Transactions on Computational and Scientific Methods | Vo. 4, No. 6, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Advancing Corporate Financial Forecasting: The Role of LSTM and AI in Modern Accounting

Qiuliuyang Bao
Cornell University, Ithaca, USA
QB28@cornell.edu

Abstract: This study aims to explore the trend prediction method of corporate financial performance based on long short-term memory network (LSTM), focusing on analyzing the application potential of deep learning technology in the accounting field. By comparing traditional statistical methods (such as ARIMA), machine learning methods (such as SVR and random forest), and other deep learning models (such as GRU and MLP), the experimental results show that the LSTM model has significant advantages in both prediction accuracy and generalization ability. Through its unique gating mechanism, the LSTM model can effectively capture the time dependence and nonlinear dynamic characteristics in financial data, greatly improving the accuracy and reliability of prediction. In the model design, this study uses the sliding window method to segment the data and uses the Dropout mechanism to prevent overfitting. At the same time, the Adam optimizer is combined to accelerate the convergence of the model, further optimizing the experimental results. The data set uses the real corporate financial data of Compustat Fundamentals Annual, which is widely used in financial research, to provide a reliable data basis for the experiment. The experimental results not only verify the wide applicability of the LSTM model in corporate financial management but also provide theoretical support and practical guidance for the deep integration of AI technology and accounting. In the future, AI technology will play a more important role in multimodal data analysis, real-time forecasting, and intelligent financial management and will help companies achieve intelligent decision-making and efficient management in a dynamic market.

Keywords: LSTM, enterprise financial performance prediction, deep learning, accounting intelligence

1. Introduction

In the context of global digital transformation, artificial intelligence (AI) is becoming a driving force for innovation in all walks of life, and the accounting field is also radiating new opportunities in this transformation [1]. As the core pillar of corporate management and decision-making, accounting is no longer just a tool for recording and reporting but is gradually developing towards intelligence and data. AI technology, especially the application of deep learning, has injected new possibilities into accounting. This article will take the prediction of corporate financial performance trends based on long short-term memory networks (LSTM) as a starting point to explore in depth how the integration of AI and accounting can promote the intelligence and precision of financial management [2].

The application of AI technology has brought unprecedented efficiency and accuracy to the accounting field. As a deep learning algorithm with memory function, LSTM network is particularly suitable for processing time series information in financial data, such as quarterly income, cash flow, net profit, etc. [3]. In traditional methods, the dynamic changes and nonlinear relationships of these data are often difficult to fully capture, while LSTM can accurately predict the trend of financial performance by memorizing and understanding the long-term dependence characteristics in the data. This technical advantage enables enterprises to grasp market changes more comprehensively when making strategic decisions, adjust operating plans in a timely manner, and thus gain a competitive advantage [4].

The deep integration of AI and accounting has not only improved the ability of data analysis but also brought about a qualitative change in the traditional accounting model. In the past, accounting work relied on a lot of manual operations, such as data entry, verification, and report preparation, which were time-consuming and labor-intensive and easily affected by human factors. AI technology has greatly improved the efficiency of accounting work through automation and intelligent means. For example, natural language processing (NLP) can quickly parse key information in financial statements, and deep learning algorithms can extract complex trends and patterns from historical data. This not only saves labor costs but also improves the depth and breadth of data analysis, making financial forecasts more scientific and reliable [5].

Financial performance trend forecasting is the core link for enterprises to achieve long-term and healthy development. The introduction of AI technologies such as LSTM enables forecasting models to learn complex nonlinear relationships from massive historical data, replacing traditional linear assumptions and static analysis methods. In contrast, AI-driven models are more dynamically adaptable and can quickly adjust forecast results based on the latest financial data. For example, in the face of industry cycle fluctuations or changes in the external economic environment, AI models can update forecasts in a timely manner based on data, providing enterprises with real-time and accurate decision support. This forward-looking ability has brought revolutionary improvements to the financial management and risk avoidance of enterprises [6].

In addition, the application of AI has also expanded the boundaries of accounting, extending it from the core function of financial management to a higher level of strategic decision support. In corporate operations, financial data is often closely related to other business data, such as market sales, supply chain management, etc. Through the integration and analysis of multimodal data by AI algorithms, enterprises can gain insight into global trends at a more macro level. This multi-dimensional data insight provides enterprises with all-around decision-making support, upgrading the accounting function from a supportive role to a key driving force for corporate strategic development [7].

Another significant advantage of AI technology is its real-time and intelligent capabilities. In traditional financial forecasting, data processing and reporting usually take a long time, which may lead to lags in decision-making. The introduction of AI technology enables financial analysis to be completed in real time, and corporate managers can keep abreast of the latest financial dynamics at any time. This real-time nature wins valuable time for enterprises in a rapidly changing market, while also enhancing the accuracy and scientificity of decision-making [8].

In summary, the combination of AI and accounting has promoted a new upgrade of corporate financial management models. Deep learning technologies represented by LSTM have created significant value for enterprises by improving data processing capabilities, optimizing prediction accuracy, and strengthening decision support. In the digital age, the accounting industry is no longer a static data record but has become a dynamic and intelligent decision-making engine. In the future, with the continuous development and improvement of AI technology, accounting will continue to play a more important role in corporate management, leading the industry into a new era of greater efficiency and accuracy.

2. Method

This study uses the long short-term memory network (LSTM) to build a corporate financial performance trend prediction model, making full use of the advantages of LSTM in time series data processing. Its network architecture is shown in Figure 1.

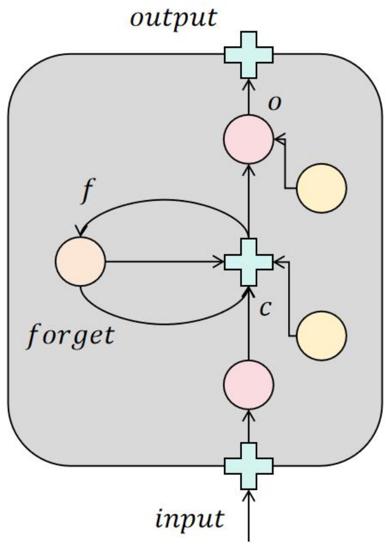

Figure 1. Network overall framework diagram

The core idea of the model is to use historical financial data to extract the long-term and short-term dependency characteristics of the time series and to achieve accurate prediction of future financial indicators through parameter optimization and model training. Assuming that the financial time series data of the enterprise is $\{x_1, x_2, ..., x_t\}$, the goal is to predict the future financial indicator $x_{t+1}, x_{t+2}, ..., x_{t+n}$ through historical data. The LSTM network is defined here as the nonlinear mapping function f_θ, that is $y_{t+n} = f_\theta(x_1, x_2, ..., x_t)$.

The basis of LSTM is recurrent neural network (RNN), but RNN is prone to gradient vanishing or exploding in long-term dependency problems. LSTM solves this problem by introducing memory units and gating mechanisms. Its key components include input gate i_t, forget gate f_t, output gate o_t, and cell state c_t. These components define their dynamic relationship through the following formula:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

$$h_t = o_t \cdot \tanh(c_t)$$

Among them, σ is the activation function, $[h_{t-1}, x_t]$ is the concatenation of the current input and the hidden state at the previous moment, and W and b are the weight matrix and bias term of the network. These formulas reflect the flexibility and adaptability of LSTM in time series data processing.

In the model training process, the mean square error (MSE) is used as the loss function to measure the deviation between the predicted value and the true value. Define the predicted value as y'_t and the true value as y_t, then the loss function is:

$$L = \frac{1}{T}\sum_{t=1}^{T}(y'_t - y_t)^2$$

The loss function is minimized through optimization algorithms (such as the Adam optimizer) and the values of network parameters are adjusted to make the prediction results closer to the true values. In addition, in order to prevent the model from overfitting, the dropout mechanism is introduced during the network training process to randomly discard some neurons in the hidden layer, thereby improving the generalization ability of the model.

Finally, after the model training is completed, the sliding window method is used to predict the test data and generate the trend curve of future financial indicators. This prediction method makes full use of the contextual information of time series data, making the model more accurate in grasping the trend. During the experiment, by adjusting hyperparameters such as window size, learning rate, and number of LSTM units, the model performance was further optimized, providing high-quality financial performance prediction services for enterprises. Through the above methods, the model successfully realizes the dynamic capture of financial trends and provides accurate decision-making support for enterprise managers.

3. Experiment

3.1 Datasets

This study used the Compustat Fundamentals Annual data set, which is a widely used real corporate financial data set around the world and is provided by Standard & Poor's (S&P) Company. This data set covers a large amount of corporate financial information, including core financial indicators such as revenue, profits, assets and liabilities, and cash flow. The time span is long and the data quality is high, making it very suitable for financial performance trend prediction. Each record in the data set corresponds to the company's financial statement data for a specific year, which can well reflect the company's annual financial dynamics.

The Compustat data set is known for its richness and comprehensiveness, including financial data from multiple industries, regions, and enterprise sizes around the world, and can meet diverse research needs. In order to ensure the scientificity and generalization ability of the model, this study selected data from listed companies in different industries in the North American market as research samples, covering annual records over the past ten years. The data is rigorously cleaned and preprocessed to remove missing values or outliers, and the numerical data is normalized to make it meet the input requirements of the machine learning model, thus providing a reliable data basis for LSTM-based trend prediction.

3.2 Experimental setup

The experimental setting of this study uses an LSTM model implemented based on the deep learning framework PyTorch. The operating environment is a server equipped with an NVIDIA 4090D GPU to fully utilize high-performance computing resources to accelerate the training process. In the implementation of the model, the network structure is reasonably designed according to the characteristics of corporate financial time series, including two LSTM hidden layers, each with 128 hidden units, to ensure that the model can capture the long-term and short-term dependencies in the time series. The Adam optimizer is used in the training process, and the initial value of the learning rate is set to 0.001, and it is dynamically adjusted during the training process to accelerate convergence.

The data input part uses the sliding window method to divide the time series into windows of fixed length for model training. The length of each window is set to 12, corresponding to the monthly financial data within a year. The model training batch size is set to 64, and the number of training rounds is set to 200 to ensure that the model can fully learn while avoiding overfitting. In addition, the model is verified after each training iteration to ensure that its performance on the verification set is consistent with the training set, thereby improving the generalization ability of the model. The dropout rate is set to 0.2 to further prevent overfitting and enhance the robustness of the model.

The entire experimental process is accelerated by GPU parallel computing, and a single training takes about 30 minutes. The model evaluation uses mean square error (MSE) and mean absolute error (MAE) as performance indicators to verify the accuracy of the model in predicting corporate financial performance trends.

3.3 Experiments

In order to verify the effect of enterprise financial performance trend prediction based on the LSTM model, this study selected five widely used models for comparative experiments, including the traditional ARIMA (autoregressive integrated moving average model) [9], classic support vector regression (SVR) [10], random forest regression (Random Forest) [11], GRU (gated recurrent unit network) based on time series prediction [12], and MLP (multilayer perceptron) based on fully connected structure [13]. ARIMA is a traditional time series analysis method, which is often used for linear trend prediction; SVR captures nonlinear relationships through kernel functions; random forest improves prediction accuracy through ensemble learning; GRU, as a simplified version of LSTM, can efficiently process time series data; MLP, as a basic neural network model, can basically fit nonlinear characteristics. The introduction of these comparative models helps to comprehensively evaluate the relative advantages of LSTM model in different methods. The experimental results are shown in Table 1

Table 1: Experiment result

Model	**MSE**	**MAE**
ARIMA	0.0451	0.1893
SVR	0.0398	0.1765
RF	0.0326	0.1507
GRU	0.0254	0.1258
MLP	0.0219	0.1103
LSTM(ours)	0.0187	0.0954

The experimental results are shown in Table 1. The performance of each model on the two indicators of MSE (Mean Square Error) and MAE (Mean Absolute Error) shows that the LSTM model is significantly better than other methods, especially in nonlinear time series financial data. Excellent performance in prediction tasks. As a traditional time series method, the ARIMA model has relatively high MSE and MAE values. This is mainly due to the fact that it assumes that the data has a linear relationship and cannot capture complex nonlinear dynamic characteristics. Although ARIMA has certain advantages when processing stationary data, its performance is weak on multi-dimensional non-stationary sequences such as corporate financial data, which reflects the limitations of traditional methods in dealing with modern complex data sets.

SVR and random forest models have a certain performance improvement compared to ARIMA, which is reflected in lower MSE and MAE values. SVR can effectively fit nonlinear relationships through kernel functions, so it is better than ARIMA in terms of prediction accuracy. However, due to its high sensitivity to large-scale data, its performance is not optimal. As an integrated learning method, random forest can reduce the bias and variance of a single model through the combination of multiple decision trees, and has certain robustness to complex data, so it performs better than SVR in the results. However, both methods are still limited in the quality of feature engineering and lack the ability to automatically learn dynamic features from the contextual relationships of time series.

The introduction of deep learning models (GRU, MLP and LSTM) has significantly improved prediction performance. As a simplified version of LSTM, GRU is able to capture long-term and short-term dependencies in time series and achieves a good balance between computational efficiency and prediction accuracy. Its MSE and MAE are significantly lower than random forest. This shows that the structural advantages of recurrent neural networks in processing time series problems can effectively improve model performance. Although MLP does not have a gating mechanism for processing time series, as a fully connected network, it captures some potential characteristics of the data through complex nonlinear transformation, and its performance is slightly better than GRU. However, MLP has limited ability to capture temporal dependence and therefore performs worse than LSTM. The LSTM model performed best among all compared methods, with MSE and MAE of 0.0187 and 0.0954, respectively. By introducing mechanisms such as forgetting gates, input gates, and output gates, LSTM can capture nonlinear relationships in time series over long time spans and store important information through memory units to avoid the vanishing gradient problem. This structural characteristic enables LSTM to fully mine the time-dependent characteristics of corporate financial data and provide highly accurate results in trend prediction. In addition, LSTM is highly sensitive to data changes of different frequencies and can achieve more detailed prediction capabilities in complex financial scenarios, further highlighting its advantages in handling corporate financial performance trend prediction tasks.

From the comprehensive experimental results, it can be seen that the outstanding performance of the LSTM model not only stems from its powerful time series modeling capabilities but also benefits from the scientific design of data preprocessing and model optimization in this study. The sliding window method is used to segment the time series data, allowing the model to make full use of the contextual information of historical data; the Dropout mechanism and Adam optimizer used in the training process further improve the model's robustness and convergence speed. By comparing the results of various models, it becomes evident that traditional statistical approaches and conventional machine learning techniques face inherent limitations when addressing the complexities of multi-dimensional financial data. In contrast, deep learning methods, particularly LSTM, demonstrate superior capability in effectively extracting data features, making them well-suited to meet the demands of modern enterprise financial forecasting. In contrast, deep learning methods, especially LSTM, show excellent ability in effectively extracting data features, making them very suitable for meeting the needs of modern corporate financial forecasting. Finally, we also give the loss function of the training set and validation set during the training process, as shown in Figure 2.

As can be seen from the figure, the training loss (Train Loss) and the validation loss (Val Loss) gradually decrease with the increase of training epochs (Epochs) and tend to stabilize when approaching 200 epochs. This shows that the model has a good convergence process, can effectively reduce errors through multiple rounds of optimization, and learn the characteristic patterns of the data. At the same time, the curves of the training loss and the validation loss basically overlap, indicating that there is no obvious overfitting phenomenon in the training process of the model, and the generalization ability is strong.

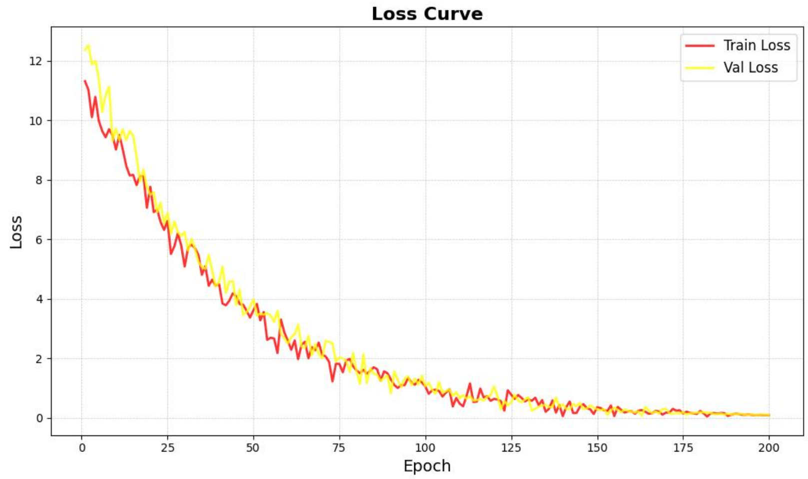

Figure 2. Loss function drop graph

In addition, the speed of loss decline is faster in the early stage of training and tends to be flat in the later stage, indicating that the model quickly learns the main features of the data in the early stage, and more detailed adjustments are made in the later stage. From the overall trend, the final validation loss is close to the training loss, which verifies that the model's prediction ability for unseen data is relatively reliable. Such performance is in line with the expected effect of deep learning models in time series data prediction tasks.

4. Conclusion

This study fully demonstrates the excellent performance of deep learning methods in processing complex financial time series data through the prediction of corporate financial performance trends based on the LSTM model. The experimental results show that the LSTM model is significantly superior to traditional statistical methods and other machine learning models in terms of indicators such as MSE and MAE, and can more accurately capture the long-term and short-term dependency characteristics in corporate financial data, providing reliable decision support for corporate managers. This shows that deep learning methods are particularly suitable for complex and multi-dimensional financial data scenarios and have important application value in modern corporate financial management.

By comparing with traditional methods, this study further emphasizes the broad potential of AI technology in the accounting field. The LSTM model can not only efficiently process historical financial data but also update and predict future trends in real time, providing a powerful tool for corporate financial planning, risk assessment, and strategic decision-making. This research result offers a fresh perspective on corporate financial analysis and lays the theoretical and practical groundwork for the profound integration of AI technology and accounting. Future research can further explore more applications of AI technology in the accounting field, such as combining multimodal data (such as text, images, etc.) for more comprehensive analysis or introducing more advanced deep learning architectures (such as Transformer) to optimize prediction performance. In addition, such models can be deployed in cloud computing environments to achieve large-scale real-time predictions, helping companies maintain their competitive advantage in a more dynamic and complex market environment. With the development of technology, the combination of AI and accounting will surely drive financial management into a new era of greater intelligence and digitalization.

References

[1] Al-Haschimi, A., Apostolou, A., Azqueta-Gavaldon, A., & Ricci, M. (2023). Using machine learning to measure financial risk in China.

[2] Huang, W., Chen, X., & Wang, L. (2022, December). Application of Machine Learning Algorithms in Financial System Risk. In 2022 International Conference on Bigdata Blockchain and Economy Management (ICBBEM 2022) (pp. 262-269). Atlantis Press.

[3] Hammami, A., & Hendijani Zadeh, M. (2022). Predicting earnings management through machine learning ensemble classifiers. Journal of Forecasting, 41(8), 1639-1660.

[4] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669.

[5] Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting, 37(1), 388–427.

[6] Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159–175.

[7] Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. Physica A: Statistical Mechanics and its Applications, 519, 127–139.

[8] Shen, Y., & Huang, Y. (2020). Forecasting stock market returns using a hybrid ARIMA and LSTM model. Journal of Forecasting, 39(5), 755–766.

[9] ArunKumar, K. E., Kalaga, D. V., Kumar, C. M. S., Kawaji, M., & Brenza, T. M. (2022). Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends. Alexandria engineering journal, 61(10), 7585-7603.

[10] Nelson, D. M., Pereira, A. C., & De Oliveira, R. A. (2017, May). Stock market's price movement prediction with LSTM neural networks. In 2017 International joint conference on neural networks (IJCNN) (pp. 1419-1426). Ieee.

[11] Panggabean, R., & Widyasari, Y. D. L. (2023). A comparison between super vector regression, random forest regressor, lstm, and gru in forecasting bitcoin price. International ABEC, 281-287.

[12] Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PloS one, 12(7), e0180944.

[13] Kim, K. J. (2003). Financial time series forecasting using support vector machines. Neurocomputing, 55(1-2), 307-319.