to continuously evaluate the performance and effectiveness of the model, and adjust and optimize it based on actual needs and feedback information to ensure that the model can accurately identify and predict fraudulent behavior, safeguarding the stability of the financial market and the rights and interests of users.

3.4 Monitoring and Early Warning

Through monitoring and early warning systems, real-time data flow processing technology is used to continuously monitor financial transaction data and user behavior, timely detect abnormal transaction behavior, and implement early warning measures to prevent fraud risks. The corresponding monitoring and early warning systems combine machine learning and deep learning technologies, which can identify abnormal patterns and patterns in the data flow in real time, and provide early warning for possible fraudulent behavior. Once the system detects suspicious transactions or abnormal behavior, it will immediately trigger an alert mechanism, send alert notifications to relevant personnel, so that they can take timely measures to investigate and handle. The implementation of monitoring and early warning systems can not only help financial institutions detect and prevent fraudulent behavior in a timely manner, protect customer asset security, but also improve the transparency and stability of the financial market, enhance market confidence, and provide strong guarantees for the sustainable development of the financial industry.

4. The Application of Big Data Technology in Financial Fraud Detection

4.1 Distributed Data Storage

In financial fraud detection, the application of big data technology is crucial, and distributed data storage is an important component. The financial field has generated a large amount of transaction data, user behavior data, etc. These data need to be efficiently stored and managed to support subsequent analysis and mining work. Distributed data storage technology can meet the storage needs of the financial industry for massive amounts of data, and has the characteristics of high reliability, high scalability, and high performance[7].

Firstly, data distributed storage technology can fully utilize the storage space of each server in the cluster by storing data in a distributed manner on multiple servers, thus solving the problem of limited storage capacity on a single server. This storage architecture can support the storage of large-scale data and meet the storage needs of financial institutions for massive amounts of data. Secondly, distributed data storage technology has high reliability. In distributed storage systems, data is usually replicated in multiple copies and stored on different nodes. When a node fails, the system can still obtain data from other nodes, ensuring the reliability and availability of data, which is crucial for the financial industry. In addition, data distributed storage technology also has high scalability. With the continuous growth of financial data, distributed storage systems can dynamically expand storage capacity according to demand without the need for downtime or data migration, ensuring the continuous operation and stable performance of the system.

4.2 Cloud storage resource invocation

In financial fraud detection, the application of big data technology is not limited to distributed storage of data, but also includes the invocation of cloud storage resources. Cloud storage resources refer to data storage services stored on cloud computing platforms, such as Amazon S3, Google Cloud Storage, Microsoft Azure Blob Storage, etc. The related cloud storage services provide highly reliable, scalable, and flexible storage solutions, providing convenient data storage and management methods for financial institutions[8].

On the one hand, the invocation of cloud storage resources enables financial institutions to store data in the cloud without the need to purchase and maintain expensive hardware equipment, reducing storage costs. Financial institutions can dynamically adjust storage capacity according to actual needs and flexibly expand according to the growth of data volume, avoiding resource waste or insufficient problems caused by changes in storage demand. On the other hand, cloud computing service providers typically adopt multi copy backup and data redundancy technologies to ensure the security and reliability of data. In addition, cloud storage services typically have high availability SLAs (Service Level Agreements), ensuring that financial institutions can access and call data stored in the cloud anytime and anywhere, ensuring the continuous operation of the business. In addition, the invocation of cloud storage resources also enables financial institutions to utilize the elastic and distributed computing capabilities provided by cloud computing platforms for large-scale data processing and analysis.

5.　Conclusion

In summary, from distributed data storage to cloud storage resource utilization, big data technology provides financial institutions with powerful data storage, management, and processing capabilities, while artificial intelligence technology can achieve in-depth analysis and data mining. Through distributed storage technology, financial institutions can effectively process massive amounts of data, achieve high reliability and scalability of data, and call cloud storage resources to further improve the data storage efficiency and flexibility of financial institutions. At the same time, it endows them with elastic and distributed computing capabilities, thereby strengthening the monitoring and prevention of fraudulent behavior. The called data is identified and analyzed based on artificial intelligence models, accurately and efficiently identifying financial fraud behavior, providing data-driven intelligent solutions for financial institutions.

REFERENCES

[1]　Li, J., Ye, Z., & Zhang, C. (2022). Study on the interaction between big data and artificial intelligence. Systems Research and Behavioral Science, 39(3), 641-648.

[2]　Bao, Y., Hilary, G., & Ke, B. (2022). Artificial intelligence and fraud detection. Innovative Technology at the Interface of Finance and Operations: Volume I, 223-247.

[3]　Ghimire, A., Thapa, S., Jha, A. K., Adhikari, S., & Kumar, A. (2020, October). Accelerating business growth with big data and artificial intelligence. In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 441-448). IEEE.

[4]   Choi, D., & Lee, K. (2018). An artificial intelligence approach to financial fraud detection under IoT environment: A survey and implementation. Security and Communication Networks, 2018.

[5]   Zhou, H., Sun, G., Fu, S., Fan, X., Jiang, W., Hu, S., & Li, L. (2020). A distributed approach of big data mining for financial fraud detection in a supply chain. Comput Mater Continua, 64(2), 1091-1105.

[6]   Zhou, H., Sun, G., Fu, S., Wang, L., Hu, J., & Gao, Y. (2021). Internet financial fraud detection based on a distributed big data approach with node2vec. IEEE Access, 9, 43378-43386.

[7]   Yu, T. R., & Song, X. (2021). Big data and artificial intelligence in the banking industry. In Handbook of financial econometrics, mathematics, statistics, and machine learning (pp. 4025-4041).

[8]   Ozili, P. K. (2021). Big data and artificial intelligence for financial inclusion: benefits and issues. In Artificial Intelligence, Fintech, and Financial Inclusion (pp. 1-10). CRC Press.