J. COMBIN. MATH. COMBIN. COMPUT. 123 (2024) 343–355

Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Deep Learning Model Based Research on Anomaly Detection and Financial Fraud Identification in Corporate Financial Reporting Statements

Wenjuan Li[1,✉], Xinghua Liu[2], Shiyue Zhou[1]

[1] *Management Science and Engineering School of Shandong University of Finance and Economics, Jinan, Shandong, 250000, China*

[2] *Suffolk County, New York, 11790, USA*

ABSTRACT

Financial frauds, often executed through asset transfers and profit inflation, aim to reduce taxes and secure credits. To enhance the accuracy and efficiency of accounting data auditing, this study proposes an anomaly detection scheme based on a deep autoencoder neural network. Financial statement entries are extracted from the accounting information system, and global and local anomaly features are defined based on the attribute values of normal and fraudulent accounts, corresponding to individual and combined anomaly attribute values. The AE network is trained to identify anomalies using account attribute scores. Results demonstrate classification accuracies of 91.7%, 90.3%, and 90.9% for sample ratios of 8:2, 7:3, and 6:4, respectively. The precision, recall, and F1 score reach 90.85%, 90.77%, and 90.81%, respectively. Training takes 95.81ms, with recognition classification requiring only 0.02ms. The proposed deep neural network achieves high recognition accuracy and speed, significantly improving the detection of financial statement anomalies and fraud.

Keywords: Financial fraud, Accounting data auditing, Financial statement, Deep neural network, Financial statement anomaly

1. Introduction

Financial statements are accounting statements reflecting the status of funds and profits of an enterprise or budget unit for a certain period of time. The state requires enterprises to prepare and

✉ Corresponding author.
E-mail addresses: wenjuanli_sdufe@163.com (Wenjuan Li).

Received 10 June 2024; accepted 20 August 2024; published 31 December 2024.

DOI: 10.61091/jcmcc123-24

report uniformly in strict accordance with the requirements of the Accounting Standards and related accounting systems, and disclose them to the public for stakeholders' decision-making reference [9]. At present, regardless of the size of the enterprise, whether private or state-owned enterprises, the different interests of the starting point of the triggers have led to the existence of financial statement fraud phenomenon more or less. Financial statement fraud not only harms the legitimate interests of investors and creditors, but also seriously affects the national macro-control of economic operations and the stable and sustainable development of the national economy [3]. Need to strengthen the supervision of financial statement fraud phenomenon, standardization, to avoid expanding the capital market disruption damage to the national interest, undermine the professional integrity of the financial industry [5].

In recent years, relevant tax units began to focus on corporate financial statements, and a large number of scholars began to research on financial statement anomaly detection.Soltani, M et al. constructed a fuzzy neural network for financial forgery detection, and found that its performance is better than ordinary statistical models and ordinary artificial neural networks [19]. Aftabi, S. Z conducted financial forgery identification for 76 Greek listed companies and research and found that Bayesian network outperforms neural network and ID3 decision tree and financial distress, leverage level, profitability, sales performance, and solvency are significantly correlated with financial fraud [1]. Ashtiani, M. N et al. used multi-layer feed-forward neural network, support vector machine, genetic programming, data processing grouping method, logistic regression, and probabilistic neural network on the mining 35 features of 101 counterfeiting companies listed on the stock exchange with 101 non-counterfeiting companies. Using t-test for feature selection, it was found that genetic programming and probabilistic neural network performed best in the feature-selected dataset, and probabilistic neural network performed best in the full dataset [2]. Schultz, M et al. used a sequential combination of supervised and unsupervised learning methods to firstly classify the companies into counterfeiting, non-counterfeiting, and suspicious using K-means. Then supervised learning method was used to predict the fraudulent company financial statements and it was verified that multilayer feed forward neural network performed the best [16]. Shahana, T used an artificial bee colony algorithm to extract a suitable subset of features and used multiple classifiers to detect fraudulent behaviors in 83 Egyptian companies [17]. Al-Hashedi, K. G constructed two sets of data, Group A for M-score model with 8 ratio indicators, and Group B is the difference of year-to-year changes in the indicators of the initial input variables of the M-score. Logistic regression and linear discriminant analysis were also utilized to find the determinants of unintentional accounting errors leading to financial restatements and to construct a detection model, and it was found that the model developed for Group B data was more accurate and the logistic regression model outperformed the linear discriminant model [7]. Pinto, S. O used a combination of random forest, gradient ascent tree, neural network, and self-coding network methods with benchmarking methods, stepwise regression logistic comparison and found that gradient ascent, neural network, and random forest perform better than the other two for internal abuse and fraud prediction in Bank of America [14].

Financial statements are an important basis for accounting information users to understand the reality of enterprises and make decisions. While improving the auditors' own technical level and professional ethics, big data analytics should also be utilized to empower auditing and improve the ability to quickly detect financial statement fraud. The purpose of this paper is to study the detection of abnormalities in corporate financial reports and the identification of financial fraud, and to bring in deep neural network methods in the research process to promote the rapid identification and classification of abnormal reports and financial fraud. Firstly, it describes the common means

of financial fraud and the motivation of financial statement fraud, and then selects samples and analyzes the characteristic indexes of financial statement fraud, so as to set variable indexes. Finally, if the anomaly score AS exceeds the threshold, the financial statement information is labeled as anomalous. For the future enterprise financial statement anomaly detection and financial fraud identification, provide guidance program.

2. Financial Statement Anomaly Detection

2.0.1. Common means of financial fraud. Routinely, financial fraud manifests itself in the following forms:

(a) Adjustment of profits using inappropriate related transactions, which refers to the direct transfer of resources, labor, or obligations between related parties, regardless of whether or not payments are received. Its main ways are shifting the burden of expenses, transferring assets, leasing of related assets, asset replacement, and non-fair purchase and sale of business [23].

(b) False recognition of revenues, liabilities and expenses mainly includes early recognition of revenues, fictitious sales activities, recognition of revenues when there is uncertainty in assets, fraudulent use of accounts payable and production costs, excessive capitalization of expenses, non-correspondence of debts and liabilities, incomplete and illegal formalities, and failure to recognize them in a timely manner.

(c) One of the forms of false recognition of assets is the use of asset evaluation to eliminate latent losses, in accordance with the provisions of accounting standards and the requirements of the principle of prudence. The potential losses of enterprises should be reflected in the income statement in accordance with legal procedures. However, in many enterprises, especially state-owned enterprises, often through asset evaluation, bad debts, long-term investment losses, fixed asset losses, slow-moving and damaged inventory, and deferred assets and other latent losses are recognized as assessment of impairment to offset capital surplus, so as to achieve the purpose of inflated profits [4].

Due to the diversity of economic activities, many companies inflate their income by fictionalizing business credits in order to regulate their profits. Enterprises may also not account for amounts that are recouped but have been treated as bad debt losses, or write off amounts that can be recovered.

(d) Measurement is based on historical cost, but assets can be impaired and require estimation by accountants. Estimates are subject to error, and accountants can use estimates to commit fraud [13].

2.0.2. Motives for financial statement fraud. Understanding the motives for financial statement fraud helps to identify anomalies and fraud in a more timely and accurate manner, and the common motives are as follows:

(a) Performance assessment indicators mainly include sales revenue, production value, return on investment, sales margin, asset turnover, sales revenue, total profit and total assets, etc., the calculation of which will use the data in the accounting statements.

(b) In China, many enterprises may make creditors extend credit to them through accounting falsification when they are short of funds and cannot meet the credit conditions [6].

(c) In order to meet the conditions for issuing shares and to issue shares at a higher price, companies often package their accounting statements when designing share reform programs. Many companies also whitewash their accounting statements when they increase their share allotments.

(d) Reduce tax incentives, income tax amount is the accounting profit adjusted to taxable income and then multiplied by the income tax rate. In order to achieve the purpose of tax leakage, tax evasion, reduction or postponement of taxes, companies often whitewash their financial statements.

(e) In recent years, creating good business performance of state-owned enterprises has become a political task for enterprise leaders. For the leaders of enterprises, accomplishing this task well may mean a bright future, or else the future may be dark. Some state-owned companies are likely to whitewash their accounting statements under political pressure [21, 22].

(f) The share price and borrowing capacity of an enterprise are largely determined by whether the profit level of the enterprise is stable or not, so in order to make the operating performance of the enterprise look stable, the management of the enterprise will whitewash the accounting statements.

2.1. Anomaly detection methods

2.1.1. Sample selection and data sources. Focusing on a shorter time period and companies in the same market can provide a higher quality model for detecting financial statement fraud. Therefore, considering the comparability between the data, this paper selects the A-share pharmaceutical manufacturing companies listed in China's Shenzhen and Shanghai markets from 2016 to 2020 as the research object, and the remaining valid samples after eliminating the samples with missing values total 1,081 items [12, 20]. The information of listed companies' violation and punishment and related financial data are from Rexis database, and the violation and punishment information is from Cathay Pacific database [11].

2.1.2. Definition of variables. According to the characteristic indicators that characterize financial statement fraud, Table 1 shows the definition of variables. In this paper, a total of eight fields, including inventory turnover X_1, accounts receivable turnover X_2, total asset turnover X_3, working capital ratio X_4, gearing ratio X_5, return on net assets X_6, net interest rate X_7, and Z-value X_8, are selected as variables for anomaly detection [15].

2.1.3. Model construction. This paper adopts unsupervised learning method to construct the abnormal detection model of listed company's financial statement fraud, based on multi-dimensional spatial clustering and calculating the distance between sample points and data clusters, and realizes the diagnosis of outliers and its cause analysis based on the judgment of distance. The analysis process includes three steps: clustering, anomaly measurement and anomaly diagnosis:

(a) Based on the variables characterizing financial statement fraud defined in the previous section, a two-step clustering method is used to cluster the sample of 1109 public companies into several classes and find the center of each class [18]. For all data, the mean and variance of K^A numeric clustering variable k are computed $\hat{\sigma}_k^2$. For each category, the sample size of class v is computed N_v, the mean and variance of K^A numeric clustering variable k is computed $\hat{\sigma}_{vk}^2$, and the sample size of K^B subtyped clustering variable lth category is computed N_{vkl}.

Type of variable	Variable name	Symbolic	Explanation of variables
Operational capacity	Accounts Receivable Turnover Ratio	X_1	Operating Income/Accounts Receivable Balance
	Inventory turnover	X_2	Operating Costs/Inventory Balance
	Total Assets Turnover	X_3	Operating Cost/Total Assets
Asset liquidity	Working Capital Ratio	X_4	(Current Assets - Current Liabilities) / Total Assets
Solvency	Gearing ratio	X_5	Total liabilities/total assets
Profitability	Return on net assets	X_6	Net Profit/Net Assets
	Net interest rate	X_7	Net Profit / Operating Income
Insolvency index	Z-score	X_8	Z = 1.2 x net working capital/total assets + 1.4 x total retained earnings assets + 3.3 x EBITDA/total assets + 0.6 x (market value of preferred and common stock)/total liabilities + 0.99 x sales/total assets

Table 1. Definition of variables

(b) On the basis of clustering, calculate the anomaly measures for all listed company samples. Find the class v to which the sample S points belong and calculate the log-likelihood distance between the sample points S and the class v called the group difference index. This index is denoted by GDI_S and its calculation formula can be expressed as:

$$GDI_S = d(v,s) = \varepsilon_v + \varepsilon_s - \varepsilon_{\langle v,s\rangle} = \varepsilon_v - \varepsilon_{\langle v,s\rangle}. \tag{1}$$

In the formula, $d(v,s)$ denotes the log-likelihood distance between sample point S and category v, ε_v, ε_s, and $\varepsilon_{\langle v,s\rangle}$ denote the log-likelihood distance of category v, the log-likelihood distance of sample S, and the log-likelihood distance of samples v to S, respectively. The formula for the log-likelihood distance for category v can be expressed as follows:

$$\varepsilon_v = -N_v \left(\sum_{k=1}^{K^A} \frac{1}{2} ln\left(\hat{\sigma}_k^2 + \hat{\sigma}_{vk}^2\right) + \sum_{k=1}^{K^B} \hat{E}_{vk} \right), \tag{2}$$

where $\hat{E}_{vk}$ denotes the rate of difference between the sample size of the calculated lnd category and the sample size of category v, so that this rate of difference can be expressed by the formula:

$$\hat{E}_{vk} = -\sum_{N_v}^{L_k} \frac{N_{vkl}}{N_v} ln\left(\frac{N_{vkl}}{N_v}\right), \tag{3}$$

where S is a sample point whose internal variation is 0, i.e., $\varepsilon_s = 0$. The group variance index reflects the incremental amount of variation within class v caused by the addition of sample point S to class v.

Calculate the variable difference index for the clustered variable k, which is denoted by VDI. For numerical clustering variables, the variable difference index is defined as:

$$VDI_k = \frac{1}{2} ln\left(\hat{\sigma}_k^2 + \hat{\sigma}_{vk}^2\right). \tag{4}$$

For taxonomic clustering variables, the variable difference index is defined as:

$$VDI_k = -\sum_{l=1}^{L_k} \frac{N_{vkl}}{N_v} ln\left(\frac{N_{vkl}}{N_v}\right). \tag{5}$$

The Variable Difference Index is the difference between the individual additive parts of the sample point S after joining class v and before entering, GDI and reflects the magnitude of the contribution of each clustered variable in the incremental amount of difference within class v caused by the sample point S joining class v.

Next, the anomaly index is calculated. For sample point S, the calculation of the anomaly index can be expressed as AI, which is calculated as:

$$AI = \frac{\frac{GDI_S}{1}}{N \sum_{i=1}^{N_v} GDI_i}. \tag{6}$$

The outlier index is a relative indicator of the ratio of the within-class variance caused by sample point S to the mean of the variances caused by other sample points within class v. The greater the value, the greater the certainty that sample point S is considered an outlier [8].

Samples of all listed companies are ranked according to the degree of abnormality to identify possible outliers and analyze the causes of financial anomalies. The anomaly index AI is sorted in descending order, and the listed companies in the first m positions are outliers, which may have financial statement fraud, and the anomaly index in the m position is the criterion for determining the outliers, and those larger than this value are outliers, and those smaller than this value are non-outliers. Meanwhile, for the outliers, the variable difference index VDI is ranked in descending order, and the variables in the first l positions are the main reasons for the possible financial anomalies of the listed company.

3. Deep Neural Network-Based Financial Statement Anomaly Detection

3.1. Deep self-encoder neural networks

AE neural networks are specific types of forward multilayer neural networks that can be trained to reconstruct inputs. The difference between the original input and the reconstruction result is defined as the reconstruction error. Figure 1 shows the system framework of an AE neural network, which usually consists of two nonlinear mappings called encoder network l and decoder network g_θ. A symmetric encoder-decoder structure is usually used, which contains multilayer neurons, nonlinear functions, and a shared parameter θ. The system framework of an AE neural network is shown in Figure 1.

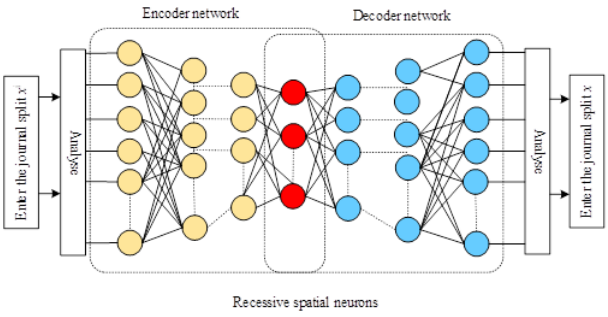

Fig. 1. System framework of AE neural network

The encoder $f_\theta(\cdot)$ maps the input vector X^i to a compressed representation Z in the implicit space Z, and thereafter, the decoder $g_\theta(\cdot)$ maps the implicit representation Z back to the reconstructed

vector $\hat{x}^i$ in the original input space, and the nonlinear encoder mapping of an AE containing multiple layers of neurons can be defined as:

$$f_\theta^l(\cdot) = \sigma^l\left(W^l\left(f_\theta^{l-1}(\cdot)\right) + b^l\right). \tag{7}$$

The decoder mapping can be defined as:

$$g_\theta^l(\cdot) = \sigma'^l\left(W'^l\left(f_\theta^{l-1}(\cdot)\right) + d^l\right), \tag{8}$$

where σ and σ' are nonlinear activation functions, θ denotes the model parameters $\{W, b, W', d\}$, $W \in R^{d_x \times d_z}$ and $W' \in R^{d_y \times d_z}$ are weight matrices, $b \in R^{d_z}$ and $b' \in R^{d_y}$ are bias vectors, and l is the number of hidden layers. In this paper, we train the self-encoder neural network using a collection of financial statement information entries $x^i = \{x_1^i, x_2^i, ..., x_k^i\}$, each entry X^i contains an array of K attributes $x^i = \{x_1^i, x_2^i, ..., x_k^i\}$, x_j^i is the jth attribute of the ith entry. Attribute x^j contains account-specific details of the entry, such as bookkeeping type, bookkeeping time, amount, general ledger, etc. [10]. In addition, the order n_j^i represents the value X_j of a particular attribute, e.g., the number of occurrences of a particular voucher type.

Train the AE network to learn the optimal set of model parameters for the encoder-decoder θ^* that minimizes the difference between a given entry X^i and its reconstruction result $\hat{x}^i = g_\theta\left(f_\theta\left(x^i\right)\right)$ to achieve $\hat{x}^i = x^i$. Thus, the optimization objective of AE training for all bookkeeping entries is:

$$\arg min \left\| X - g_\theta\left(f_\theta\left(X^i\right)\right)\right\|. \tag{9}$$

In network training, the cross-entropy loss is utilized and the loss function L_θ is defined as:

$$L_\theta\left(\hat{x}^i; x^i\right) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k} ln\left(x_j^i\right) + \left(1 - x_j^i\right) ln\left(1 - x_j^i\right), \tag{10}$$

where X^i, $i = 1, ..., n$ are the set of n accounting entries and $\hat{x}^i$ is the reconstruction result on all accounting entry attributes $j = 1, ..., k$. In this paper, x and $\hat{x}$ are treated as two independent multivariate Bernoulli distributions using binary encoded attribute values, and the deviation between them is measured by $L_\theta\left(x^i; \hat{x}^i\right)$.

To prevent overfitting, the number of neurons in the hidden layer of the network is limited such that $R^{d_x} R^{d_z}$, i.e., a bottleneck structure is used. By imposing this constraint in the hidden layer of the network, the AE learns the optimal set of parameters θ^* and obtains the most dominant compression model of the distribution of binned attribute values and their dependencies.

3.2. Accounting data anomaly scoring

Based on the above analysis, a new anomaly scoring mechanism is proposed to check global and local anomalies in the actual accounting dataset. The proposed scoring mechanism takes into account two observational properties:

(a) Any uncommon attribute value record that corresponds to a global anomaly.

(b) Any uncommon attribute value co-occurring record corresponds to a localized anomaly. Characterizing the occurrence of unusual attribute values, in order to mine uncommon attribute values from observations, the probability of occurrence of each value X_j in all ledgers is first defined as n_j^i/N and N as the total number of financial statement information journal entries, e.g., the probability that a particular bookkeeping code will be entered in X. In addition, the

sum of log probabilities of individual attribute values on all j attributes is computed for each ledger X^i:

$$P\left(x^i\right) \sum_{j=1}^{k} ln\left(1+\frac{n_j^i}{N}\right). \tag{11}$$

Finally, the normalized probability score AP is calculated for the attribute values:

$$AP\left(x^i\right) = \frac{P\left(x^i\right) - P_{min}}{P_{max} - P_{min}}. \tag{12}$$

Anomalous attribute-value combination occurrence characterization is to mine the attribute-value combinations of anomalous occurrences in observations and locate local anomalies, such as the probability of combining a specific general ledger account with a specific type of posting within all financial statement information journal entries X, and derive the reconstruction error of the financial statement information journal entries through the training of deep AE neural networks. Anomalous co-occurrences are difficult to detect, and therefore difficult to reconstruct effectively in low-dimensional implicit representations, and can result in high reconstruction errors. For this reason, the reconstruction error E of the trained AE network under the optimal model parameters θ^* is derived for the financial statement information entries X^i and the corresponding reconstruction results $\hat{x}^i$:

$$E_\theta\left(x^i;\hat{x}^i\right) = \frac{1}{k}\sum_{j=1}^{k}\left(x_j^i - x_j^i\right)^2. \tag{13}$$

Finally, the normalized reconstruction error RE is computed as:

$$RE_\theta\left(x^i;\hat{x}^i\right) = \frac{E_\theta\left(x^i;\hat{x}^i\right) - E_{\theta^*,min}}{E_{\theta^*,max} - E_{\theta^*,min}}, \tag{14}$$

where E_{min} and E_{min} denote the minimum and maximum values of the reconstruction error derived through E_{θ^*}, respectively.

Accounting data anomaly scoring is based on attribute value occurrence characteristics and combination occurrence characteristics in individual entries, which can infer whether the entry is anomalous or not, and whether the creation of the entry originates from normal business activities. To detect global and local accounting anomaly data in the practice statistics task, for each entry X^i, a score AS is computed X^i based on normalized attribute probabilities AP and reconstruction errors RE using the optimal model parameter θ^*:

$$AS\left(x^i;\hat{x}^i\right) = \alpha \times RE_{\theta^*}\left(x^i;x^i\right) + (1-\alpha) \times AP\left(x^i\right), \tag{15}$$

where α is a balancing factor to balance the occurrence characteristics and combination of occurrence characteristics of attribute values. Based on the anomaly scoring, a threshold parameter β is defined, for each entry X^i under the optimal model parameter θ^*. If the anomaly scoring AS exceeds the threshold β, the financial statement information entry is marked as anomalous.

3.3. Evaluation indicators

The evaluation metrics used in this study are accuracy, precision, recall, and F1 value formulas respectively:

$$precision = \frac{TP}{TP+FP}, \tag{16}$$

$$recall=\frac{TP}{TP+FN}, \tag{17}$$

$$F1=\frac{2\cdot precision\cdot recall}{precision+recall}, \tag{18}$$

where T is the number of correct predictions, N is the total number, TP is the number of positive predictions, FP is the number of negative predictions, and FN is the number of another negative prediction.

4. Data Acquisition

The classification experiments on data-driven and knowledge-driven models are conducted to compare the effectiveness of financial forgery recognition. Among them, the data-driven financial forgery recognition model is a typical binary classification problem, including the benchmark model represented by support vector machine, logistic regression and decision tree, and the cost-sensitive lightweight gradient boosting tree model incorporating cost-sensitive learning. Learning prediction is performed for a dataset consisting of 2,842 Chinese companies in 2015, and the ratio of normal samples to counterfeiting samples based on the financial counterfeiting labels on December 31, 2022 is 2,815:27. The training and testing sets are divided by 4:1 while maintaining the class distribution ratio, i.e., learning training is performed on 2,273 samples, of which 22 are counterfeiting companies, and 569 companies are for financial counterfeiting identification, of which 5 counterfeiting companies.

In order to select the most appropriate proportion of data set division and related algorithms, the different proportion of training and test sets are verified, and 60%-80% of the samples in the data set are usually selected as the training set in practical applications. , and logistic regression, support vector machine, decision tree, and deep neural network are used for classification prediction, and 70% of the samples in the relevant studies and others are selected as the training set.

5. Anomaly Detection and Fraud Identification

5.1. Analysis of classification accuracy

In this paper, on the other hand, the three ratios of 8:2, 7:3, and 6:4 were selected for classification using logistic regression, support vector machine, decision tree, and deep neural network with default parameters, and the results of classification accuracy are shown in Table 2. The decision tree model has the worst classification effect, which classifies the samples with the three ratios of 8:2, 7:3, and 6:4, and the accuracy is the lowest among all the compared models, which are 70.6%, 65.0%, and 66.7%, respectively. The classification effect of the logistic regression model is relatively more satisfactory, and the model classifies samples with the ratios of 8:2, 7:3, and 6:4 with accuracies of 74.6%, 68.7%, and 72.0%, respectively. Among all the models, the model in this paper has the most satisfactory classification results. Classifying the samples with the three ratios of 8:2, 7:3, and 6:4, the accuracy rates are 91.7%, 90.3%, and 90.9%, respectively, which is the best classification effect.

5.2. Analysis of rating indicators

In order to verify the effectiveness of the prediction model in this paper, four models, namely, logistic regression, support vector machine, decision tree, and deep neural network, were used to predict and classify the fraud behaviors in the samples, and finally, the accuracy, the recall, and the F1 value

Model name	Proportion of sample division	Accuracy %
Logistic regression model	8:2	74.6
	7:3	68.7
	6:4	72.0
Support vector machine model	8:2	71.8
	7:3	66.4
	6:4	69.9
Decision tree model	8:2	70.6
	7:3	65.0
	6:4	66.7
Model of this paper	8:2	91.7
	7:3	90.3
	6:4	90.9

Table 2. Classification accuracy results

were used as the final comparative analysis indexes, which were calculated through Eq. (16)-Eq. (18), and the evaluation indexes of different models were shown in Table 3. In terms of classification precision of all models, the model of this paper has the highest classification precision of 90.85%, followed by the support vector machine model of 83.68%, and the lowest precision is the decision tree model. In terms of recall, this paper's model has the highest recall of 90.77%, while the lowest is the logistic regression model with only 76.22%. In terms of F1 value, this paper's model has the highest F1 value of 90.81%, while the lowest F1 value is logistic regression model with only 78.48%. It can be seen that this paper's model shows ideal results in terms of precision, recall and F1 value, while the LR model is the least suitable for the detection of anomalies in corporate financial statements and the identification of financial fraud, and the recall and the F1 value are the least ideal among all the models.

Model name	Precision (%)	Recall (%)	F1 (%)
Logistic regression model	80.87	76.22	78.48
Support vector machine model	83.68	77.46	80.45
Decision tree model	79.12	88.35	83.49
Model of this paper	90.85	90.77	90.81

Table 3. Evaluation indicators for different models

5.3. Time spent on identification and classification

In order to verify the advantages of the models in this paper in terms of identification and classification time, four models, namely, logistic regression, support vector machine, decision tree, and deep neural network, are used for training, and four models are used to identify and classify the anomaly detection and financial fraud in the sample set. The model training time and identification and classification time are recorded for comparing the running time of the four models, and Figure 2 shows the training time and classification time of multiple models. The model with the highest training time is the decision tree model, which has a training time of 504.83 ms, but at the same time, the classification and identification time under this model is <0.01 ms. Thus, there is a polarization between the

training time and the identification and classification time, and the decision tree model is poorly trained with respect to the sample set. The model with the second highest training time is the logistic regression model, with a training time of 446.76ms, while the recognition and classification time is 0.12ms, which is not dominant in terms of training time and recognition and classification time. The sample set training time under the support vector machine model is 342.39ms and the recognition and classification time is 0.07ms, while the model in this paper takes the least time to train the sample set as well as to recognize and classify the sample set, the training time is 95.81m and the recognition and classification time is only 0.02ms. The model in this paper has good performance in terms of the training time and the recognition and classification time, which is the best performance.

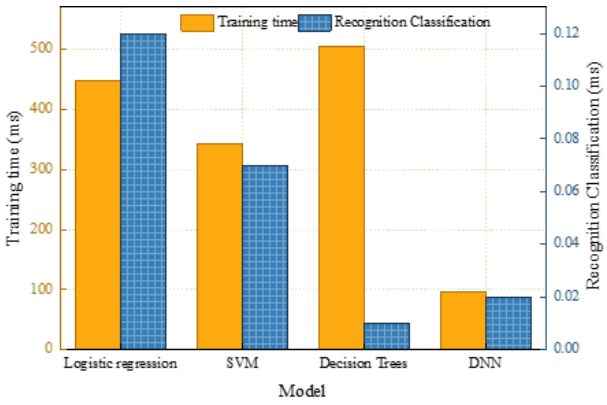

Fig. 2. Training time and classification time for multiple models

6. Conclusion

Under the background of big data, this paper brings deep neural network into financial statement abnormality recognition and classification, and solves the problem of low efficiency of financial statement abnormality recognition and classification through the advantage of deep neural network. The effectiveness of this paper's model is discussed from three aspects: analysis of classification accuracy, analysis of rating indexes, and analysis of time spent on recognition and classification. The results show that this paper's model classifies samples with 8:2, 7:3, and 6:4 ratios with accuracies of 91.7%, 90.3%, and 90.9%, respectively, and the precision, recall, and F1 value of the recognition and classification are the highest among the compared models. The time taken to train the sample set and to recognize and classify the samples is 95.81m and 0.02ms respectively, which is the most ideal among all models. The deep neural network based accounting data anomaly detection method can effectively reduce the auditor's workload and reduce the risk of financial statement fraud.

References

[1] S. Z. Aftabi, A. Ahmadi, and S. Farzi. Fraud detection in financial statements using data mining and gan models. *Expert Systems with Applications*, 227:120144, 2023. https://doi.org/10.1016/j.eswa.2023.120144.

[2] M. N. Ashtiani and B. Raahemi. Intelligent fraud detection in financial statements using machine learning and data mining: a systematic literature review. *Ieee Access*, 10:72504–72525, 2021. https://doi.org/10.1109/ACCESS.2021.3096799.

[3] A. Bakumenko and A. Elragal. Detecting anomalies in financial data using machine learning algorithms. *Systems*, 10(5):130, 2022. https://doi.org/10.3390/systems10050130.

[4] T. Chiu, Y. Wang, M. A. Vasarhelyi, et al. The automation of financial statement fraud detection: a framework using process mining. *Journal of Forensic and Investigative Accounting, 12 (1), 86*, 108, 2020.

[5] P. Craja, A. Kim, and S. Lessmann. Deep learning for detecting financial statement fraud. *Decision Support Systems*, 139:113421, 2020. https://doi.org/10.1016/j.dss.2020.113421.

[6] S. Gupta and S. K. Mehta. Data mining-based financial statement fraud detection: systematic literature review and meta-analysis to estimate data sample mapping of fraudulent companies against non-fraudulent companies. *Global Business Review*, 25(5):1290–1313, 2024. https://doi.org/10.1177/0972150920984857.

[7] K. G. Al-Hashedi and P. Magalingam. Financial fraud detection applying data mining techniques: a comprehensive review from 2009 to 2019. *Computer Science Review*, 40:100402, 2021. https://doi.org/10.1016/j.cosrev.2021.100402.

[8] H. Hemati, M. Schreyer, and D. Borth. Continual learning for unsupervised anomaly detection in continuous auditing of financial accounting data. *arXiv preprint arXiv:2112.13215*, 2021. https://doi.org/10.48550/arXiv.2112.13215.

[9] W. Hilal, S. A. Gadsden, and J. Yawney. Financial fraud: a review of anomaly detection techniques and recent advances. *Expert systems With applications*, 193:116429, 2022. https://doi.org/10.1016/j.eswa.2021.116429.

[10] A. Javadian Kootanaee, M. Hosseini Shirvani, et al. A hybrid model based on machine learning and genetic algorithm for detecting fraud in financial statements. *Journal of Optimization in Industrial Engineering*, 31(2):169, 2021. https://doi.org/10.22094/joie.2020.1877455.1685.

[11] M. R. MOHAIMIN, M. Sumsuzoha, M. A. H. Pabel, and F. Nasrullah. Detecting financial fraud using anomaly detection techniques: a comparative study of machine learning algorithms. *Journal of Computer Science and Technology Studies*, 6(3):01–14, 2024. https://doi.org/10.32996/jcsts.2024.6.3.1.

[12] M. Mohammadi, S. Yazdani, M. H. Khanmohammadi, and K. Maham. Financial reporting fraud detection: an analysis of data mining algorithms. *International Journal of Finance & Managerial Accounting*, 4(16):1–12, 2020.

[13] W. T. Mongwe and K. M. Malan. A survey of automated financial statement fraud detection with relevance to the south african context. *South African Computer Journal*, 32(1):74–112, 2020.

[14] S. O. Pinto and V. A. Sobreiro. Literature review: anomaly detection approaches on digital business financial systems. *Digital Business*, 2(2):100038, 2022. https://doi.org/10.1016/j.digbus.2022.100038.

[15] M. Riskiyadi. Detecting future financial statement fraud using a machine learning model in indonesia: a comparative study. *Asian Review of Accounting*, 32(3):394–422, 2024. https://doi.org/10.1108/ARA-02-2023-0062.

[16] M. Schultz and M. Tropmann-Frick. Autoencoder neural networks versus external auditors: detecting unusual journal entries in financial statement audits, 2020.

[17] T. Shahana, V. Lavanya, and A. R. Bhat. State of the art in financial statement fraud detection: a systematic review. *Technological Forecasting and Social Change*, 192:122527, 2023. https://doi.org/10.1016/j.techfore.2023.122527.

[18] A. K. Singh Yadav and M. Sora. Unsupervised learning for financial statement fraud detection using manta ray foraging based convolutional neural network. *Concurrency and Computation: Practice and Experience*, 34(27):e7340, 2022. https://doi.org/10.1002/cpe.7340.

[19] M. Soltani, A. Kythreotis, and A. Roshanpoor. Two decades of financial statement fraud detection literature review; combination of bibliometric analysis and topic modeling approach. *Journal of Financial Crime*, 30(5):1367–1388, 2023. https://doi.org/10.1108/JFC-09-2022-0227.

[20] M. Tragouda, M. Doumpos, and C. Zopounidis. Identification of fraudulent financial statements through a multi-label classification approach. *Intelligent Systems in Accounting, Finance and Management*, 31(2):e1564, 2024. https://doi.org/10.1002/isaf.1564.

[21] J. Vanhoeyveld, D. Martens, and B. Peeters. Value-added tax fraud detection with scalable anomaly detection techniques. *Applied Soft Computing*, 86:105895, 2020. https://doi.org/10.1016/j.asoc.2019.105895.

[22] P. Vanini, S. Rossi, E. Zvizdic, and T. Domenig. Online payment fraud: from anomaly detection to risk management. *Financial Innovation*, 9(1):66, 2023. https://doi.org/10.1186/s40854-023-00470-w.

[23] G. Wang, J. Ma, and G. Chen. Attentive statement fraud detection: distinguishing multimodal financial data with fine-grained attention. *Decision Support Systems*, 167:113913, 2023. https://doi.org/10.1016/j.dss.2022.113913.