Profit Estimation Model and Financial Risk Prediction Combining Multi-Scale Convolutional Feature Extractor and BGRU Model

Yangyan Liu[1*], Bolin Pan[2]
[1]School of Economics and Management, Guangxi Science & Technology Normal University, Laibing 546199, China
[2]School of Economic Management, Ningxia Institute of Science and Technology, Shizuishan 753000, China
E-mail: liuyangyan001@163.com
[*]Corresponding author

Keywords: convolutional neural network, feature extraction, bidirectional gated recurrent unit, profit estimate, financial risk

Received: March 21, 2024

In response to the inaccuracy of financial risk prediction and profit prediction for enterprises, a financial risk prediction model based on the graph networks was designed. This experiment combined multi-scale feature extraction and sequence analysis methods. In addition, the model adopted a structurally concise and effective bidirectional gated recurrent unit to capture temporal relationships in time series data. The profit prediction model was combined multi-scale advantages and attention mechanisms. The latter enhanced the recognition and utilization of influential features, which improved predictive ability and practical value. These results confirmed that the accuracy of this model significantly improved to 98.03% after iterative training. The F1 score of the financial risk prediction model reached 0.98, demonstrating an excellent performance. The profit prediction model performed better than other models in both regression and classification problem indicators, with an error close to 0 and a mean square error of 0.0232. This indicated that the model had extremely high prediction accuracy. Therefore, both models have strong predictive ability and have practical application significance.

Povzetek: Razvit je finančni model, ki združuje multi-lestvični konvolucijski ekstraktor lastnosti in BGRU model za natančno napovedovanje dobička in finančnih tveganj podjetij. Model z uporabo pozornostnih mehanizmov in sekvenčne analize dosega visoko zmogljivost in praktično vrednost v finančni analizi ter obvladovanju tveganj.

1 Introduction

In the current globalized economic environment, the estimation of corporate profitability and the prediction of financial risk becomes the important links in financial analysis and investment decision-making. On the one hand, it is crucial to accurately estimate a company's profit prospects for investors. Because it directly relates to the investment and return of capital. On the other hand, financial risk prediction is particularly important for corporate management. Because it can help enterprises identify potential financial problems in a timely manner, formulate corresponding countermeasures, and ensure the stable development of the enterprise [1-3]. As information technology develops, data analysis technology becomes a key force driving innovation, especially the application of artificial intelligence in the financial field. In recent years, deep learning technologies have emerged and developed. These technologies have made significant achievements in multiple fields such as image recognition, speech processing, and natural language processing. In this context, the application of deep learning techniques to profit estimate and financial risk prediction begins to receive widespread attention from academia and industry. Specifically, Convolutional Neural Network (CNN) shows great potential in processing time-series data due to its excellent feature extraction capabilities [4-6]. Gated Recurrent Unit (GRU) is an effective variant of recurrent neural networks, which has value in processing sequence information [7]. The financial data of enterprises are high-dimensional and contain rich time-series characteristics. However, traditional machine learning techniques are inadequate in dealing with such problems. In addition, there are many factors that need to be considered due to the complexity of financial data. How to effectively integrate these factors has become a significant challenge [8-10]. In this context, this study aims at combing multi-scale convolutional feature extractions with bio-heuristic for a bidirectional gated recurrent single model. A novel profit estimate and financial risk prediction framework is constructed, aiming to achieve more accurate and robust performance in predicting future profits and risks. The study first proposes the research purpose. Secondly, the financial risk prediction and profit estimate models are designed. Then the model performance is validated. Finally, the research conclusion is obtained.

In recent years, academic discussions on financial prediction automation continue to increase. Peng et al. proposed a novel predictive tool that utilized a model to describe the volatility of financial returns as numerous potential statistical distributions. Meanwhile, the predictive accuracy was optimized by analyzing the best case of these distributions. These results confirmed that the model performed better than most traditional prediction methods [11]. Cao et al. analyzed and predicted the financial crisis of listed companies from the perspective of constructing financial risk early warning systems for e-commerce enterprises, using methods based on deep learning technology. Meanwhile, they designed a financial stability warning system to achieve the early detection and handling of crises [12].

In profit forecasting, Sirisha et al. fused three neural network models to form a comprehensive model. This model was used to process past financial data to predict the future economic situation of the enterprise. These results confirmed that the model exhibited an accuracy prediction rate of over 90% [13]. Kim et al. developed a multivariate model under self-attention mechanism, aiming to plot the Bitcoin prices. These results confirmed that the prediction results of this model had high accuracy [14]. Shajalal et al. proposed an innovative deep learning network architecture aimed at predicting the stock risk of market products. They used a series of data balancing techniques to train data balance. Then the balanced data were used as the basis for training deep neural networks. These results confirmed the effectiveness and accuracy of the model in predicting product shortages [15].

He et al. developed two flood forecasting models in the application of Bidirectional Gated Recurrent Unit (BGRU). These two models were based on BGRU and Backpropagation Neural Network (BPNN), as well as bidirectional Long Short-Term Memory (LSTM) and BNN, respectively. These two models surpassed both unidirectional and bidirectional models in terms of prediction accuracy and reliability [16]. Dangut et al. developed a rare fault recognition model that relied on deep hybrid learning methods. They conducted empirical analysis using large-scale records from the central maintenance system of large aircraft. These experiments confirmed that this model improved accuracy by 18% and recall by 5%, effectively identifying rare faults [17]. Li et al. created a deep learning model that integrated end-to-end CNN and BGRU for extracting and classifying spatiotemporal information. These experiments confirmed that the model could effectively distinguish healthy rats from diseased rats, with a classification accuracy of up to 98.73%. This study had important reference significance for the clinical diagnosis of Parkinson's disease [18].

In summary, in recent years, researchers have increasingly focused on building efficient, accurate, and stable early warning systems in automated financial forecasting. The improved neural networks and deep learning-based methods are important. These methods play important roles in enterprise financial risk prediction and the development of financial stability warning systems. There are various prediction tools, such as self-attention mechanisms based on multivariate long short-term models. From the perspective of prediction, the proposal of these tools undoubtedly opens up new avenues for financial prediction. This study combines the method of multi-scale convolutional feature extraction, which has advantages over traditional models in information acquisition. The addition of BGRU enables the prediction model to better capture time series information, effectively improving the accuracy of profit estimate and financial risk predictions. The literature summary is shown in Table 1.

Table 1: Literature summary

Author	Literature dimension	Content	Conclusion
Li et al.	Risk prediction	Design and improve a neural network model, using data from listed companies from 2017 to 2020 as the basic dataset for empirical analysis	The prediction accuracy of the improved model exceeds 80%
Peng et al.		Design a risk prediction model based on financial volatility as a statistical distribution	The model performs better than most traditional prediction tools
Cao et al.		Construct an early warning model for e-commerce enterprise financial risk based on deep learning	The model can achieve early risk detection and risk handling
Sirisha et al.	Profit forecast	Design a comprehensive model that integrates three types of neural networks to predict the future profit situation of enterprises	The model accuracy prediction rate exceeds 90%
Kim et al.		Design a multivariable model based on self-attention mechanism to	The prediction results of the model have high accuracy and superiority

		predict the profitability of Bitcoin	
Shajalal et al.		Propose a deep learning network architecture to predict the risk of product shortage and product profit in the market	The effectiveness and accuracy of the model are superior
He et al.		Propose a flood prediction model that integrates BGRU and BNN	The model has advantages over traditional prediction methods in terms of prediction accuracy and reliability
Dangut et al.	Application of BGRU model	Design a rare fault recognition model for large aircraft based on deep hybrid learning method	This model improves prediction accuracy by 18% and recall by 5%
Li et al.		Design a deep learning model that integrates end-to-end CNN and BGRU for clinical diagnosis of Parkinson's disease	The model can effectively distinguish healthy and diseased rats, with a classification accuracy of up to 98.73%

Current research has predicted enterprise risks and profits from multiple perspectives. However, all methods are single risk or profit predictions, which have not formed an integrated system for predicting risks and profits simultaneously. This study presents a concise and effective bidirectional GRU and attention mechanism. This method not only improves the predictive ability of the model, but also integrates risk prediction and profit prediction. Then enterprises can make comprehensive decisions and avoid information conflicts caused by different prediction systems.

2 Design of financial risk and profit prediction models

A novel financial risk prediction model, namely Multi-Scale Bidirectional Gated Recurrent Unit (M-BU), is designed based on BGRU. In addition, a profit estimate model is designed, which is the Multi-Scale Adaptive Bidirectional Gated Recurrent Unit (M-AU). The financial risk prediction model is based on the graph networks and adopts a strategy combining multi-scale feature extraction technology and sequence analysis methods. Convolutional layers with different sizes of convolutional kernels are used to extract rich features. Meanwhile, advanced techniques such as Batchnorm and Mish functions are used to prevent overfitting while maintaining the model's generalization ability. BGRU can effectively capture temporal relationships in time series data. The profit estimate model combines multi-scale advantages and attention mechanisms. The latter enhances the recognition and utilization of influential features by assigning different weights to different features.

2.1 Design of financial risk prediction model

The first step is the analysis and reorganization of the sample data in all financial risk prediction model designs. The essential step before data analysis is to determine which variables should be included in the research scope [19-21]. The study sets a criterion to exclude variables with a missing degree exceeding one-third in the dataset. Because these variables can provide insufficient information to facilitate the analysis. The missing data will be filled in. The dynamic data imputation method is adopted instead of simply replacing the mean or minimum values [22-24]. The study observes that the values of many financial variables mostly fluctuate around 30% of their mean after in-depth analysis of the data distribution of various financial variables. Therefore, the strategy for handling missing data is as follows: a random number based on the fluctuation range is used to replace the average value of the variables. The calculation method is represented by formula (1).

$$\begin{aligned}&\text{Missing value} =\\&\text{Mean value} + \text{Randomness} \times \text{Mean value}\end{aligned} \tag{1}$$

A distortion rate is defined higher than 20% when encountering samples with severe data missing in generating a dataset. This means that the information of these samples is not sufficient to support a comprehensive analysis. Samples containing numerous missing values will be excluded under these conditions. There may still be situations where data is scarce despite going through layers of screening. For example, some corporate samples may only cover some financial variables. The study also decides to delete the incomplete data samples. Dealing with outliers is equally essential. Then this study applies the rule based on triple standard deviation to identify and remove outliers. The study uses the upper and lower quartiles to define and handle outliers based on the distribution of data from each year given the differences in data from different years. The key to this step is to eliminate data points that may distort the analysis results. The final stage of data processing is quantitative standardization. Different variables may have different scales and distribution intervals. Therefore, the study uses the min-max method to transform all financial variables into a unified standard interval of 0 to 1. Min-max standardization refers to adjusting each variable to a uniform size. This can eliminate the impact of different measurement units and maintain comparability between different variables. Standardization involves subtracting the minimum value of each variable from its value across all data points. Meanwhile, the variable is divided by the amplitude of the variable range to calculate the standardized data points [25-27]. The calculation method is represented by formula (2).

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

In formula (2), $\min(x)$ and $\max(x)$ represent the minimum and maximum values of the financial index, respectively. A completely new pattern is designed. This pattern enhances the ability to mine important information in financial data by integrating two main methods: multi-level feature recognition and sequence analysis. Firstly, a feature extraction module with multi-scale perception capability is designed to capture risk conditions associated with numerous financial indicators. The innovative inspiration for this strategy comes from the concept of graph networks. This module can integrate convolution kernels of various sizes and successfully extract rich feature levels from financial data. Figure 1 shows the structure of convolution and feature recognition.

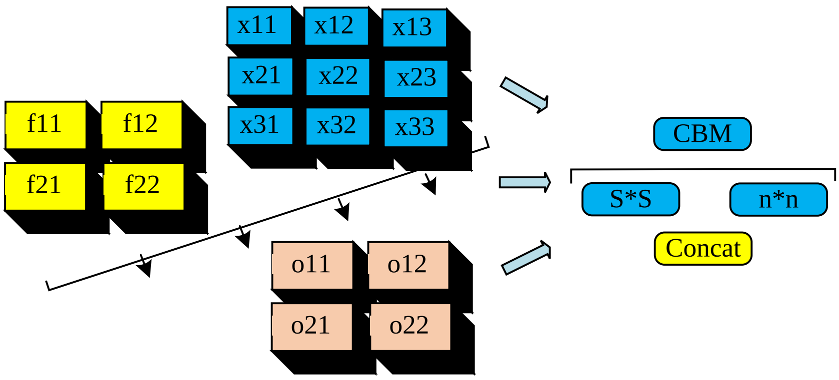

Figure 1: Convolutional and feature recognition structures

Small-scale convolution kernels distinguish subtle local patterns, while large-scale convolution kernels help to interpret external environmental correlations. This model can capture and apply the complex interactions between various financial indicators. The Batchnorm method is utilized to prevent overfitting of the model. Firstly, the mean is calculated using formula (3).

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i \quad (3)$$

Afterwards, the data variance is calculated using formula (4).

$$\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m}\left(x_i - \mu_B\right)^2 \quad (4)$$

Normalization is represented by formula (5).

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \quad (5)$$

The final calculation result can be obtained, represented by formula (6).

$$y_i = \gamma \hat{x}_i + \beta \quad (6)$$

In formula (6), γ represents the scaling variable. β is a translation variable. The Mish function is utilized as the activation function, represented by formula (7).

$$Mish = x \cdot \tanh\left(\ln\left(1 + e^x\right)\right) \quad (7)$$

In multi-scale feature extraction, convolution operation is first required to obtain the feature representation vector, represented by formula (8).

$$\begin{cases} s \leftarrow conv_s(x) \\ s \leftarrow conv_m(x) \end{cases} \quad (8)$$

In formula (8), x represents the eigenvector. Then the feature representation vector can be output using formula (9).

$$x' = concat(s, m) \quad (9)$$

The model also includes a pooling layer, aiming at screening and enhancing favorable features. This model has the ability to select the main information in the feature vectors and exclude non-deterministic interference information. The ability is achieved by applying filters of various sizes in the max pooling operation of multiple feature images [28-30]. This approach not only reduces the dimensionality of the data, but also reduces the overall complexity of the model to accelerate the training objectives. The key to model design is to adopt sequence learners after completing feature extraction. Figure 2 shows the pooling structure.

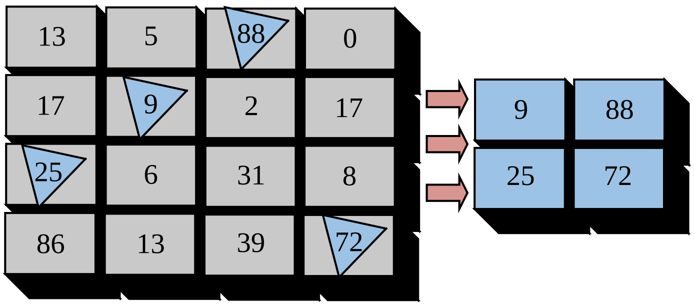

Figure 2: Pooled structure

The study utilizes the BGRU structure, which can capture temporal continuity in financial data and simultaneously learn past and future information. As a result, more accurate risk prediction can be achieved. The design of BGRU is concise and effective, reducing the parameter complexity of LSTM, but still retaining the ability to learn long-range dependencies. It is possible to determine when to retain or discard information by judging the gating mechanism. As a result, the efficiency of neural networks can be improved in processing time series data. Figure 3 shows the structure of the financial risk prediction model.

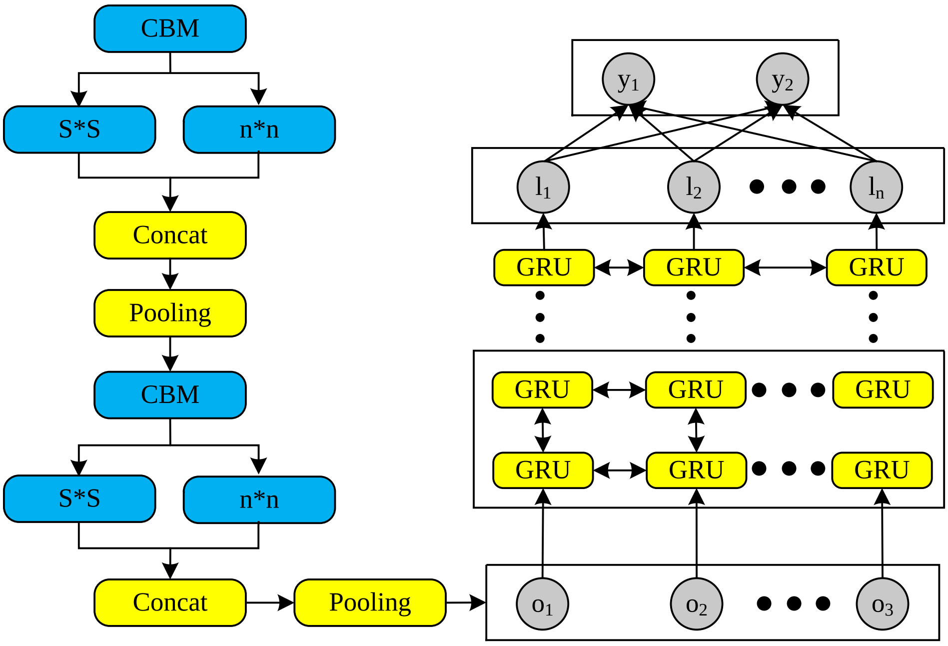

Figure 3: Structure of financial risk prediction model

The model reset gate is represented by formula (10).

$$Z_t = \sigma\left(X_t W_{xr} + H_{t-1} W_{hr} + b_r\right) \quad (10)$$

In formula (10), σ represents the activation function. H_{t-1} is the implicit state of the previous node. X_t refers to the current input. W means weight. The update gate is represented by formula (11).

$$Z_t = \sigma\left(X_t W_{xs} + H_{t-1} W_{hs} + b_s\right) \quad (11)$$

The memory gate is represented by formula (12).

$$\tilde{H}_t = \tanh\left(X_t W_{xh} + \left(R_t \sim H_{t-1}\right) W_{hh} + b_h\right) \quad (12)$$

Finally, M-BU combines multi-level spatial features with the temporal features learned by BGRU. As a result, a comprehensive prediction of the company's financial risk can be achieved. The dual advantages of feature mining and sequence analysis provide strong data support for model decision-making. Firstly, a multi-level convolutional network performs feature extraction by using convolutional kernels of different sizes. Therefore, clear hierarchical financial data features can be extracted. Next, the pooling operation further optimizes the feature vectors, reduces dimensionality, and removes noise. Finally, these features are integrated into a financial risk foresight model that takes into account the time dimension through the BU network. Table 2 shows the parameters of the network structure.

Table 2: Structural parameters of financial risk prediction model

Number of layers	Layer name	Nuclear size	Output size	Parameter quantity
0	Input layer	-	(38, 1)	0
1	One-dimensional convolutional layer 1	5×5	(38, 4.2)	25.2
1	One-dimensional convolutional layer 2	3×3	(38, 4.1)	16.8
2	One-dimensional convolutional layer 3	5×5	(38, 8.4)	168.4
2	One-dimensional convolutional layer 4	3×3	(38, 8.2)	104.8
3	Merge layers	-	(38, 16.6)	0
4	Bidirectional GRU layer 1	64	(16, 128.4)	39600
5	Bidirectional GRU layer 2	64	(16, 128.2)	74160
6	Dropout	-	(16, 128)	0
7	Leveling layer	-	(1, 2048.8)	0
8	Fully connected layer	-	(1, 2)	4100

2.2 Design of financial profit prediction model

The initial step involves collecting economic report data from numerous listed entities in the stock market in the design of the financial profit estimate model. Incomplete or unrelated information is discarded in the screening process to maintain the accuracy of the developed model. Economic data contain many core indicators, such as the proportion of corporate capital benefits and fixed asset profitability. They undoubtedly have analytical significance, but not all indicators have equal influence on predicting economic benefits. The M-AU model has been designed as a solution in response to this situation. This model is committed to improving and enhancing the efficiency of classical methods. Traditional BGRU often assumes that different features have the same influence when processing time series data. In view of this, the study decides to prioritize the importance of features to enhance the model's discriminative ability. M-AU has allocated weights to different features by inheriting the attention mechanism and improving the model. Meanwhile, more computing power is applied to features that have a significant impact on expected results, thereby enhancing the model's ability to identify and evaluate impacts. Figure 4 shows the structure of the profit estimate model.

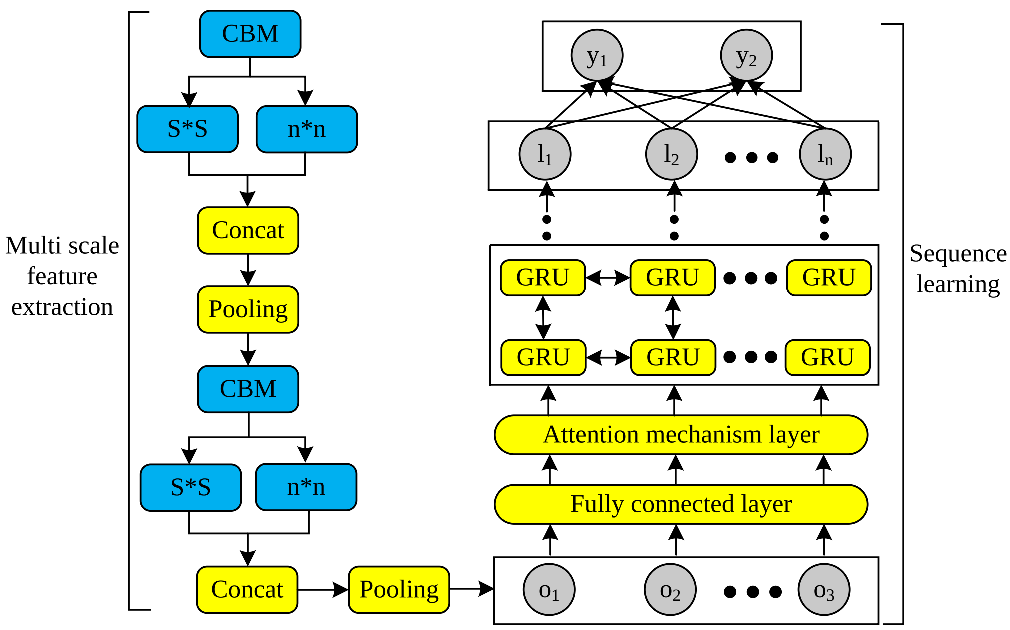

Figure 4: Structure of profit forecast model

This mechanism does not exist alone, but can be combined with other network structures. The identification rate of key information can be improved by applying it in various stages of model processing. More precisely, the typical design of M-AU includes using filters of various scales to perform deep feature mining of financial data and identify data characteristics of different scales. Subsequently, this model is able to learn the connections between time series data by utilizing the bidirectional cyclicity of the BGRU structure. The attention mechanism introduced on this basis assigns specific weights to each hidden layer state, represented by formula (13).

$$a_t = score\left(h_t\right) \qquad (13)$$

In formula (13), h_t represents the output of the multiscale extractor. These weights are normalized through the softmax function, representing the weighted impact of each feature on the final estimation result. This model integrates weighted features in subsequent stages and obtains the final prediction result through the fully connected layer and softmax function. The normalized weights are represented by formula (14).

$$p_t = \exp\left(a_t\right) / \sum_{t=1}^{n} a_t \qquad (14)$$

Finally, resource allocation is based on the attention mechanism, represented by formula (15).

$$y_t = \sum_{i=1}^{n} p_t h_t \qquad (15)$$

Overall, M-AU is committed to improving the accuracy of financial data analysis through exquisite structural selection and algorithm integration. This provides a more reliable scientific basis for financial decision-making. Figure 5 shows the model training process.

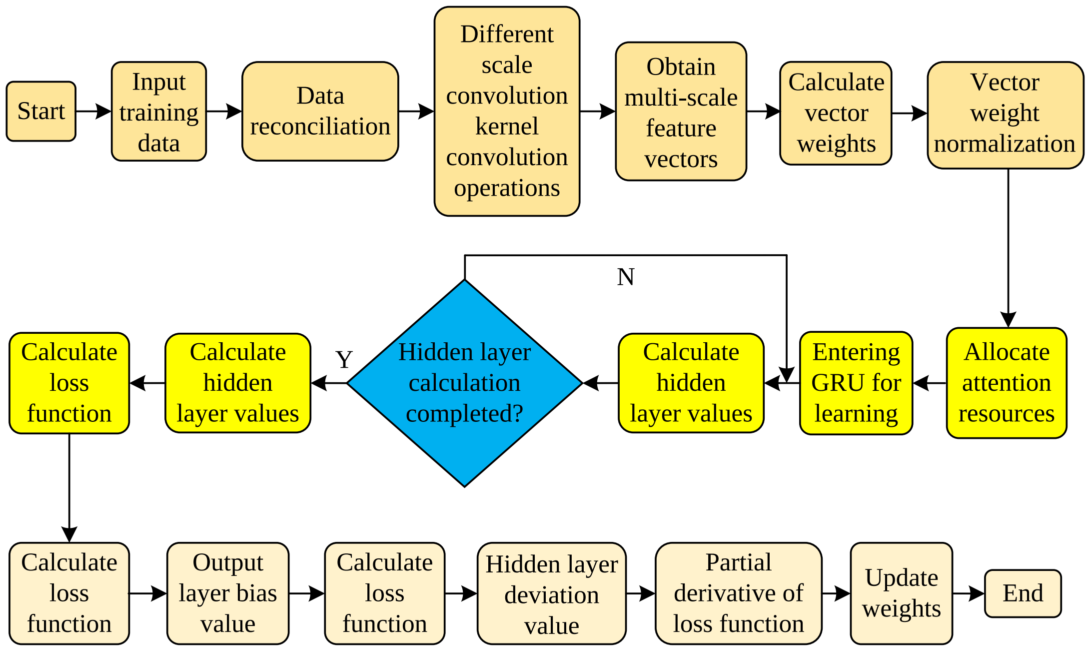

Figure 5: Model training process

The M-AU architecture innovatively integrates hierarchical financial risk tools and focusing techniques in the latest financial data analysis and expectation framework. This model has new insights in financial information assessment and trend prediction. This mode focuses on decoding continuous data and can identify the main characteristics closely related to prediction performance, thereby improving prediction accuracy and reliability. M-AU processes financial information with meticulous techniques under this approach and can tailor precise enterprise value evaluations for fund investors. M-AU adopts a dual scale filter and analyzes the input financial information layer by layer in the structural design. Therefore, the richness and complexity of the data can be ensured while mining the information. This stage is crucial in capturing patterns that may be overlooked, such as using filters of different scales to simultaneously reveal short-term and long-term financial trends, adding a three-dimensional perspective to the decision-making process. BGRU is used to analyze historical and potential trends after the financial risk analysis. BGRU provides a comprehensive analytical perspective for understanding and predicting complex financial time series and successfully captures the temporal dependencies of data. The introduction of a focus layer can improve the pattern prediction function. Hierarchy, as a filter in the network, emphasizes the recognition of important features by weighting the feature vectors of convolutional layers. This indicates that in decision-making, patterns can automatically identify the most relevant data without relying on artificially set feature importance. Therefore, the mode's ability to adjust and formulate strategies can be significantly enhanced in dealing with complex backgrounds. The network architecture, as an ultimate output, presents the superimposed characteristics as predicted values through the combination of compact layers and softmax incentive function. Then the values are further transformed into net profit prediction for the next cycle. The output results provides potential economic performance predictions for fund operators and decision-makers. Meanwhile, potential risks can be quantitatively measured and returned on funds. Table 3 shows the structural parameters.

Table 3: Structural parameters of profit prediction model

Number of layers	Layer name	Nuclear size	Output size	Parameter quantity
0	Input layer	-	(38, 1)	0
1	One-dimensional convolutional layer 1	5×5	(38, 4)	24.5
1	One-dimensional convolutional layer 2	3×3	(38, 4)	16.8
2	One-dimensional convolutional layer 3	5×5	(38, 8)	169.2
2	One-dimensional convolutional layer 4	3×3	(38, 8)	105.6
3	Merge layers	-	(38, 16)	0
4	Bidirectional GRU layer 1	64	(16, 128)	39608
5	Bidirectional GRU layer 2	64	(16, 128)	74256
6	Fully connected layer 1	-	(128, 16)	275.5
7	Attention mechanism layer	-	(16, 128)	0
8	Flatten layer	-	(1, 2048)	0
9	Fully connected layer 2	-	(1, 1)	2051.3

The exquisite design of M-AU provides efficient analytical tools for financial analysis professionals at the practical application level. Therefore, a thorough understanding of the economic situation and revenue potential of enterprises can be enhanced. Overall, the economic benefit estimation framework based on M-AU continuously analyzes and practices the latest data analysis technologies to enhance the certainty of economic analysis. A framework is formed by combining multi-level convolutional networks, bidirectional control loops, and focusing mechanisms. This framework can efficiently process time series data, identify complex economic patterns, and accurately predict the company's future economic performance. The M-AU prediction mechanism is expected to transform into a disruptive tool in the financial industry with the increase of data and the continuous refinement of patterns. This provides theoretical support for business strategy formulation and opens up new perspectives for financial investors in risk and return assessment.

First, the data are divided into training and test sets according to 70% and 30% arm strength to ensure that the model has sufficient generalization ability for unseen data. In addition, the study introduces L2 regularization technique to reduce the risk of overfitting. Regularization works by adding a penalty term to the loss function, which can limit the weight of the model and avoid excessive model complexity. The early stop method is used to control overfitting. This means that if the loss of the validation set does not improve within 10 consecutive training cycles, the training process will be stopped.

Indicators closely related to the recent financial health and liquidity ability of the enterprise are selected in the risk prediction, such as asset liability, current, and quick ratios. Because these indicators can fully reflect the solvency and cash flow status of the enterprise. These indicators are closely related to enterprise risks, which can reflect the enterprise's risk situation in real-time. Indicators directly related to the company's profitability are selected in the profit forecasting, such as revenue growth rate, gross profit margin, net profit margin, etc. These indicators can directly reflect the company's profitability and potential, which can predict the company's future profitability performance.

3 Effectiveness of financial risk prediction and profit prediction models

The performance analysis of financial risk prediction and profit estimate model included three parts. Firstly, the performance of financial risk prediction was analyzed. Secondly, the performance of the profit estimate model was analyzed. Finally, a comprehensive prediction effect analysis was conducted using actual enterprise cases.

3.1 Financial risk prediction effect

The aim of this experiment is to analyze and predict data from Special Treatment (ST) and non-ST companies using machine learning models. The study divided the dataset into training and testing sets based on standard proportions to effectively evaluate the performance of the model. Table 4 shows the experimental setup.

Table 4: Structural parameters of profit prediction model

Dataset settings			
Dataset categories	Number of ST companies	Number of non-ST companies	Total
Training set	252	11500	11752
Test set	85	4760	4845
Total	335	16260	16595

Parameter settings	
Set project	Detail settings
Dataset partitioning ratio	Training set: Test set=7:3
Learning rate	0.001
Optimizer	Adam
Training rounds	50
Dropout application layer	Fully connected layer

In Table 4, the training set includes 252 data points from ST company and 11500 data points from non-ST company, totaling 11752 samples. The test set includes 85 data points from ST company and 4760 data points from non-ST company, totaling 4845 samples. The entire dataset consists of 335 ST company data points and 16260 non-ST company data points, totaling 16595 samples. In addition, the Dropout technique was introduced in the fully connected layer of the model to improve the generalization ability of the model and prevent overfitting. This means that a certain proportion of neurons will be randomly discarded during training. This can increase the sparsity of the model structure, forcing the model to learn effectively through a small number of features, thereby improving its performance on unseen data. The models used for comparison are as follows: CNN, GRU, Multi-Scale Gated Recurrent Unit (M-GU), Support Vector Machine (SVM), FastText, Deep Long Short-Term Memory (D-LSTM), Character-Level Convolutional Neural Network (Char-CNN), Character-Level Convolutional Recurrent Neural Network(Char-CRNN), Word-Level Convolutional Neural Network with Attention(Word-CNN-Att), Deep Recurrent Neural Network (DRNN), and Deep Bidirectional Gated Recurrent Unit (D-BU).

First, the study collected data from multiple financial databases and publicly available company information for ST and non-ST companies in the experimental process. The focus was on financial indicators during the data collection process, such as debt ratio and current ratio, and data cleaning, preprocessing. Meanwhile, feature selection steps were carried out. Then the experiment began by using the Adam optimizer and setting appropriate learning rates and training rounds. The experiment was divided into three stages. Firstly, the learning mode and generalization ability of the model were trained and tested, followed by financial risk prediction. Secondly, financial indicators were accurately predicted and profit forecasts were made through parameter adjustments and test data validation. Finally, J company was chosen as the empirical research object to evaluate the actual effectiveness of predictions.

Figure 6 shows the parameter variation effect.

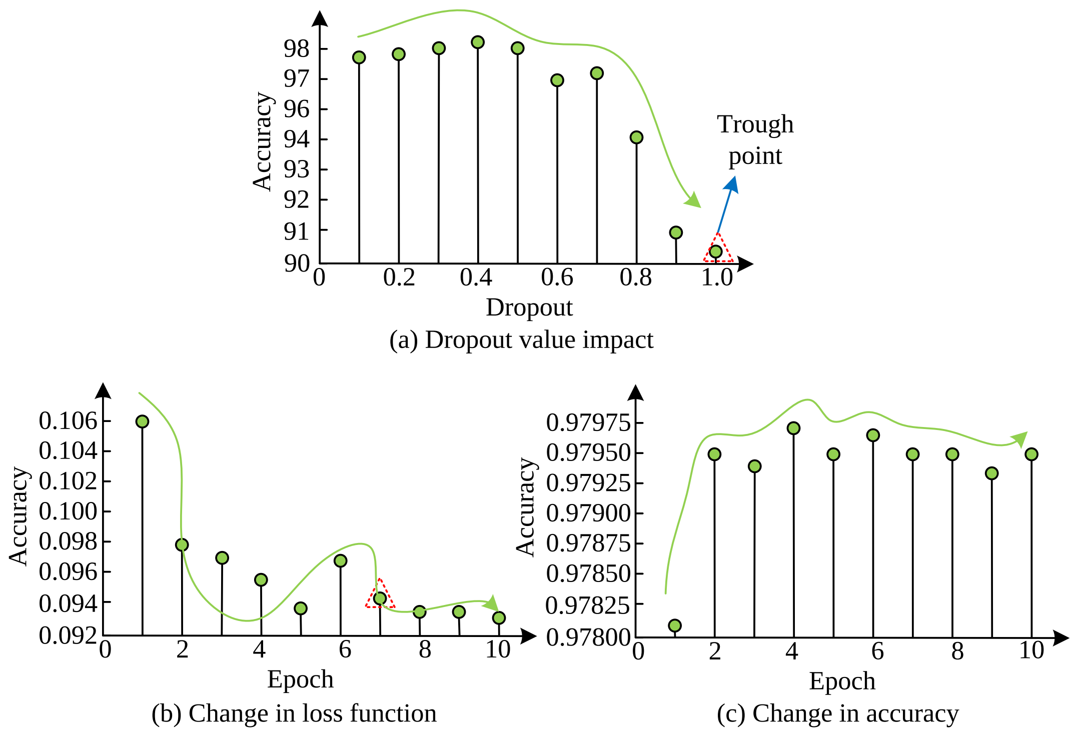

Figure 6: Parameter variation effect

In Figure 6, the loss function values in the database showed a significant decreasing trend as the iteration increased. This indicated that the model's predictive ability gradually improved in consistency with the actual situation. Meanwhile, the accuracy of the model also significantly improved as the loop tests increased. The proportion of values close to 1 continued to increase, further confirming the improvement in the performance of the research model. In addition, the classification accuracy was the highest when the dropout value was 0.4, reaching 98.03%. Table 5 shows a comparison of vertical improvements.

Table 5: Vertical improvement comparison

Model	Data set	Accuracy	F1 score	Number of parameters	Training time	Memory footprint
SVM	A	0.9498	0.945	15000	100.50s	500MB
SVM	B	0.9376	0.938	15000	95.75s	500MB
CNN	A	0.9672	0.967	250000	25.09s	1.2GB
CNN	B	0.9549	0.955	250000	30.17s	1.2GB
GRU	A	0.9532	0.953	80000	197.78s	750MB
GRU	B	0.9485	0.948	80000	205.89s	750MB
BGRU	A	0.9702	0.97	120000	349.05s	1GB
BGRU	B	0.9667	0.966	120000	360.11s	1GB
M-GU	A	0.9742	0.974	100000	237.45s	900MB
M-GU	B	0.9689	0.968	100000	245.96s	900MB
M-BU	A	0.9803	0.98	150000	377.93s	1.1GB
M-BU	B	0.9757	0.975	150000	390.07s	1.1GB

In Table 5, M-BU achieved an accuracy of 98.03% on dataset A, which was superior to other models such as BGRU's 97.02%, GRU's 95.32%, and the classical CNN model's 96.72%. M-BU also performed well on dataset B, achieving an accuracy of 97.57%. M-BU still maintained a leading position compared to BGRU's 96.67%, GRU's 94.85%, and CNN's 95.49%. This high accuracy meant that M-BU more accurately identified potential risk points when predicting financial risk. In addition, M-BU also performed excellently in F1 scores, reaching 0.980 and 0.975 on datasets A and B, respectively, significantly higher than other models. This indicated that M-BU had higher reliability in predicting positive and negative samples. The training time of M-BU was longer compared to other models, with 377.93 seconds and 390.07 seconds on datasets A and B, respectively. This reflected the additional time investment required for model training, partly due to the predictable complexity of the M-BU structure compared to general GRU. However, this additional time investment was reasonable and beneficial considering that training was a one-time cost and the model was used multiple times after deployment. M-BU had 150000 parameters in terms of parameter quantity, slightly more than other competitive models such as GRU's 80000. The larger number of parameters corresponded to the learning ability of the model, increasing the risk of overfitting and the demand for computing resources. M-BU was approximately 1.1GB in terms of memory usage. This indicated that M-BU could perform effective training with certain computing resources. Figure 7 shows a horizontal comparison.

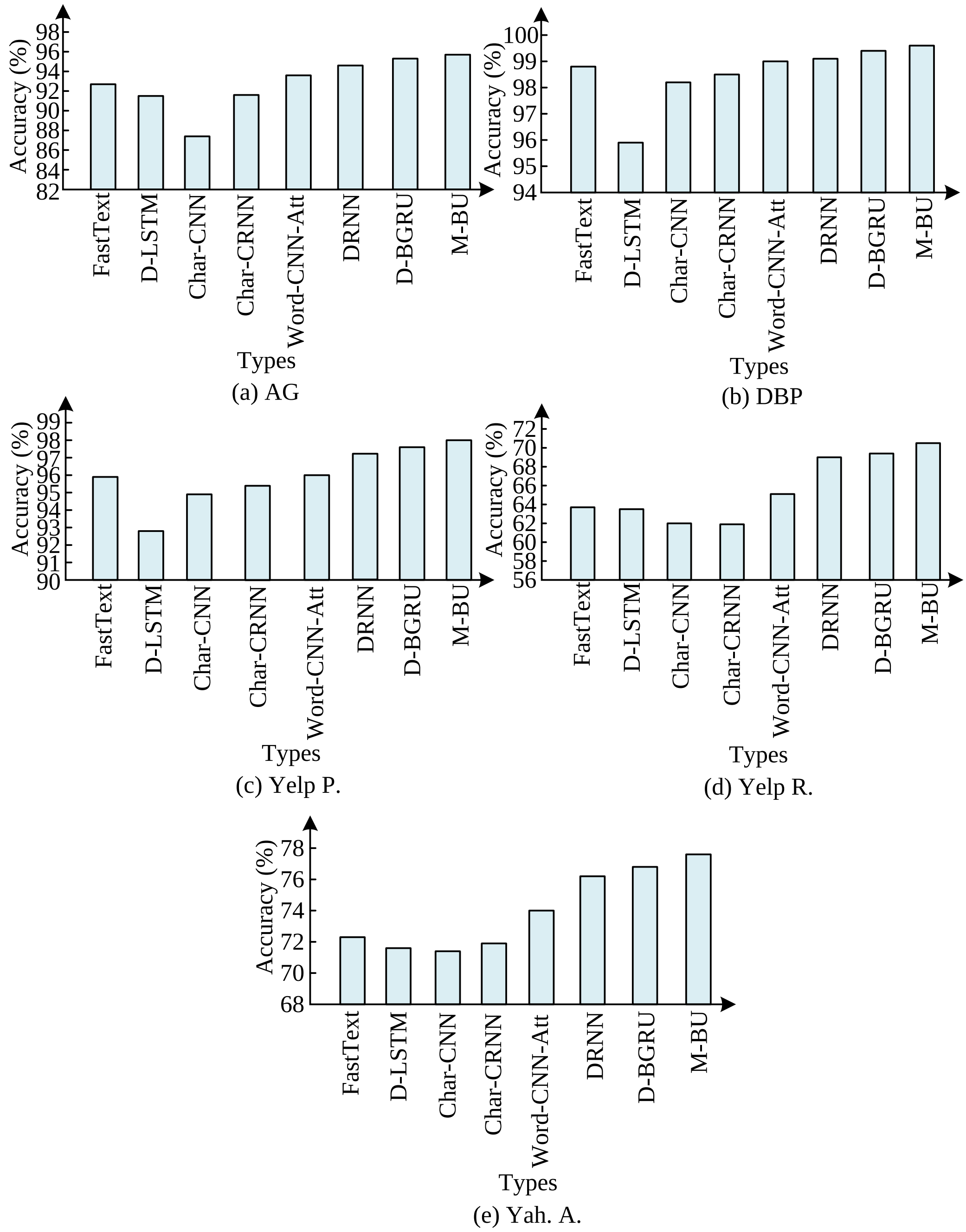

Figure 7: Horizontal comparison

In Figure 7, the performance of M-BU was excellent on different datasets in a series of recent horizontal comparison analysis of models. M-BU not only demonstrated excellent processing capabilities compared to other cutting-edge technologies in comparative studies, but also reflected unique insights into prediction accuracy.

M-BU achieved a classification accuracy of 95.7% in the AG dataset, surpassing the 95.2% of D-BU in the same series. M-BU had shown enormous potential in the DBP dataset, leading with an accuracy of 99.6%, which was more accurate than D-BU's 99.5%. M-BU achieved an accuracy of 98.0% in the Yelp P. dataset, which was the highest among all comparison models and significantly better than the DRNN model's 97.3%. The improvement of this data indicated that M-BU had better recognition ability when dealing with complex text content. M-BU also demonstrated strong predictive performance in the Yelp R. dataset, with a classification accuracy of 70.5%, surpassing D-BU's 69.3%. This indicated that M-BU could more accurately capture semantic details and user emotional tendencies in handling user comments. M-BU still maintained a leading position in the Yah. A. dataset, with an accuracy of 77.6% being the best performance among all comparison models. This ratio was significantly higher than D-BU's 76.7%, demonstrating its high applicability and stability in a wide range of fields.

3.2 Profit forecast effect

The study first analyzed the effect of parameter changes on M-AU in the profit estimate effect analysis. Figure 8 shows the variation pattern of the loss function value with increasing iteration.

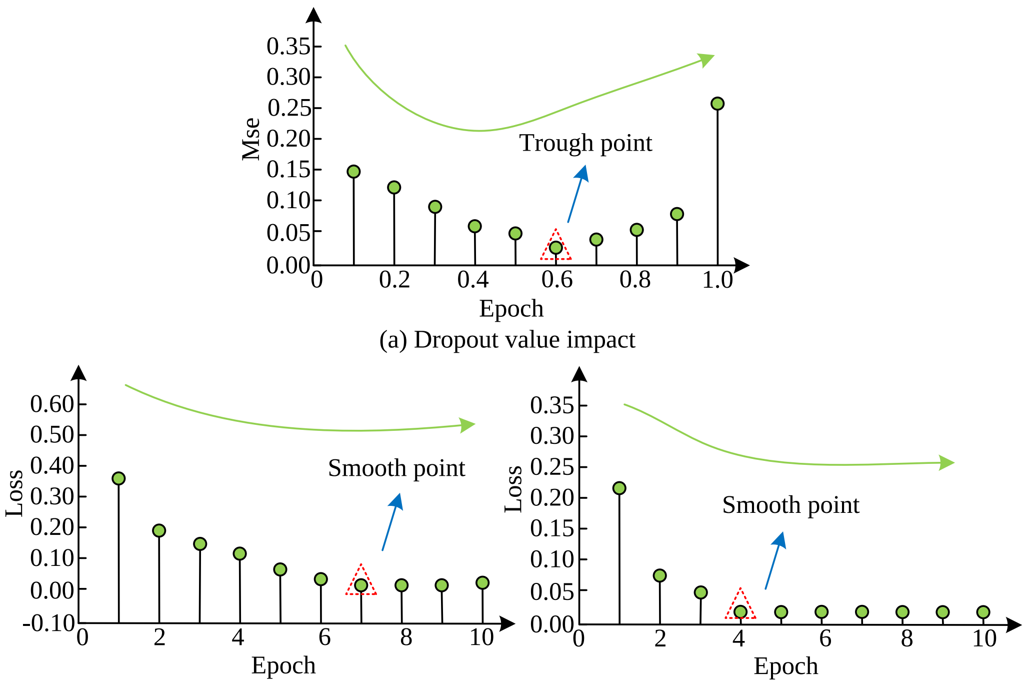

Figure 8: The parameter variation effect of M-AU

Figure 8 shows the variation pattern of the loss function value with increasing iteration. These data confirmed that as iterations increased, the loss function value significantly decreased, approaching a small value close to 0. This change indicated that this model was increasingly approaching the optimal fit to the training data. A Mean Square Error (MSE) of 0.0232 was obtained by evaluating the performance of the model on the test set. This provided an important basis for studying and evaluating the prediction accuracy of the model. On the other hand, experiments with different dropout values showed that they had a significant impact on the prediction error of the model. The model reached the lowest MSE on the test set especially when the dropout value was set to 0.6, which was 0.0349. This result suggested that the model could better generalize to unseen data and avoid overfitting under this setting. Therefore, it was recommended to set dropout to 0.6 to optimize model performance under the current experimental conditions. Figure 9 shows the horizontal comparison of M-AU.

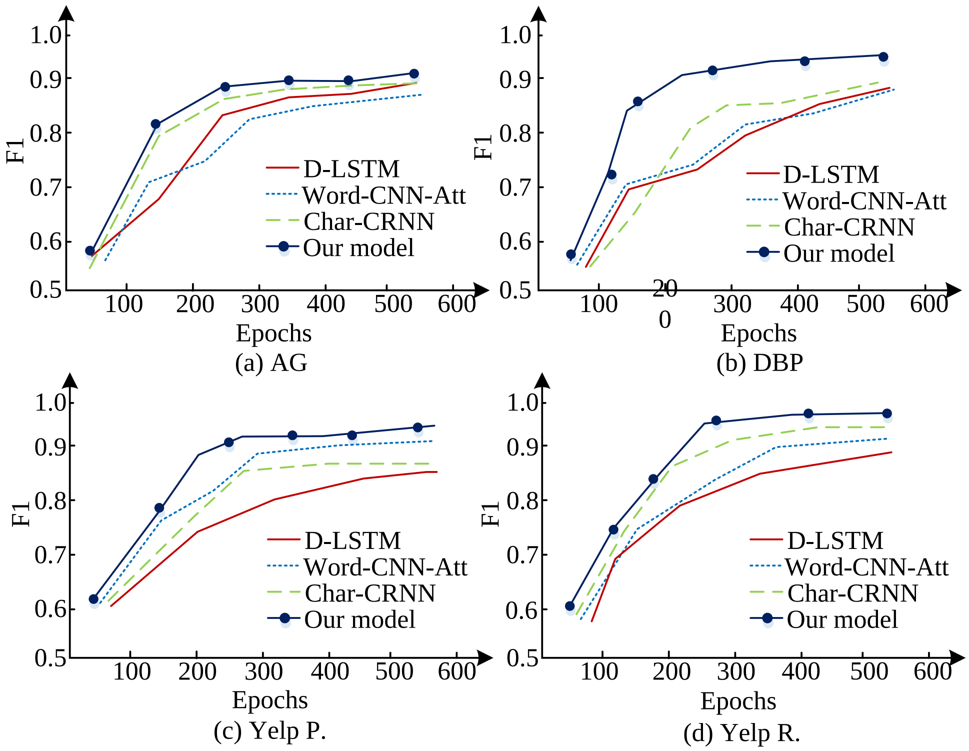

Figure 9: Horizontal comparison of M-AU

In Figure 9, the design model consistently maintained an advantage in F1 values among these four subgraphs representing different datasets, with the curve located above the F1 value curve of the other models. Therefore, the design model performed better on commonly used datasets. Table 6 shows a comparison of the vertical improvements of M-AU.

Table 6: Comparison of M-AU vertical improvements

Dataset name	Model	MSE	MAE	MAD	RMSE	R^2 score	Accuracy	Recall	F1 score
Financial dataset A	CNN	0.1978	0.5684	0.5044	0.4447	0.81	0.75	0.72	0.735
Financial dataset A	GRU	0.1898	0.5447	0.484	0.4356	0.83	0.78	0.74	0.76
Financial dataset A	BGRU	0.1547	0.4439	0.3944	0.3933	0.86	0.82	0.79	0.805
Financial dataset A	M-BU	0.1089	0.3125	0.2777	0.33	0.9	0.85	0.83	0.84
Financial dataset A	M-AU	0.0232	0.1191	0.1058	0.1523	0.96	0.92	0.91	0.915
Financial dataset B	CNN	0.2056	0.5792	0.5164	0.4533	0.78	0.73	0.7	0.715
Financial dataset B	GRU	0.1945	0.5512	0.4998	0.441	0.81	0.76	0.73	0.745
Financial dataset B	BGRU	0.1602	0.4589	0.4129	0.4002	0.85	0.81	0.78	0.795
Financial dataset B	M-BU	0.1137	0.3291	0.2905	0.3372	0.88	0.83	0.81	0.82
Financial dataset B	M-AU	0.0263	0.1256	0.1128	0.1621	0.94	0.9	0.88	0.89

In Table 6, the comparison indicators are MSE, Mean Absolute Error (MAE), Mean Absolute Deviation (MAD), and Root Mean Squared Error (RMSE). M-AU was only 0.0232 on MSE on financial dataset A, which was about 88% lower than CNN's 0.1978. This demonstrated its high accuracy in estimating true profit values. The MAE also significantly decreased from CNN's 0.5684 to 0.1191, further confirming that M-AU had higher prediction accuracy at the average level. Similarly, M-AU led the MAD index at 0.1058, while CNN was 0.5044. This trend was equally evident in RMSE and R2 scores. The RMSE of M-AU reached 0.1523, which was much lower than CNN's 0.4447, reflecting that M-AU had smaller errors in predicting various data points. Meanwhile, the R2 score of M-AU reached 0.96, close to perfect predictive ability, far exceeding CNN's 0.81. This demonstrated that M-AU was significantly better than other models in overall data fitting. In terms of financial dataset B, the performance of all models has decreased compared to dataset A. This might be attributed to the complexity of dataset B, but M-AU continued to maintain its superiority. The MSE and MAE of M-AU were 0.0263 and 0.1256, respectively, slightly increasing compared to dataset A, but still leading other models. M-AU demonstrated strong performance in classification performance indicators such as accuracy, recall, and F1 score on both datasets. The accuracy of M-AU reached 0.92 and 0.9, recall rates were 0.91 and 0.88, and F1 scores were 0.915 and 0.89, respectively. These were all higher than other models. For example, the accuracy of CNN on dataset A was 0.75, the recall was 0.72, and the F1 score was 0.735, which highlighted the significant improvement of M-AU in classification accuracy.

3.3 Comprehensive case analysis

The J company selected for the study is listed electronic technology Co., Ltd., which focuses on the manufacturing of high-end consumer electronics products and has approximately 200 employees. This company's main products include smartphones and wearable devices to meet the market's demand for high-quality electronic devices. Table 7 shows the case results.

Table 7: Case results

		Financial risk prediction		
Risk indicators	Description	Estimation	Prediction interval	Remarks
Financial risk level	Model-based risk classification	Low risk	-	Risk is effectively managed
Asset liability ratio	The percentage of liabilities to assets	0.45	43%-47%	Control at a safe level
Current ratio	Display short-term solvency	2.1	2.0-2.2	Good short-term liquidity
Quick ratio	Exclude the solvency of inventory	1.5	1.4-1.6	Quick assets in good condition
Interest coverage ratio	Operate profit before interest expenses	6	5.5-6.5	Interest payment without pressure
Cash flow ratio	Cash flow to current debt ratio	1.2	1.1-1.3	Good cash flow situation
		Profit forecast		
Profit indicators	Description	Predictive value	Prediction interval	Remarks
Revenue growth rate	Expected annual revenue growth rate	0.05	4%-6%	Stable business growth
Return on equity	Return on equity to shareholders	0.1	9%-11%	Good investment return
Total asset turnover rate	The efficiency of generating revenue from total assets	0.8	0.7-0.9	Reasonable asset utilization efficiency
Net profit attributable to shareholders of the parent company	Actual net profit attributable to shareholders	32 million yuan	31 million to 33 million yuan	Expected good profitability
Operating cost ratio	The proportion of operating costs to operating revenue	0.85	80%-90%	Reasonable gross profit margin
Earnings per share	Earnings per share	1.5yuan	1.4-1.6 yuan	Good earnings per share

In Table 7, the financial risk forecast showed a debt level of 45%, which was at the center of the forecast range of 43%-47%. This indicated that companies were cautious in the application of financial leverage, with lower risks, which were effectively managed. The current ratio and quick ratio were 2.1 and 1.5, respectively, both within their predicted ranges of 2.0-2.2 and 1.4-1.6. This indicated that both short-term liquidity and quick assets were in good condition, indicating that the company had sufficient ability to cope with short-term debt. In addition, the interest coverage ratio was 6 times, safely within the predicted range of 5.5-6.5. This meant that the enterprise could easily pay interest expenses with operating profits. The cash flow ratio was 1.2, which fell within the predicted range of 1.1-1.3. This indicated that the enterprise had healthy cash flow and was conducive to ensuring smooth daily operations and investment activities. The predicted revenue growth rate was 5% in the profit estimate, which was consistent with the predicted range of 4%-6%. This demonstrated stable business growth for the enterprise. The return on equity was 10%, accurately falling within the expected range of 9%-11%, proving that investors could expect a good return on investment. The total asset turnover rate was 0.8, which was within the predicted range of 0.7-0.9. This indicated that the enterprise had a reasonable asset utilization efficiency. The predicted value of 32 million reflected the expected good profitability of the company for the net profit attributable to shareholders of the parent company. This prediction was close to the edge of its prediction range of 31 million to 33 million. The operating cost rate was controlled at 85%, with a predicted range of 80%-90%, maintaining a reasonable gross profit margin. Finally, the earnings per share forecast was 1.5 yuan, within the forecast range of 1.4 to 1.6 yuan, indicating a good earnings per share. In summary, the financial forecasting model applied provided excellent and reliable predictions of financial risk and profitability. These predicted values matched the actual indicators and were mostly within a safe prediction range. This indicated that the model had high accuracy and reliability in understanding the financial health status of enterprises. Enterprises could adjust their strategies in a timely manner, manage potential risks, and ensure sustained positive growth through these data-driven insights.

Table 8: Cost calculation and scalability analysis

Characteristic	Dataset size	Lightweight transformer	Model
Training time	Small (~10000 samples)	2 hours	1 hour
	Medium (~100000 samples)	12 hours	6 hours
	Large (>1000000 samples)	>24 hours	12 hours
Resource requirements	CPU core	4-core	2-core
	RAM	8GB	4GB
	Recommended GPU	NVIDIA GTX 1080	NVIDIA GTX 1060

From Table 8, the designed model had significant advantages in training time compared to the lightweight transformer. Meanwhile, this model achieved efficient operation on lower computational configurations, which was more advantageous.

4 Discussion

Li X et al. used an optimized BPNN to predict the financial risks of listed companies. The optimized BPNN improved the accuracy of the prediction, with a prediction accuracy rate of over 80% [31]. This study is also based on neural networks. Moreover, more complex data can be handled using bidirectional gated recursive units to capture temporal relationships in time series data. Uddin MS et al. used an optimized random forest method to predict the credit risk of micro enterprises, which was more targeted but less common [32]. Meanwhile, this study combined multi-scale feature extraction methods to effectively extract and predict various types of enterprise information, which was more universal. Barra S et al. focused on the temporal nature of information similar to the approach of this study. Time-series encoding methods combined with CNN were used to transform time-series data into Gram angle field images for analysis [33]. The designed model had a simpler execution and stronger robustness. Uthayakumar J et al. focused on predicting financial crises by combining improved K-means clustering and chaotic genetic ant colony algorithm [34]. On the other hand, Bianchi D focused on bond risk by combining extreme tree and neural network models to analyze the likelihood of bond returns [35]. Both studies focused on one aspect of corporate risk, such as financial risk. On the side of bond risk, this study takes the perspective of enterprises as the main body to conduct integrated prediction analysis of enterprise risk and profit. The designed model performed better in classification and efficiency indicators than other models compared with the model of Li X et al. and Uddin MS et al. This demonstrated higher prediction accuracy, as well as more comprehensive and stable information processing.

5 Conclusion

A novel financial risk prediction model based on graph network concept was designed, which combined multi-scale feature extraction technology with sequence analysis methods. A model based on multi-scale advantage and attention mechanism was also proposed for profit estimate. These results confirmed that the prediction accuracy of these two models was significantly improved compared to traditional financial prediction models. The model's accuracy had significantly improved after 50 rounds of training, with the highest classification accuracy reaching 98.03%. The financial risk prediction model achieved an accuracy of 98.03% and an F1 score of 0.980 on dataset A. The error was close to 0 on the profit estimate model, and the MSE was 0.0232. This indicated that the model could more accurately identify potential risks and profit points when predicting risks and profits. In the case study, the debt level was 45% for the financial risk forecast, the current ratio and quick ratio were 2.1 and 1.5, respectively. The actual interest coverage ratio and cash flow ratio were within the forecast range. The net profit attributable to the shareholders of the parent company was expected to be 32 million in terms of profit estimate, which almost matched the actual indicators. Therefore, the designed comprehensive model is an efficient and feasible method.

The designed model integrates risk prediction and profit prediction has strong predictive ability. A prediction system is formed for multi-feature data and their time characteristics. However, there are many differences in economic structure, market mechanisms, policy environment, etc. between different regions in practical applications. Therefore, the specific environment in which the model is applied may also vary, leading to prediction bias in universal models under specific conditions. Therefore, it is necessary to modularize the different functions of the prediction model. Meanwhile, targeted parameter setting and application solutions can be developed for different environments in the future.

References

[1] A. Lin, N. Manral, P. McElhinney, K. Aditya, H. Matsumoto, and J. Kwiecinski, "Deep learning-enabled coronary CT angiography for plaque and stenosis quantification and cardiac risk prediction: an international multicentre study," The Lancet Digital Health, vol. 4, no. 4, pp. e256-e265, 2022. https://doi.org/10.1016/S2589-7500(22)00022-X.

[2] W. Gleißner, T. Günther, and C. Walkshäusl, "Financial sustainability: measurement and empirical evidence," Journal of Business Economics, vol. 92, no. 3, pp. 467-516, 2022. https://doi.org/10.1007/s11573-022-01081-0.

[3] R. Patel, S. Tanwani, and C. Patidar, "Relation extraction between medical entities using deep learning approach," Informatica, vol. 45, no. 3, pp. 359-366, 2021.https://doi.org/10.31449/inf.v45i3.3056.

[4] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting," Annals of Data Science, vol. 10, no. 1, pp. 183-208, 2023. https://doi.org/10.1007/s40745-021-00344-x.

[5] J. Kong, S. Oh, B. O. Kang, and J. Jung. "Development of an incentive model for renewable energy resources using forecasting accuracy in South Korea," Energy Science & Engineering, vol. 10, no. 9, pp. 3250-3266, 2022. https://doi.org/10.1002/ese3.1020.

[6] M. Salb, M. Zivkovic, N. Bacanin, A. Chhabra, and M. Suresh, "Support vector machine performance improvements for cryptocurrency value forecasting by enhanced sine cosine algorithm," Computer Vision and Robotics, pp. 527-536, 2022. https://doi.org/10.1007/978-981-16-8225-4_40.

[7] F. Liu, Y. Wu, N. Almarri, M. Habibollahi, H. T. Lancashire, B. Bryson, L. Greensmith, D. Jiang, and A. Demosthenous, "A fully implantable opto-electro closed-loop neural interface for motor neuron disease studies," IEEE Transactions on Biomedical Circuits and Systems, vol. 16, no. 5, pp. 752-765, 2022. https://doi.org/10.1109/TBCAS.2022.3202026.

[8] V. K. Tewari and A. Verma, "Design and application of double-gate MOSFET in two loops controlled multi-input DC-DC converter," Silicon, vol. 14, no. 12, pp. 7321-7333, 2022. https://doi.org/10.1007/s12633-021-01469-7.

[9] P. Preethi and H. R. Mamatha, "Region-based convolutional neural network for segmenting text in epigraphical images," Artificial Intelligence and Applications, vol. 1, no. 2, pp. 119-127, 2023. https://doi.org/10.47852/bonviewAIA2202293.

[10] X. Li, J. Wang, and C. Yang, "Risk prediction in financial management of listed companies based on optimized BP neural network under digital economy," Neural Computing and Applications, vol. 35, no. 3, pp. 2045-2058, 2023. https://doi.org/10.1007/s00521-022-07377-0.

[11] S. Peng, S. Yang, and J. Yao, "Improving value-at-risk prediction under model uncertainty," Journal of Financial Econometrics, vol. 21, no. 1, pp. 228-259, 2023. https://doi.org/10.1093/jjfinec/nbaa022.

[12] Y. Cao, Y. Shao, and H. Zhang, "Study on early warning of E-commerce enterprise financial risk based on deep learning algorithm," Electronic Commerce Research, vol. 22, no. 1, pp. 21-36, 2022. https://doi.org/10.1007/s10660-020-09454-9.

[13] U. M. Sirisha, M. C. Belavagi, and G. Attigeri, "Profit prediction using Arima, Sarima and LSTM models in time series forecasting: A Comparison," IEEE Access, vol. 10, pp. 124715-124727, 2022. https://doi.org/10.1109/ACCESS.2022.3224938.

[14] G. Kim, D. H. Shin, J. G. Choi, and S. Lim, "A deep learning-based cryptocurrency price prediction model that uses on-chain data," IEEE Access, vol.

10, pp. 56232-56248, 2022. https://doi.org/10.1109/ACCESS.2022.3177888.

[15] M. Shajalal, P. Hajek, and M. Z. Abedin, "Product backorder prediction using deep neural network on imbalanced data," International Journal of Production Research, vol. 61, no. 1, pp. 302-319, 2023. https://doi.org/10.1080/00207543.2021.1901153.

[16] S. He, X. Sang, J. Yin, Y. Zheng, and H. Chen, "Short-term runoff prediction optimization method based on bgru-bp and blstm-bp neural networks," Water Resources Management, vol. 37, no. 2, pp. 747-768, 2023. https://doi.org/10.1007/s11269-022-03401-z.

[17] M. D. Dangut, I. K. Jennions, S. King, and Z. Skaf, "A rare failure detection model for aircraft predictive maintenance using a deep hybrid learning approach," Neural Computing and Applications, vol. 35, no. 4, pp. 2991-3009, 2023. https://doi.org/10.1007/s00521-022-07167-8.

[18] H. Li, Q. He, and L. Wu, "Detection of brain abnormalities in parkinson’s rats by combining deep learning and motion tracking," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, no. 1, pp. 1001-1007, 2023. https://doi.org/10.1109/TNSRE.2023.3237916.

[19] P. Weston and M. Nnadi, "Evaluation of strategic and financial variables of corporate sustainability and ESG policies on corporate finance performance," Journal of Sustainable Finance & Investment, vol. 13, no. 2, pp. 1058-1074, 2023. https://doi.org/10.1080/20430795.2021.1883984.

[20] S. Wahab, M. Imran, A. Safi, and D. Kirikkaleli, "Role of financial stability, technological innovation, and renewable energy in achieving sustainable development goals in BRICS countries," Environmental Science and Pollution Research, vol. 29, no. 32, pp. 48827-48838, 2022. https://doi.org/10.1007/s11356-022-18810-1.

[21] H. N. H. Al-Hashimy, T. T. Y. Alabdullah, E. Ries, M. A. Ahmed, M. I. Nor, and K. A. M. Jamal, "The impact of financial management elements and behavioral intention on the financial performance," International Journal of Scientific and Management Research, vol. 5, no. 12, pp. 117-149, 2022. https://doi.org/10.37502/JSMR.2022.51210.

[22] Y. Si, S. Heeringa, D. Johnson, R. J. A. Little, W. Liu, F. Pfeffer, and T. Raghunathan, "Multiple imputation with massive data: An application to the panel study of income dynamics," Journal of Survey Statistics and Methodology, vol. 11, no. 1, pp. 260-283, 2023. https://doi.org/10.1093/jssam/smab038.

[23] Z. Zhao, L. Tang, M. Fang, X. Yang, C. Li, and Q. Li, "Toward urban traffic scenarios and more: a spatio-temporal analysis empowered low-rank tensor completion method for data imputation," International Journal of Geographical Information Science, vol. 37, no. 9, pp. 1936-1969, 2023. https://doi.org/10.1080/13658816.2023.2234434.

[24] F. Conceição, C. H. Antunes, M. Gomes, V. Silva, and R. Dinis, "Max-min fairness optimization in uplink cell-free massive MIMO using meta-heuristics," IEEE Transactions on Communications, vol. 70, no. 3, pp. 1792-1807, 2022. https://doi.org/10.1109/TCOMM.2022.3144989.

[25] Y. K. Saheed, A. I. Abiodun, S. Misra, A. S. Albahri, A. H. Alamoodi, B. B. Zaidan, S. Qahtan, H. A. Alsatar, M. S. Al-Samarraay, and A. N. Jasim, "A machine learning-based intrusion detection for detecting internet of things network attacks," Alexandria Engineering Journal, vol. 61, no. 12, pp. 9395-9409, 2022. https://doi.org/10.1109/JBHI.2022.3167256.

[26] V. Gómez-Escalonilla, P. Martínez-Santos, and M. Martín-Loeches, "Preprocessing approaches in machine-learning-based groundwater potential map: an application to the Koulikoro and Bamako regions, Mali," Hydrology and Earth System Sciences, vol. 26, no. 2, pp. 221-243, 2022. https://doi.org/10.5194/hess-26-221-2022.

[27] Y. Shi, X. Nie, X. Liu, L. Yang, and Y. Yin, "Zero-shot hashing via asymmetric ratio similarity matrix," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 5, pp. 5426-5437, 2022. https://doi.org/10.1109/TKDE.2022.3150790.

[28] R. Nirthika, S. Manivannan, A. Ramanan, and R. Wang, "Pooling in convolutional neural networks for medical image analysis: a survey and an empirical study," Neural Computing and Applications, vol. 34, no. 7, pp. 5321-5347, 2022. https://doi.org/10.1007/s00521-022-06953-8.

[29] J. J. Liu, Q. Hou, Z. A. Liu, and M. M. Cheng, "Poolnet+: Exploring the potential of pooling for salient object detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 1, pp. 887-904, 2022. https://doi.org/10.1109/TPAMI.2021.3140168.

[30] U. Nandi, A. Ghorai, M. M. Singh, C. Changdar, S. Bhakta, and R. K. Pal, "Indian sign language alphabet recognition system using CNN with diffGrad optimizer and stochastic pooling," Multimedia Tools and Applications, vol. 82, no. 7, pp. 9627-9648, 2023. https://doi.org/10.1007/s11042-021-11595-4.

[31] X. Li, J. Wang, and C. Yang, "Risk prediction financial management of listed companies based on optimized BP neural network under digital economy," Neural Computing and Applications, vol. 35, no. 3, pp. 2045-2058, 2023. https://doi.org/10.1007/s00521-022-07377-0.

[32] M. S. Uddin, G. Chi, M. A. M. Al Janabi, and T. Habib, "Leveraging random forest in micro-enterprises credit risk modelling for accuracy and interpretability," International Journal of

Finance & Economics, vol. 27, no. 3, pp. 3713-3729, 2020. https://doi.org/10.1002/ijfe.2346.

[33] S. Barra, S. M. Carta, A. Corriga, A. S. Podda, and D. R. Recupero, "Deep learning and time series-to-image encoding for financial forecasting," IEEE/CAA Journal of Automatica Sinica, vol. 7, no. 3, pp. 683-692, 2020. https://doi.org/10.1109/JAS.2020.1003132.

[34] J. Uthayakumar, N. Metawa, K. Shankar, and S. K. Lakshmanaprabu, "Intelligent hybrid model for financial crisis prediction using machine learning techniques," Information Systems and e-Business Management, vol. 18, pp. 617-645, 2018. https://doi.org/10.1007/s10257-018-0388-9.

[35] D. Bianchi, M. Büchner, and A. Tamoni, "Bond risk premiums with machine learning," The Review of Financial Studies, vol. 34, no. 2, pp. 1046-1089, 2020. https://doi.org/10.1093/rfs/hhaa062.