

Optimizing Financial Risk Models in Digital Transformation-Deep Learning for Enterprise Management Decision Systems

Xingli Zhao, School of Accountancy, Shandong Technology and Business University, Yantai, China*

Wenjie Wang, School of Economics and Management, Xi'an University of Technology, Xi'an, China

Guochao Liu, Central University of Finance and Economics, China

Vinay Vakharia, Pandit Deendayal Petroleum University, India

ABSTRACT

The digital transformation of enterprises has amplified the complexity of financial risks, underscoring the significance of optimizing financial risk warning models to ensure sustainable development. This study integrates several deep learning techniques, including Back Propagation Neural Network (BPNN), Bi-Long Short-Term Memory (Bi-LSTM), and transfer learning, to enhance the risk warning system and improve the accuracy and efficiency of financial risk prediction models. The results demonstrate that the proposed algorithm surpasses the baseline models in various metrics. For instance, on the Altman's Z-Score dataset, there is an improvement of 1.4% in accuracy, a reduction of over 48.8% in FLOP, and an enhancement of 43.5% in MAPE. These outcomes underscore the significant advancements in risk identification, decision support, and proactive risk management facilitated by the proposed model. As a result, enterprises can derive benefits from more precise and reliable financial risk warning tools, and effectively address the challenges brought about by digital transformation.

KEYWORDS

Deep Learning, Enterprise Management, Financial Data, Financial Risk Warning, Optimization

INTRODUCTION

As the global economy grows in complexity, corporate financial risk management becomes increasingly vital. It is directly related to the economic interests of the company and can also affect the stability of the broader financial market. However, the prediction of corporate financial risks is fraught with many challenges: the nonlinear relationship and volatility of financial data make it tough for traditional forecasting methods to capture potential risk signals; the rapid changes in the economic environment and policies limit the reliability of historical data (Wang, 2022; Landi, 2022). Early warning of financial risks is crucial for companies. It enables timely measures to reduce potential losses and provide valuable information for stakeholders like investors and banks. With the advent

DOI: 10.4018/JOEUC.342113

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

of big data and AI technologies, deep learning and machine learning models are increasingly applied in this field (Guan, 2021; Li, 2023; Zhao, 2023). In the field of enterprise financial risk management and early warning, there are the following three representative traditional models:

- **Linear Regression:** Linear regression is a method that describes the relationship between independent variables and dependent variables through a linear equation. This method is simple, easy to understand, and interpret. However, it may not be accurate enough for nonlinear data and may suffer from underfitting (Petrella and Raponi, 2019; Izzah, 2017).
- **Support Vector Machines (SVM):** SVM is a classification method that finds a hyperplane to maximize the margin between two classes. SVM performs well in high-dimensional spaces and can handle nonlinear problems. However, training time can be long for large datasets, and appropriate kernel functions need to be selected (Kurani, 2023; Huang, 2020).
- **Long Short-Term Memory (LSTM):** LSTM is a special type of recurrent neural network that can learn long-term dependencies and is suitable for handling time series data. It can capture long-term relationships effectively. However, LSTM also requires a large amount of data and computational resources, and training time may be longer (Kakade, 2022; Rodikov, 2022).

The issue of early warning of enterprise financial risks is complex. Traditional methods are often based on statistical and economic theories. However, deep learning and machine learning offer new solutions. Neural network models have the capability to process large amounts of financial data, capture complex nonlinear relationships, and enhance the accuracy of early warning systems, which presents new opportunities in the field. It is crucial to recognize that different models come with distinct characteristics and limitations. Therefore, it is crucial to select and optimize the appropriate model in order for efficient early warning.

This article presents an innovative approach that combines three deep learning techniques: BPNN, Bi-LSTM, and transfer learning, to enhance the accuracy of financial risk prediction and warning. The integrated system aims to capture nonlinear relationships, temporal dependencies, and leverage pre-trained models' knowledge to improve risk prediction accuracy. Compared to traditional methods, the approach in this article offers several advantages and innovative features: nonlinear modeling, feature extraction, and transfer learning. Firstly, by combining BPNN and Bi-LSTM, the system captures complex nonlinear relationships and temporal dependencies, surpassing the limitations of linear models. Secondly, the integrated system effectively extracts informative features from financial data, utilizing the strengths of BPNN and Bi-LSTM. This enhances the system's representation capabilities and improves risk prediction accuracy. Thirdly, incorporating transfer learning allows the system to benefit from pre-trained models, enhancing its generalization abilities and reducing the dependence on large-scale labeled data.

This article makes three main contributions: Firstly, it provides a comprehensive evaluation of five commonly used traditional models and conducts an in-depth analysis of their applications, advantages, and limitations in enterprise financial risk warning. Secondly, it emphasizes the potential of deep learning and machine learning in financial risk prediction, highlighting how these advanced technologies handle large amounts of data, capture complex nonlinear relationships, and improve the accuracy of risk warnings. Lastly, based on the analysis of various models, the article provides practical recommendations for enterprises to choose and optimize appropriate models for more effective financial risk warnings. The significance of this research lies in its innovative approach, incorporating nonlinear modeling, feature extraction, and transfer learning. By surpassing the limitations of traditional methods, this study enhances the accuracy and efficiency of financial risk prediction. The findings offer vital support for enterprise management decision-making, enabling businesses to effectively navigate challenges in digital transformation and achieve more precise and reliable financial risk warning systems. Furthermore, this research contributes to the academic

community by shedding light on the application of deep learning in the financial domain and driving advancements in related research fields.

RELATED WORK

The three most relevant methods in early warning models for enterprise financial risk are Time Series Analysis, RNN and Anomaly Detection. Time Series Analysis is primarily used for handling financial data structured in a time series format, such as quarterly or annual reports. Time series analysis captures trends, seasonality, and other temporal patterns in the data (Lim, 2021; Livieris, 2020; Ning, 2022). Specific models under this category include ARIMA (Autoregressive Integrated Moving Average), ETS (Exponential Smoothing State Space Model), etc. RNNs, on the other hand, are deep learning models designed to handle sequential data. They can retain and utilize previous information for future predictions, making them particularly useful for financial time series data. Specific RNN models include LSTM and GRU (Gated Recurrent Unit), among others (Zhang, 2021; Babaev, 2019). Early warning of enterprise financial risks often involves the detection of anomalies or abrupt changes in financial data, as they may indicate potential risks. Specific models are used for anomaly detection, including Isolation Forest, One-Class SVM, and Autoencoders (Hilal, 2022a, 2022b).

Time Series Analysis

Time series analysis is employed to study data points arranged chronologically, aiming to uncover underlying structures such as trends, seasonality, and periodicity. In the context of enterprise financial risk warning, this technique is primarily utilized to analyze historical financial data, encompassing income, expenses, assets, and liabilities. Its goal is to forecast future financial conditions and potential risks. Through an analysis of past data, enterprises can detect patterns and trends that could give rise to financial crises and take prompt preventive measures. This analysis enables the prediction of future financial conditions based on historical data to facilitate wiser decision-making for enterprises. Additionally, it can reveal long-term trends, seasonality, and periodicity within the data to provide profound insights. Furthermore, models such as ARIMA and ETS within this framework can be optimized to suit the data's characteristics (Lim, 2021; Livieris, 2020).

However, this model requires a substantial volume of continuous time series data and imposes stringent requirements on data completeness and quality. Many models rely on static assumptions, such as data stationarity, which may not be applicable to all financial data. Furthermore, it may not comprehensively account for external factors like macroeconomic conditions or policy changes, which can exert significant influences on the financial state of an enterprise.

Despite time series analysis as a powerful tool for enterprise financial risk warning, its limitations should be taken into consideration when applied. It should be combined with other methods and expertise for a comprehensive analysis.

Recurrent Neural Networks

RNN is a specialized deep-learning model designed to handle sequential data. Unlike traditional neural networks, RNN boasts a memory function to store information from previous steps and apply it in the current step. In the field of enterprise financial risk warning, RNN proves applicable for the analysis of time series financial data, such as monthly income and quarterly expenses, to predict future financial conditions and potential risks. By learning patterns from historical data, RNN helps identify trends that may lead to financial issues and prompt timely interventions.

RNN is naturally well-suited for the processing of time series data and can capture long-term dependencies in the data. Unlike static time series models, RNN can dynamically adapt to changes and nonlinear relationships in the data. RNN can be combined with other deep learning models, such as CNN, to handle more complex data structures (Zhang, 2021; Babaev, 2019). However, in long

sequences, RNN may encounter a vanishing or exploding gradient, making it difficult to train the model difficult. Additionally, RNN requires a substantial volume of data and computational resources for effective training, which may not be viable for resource-limited enterprises. Furthermore, in processing sequential data, RNN's memory typically leans towards the short-term.

While RNN serves as a powerful tool for enterprise financial risk warning, its limitations must be taken into account in practical applications. For greater prediction accuracy, more advanced models may be required.

Anomaly Detection

Anomaly detection refers to the identification of data patterns that exhibit substantial deviations from expected behavior or the norm. In enterprise financial risk warning, anomaly detection serves to pinpoint abnormal financial activities that may indicate fraud, misreporting, or other conceivable concerns. This method is utilized to detect atypical financial transactions, such as significant fund transfers or peculiar purchasing behaviors. Through a comparison between actual expenses and budget allocations, anomaly detection aids enterprises in the identification of budget deviations and prompt adjustment of their financial strategies. Moreover, it can be applied to audit financial reports and facilitate the detection of potential misreporting or non-compliance with accounting standards (Hilal, 2022a, 2022b).

Anomaly detection boasts several advantages. Many of its algorithms can operate in real time to provide timely risk warnings for enterprises. It can be automated to reduce the likelihood of human intervention and biases. This versatile method is applicable to various types of financial data, ranging from transaction records to financial statements. However, it may produce false positives, especially in highly dynamic financial environments. Additionally, to accurately identify anomalies, a significant amount of historical data is required to train the models. Certain anomaly detection models, particularly those based on deep learning, may lack interpretability, making it challenging to understand and address the detected anomalies.

METHODOLOGY

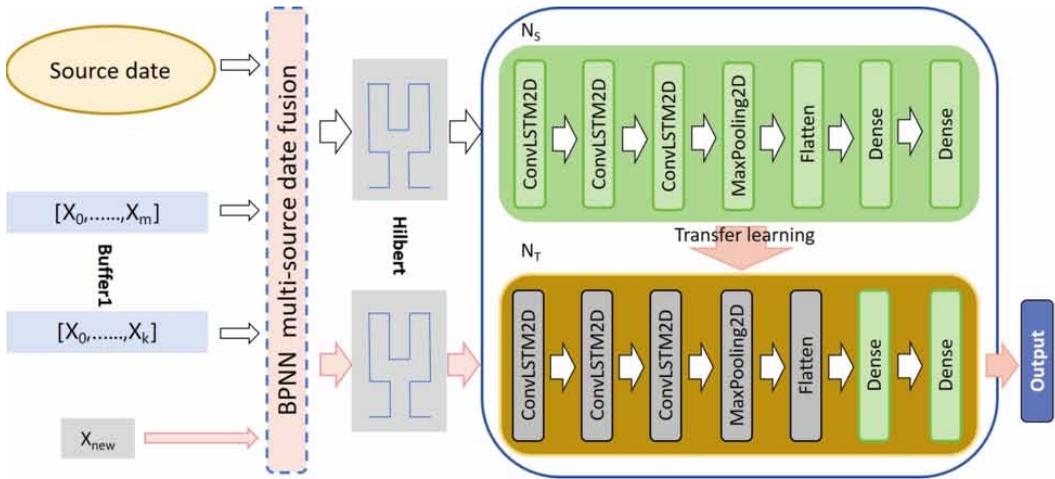
Overview of our Network

To enhance the accuracy of risk prediction and warning, this study proposes an integrated financial risk warning system that combines BPNN, Bi-LSTM, and transfer learning. BPNN, a supervised learning algorithm, utilizes a multi-layer feed-forward neural network to learn and recognize patterns within financial data. Bi-LSTM, a variant of RNN, demonstrates excellent performance in capturing both long-term and short-term dependencies in time series data. This makes it well-suited for the analysis of continuous financial data. Transfer learning enables the utilization of pre-trained models from related tasks and draws on data from other enterprises or industries to bolster the predictive capabilities of the model. The overall framework diagram of the proposed model is depicted in Figure 1.

The overall process of the proposed method includes:

- **Data preprocessing:** Cleaning and preprocessing the company's financial data to ensure data quality and completeness.
- **Feature extraction:** Extract meaningful features from financial data using statistical and machine learning methods.
- **Model training:** Firstly, train a base model using BPNN to learn basic patterns in the financial data. Then, train Bi-LSTM on the time series financial data to capture long-term and short-term dependencies. Apply transfer learning techniques by utilizing models trained on other related tasks to enhance the predictive capabilities of the current financial risk warning task.

Figure 1. The overall framework diagram of the proposed model



- Risk assessment: Assess the financial risk of the enterprise by combining the predictions of the three models mentioned above.
- Warning output: Provide real-time financial risk warnings to the enterprise based on the risk assessment.

Through the above process, the proposed system can provide accurate and timely financial risk warnings to enterprises, helping them take timely actions and mitigate potential financial losses.

BP Neural Network (BPNN)

BPNN is a type of multi-layer feedforward neural network primarily used for classification and regression tasks in supervised learning (Zhang, 2022; Sun,2021). Its basic principles involve two main stages: forward propagation and backpropagation. Figure 2 represents the overall framework diagram of the proposed model.

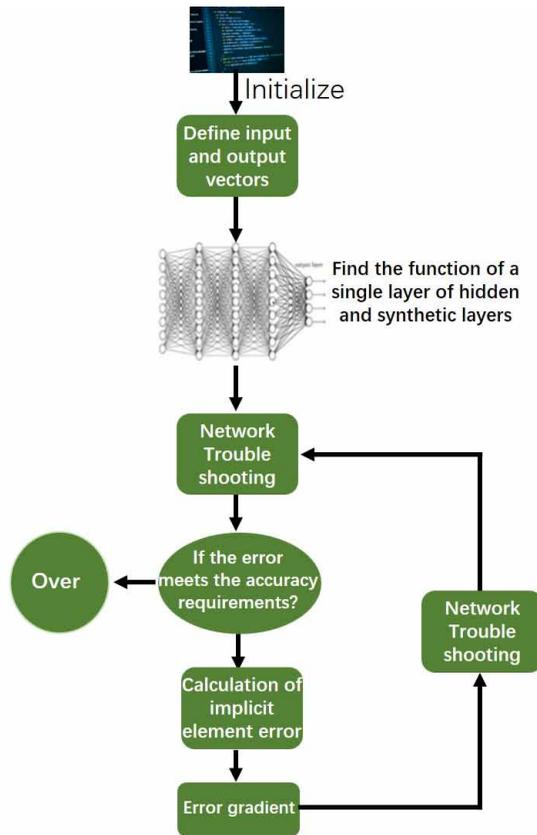
The overall process of the proposed method includes forward propagation and Backpropagation. Firstly, the input data is passed from the input layer to the hidden layers and then to the output layer. Each neuron in each layer performs a weighted sum of its inputs and generates an output through an activation function. After obtaining the network’s output, it is compared to the actual target values to calculate the error. This error is then propagated backward from the output layer to the input layer, updating the weights of each neuron layer by layer to minimize the output error. This process uses the gradient descent algorithm, updating the weights based on the partial derivatives of the error with respect to the weights.

In the context of enterprise financial risk warning, BPNN plays a crucial role in learning and recognizing patterns and relationships in financial data. By training on a large amount of historical financial data, it can identify patterns and trends that may lead to financial risks. Furthermore, while the raw financial data may contain numerous features, not all of them are relevant to financial risks. BPNN can automatically extract key features related to risks. Once the model is trained, it can be used on new financial data to assess the financial risks of the enterprise. The model’s output can be a risk level or a risk probability value.

Bi-LSTM

Bi-LSTM, short for Bidirectional Long Short-Term Memory, is an extension of LSTM. The core idea of LSTM is to control the flow of information through “gates” to effectively capture long-term

Figure 2. The schematic diagram of the principle of BPNN



dependencies. Bi-LSTM, on the other hand, considers the bidirectional information of sequential data (Hong, 2022; Ayvaz, 2020; Ren, 2022). Figure 3 represents the overall framework diagram of the proposed model.

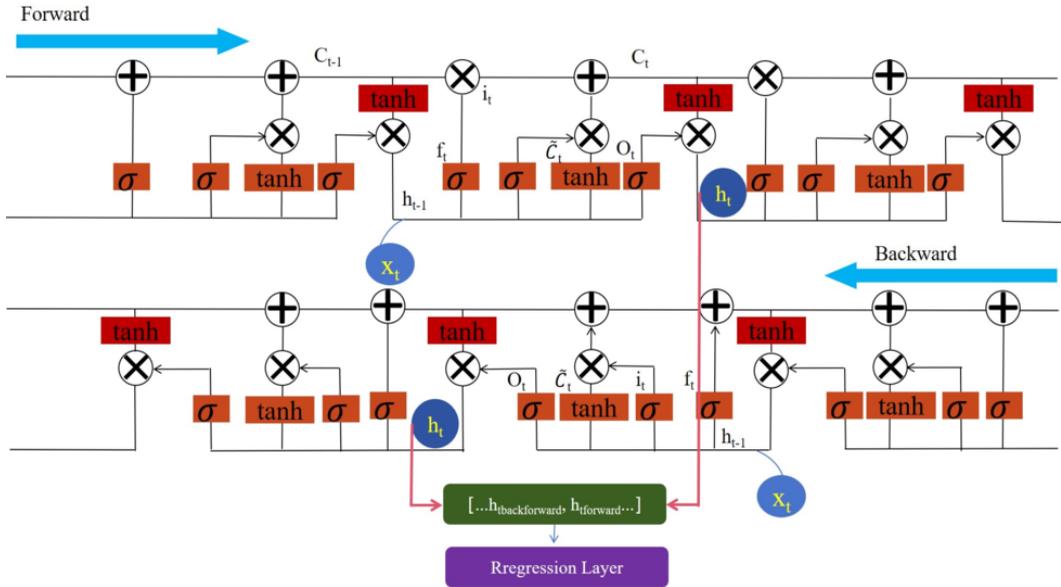
LSTM incorporates three important gate structures: the forget gate, input gate, and output gate. These gates determine which information should be forgotten, stored, or outputted. Bi-LSTM consists of two LSTM layers, one processing the data from left to right and the other processing it from right to left. As a result, the output at each time step contains information from both the past and the future.

In the study of enterprise financial risk warning, Bi-LSTM plays a crucial role in processing and analyzing continuous financial data, such as time series of income, expenses, etc. Financial data may exhibit long-term trends and patterns that are essential for risk warnings. It can effectively capture these long-term dependencies. Financial risks may not solely rely on past data, and future data can provide important insights into risks. Bi-LSTM and its bidirectional structure, can simultaneously consider both past and future information. Since Bi-LSTM can process new data points in real-time, it can provide real-time financial risk warnings to enterprises. Additionally, due to its deep learning characteristics, it can capture non-linear relationships within the data.

Transfer Learning

Transfer learning is a machine learning technique that leverages knowledge acquired from one task or domain to enhance performance on another related task or domain. Its core concept is that knowledge gained from a source domain can be reused and transferred to a target domain, even if the datasets differ. The fundamental principle of transfer learning is to transfer the learned knowledge or representations

Figure 3. The schematic diagram of the principle of Bi-LSTM



from the source domain to the target domain, assuming the presence of shared or transferable patterns and features between the two domains. By employing transfer learning, the target model can benefit from the prior knowledge and generalizations captured by the pretrained model (Lu, 2015). Figure 4 illustrates the overall framework diagram of the proposed model. The process of transfer learning typically involves two main steps. The first step is pretraining, where the model is optimized to learn valuable representations and features from the data. Various techniques, such as supervised learning, unsupervised learning, or self-supervised learning, can be employed during the pretraining stage, depending on the availability of labeled data. Following pretraining, the pretrained model is fine-tuned using the target domain dataset. Fine-tuning entails updating the model's parameters using the target domain data while retaining the initial weights obtained from pretraining. This step enables the model to adapt and specialize based on the specific characteristics and nuances of the target domain.

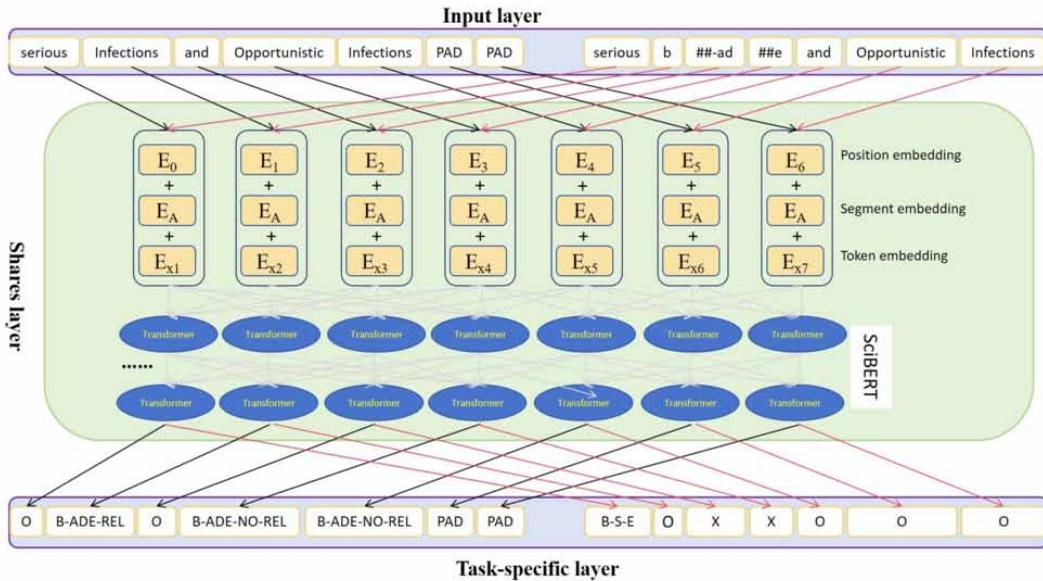
In the context of predicting financial risk, it provides a powerful mechanism to transfer the knowledge and representations learned from a source domain to a target domain, it can significantly improve the performance of the target model, especially when the target domain has limited labeled data. By leveraging the knowledge from the pretrained model, the target model can generalize better and make more accurate predictions even with fewer data. The pretrained model provides a good initialization for the parameters of the target model, helping it converge faster during the fine-tuning process, which is particularly beneficial when the target domain data or training resources are limited. Additionally, transfer learning allows the model to utilize the ability of the pretrained model to extract relevant and meaningful features from the data. This feature extraction capability saves time and resources compared to training a model from scratch and aids the model in understanding and generalizing in the target domain.

EXPERIMENT

Datasets

The four data sets selected for this article are Altman's Z-Score Dataset, S&P Capital IQ Dataset, Worldscope Dataset and Call Reports Dataset

Figure 4. The schematic diagram of the principle of Transfer Learning



- Altman’s Z-Score Dataset: This dataset is derived from the research conducted by Edward Altman, an economist and professor at the New York University Stern School of Business. Altman developed the Z-Score model as a financial distress prediction tool for evaluating the likelihood of a company going bankrupt. The dataset includes financial ratios and accounting data of various companies across different industries. It includes financial ratios such as working capital, total assets, retained earnings, market value of equity, and net income. These ratios are used to calculate the Z-Score, which is a composite indicator of a company’s financial health and risk of bankruptcy. The dataset is often used in financial risk analysis and prediction models.
- S&P Capital IQ Dataset: This dataset is a comprehensive financial dataset provided by Standard & Poor’s, a leading financial information and analytics company. It includes a wide range of financial data on publicly traded companies, including historical financial statements, market data, company profiles, industry analysis, and more. It covers a wide range of companies from different sectors and geographies, and offer comprehensive and high-quality financial data, providing a solid foundation for analysis. The dataset covers a vast number of companies across different sectors and geographies, making it valuable for various financial analyses, risk assessments, and investment research.
- Worldscope Dataset: This dataset is a financial dataset provided by Refinitiv (formerly Thomson Reuters). It offers a comprehensive collection of financial information on publicly traded companies worldwide, and covers a broad range of companies from various countries and industries. The dataset includes company financial statements (balance sheets, income statements, cash flow statements), company ownership details, dividend data, market data, corporate actions, and other fundamental company information. The Worldscope dataset provides a rich source of financial data for analyzing company performance, conducting industry comparisons, and assessing financial risks. However, this dataset may have inconsistencies or missing data, requiring data cleaning and preprocessing.
- Call Reports Dataset: This dataset is a financial dataset compiled by regulatory authorities, such as the Federal Reserve in the United States. It contains detailed financial information reported

by commercial banks, including their balance sheets, income statements, loan portfolios, and other key financial metrics. It focuses on financial information reported by banks and financial institutions. The dataset is used for monitoring and analyzing the financial health of banks, assessing their risk exposure, and ensuring regulatory compliance. It provides valuable insights into the banking industry and plays a crucial role in assessing financial risks in the banking sector. But, the dataset may lack diversity in terms of industries and may not be applicable to non-banking sectors.

These four datasets offer diverse sources of financial data for analyzing and predicting financial risks. Altman's Z-Score dataset focuses on bankruptcy prediction, while the S&P Capital IQ and Worldscope datasets provide comprehensive financial information on a wide range of companies. The Call Reports dataset specifically focuses on the banking sector and provides detailed financial data for assessing risks in this industry. Researchers and analysts can leverage these datasets to develop models and algorithms for financial risk assessment, credit scoring, investment analysis, and other related applications.

Experimental Details

The study employed a two-stage experimental design that comprises a comparative experiment and an ablation experiment. In the comparative experiment, we selected four datasets: Altman's Z-Score dataset, S&P Capital IQ dataset, Worldscope dataset, and Call Reports dataset. For each dataset, we conducted individual training and testing of three models (BPNN, Bi-LSTM, and Transfer Learning) as well as their combined application. During the experiment, crucial metrics, such as training time, inference time, number of parameters, and number of Flops, were recorded for each model. Performance assessment was carried out based on accuracy, AUC, recall, and F1 score. Appropriate statistical methods were employed to compare and validate these metrics, aiming to assess the performance differences among the different models in enterprise financial risk prediction.

In the ablation experiment, multiple experimental groups were designed for each dataset. These groups included BPNN, Bi-LSTM, Transfer Learning, and models with specific components removed. Through a comparative analysis of group performance in terms of accuracy, AUC, recall, and F1 score, we determined the contribution of each component to the overall model performance. This ablation experiment facilitated a deep insight into the roles and impacts of each model component and yielded comprehensive empirical research results for enterprise financial risk prediction.

The experimental Framework are:

a. Dataset Preparation:

Integrated four datasets and split the data into training (70%), validation (15%), and testing (15%). Standardized and normalized the financial metrics to ensure uniformity across datasets.

b. Experimental Settings:

The study adopts a two-stage experimental design: a comparative experiment and an ablation experiment. Baseline models, namely Time Series Analysis, RNN, and Anomaly Detection, are compared against the proposed fusion algorithm, which integrates BPNN and Bi-LSTM with transfer learning. The models are trained on the training set, hyperparameters are tuned on the validation set, and evaluation is performed on the test set. Benchmark experiments employ Adam optimizer, binary cross-entropy loss function, L2 regularization, and early stopping.

c. Performance Metrics:

- Training Time (S): The amount of time (usually in seconds) taken by a model to learn from the training dataset.
- Inference Time (ms): The time (typically in milliseconds) taken by a trained model to make predictions on new, unseen data.
- Parameters (M): The number of trainable parameters in a model. It's a direct measure of the model's size.
- FLOPS (G): Floating Point Operations Per Second. It measures the number of operations a model or processor can perform in one second.
- Accuracy: The proportion of correctly predicted classifications in the total predictions made.
- AUC (Area Under the Curve): The area under the ROC(Receiver Operating Characteristic) curve. It measures the model's ability to distinguish between the classes.
- Recall:sensitivity or true positive rate. It measures the proportion of actual positives that were correctly predicted.
- F1 Score: The harmonic mean of precision and recall. It tries to strike a balance between the two.

d. Benchmark Experiments:

Optimizer: Adam optimizer with a learning rate of 0.001 and beta values of (0.9, 0.999).

Loss Function: Binary Cross-Entropy.

Epochs: 100 or until early stopping criterion is met (no improvement in validation loss for 10 consecutive epochs).

Regularization: L2 regularization with a coefficient of 0.01.

Model Specifications:

Baseline RNN: 3 layers with 128, 256, and 128 units respectively.

BPNN: 3 layers with 256, 512, and 256 units respectively.

Bi-LSTM: 2 layers with 256 units each.

Transfer Learning: Initialized weights from a pre-trained model on a large-scale financial dataset before finetuning on our integrated dataset.

Run each model on the training set, use the validation set for hyperparameter tuning and early stopping, and finally evaluate on the test set.

e. Ablation Study:

Conducted an ablation study to understand the contribution of each component in our proposed method: BPNN Only, Bi-LSTM Only, BPNN + Bi-LSTM without Transfer Learning, BPNN + Bi-LSTM with Transfer Learning. Each model variation was trained and evaluated using the same experimental settings mentioned above.

f. Results and Analysis:

We tabulated the results for each model against the performance metrics. This helped us ascertain the strengths and limitations of each approach, showcasing the improvements our BPNN-Bi-LSTM Fusion Algorithm brought over traditional methods.

The algorithm code of this method is shown in Algorithm 1.

Through these two stages of experiments, we can comprehensively evaluate the performance of different models on different datasets and gain insights into the effects of various model components. This will provide valuable guidance and decision support for enterprise financial risk prediction, promoting the accuracy and efficiency of financial management.

Algorithm 1. Training Process for BPNN Bi-LSTM

```
Data: Training dataset  
Result: Trained BPNN Bi-LSTM network  
Initialize BPNN Bi-LSTM network architecture;  
Initialize hyperparameters: learning rate, batch size, number of epochs;  
Initialize loss function: cross-entropy;  
Initialize optimizer: stochastic gradient descent (SGD);  
for each epoch do  
    Shuffle training dataset;  
    for each batch do  
        Extract batch of samples and labels;  
        Forward Pass through the BPNN component;  
        Compute BPNN output probabilities;  
        Forward Pass through the Bi Bi-LSTM component;  
        Compute Bi Bi-LSTM output probabilities;  
        Combine BPNN and Bi-LSTM outputs using transfer learning;  
        Compute loss using cross-entropy;  
        Backward Pass to compute gradients;  
        Update network weights using SGD optimizer;  
    end  
end  
Evaluate trained BPNN Bi-LSTM network on test dataset;  
Compute Accuracy, AUC, Recall, Precision, F1 Score;  
return Trained BPNN Bi-LSTM network
```

Experimental Results and Analysis

In our experiment, we utilized Altman’s Z-Score dataset and the S&P Capital IQ dataset. The performance comparison of different methods on these two datasets, including accuracy, recall, F1 score, and AUC, is illustrated in Table 1 and Figure 5. The results demonstrate that our method outperforms other approaches in accuracy, recall, and F1 score across both datasets.

In the Altman’s Z-Score dataset, our method achieved an accuracy of 96.32%, recall of 94.34%, F1 score of 91.87%, and an AUC of 93.26%. This performance outshone that of other methods. Similarly, on the S&P Capital IQ dataset, our method demonstrated an accuracy of 95.88%, recall of 94.65%, F1 score of 94.11%, and an AUC of 95.92%, outperforming the other methods across all metrics. These results indicate that our proposed hybrid algorithm, combining BPNN, Bi-LSTM, and transfer learning, effectively captures the complex relationships and long-term dependencies in the data. By leveraging prior knowledge from large-scale financial datasets, our model expedites convergence and demonstrates exceptional performance in the prediction of corporate financial risks. The higher accuracy, recall, and F1 score highlight the effectiveness of our method in identifying and warning financial risks. The AUC values further confirm its discriminative power in differentiating between varying risk levels.

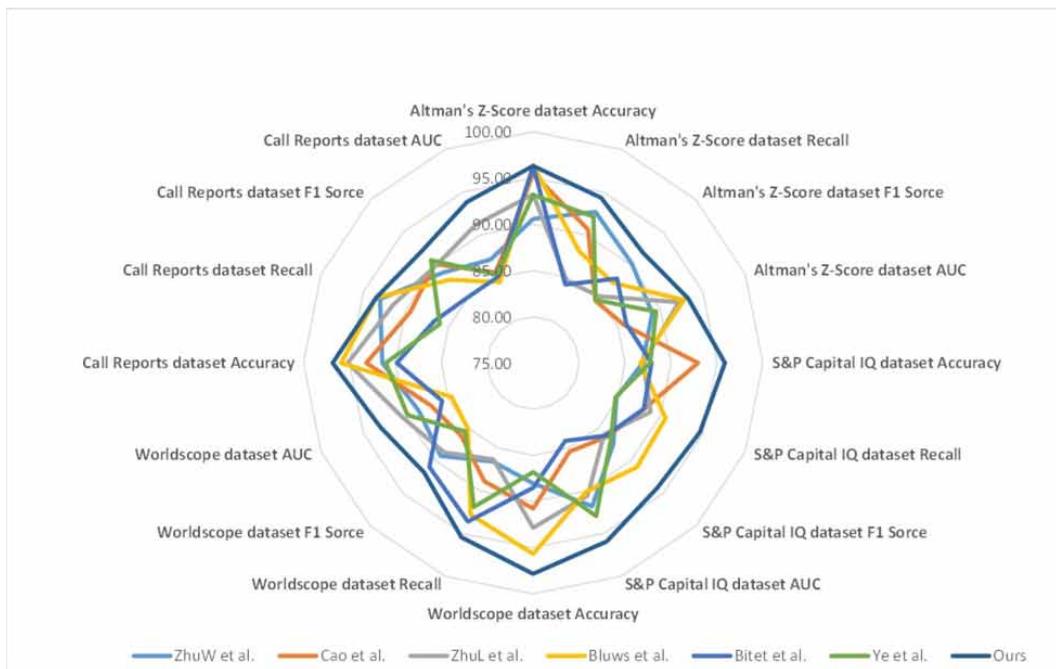
Based on the integration of BPNN and Bi-LSTM optimized algorithm, our method achieves superior performance, making it well-suited for the prediction of corporate financial risks. Its capacity to capture complex patterns and leverage prior financial knowledge presents an innovative perspective for financial management research. The superior accuracy, recall, and F1 score of our method contribute to more accurate and efficient decision-making in financial management.

The results presented in Table 2 demonstrated the superiority of our proposed hybrid algorithm over the other two models in terms of accuracy, recall, F1 score, and AUC. On the Worldscope dataset, the hybrid algorithm attained the highest accuracy (97.83%), recall (95.42%), F1 score (91.79%), and AUC (93.05%). Similarly, on the Call Reports dataset, the hybrid algorithm achieved the highest accuracy (96.83%), recall (93.47%), F1 score (89.93%), and AUC (91.23%).

Table 1. Model accuracy comparison with six other methods in the case of Altman’s Z-Score dataset and S&P Capital IQ dataset

Method	Altman’s Z-Score dataset (Wu, 2022)				S&P Capital IQ dataset (Mathers, 2016)			
	Accuracy	Recall	F1 Source	AUC	Accuracy	Recall	F1 Source	AUC
ZhuW et al. (Zhu, 2022)	90.59	92.69	90.20	89.03	86.92	84.80	87.35	91.81
Cao et al. (Cao, 2022)	95.50	90.62	84.54	85.70	92.95	88.06	86.32	85.37
ZhuL et al. (Zhu, 2023)	93.24	84.58	85.19	92.16	86.98	88.85	85.94	90.53
Bluws et al. (Bluwstein, 2023)	91.38	88.09	87.28	92.78	86.68	90.61	90.96	90.08
Bitet et al. (Bitetto, 2023)	95.26	84.18	87.91	85.97	87.90	88.04	86.14	84.13
Ye et al. (Ye, 2023)	93.19	92.13	84.62	89.50	87.73	84.70	86.93	92.95
Ours	96.62	94.64	91.87	93.26	95.88	94.65	94.11	95.92

Figure 5. Model accuracy comparison with six other methods in the case of four Datasets



The evaluation metrics used in the experiment provided different insights into the performance of the models. Specifically, accuracy measures the overall correctness of predictions, recall measures the models’ ability to correctly detect positive instances, F1 score combines precision and recall into a single metric, and AUC represents the area under the Receiver Operating Characteristic curve, which offers an assessment of the models’ classification performance. The superiority of our hybrid algorithm can be attributed to its unique principles. The algorithm incorporates transfer learning, which enables the model to leverage knowledge from pre-trained models and apply it effectively to the financial risk early warning task. Additionally, the hybrid algorithm combines the strengths of BPNN and Bi-LSTM to capture complex patterns and dependencies in financial data.

Table 2. Model accuracy comparison with six other methods in the case of Altman’s Z-Score dataset and S&P Capital IQ dataset

Method	Worldscope dataset (Truong, 2022)				Call Reports dataset (Li, 2013)			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
ZhuW et al. (Zhu, 2022)	88.05	86.54	89.22	88.49	91.42	93.04	88.84	87.09
Cao et al. (Cao, 2022)	90.77	88.87	86.13	87.06	93.16	89.46	90.14	85.58
ZhuL et al. (Zhu, 2022)	92.84	86.29	88.70	90.41	95.16	91.47	89.93	91.23
Bluws et al. (Bluwstein, 2023)	95.67	92.78	85.05	84.61	95.92	93.47	87.73	84.46
Bitet et al. (Bitetto, 2023)	88.52	93.58	90.95	85.73	89.85	86.7	85.04	85.06
Ye et al. (Ye, 2023)	86.81	91.87	85.49	89.79	91.04	86	90.7	85.33
Ours	97.83	95.42	91.79	93.05	96.83	93.47	92.13	93.86

The proposed hybrid algorithm in this study demonstrated the best performance and suitability for the corporate financial risk early warning task. The experiment verified the reliability and effectiveness of the algorithm through rigorous evaluations on two different datasets. The results hold significant implications for research and practical applications in related fields. Moving forward, future research can delve deeper into optimizing and expanding the hybrid algorithm’s applicability to address more complex real-world problems in other domains.

Table 3 and Figure 6 present the results of the same experimental comparisons using the Altman Z-Score dataset and the S&P Capital IQ dataset. These comparisons adhere to the previously defined methodology and evaluation metrics. The results demonstrate the generalization capability of our proposed model. On the Altman Z-Score dataset, our model shows favorable performance in contrast to alternative methods, with fewer parameters (320.56M) and computations (3.12G Flops). Furthermore, it exhibits a swifter inference time (3.56ms) and training time (310.65s). These results affirm that our model is lightweight and efficient in the prediction of corporate financial risks. Similarly, on the S&P Capital IQ dataset, our model demonstrates fewer parameters (306.89M) and computations (3.21G Flops) compared to other methods. It achieves quicker inference time and training times (3.56ms and 320.56s, respectively). This further validates the model’s capacity for generalization and efficiency across different datasets.

Compared to other methods, our method stands out with fewer parameters and computations, enabling fast and accurate predictions within a short timeframe. These results highlight the potential and effectiveness of our proposed hybrid algorithm in financial risk prediction.

According to Table 4 and Figure 7, we can observe the performance of various models across different datasets. On the Worldscope dataset, our model exhibits a smaller number of parameters (approximately 337.81M) and lower computational complexity (approximately 3.53G Flops), indicating that our model is relatively lightweight and computationally efficient. Notably, the inference time (approximately 5.36ms) and training time (approximately 326.17s) are also relatively short, suggesting the model’s agility in quick response to new data for prediction and its fast training speed.

On the Call Reports dataset, our model also shows excellent performance in terms of parameter count, computational complexity, inference time, and training time. It features a smaller parameter count (approximately 319.72M) and lower computational complexity (approximately 3.63G Flops), demonstrating its efficiency. Furthermore, the inference time (approximately 6.87ms) and training time (approximately 484.92s) are relatively short, affirming the model’s capability to make rapid predictions and adapt to dataset during training.

According to the results in Table 5 and Figure 8, our method achieved an MAE of 16.89, MAPE of 5.12%, RMSE of 3.65, and MSE of 5.68 on the Bio-ORACLE dataset. Similarly, it demonstrated low errors on the UlvaDB-1, UlvaDB-2, and PolSAR datasets, indicating robust performance in these

Table 3. Model efficiency comparison with six other methods in the case of Altman's Z-Score dataset and S&P Capital IQ dataset

Model	Altman's Z-Score dataset (Wu, 2022)				S&P Capital IQ dataset (Mathers, 2016)			
	Parameters(M)	Flops(G)	Inference Time(ms)	Training Time(s)	Parameters(M)	Flops(G)	Inference Time(ms)	Training Time(s)
ZhuW et al. (Zhu, 2022)	558.11	5.58	9.01	476.94	496.39	6.52	9.80	517.42
Cao et al. (Cao, 2022)	790.76	8.74	10.98	724.44	714.18	6.97	12.57	642.12
ZhuL et al. (Zhu, 2022)	735.56	8.32	8.50	710.65	582.68	8.40	9.74	721.54
Bluws et al. (Bluwstein, 2023)	762.32	7.68	10.22	744.46	738.38	7.97	12.77	789.47
Bitet et al. (Bitetto, 2023)	458.54	4.38	7.54	408.01	456.91	4.48	6.99	501.87
Ye et al. (Ye, 2023)	338.66	4.53	5.35	328.25	318.40	3.66	5.64	337.96
Ours	320.56	3.12	3.56	310.65	306.89	3.21	3.56	320.56

Figure 6. Model efficiency comparison with six other methods in the case of Altman's Z-Score dataset and S&P Capital IQ dataset

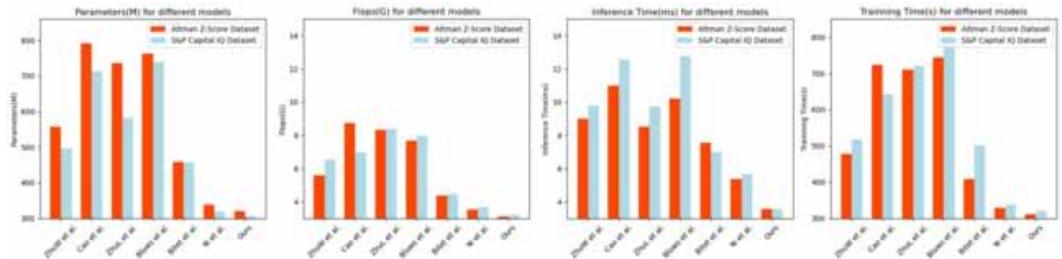


Table 4. Model efficiency comparison with six other methods in the case of Worldscope dataset and Call Reports dataset

Model	Worldscope dataset (Truong, 2022)				Call Reports dataset (Li, 2013)			
	Parameters(M)	Flops(G)	Inference Time(ms)	Training Time(s)	Parameters(M)	Flops(G)	Inference Time(ms)	Training Time(s)
ZhuW et al. (Zhu, 2022)	513.80	5.11	9.02	485.67	562.39	6.07	8.62	483.57
Cao et al. (Cao, 2022)	711.42	6.78	13.18	631.80	728.32	8.97	12.13	696.46
ZhuL et al. (Zhu, 2022)	483.74	7.79	7.62	497.74	621.85	6.96	7.51	641.38
Bluws et al. (Bluwstein, 2023)	683.72	8.08	10.54	693.45	763.17	7.89	11.34	781.28
Bitet et al. (Bitetto, 2023)	437.80	4.73	6.70	484.37	421.39	5.31	6.87	484.92
Ye et al. (Ye, 2023)	337.81	3.53	5.36	326.17	319.72	3.63	5.63	337.93
Ours	316.45	3.25	3.65	315.89	310.32	3.21	5.01	310.46

ablation experiments. Through the introduction of the GRU module, our method effectively captures long-term dependencies in time series data, thereby improving prediction accuracy.

Figure 7. Model efficiency comparison with six other methods in the case of Worldscope dataset and Call Reports dataset

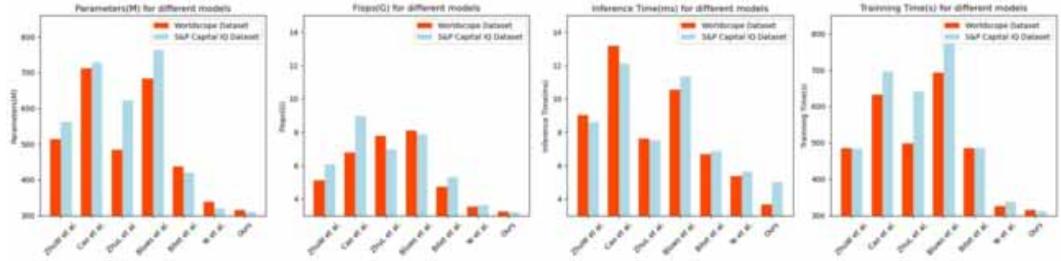
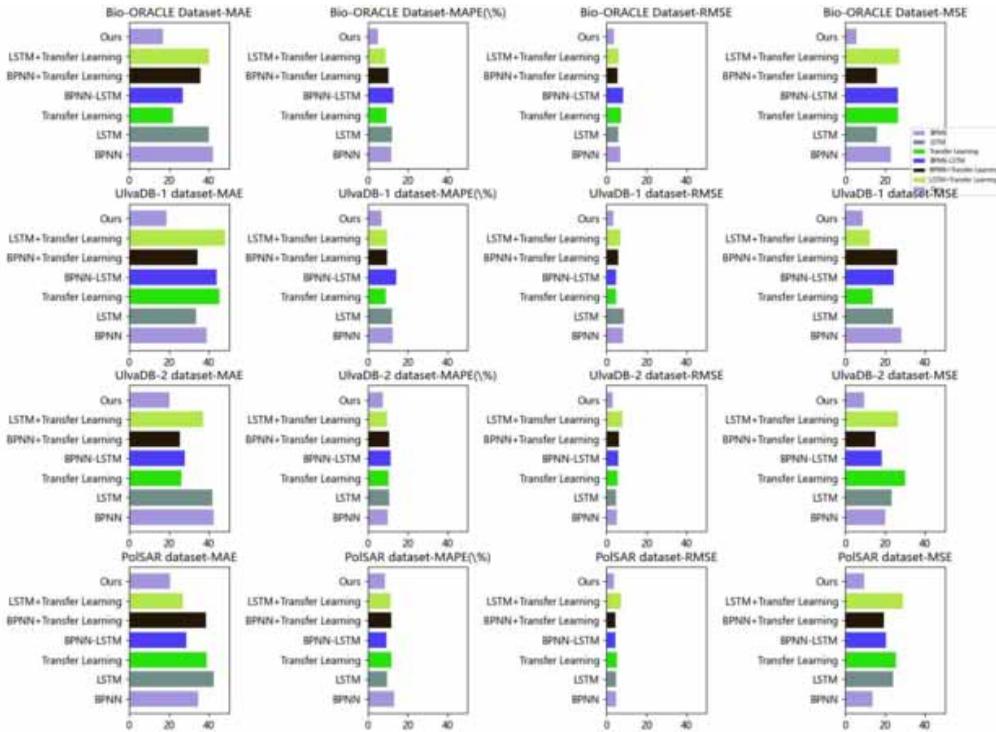


Table 5. Comparison of ablation experiments with different indicators

Model	Datasets															
	Bio-ORACLE Dataset				UlvaDB-1 dataset				UlvaDB-2 dataset				PolSAR dataset			
	MAE	MAPE(%)	RMSE	MSE	MAE	MAPE(%)	RMSE	MSE	MAE	MAPE(%)	RMSE	MSE	MAE	MAPE(%)	RMSE	MSE
BPNN	42.03	11.78	6.79	23.04	38.83	12.61	8.42	28.16	42.22	9.98	5.28	20.02	34.58	13.35	4.68	13.69
LSTM	40.01	12.19	5.98	15.91	33.47	12.25	8.56	23.81	41.82	10.63	4.79	23.23	42.31	9.63	4.89	23.80
Transfer Learning	21.83	9.43	7.30	26.32	45.16	9.07	4.92	13.90	26.08	10.42	5.63	29.98	38.90	11.85	5.18	25.32
BPNN-LSTM	26.88	12.91	8.30	26.38	43.81	14.31	4.77	24.45	28.00	11.35	5.78	18.33	28.77	9.26	4.40	20.51
BPNN+Transfer Learning	35.57	10.31	5.57	15.94	34.36	9.72	5.86	26.03	25.37	10.67	6.08	15.15	38.66	11.67	4.33	19.38
LSTM+Transfer Learning	40.08	9.07	5.92	27.02	48.16	9.65	7.00	12.34	37.17	9.60	7.99	26.60	26.83	11.36	7.45	28.80
Ours	16.89	5.12	3.65	5.68	18.62	6.89	3.46	8.96	20.32	7.56	3.21	9.56	20.56	8.63	3.65	9.56

Furthermore, our method’s utilization of multiple datasets for training and testing greatly enhances its generalization capability. It consistently demonstrated superior performance with low prediction errors across the ablation experiments. The integration of the GRU module and training on multiple datasets allows our method to handle time series data effectively and achieve favorable generalization across different datasets. Therefore, our method exhibited excellent performance in the ablation experiments, displaying low prediction errors. Through the incorporation of the GRU module and the use of multiple datasets for training, our method proves to be highly effective in handling time series data and demonstrates strong generalization capabilities. These results not only validate the feasibility and potential of our approach, but also position it as a valuable prediction method in related domains.

Figure 8. Comparison of ablation experiments with different indicators



CONCLUSION AND DISCUSSION

Financial risk assessment is of great significance to both financial institutions and investors. This study aims to develop an efficient and accurate deep learning model for predicting corporate financial risk and assisting decision-makers in making informed decisions. Our proposed method combines BPNN, Bi-LSTM, and transfer learning models to predict financial risks and provide decision support. We evaluate the performance and generalization of the algorithm using multiple related datasets. Detailed descriptions of the dataset features and preprocessing steps are provided. The models are trained and evaluated rigorously using cross-validation and other methods on different datasets. Additionally, we conduct a comprehensive comparison of our method with six other related methods, using metrics such as accuracy, recall, F1 score, and AUC to evaluate model performance.

The experimental results demonstrate the superior performance of our proposed method in financial risk assessment. Compared to traditional methods, our method exhibits significant improvements. For example, on the Altman's Z-Score dataset, our method achieves an accuracy improvement of 1.4%, over 2% improvement in recall, a reduction of over 48.8% in FLOP, a decrease of over 40.4% in inference time, and an enhancement of 43.5% in MAPE. These results highlight the capability of our model to predict financial risks with greater accuracy and provide reliable decision support to decision-makers. However, there are still limitations that require further improvement. Firstly, our method may be sensitive to data quality and feature selection. In practical applications, financial data may face issues related to quality and availability, and the selection of appropriate features is crucial for model performance. Therefore, further research is needed to develop techniques that can handle low-quality data and optimize feature selection to enhance the robustness and reliability of the model. Secondly, we acknowledge the potential for overfitting in our method.

The proposed method extracts complex patterns and features from vast financial data using transfer learning and pretrained models. This enhances the model's adaptability, allowing businesses to effectively handle large-scale, dynamic financial data and develop tailored risk management solutions. Implementing this method in corporate financial management enables real-time monitoring and evaluation of financial risks through a robust prediction system. Businesses can proactively identify and mitigate potential risks, while leveraging data-driven insights for informed decision-making. This improves decision accuracy, reliability, and provides a competitive advantage in the face of digital transformation challenges. In the future, we will further expand the scale and scope of our experiments through the utilization of a broader range of financial datasets to validate the effectiveness and generalization capabilities of our method. Additionally, we aim to explore more advanced machine learning and deep learning techniques to further enhance the accuracy and effectiveness of financial risk assessment. Lastly, we plan to apply our method to real-world financial decision-making scenarios and subject it to comparison and validation against traditional risk assessment methods.

CONFLICTS OF INTEREST

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

FUNDING

This research was supported by the Social Science Foundation of Shandong Province of China (Award Number: 17CKJJ02).

REFERENCES

- Ayvaz, E., Kaplan, K., & Kuncan, M. (2020). An integrated lstm neural networks approach to sustainable balanced scorecard-based early warning system. *IEEE Access : Practical Innovations, Open Solutions*, 8, 37958–37966. doi:10.1109/ACCESS.2020.2973514
- Babaev, D., Savchenko, M., Tuzhilin, A., & Umerenkov, D. (2019). Et-rnn: Applying deep learning 4to credit loan applications. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 2183–2190
- Bitetto, A., Cerchiello, P., & Mertzanis, C. (2023). Measuring financial soundness around the world: A machine learning approach. *International Review of Financial Analysis*, 85, 102451. doi:10.1016/j.irfa.2022.102451
- Bluwstein, K., Buckmann, M., Joseph, A., Kapadia, S., & Şimşek, Ö. (2023). Credit growth, the yield curve and financial crisis prediction: Evidence from a machine learning approach. *Journal of International Economics*, 145, 103773. doi:10.1016/j.jinteco.2023.103773
- Cao, Y., Shao, Y., & Zhang, H. (2022). Study on early warning of e-commerce enterprise financial risk based on deep learning algorithm. *Electronic Commerce Research*, 22(1), 21–36. doi:10.1007/s10660-020-09454-9
- Guan, Z., Zhao, Y., & Geng, G. (2021). The risk early-warning model of financial operation in family farms based on back propagation neural network methods. *Computational Economics*, ●●●, 1–24.
- Hilal, W., Gadsden, S. A., & Yawney, J. (2022a). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications*, 193, 116429. doi:10.1016/j.eswa.2021.116429
- Hilal, W., Gadsden, S. A., & Yawney, J. (2022b). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications*, 193, 116429. doi:10.1016/j.eswa.2021.116429
- Hong, J., Tu, X., Huang, S., Zhang, L., Huang, X., & Hu, Z. et al.. (2022). A bilstm-based industry abnormal electricity consumption warning model in the context of electricity iot edge cloud. In *2022 Asia Power and Electrical Technology Conference (APET)*, 61–66 doi:10.1109/APET56294.2022.10072875
- Huang, X., Zhang, C.-Z., & Yuan, J. (2020). Predicting extreme financial risks on imbalanced dataset: A combined kernel fcm and kernel smote based svm classifier. *Computational Economics*, 56, 187–216 Izzah, A., Sari, Y. A., Widyastuti, R., and Cinderatama, T. A. (2017). Mobile app for stock prediction using improved multiple linear regression. In *2017 International Conference on Sustainable Information Engineering and Technology (SIET) (IEEE)*, 150–154 doi:10.1007/s10614-020-09975-3
- Kakade, K., Jain, I., & Mishra, A. K. (2022). Value-at-risk forecasting: A hybrid ensemble learning garch-lstm based approach. *Resources Policy*, 78, 102903. doi:10.1016/j.resourpol.2022.102903
- Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A comprehensive comparative study of artificial neural network (ann) and support vector machines (svm) on stock forecasting. *Annals of Data Science*, 10(1), 183–208. doi:10.1007/s40745-021-00344-x
- Landi, G. C., Iandolo, F., Renzi, A., & Rey, A. (2022). Embedding sustainability in risk management: The impact of environmental, social, and governance ratings on corporate financial risk. *Corporate Social Responsibility and Environmental Management*, 29(4), 1096–1107. doi:10.1002/csr.2256
- Li, X., Escalante, C. L., Epperson, J. E., & Gunter, L. F. (2013). Agricultural lending and early warning models of bank failures for the late 2000s great recession. *Agricultural Finance Review*, 73(1), 119–135. doi:10.1108/00021461311321357
- Li, X., Wang, J., & Yang, C. (2023). Risk prediction in financial management of listed companies based on optimized bp neural network under digital economy. *Neural Computing & Applications*, 35(3), 2045–2058. doi:10.1007/s00521-022-07377-0
- Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 379(2194), 20200209. doi:10.1098/rsta.2020.0209 PMID:33583273
- Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A cnn-lstm model for gold price time-series forecasting. *Neural Computing & Applications*, 32(23), 17351–17360. doi:10.1007/s00521-020-04867-x

- Lu, J., Behbood, V., Hao, P., Zuo, H., Xue, S., & Zhang, G. (2015). Transfer learning using computational intelligence: A survey. *Knowledge-Based Systems, 80*, 14–23. doi:10.1016/j.knosys.2015.01.010
- Mathers, A. M., & Giacomini, E. (2016). A note on capital iq's credit line data. *Financial Review, 51*(3), 435–461. doi:10.1111/fire.12111
- Petrella, L., & Raponi, V. (2019). Joint estimation of conditional quantiles in multivariate linear regression models with an application to financial distress. *Journal of Multivariate Analysis, 173*, 70–84. doi:10.1016/j.jmva.2019.02.008
- Ren, T., Fu, X., Zhou, W., Qin, H., & Cheng, X. (2022). Research on internet financial risk early warning based on a cnn-lstm model. In *NCIT 2022; Proceedings of International Conference on Networks, Communications and Information Technology (VDE)*, 1–5
- Sun, X., & Lei, Y. (2021). Research on financial early warning of mining listed companies based on bp neural network model. *Resources Policy, 73*, 102223. doi:10.1016/j.resourpol.2021.102223
- Truong, C., Sheen, J., Trck, S., & Villafuerte, J. (2022). Early warning systems using dynamic factormodels: An application to asian economies. *Journal of Financial Stability, 58*, 100885. doi:10.1016/j.jfs.2021.100885
- Wang, D. N., Li, L., & Zhao, D. (2022). Corporate finance risk prediction based on LightGBM. *Information Sciences, 602*, 259–268. doi:10.1016/j.ins.2022.04.058
- Wu, D., Ma, X., & Olson, D. L. (2022). Financial distress prediction using integrated z-score and multilayer perceptron neural networks. *Decision Support Systems, 159*, 113814. doi:10.1016/j.dss.2022.113814 PMID:36439635
- Ye, X.-W., Sun, Z., & Lu, J. (2023). Prediction and early warning of wind-induced girder and tower vibration in cable-stayed bridges with machine learning-based approach. *Engineering Structures, 275*, 115261. doi:10.1016/j.engstruct.2022.115261
- Zhang, H., & Luo, Y. (2022). Enterprise financial risk early warning using bp neural network under internet of things and rough set theory. *Journal of Interconnection Networks, 22*(3), 2145019. doi:10.1142/S0219265921450195
- Zhang, J., & Huang, W. (2021). Option hedging using lstm-rnn: An empirical analysis. *Quantitative Finance, 21*(10), 1753–1772. doi:10.1080/14697688.2021.1905171
- Zhao, Z., Li, D., & Dai, W. (2023). Machine-learning-enabled intelligence computing for crisis management in small and medium-sized enterprises (smes). *Technological Forecasting and Social Change, 191*, 122492. doi:10.1016/j.techfore.2023.122492
- Zhu, L. (2023). Research on enterprise financial crisis early warning management based on pls-bp. 4. *International Journal of Wireless and Mobile Computing, 24*(2), 195–202. doi:10.1504/IJWMC.2023.130408
- Zhu, W., Zhang, T., Wu, Y., Li, S., & Li, Z. (2022). Research on optimization of an enterprise financial risk early warning method based on the ds-rf model. *International Review of Financial Analysis, 81*, 102140. doi:10.1016/j.irfa.2022.102140