International Journal of Innovative Computing, Information and Control
Volume **21**, Number **1**, February **2025**
ICIC International ©2025 ISSN 1349-4198
pp. 173–184

AN LSTM-BASED ALGORITHM FOR PREDICTING STOCK PRICE VOLATILITY FROM MULTI-FACTOR ANALYST PREDICTION DATA

Jianming Li[1], Xiaojun Chen[1], Suhao Guan[2] and Xiangpei Hu[3]

[1]School of Computer Science
[2]Finance Department
[3]School of Management and Economics
Dalian University of Technology
No. 2, Linggong Road, Ganjingzi District, Dalian 116023, P. R. China
{ lijm; guansuhao; drhxp }@dlut.edu.cn; 877540927@qq.com

Received April 2024; revised July 2024

Abstract. *In the complex financial markets, analysts' forecasts provide crucial market insights and data support for investors, serving as a vital source of information for many investment decisions. This paper thoroughly investigates the multiple factors influencing the quality of analysts' predictions, explores individual characteristics of analysts, and based on these, redesigns factors related to analysts' forecasts. We further developed stock factors closely related to analysts' predictions, together forming a multi-factor data set. Addressing the limitations of existing models in processing data sets of various types and multiple time scales that include analysts' predictions and stock factors, this paper introduces an improved Inception-ALSTM model for stock selection prediction. The model utilizes an Inception structure to effectively capture key features affecting stock prices across different time scales, thereby better adapting to the diversity of factor time scales. Additionally, to address the differences in the importance of various feature dimensions within the data set, the model incorporates an attention mechanism. Experimental data show that this method improves the accuracy of stock price movement prediction experiments, achieving an accuracy rate of 0.645 in the task, thus proving the model's effectiveness and providing a reference for future stock selection strategies using analyst data.*

Keywords: Analyst forecast, Multi-factor data set, LSTM, Self-attention, Inception

1. **Introduction.** In recent years, with the growth of the economy, an increasing number of investors have been choosing stocks as a means to expand their investment portfolios [1]. However, due to the rapid development of big data on the Internet, the securities market generates a vast amount of complex and constantly changing investment information every day, making it challenging for most investors to achieve satisfactory returns. An area of current intense research is the processing and evaluation of vast amounts of heterogeneous data from multiple sources in the financial market, utilizing deep learning and other advanced techniques to provide investors with more precise investment recommendations.

Analysts play a crucial role in providing market information. Their analyses often help explain why stock prices experience frequent and significant movements that are not solely based on company fundamentals [2]. Utilizing their expertise and access to information, analysts conduct thorough and regular evaluations of listed companies' investment potential. Beyond offering professional insights, analysts also shape investors' perceptions of individual stocks, thereby influencing stock price changes. According to a study, analyst

DOI: 10.24507/ijicic.21.01.173

research reports on earnings projections and investment ratings carry greater informational value compared to publicly available market information [3]. Research has shown that using analyst ratings and tracking their changes can substantially enhance investment returns, affirming analysts' ability to provide valuable insider information [4]. However, the impact includes both positive and negative effects. Analyst bias may increase the risk of stock price crashes [5]. These research findings indicate that analyst forecasts are among the factors influencing stock market fluctuations. By analyzing these forecasts, investors can promptly identify current trends and hidden information in the market, thereby developing investment strategies that mitigate risks and yield superior returns.

In addition to analyst forecasts, researchers are also constantly exploring how to use other data sources to find more factors that may affect the stock market changes. The multi-factor model has always been one of the main quantitative models in financial big data research because of its comprehensiveness and stability. Compared to single-factor models, multi-factor models can comprehensively consider various factors affecting stock price trends. Researchers establish a stock selection model by quantifying the influence of different factors on the stock's rate of return. In the realm of fundamental factor models, notable examples are the three-factor model proposed by Fama and French [6], as well as the four-factor model devised by Carhart [7]. With the continuous development of financial big data and artificial intelligence technology, multi-factor models are constantly evolving and improving. For example, a study introduced a factor model to aggregate volatility risk, underscoring the importance of integrating both aggregate and idiosyncratic volatility risks when analyzing stock returns across different sectors [8]. Furthermore, there are studies involving the segmentation of other dimensional data, such as industry data and the investment characteristics of Hong Kong-funded institutions, to construct a multi-factor data set combined with machine learning models for stock price prediction, achieving results that are superior to existing machine learning models [10,11]. Researchers have begun to explore machine learning models that are better suited for handling complex data and have successfully applied these models to stock price prediction. Some researchers have used daily stock data as input, employing CNN for feature extraction and LSTM to learn these extracted features in order to predict stock closing prices [12]. In 2019, Bhattacharjee and his team compared the performance of traditional statistical methods and machine learning approaches in predicting stock closing prices, finding that machine learning methods, especially neural network models, were superior in terms of prediction accuracy. Among neural network models, LSTM achieved the best results due to its unique ability to capture long-term dependencies in sequence data [13]. Overall, multi-factor data sets and models combined with machine learning technology have widespread applications in the financial domain, particularly in the area of stock price prediction.

The selection of factors in multi-factor models holds significant importance, and considering analyst forecasting is one perspective worth exploring. It has been observed that forecasts disseminated by teams display more pronounced herding behavior and less timeliness compared to forecasts issued individually [14]. This means that the individual dimension of analyst factors can be further explored and analyzed to generate a novel type of financial factor data. In addition, the data of stocks also play a role in the fluctuations of stock prices. Changes in these data points can directly or indirectly influence the ascent and descent of stock prices [15]. However, not all stock factors are useful; useless ones instead generate noise and need to be refined to closely align with analyst forecasts. Therefore, this paper considers a multi-dimensional perspective in data set construction, merging analyst forecast data with relevant historical stock data. This study also incorporates both multi-factor models and machine learning, enhancing the traditional LSTM model to better handle the constructed multi-time scale and multi-dimensional data set.

The rest of our paper is structured as follows: Section 2 discusses the construction of the multi-factor data set, Section 3 describes the model construction in detail, Section 4 provides experiments to demonstrate the effectiveness of the Inception-ALSTM model, and Section 5 summarizes this work.

2. **Factor Construction.** We obtain our data from the China Stock Market & Accounting Research (CSMAR) database. The database contains extensive securities-related data for Chinese listed companies from 2001. The data spans from September 2009 to March 2023. We also obtain the basic market data and financial indicator data of stocks from the Tushare Pro financial data interface package. Based on the aforementioned raw data, a multi-factor data set is constructed, as shown in Figure 1.

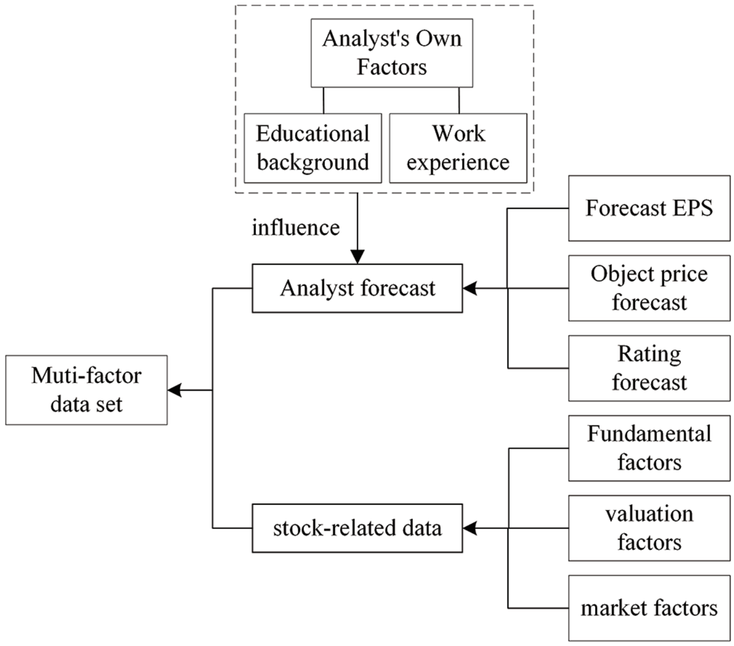


FIGURE 1. Composition of multi-factor data set

Since the accuracy and reliability of analysts' forecasts depend on individual characteristics such as their understanding of the market and industry, we considered the educational background and work experience of analysts when constructing scores for the importance of their forecasts. The analyst forecast factor is jointly constructed based on the analysts' predictions for future company earnings and their scores for the importance of these prediction indicators. It includes factors such as the analysts' forecasts for earnings per share, object price predictions, and analyst rating forecasts. The stock-related factors include fundamental factors, valuation factors, and market factors. These factors assist analysts and investors in understanding the degree of deviation in current market valuations and enable them to study stock price predictions based on these factors.

The following describes some relevant information about several important dimensions and explains the calculation methods for their key factors.

2.1. **Analyst's own factors.** Research indicates that the proficiency of securities analysts in stock recommendations is profoundly shaped by their personal attributes, including their work experience, knowledge base, and analytical competencies [16]. This study quantifies the analyst's work experience and education background, constructing a factor for predicting analyst accuracy that encompasses personal attributes.

2.1.1. *Quantification of analyst work experience and education background.* Working years are also an important factor to measure the work experience of securities analysts [17]. Given the securities market's high-risk nature and its constant evolution, analysts must continually learn to stay abreast of market developments. Being qualified for the post of securities analyst for a long time shows that the knowledge and experience of the securities analyst are recognized by brokerages and investors. Therefore, referring to the general industry standards and data set characteristics, this study categorizes the analysts' years of service into various levels, calculated by the number of quarters since the analyst made their first profit forecast until the end of this year. For example, the categories include "0-8 quarters", "8-20 quarters", "over 36 quarters", etc. These categories are then normalized as shown in Table 1. Although education background is not directly related to the accuracy of securities analysts' prediction, it not only represents the knowledge mastery ability of securities analysts at the present stage, but more importantly, it reflects the learning ability of securities analysts in the future [18]. Since the research in this paper mainly focuses on the forecasting indicators of analysts in research reports, and a research report is often analyzed by a team of several analysts, this paper deals with this situation as follows: The average educational background of the analysts within a team is used as the educational factor. The teams are categorized based on different levels of educational backgrounds, employing a scoring method that uses A, B, C, and D to represent different weights. These grades are used to weigh the average blend of the team's educational background and work experience, as shown in Table 1.

Table 1. Quantitative standards for work experience and educational background

Label	Work experience	Education background
A	Over 36	Doctor
B	20-36	Master
C	8-20	Bachelor
D	0-8	Undergraduate or below

2.1.2. *Analyst forecast importance score.* After comprehensively considering both the educational background and work experience of analysts, the importance score factor of the analyst forecast is constructed as one of the reliabilities of the analyst forecast indicators. The specific calculation method is as follows:

$$score(analyst, importance) = exp_{label} + edu_{label} \tag{1}$$

In the formula, exp_{label} is the corresponding rating of the analyst's work experience, and edu_{label} is the corresponding rating of the analyst's education. According to the Matthew effect and the theory of 'learning by doing', both significantly influence the accuracy and reliability of an analyst's predictions. The analyst prediction importance score factor proposed in this paper aims to assess the credibility of different analysts' forecast indicators. Assuming the analyst's forecasting ability remains constant over a period, and the value of this factor is updated annually.

2.2. **Analyst forecast indicators.** This study combines the analyst forecast importance factor constructed in the previous step with the analyst forecast indicator data to reconstruct the analyst forecast indicators. These include analysts' rating forecast for the stock, analysts' object price predictors for the stock, and analysts' forecast earnings per share. the construction method is as follows.

Analyst forecast rating factor refers to a predictor constructed based on analysts' evaluation of a company's financial performance and stock market performance. The following are the general steps for constructing analyst forecast rating factors: Analyst ratings can be divided into multiple levels, such as "buy", "hold", and "sell". For different grades, as shown in Table 2, different labels are given to reflect their different influences on the future performance of the stock.

Table 2. Rating quantification criteria

Label	Stock rating	Investment returns over the next six months
1	Buy	Lead the broader market index by more than 10%
2	Increase	Lead the broader market index by 4% to 10%
3	Hold	Between −2% to 4% relative to the broader market index
4	Reduce	Lag the broader market index by 2% to 10%
5	Sell	Lag the broader market index by more than 10%

Analysts' object price prediction is usually based on the analysis of the company's fundamentals and the prediction of the market environment, so it can provide a reference for investors to a certain extent. In the field of machine learning, models can be trained to predict the future trend of stock prices by analyzing historical stock price data and company fundamental data. This can help investors formulate trading strategies, while the predictions of machine learning models can also provide reference for analysts' research reports, improving the accuracy and credibility of research reports.

The specific calculation method of the analyst's object price predictor for the stock is as follows:

$$fprice = (predict - close)/close \tag{2}$$

$$sfprice = \sum_{i=1}^{n} (fprice * S_i) \Big/ \sum_{i=1}^{n} S_i \tag{3}$$

where *fprice* is the analyst's object price predictor for the stock, *predict* is the analyst's forecast data for the stock price disclosed in the research report, and *close* is the stock's closing price on the day. This factor can indicate whether the analyst is optimistic about the future of the stock and the degree of optimism. *sfprice* is the object price predictor of the stock based on the analyst's own characteristics. S_i is the prediction accuracy factor of the i analyst, and n is the number of analysts who disclosed their forecasts for the stock in a certain time window.

EPS is usually used to reflect the operating results of enterprises, measure the profit level of common shares and investment risk, and then make relevant economic decisions. If analysts predict that a stock's earnings per share will increase, it means that the company's future profitability is likely to improve, making investors more bullish on the stock, which in turn pushes up its share price, and vice versa. Therefore, this indicator has some influence on the future rise and fall of the stock.

2.3. **Stock related data.** Analysts typically consider various factors before releasing stock forecasts, including a company's financial situation, the valuation accorded by the stock market, and the stock's behavior in the market. Analysts' predictions are also subjectively influenced by market sentiment. Incorporating fundamental analysis of stocks can provide a more objective perspective. Therefore, this article deeply expands upon the economic interpretation of factors, comprehensively considering stock performance and potential from multiple dimensions. It constructs three major categories of factors from the perspectives of the financial conditions of the stock companies, stock market valuation, and stock market performance, totaling 64 factors. The factors constructed from the financial condition of stock companies include revenue, profit, and debt, etc., reflecting the company's operational efficiency and profitability. The core of the stock market valuation factors lies in quantifying the relationship between the market price of a stock and its intrinsic value, providing investors with a benchmark, including EPS and the Historical Percentile of P/E Ratio, etc. Stock market factors typically stem from trading conditions and reflect changes in stock prices and trading activity, including factors such as closing price, price fluctuations, and trading volume. Part of the factors used are shown in Table 3.

Table 3. Stock related data

Factor name	Factor action
EPS	A key indicator for measuring a company's profitability
ROE	Return on Equity (ROE) measures the effectiveness of a company's use of capital to create profits
...	...
Revenue_ps	An indicator for evaluating a company's sales ability and market share based on per share operating income

The calculation formula for the Historical Percentile of P/E Ratio factor is as follows: The P/E ratio is the percentile of the current P/E ratio in the historical P/E ratio. This concept can help investors better assess the level of a stock's valuation. Formula (4) is the most representative P/E historical percentile factor, where P_{hist} is the stock's historical P/E percentile, where Day_{lower} is the number of trading days historically below the current P/E value, and Day_{total} is the total number of historical trading days.

$$P_{hist} = \frac{Day_{lower}}{Day_{total}} \times 100\% \tag{4}$$

The P/E ratio predicted by analysts is usually based on factors such as the company's financial data, industry trends and other factors, but these factors are not necessarily reflected in the historical P/E ratio data. When considering these two indicators, investors can evaluate the valuation level of stocks from different angles. If the stock quantile is very low, indicating that the stock's current P/E level is lower than the historical P/E data, and the analyst's P/E forecast is higher, it can be considered that the stock has potential for growth and attractive investment opportunities.

3. **Overview of the Model.** This article presents an improved attention Long Short-Term Memory (LSTM) model specifically designed to fuse analyst data and stock related data called Inception-ALSTM. As shown in Figure 2, the architecture of the model consists of five layers. The input to the model consists of two parts: the analyst forecast data containing the analyst's personal characteristics and the stock related factors. These diverse factor data are converted into a vector format and combined at the input layer to form a comprehensive feature vector. Parallel Convolution and LSTM networks are used for feature extraction to obtain the local change information of factors and the global information of timing. Due to the different contribution degrees of factors, the attention mechanism will be used in the attention layer to calculate the importance of each factor, and then the weight results will be combined with h, the output of LSTM hidden layer,

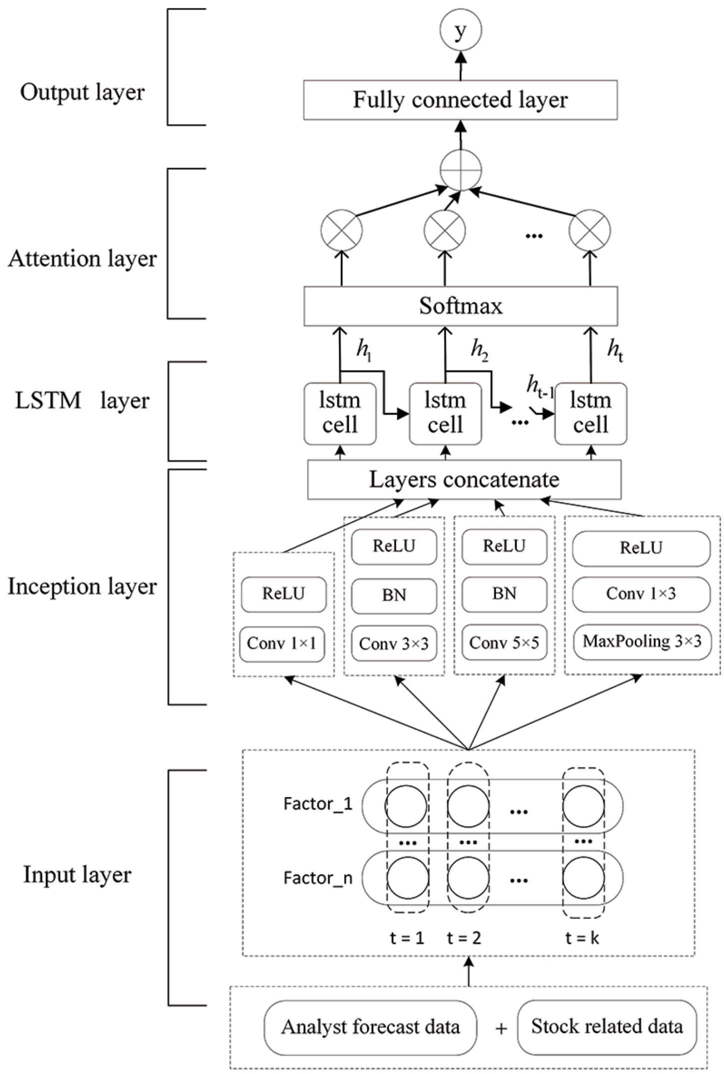


FIGURE 2. Model architecture

input the fully connected network, and finally output the forecast classification of future stock prices called y.

The main parts of the model are described below.

Input layer: The input layer processes the data by normalizing it, then converting it into NumPy arrays. The processed data set includes multiple factors, marked from Factor_1 to Factor_n, each with a sequence of data points across different time steps ($t = 1$ to $t = k$). These data points represent the time-series data for each factor.

Inception layer: A-share data is generated almost every day, but when analysts publish research reports predicting the future rise and fall of A-share, they often only forecast for a quarter, half a year or even a year. Therefore, in order to obtain the changes of the same stock trading at different time scales, convolution kernels of different sizes should be adopted for feature extraction during model training to increase the depth and width of the network. Inception module is to assemble multiple convolution or pooling operations together into a network module. When designing neural networks, the entire network structure is assembled by modules [19]. Batch Normalization (BN) is employed to accelerate convergence during the training of the network. The input data is nonlinearly transformed using ReLU (Rectified Linear Unit), allowing for the integration of features on different time scales.

LSTM layer: LSTM is a type of time-cyclic neural network suitable for processing and predicting important events in time series with relatively long intervals and delays. LSTM is a special cyclic network structure, which introduces the concept structure of "Gates" to realize long-term memory of data information [20]. The trading situation of stock data at a certain moment is affected by the previous moment and the historical multi-moment trading situation, so the LSTM model is used to correlate the useful information in the multi-factor data set at a long time interval, and the correlation of the multi-factor data set is mined.

Attention layer: With the change of economic situation and market changes, the impact of different factors on industry trends is also constantly changing. Using the LSTM model alone cannot account for changes in the importance of factors. The essence of the attention mechanism is to focus the model's attention on a few important elements from a vast amount of information, thereby ignoring most of the less important information [21]. The principle of the attention mechanism is shown in Figure 3. The process of focusing attention is essentially a process of weight calculation, where a larger weight means the corresponding value has a greater impact on the final result, making its information more important, as shown in Equation (5). Imagine the elements in the *Source* as many <*Key*, *Value*> data pairs. By calculating the relevance between the *Query* and each *Key*, we can determine the weight coefficients for the values corresponding to each *Key*. Then, by performing a weighted sum of these values, we obtain the final value of the Attention.

$$\text{Attention}(Source, Query) = \sum_{i=1}^{n} Similarity(Query, key_i) * Value_i \tag{5}$$

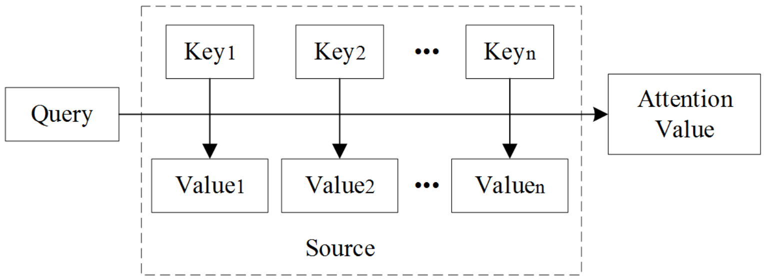


FIGURE 3. Attention mechanism

4. Experimental Analysis and Discussion.

4.1. Label calculation. Traditional LSTM is widely used in numerical prediction for continuous data, considering the characteristics of analysts predict data and the actual demand, this paper predicts the growing rate of the stock in the following period of time. So first of all, we need to classify the changes in stock prices. Since the ultimate purpose of making forecasts is to select stocks, the situation when stocks rise is divided in detail: Explosive Growth, Strong Growth, Steady Growth, Slight Change, Downtrend. It is calculated as follows.

$$label = \begin{cases} 0 & \Delta p_{price} > 35\% \\ 1 & 20\% < \Delta p_{price} \leq 35\% \\ 2 & 10\% < \Delta p_{price} \leq 20\% \\ 3 & -2\% < \Delta p_{price} \leq 10\% \\ 4 & \Delta p_{price} \leq -2\% \end{cases} \tag{6}$$

$$\Delta p = \left(Pclose_{\max\text{-}label\_days} - Pclose_t\right)/Pclose_t \tag{7}$$

$$\Delta p_{index} = \left(Pclose^*_{\max\text{-}label\_days} - Pclose^*_t\right)/Pclose^*_t \tag{8}$$

$$\Delta price = \Delta p - \Delta p_{index} \tag{9}$$

$Pclose_{\max\text{-}label\_days}$ is the highest closing price of the stock within $label\_days$; $Pclose_t$ is the closing price of the stock on day t. In order to remove the influence of the market trend on the experimental results, the rise and fall of the stock itself in the label period Δp is subtracted from the rise and fall of the index in the same period Δp_{index}, and finally the label measure $\Delta price$ is obtained. After division, the distribution of each label is relatively uniform, and the specific quantity is shown in Table 4.

Table 4. Number of labels

Label	0	1	2	3	4
Number	39045	35019	40980	50355	49542

4.2. **Experiment.** The model experiment parameters are shown in Table 5. The setting of model parameters is based on best practices.

Table 5. Parameter settings of the model

Learning rate	batch_size	dropout	hidden_size	Classification function	Categories
0.001	128	0.2	256	Softmax	5

To ensure the integrity and reliability of the data, after removing delisted stocks and stocks with incomplete analyst coverage data, a data set consisting of 300 stocks with high analyst attention was constructed. The first 70% of the factor data set was used as the training set, with the remaining data serving as the prediction set. To verify the overall effectiveness of the model proposed in this paper, comparative experiments and ablation studies were conducted using the same data set input. As a versatile convolutional neural network model, CNN has shown remarkable capabilities and achieved impressive outcomes in the financial sector. The LSTM, which excels in long-term financial time series prediction tasks, serves as the baseline model for this study. Thus, comparative experiments using CNN and LSTM were conducted. Ablation experiments were carried out on the Inception-ALSTM-1 model, which lacks the inception structure, and the Inception-ALSTM-2 model, which lacks the attention mechanism. These experiments validate the efficacy of the algorithmic improvements. Specific experimental results are presented in Table 6.

Table 6. Comparison of results from different models

Model	Accuracy	Precision	Recall	F1 macro	Kappa	HD
CNN	0.555	0.564	0.563	0.555	0.123	0.445
LSTM	0.585	0.587	0.588	0.584	0.172	0.415
Inception-ALSTM-1	0.634	0.636	0.639	0.634	0.272	0.365
Inception-ALSTM-2	0.628	0.633	0.633	0.627	0.261	0.371
Inception-ALSTM	0.645	0.652	0.653	0.645	0.299	0.354

In the test set, our model achieved good performance with an accuracy rate of 0.645. Overall, from multiple perspectives such as accuracy, precision, recall, F1 macro, Kappa, and HD, the Inception-ALSTM model exhibits the best performance in these indicators.

5. **Conclusions.** This paper comprehensively considers the characteristics of analyst prediction data and stock related data, constructs the Inception-ALSTM model for learning and training, so as to grasp the potential law between analyst predictions and stock fluctuations, and realize the stock rise and fall prediction. We found that training to treat all predictors as equally important when using analyst forecast data does not match the actual accuracy of different analysts' forecasts. In order to solve this problem, this paper introduces the attention mechanism to make the LSTM model pay more attention to the important information, so as to improve the model effect.

Compared with other research analyst forecasting methods, our attention mechanism provides more insights into the factors that affect stock prices: It can be seen from the experimental results that the stock forecasting model proposed in this study based on machine learning and multi-factor analyst forecasting data has better forecasting ability. Among them, the analyst prediction factor, the analyst's own factor and the stock related factor all contribute to the performance of the model. This shows that for the forecast of the stock market, it is necessary to consider not only the factors of the stock itself, but also the analysts' forecasts and their own factors. In addition, the Inception-ALSTM model has better performance than the traditional LSTM model. This may be because the Inception module can integrate features across different time scales, and the attention mechanism is able to learn the importance of each moment, thus improving the model's ability to generalize.

The main limitation of this study is that the characteristics used in the study only relate to analyst predictors, analyst's own factors, and stock related factors. In the future, more features, such as industry or technical factors, can be explored to improve the predictive power and stability of the model. It is also considered to further apply the stock selection algorithm to the future stock selection strategy and system.

REFERENCES

[1] H. L. Liu, Investor status of China's Stock Market, *China Finance*, no.9, pp.40-42, 2020.
[2] M. Schutte and E. Unlu, Do security analysts reduce noise?, *Financial Analysts Journal*, vol.65, no.3, pp.40-54, 2009.
[3] K. Ota, Information content of analysts' stock ratings and earnings forecasts in the presence of management earnings forecasts, *Japanese Accounting Review*, vol.2, pp.87-116, 2013.
[4] B. M. Barber, R. Lehavy and B. Trueman, Ratings changes, ratings levels, and the predictive value of analysts' recommendations, *Financial Management*, vol.39, no.2, pp.533-553, 2010.
[5] Y. Jin, P. Ji, X. Wang and Y. Xi, Analyst forecast bias, investors' heterogeneous beliefs and stock price crash risk, *Foreign Economics and Management*, vol.43, no.6, pp.90-104, 2021.
[6] E. F. Fama and K. R. French, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, vol.33, no.1, pp.3-56, 1993.
[7] M. M. Carhart, On persistence in mutual fund performance, *Journal of Finance*, vol.52, no.1, pp.57-82, 1997.
[8] A. Ang, R. J. Hodrick, Y. Xing and X. Zhang, The cross-section of volatility and expected returns, *NBER Working Papers*, vol.61, no.1, pp.259-299, 2004.
[9] L. Zhou and Y. Z. Wang, Research on the low-risk anomaly of China's Stock Market, *Financial Theory and Practice*, vol.3, pp.90-96, 2020.
[10] J. Li, G. Guo and X. Hu, Prediction algorithm of industry rotation based on attention LSTM model, *International Journal of Innovative Computing, Information and Control*, vol.18, no.6, pp.1969-1977, 2022.
[11] J. Li, T. Zhou and X. Hu, Prediction algorithm of stock holdings of Hong Kong-funded institutions based on optimized PCA-LSTM model, *International Journal of Innovative Computing, Information and Control*, vol.18, no.3, pp.999-1008, 2022.
[12] W. Lu, J. Li, Y. Li et al., A CNN-LSTM-based model to forecast stock prices, *Complexity*, pp.1-10, 2020.

[13] I. Bhattacharjee and P. Bhattacharja, Stock price prediction: A comparative study between traditional statistical approach and machine learning approach, *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*, Khulna, Bangladesh, pp.1-6, 2019.
[14] L. Ma and X. Lu, Does team make better earnings forecast?: Evidence from China, *Joint International Social Science, Education, Language, Management & Business Conference*, 2015.
[15] Y. Huang, L. F. Capretz and D. Ho, Machine learning for stock prediction based on fundamental analysis, *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2021.
[16] Z. Zhong, Z. Xie and Z. Wang, Analysis of individual characteristics of securities analysts in China, *Epochal Finance*, vol.29, no.531, pp.74-75+88, 2013.
[17] S. Keskek, L. A. Myers, T. C. Omer and M. K. Shelley, The effects of disclosure and analyst regulations on the relevance of analyst characteristics for explaining analyst forecast accuracy, *Journal of Business Finance & Accounting*, 2017.
[18] J. A. Ou and J. F. Sepe, Analysts earnings forecasts and the roles of earnings and book value in equity valuation, *Journal of Business Finance & Accounting*, 2002.
[19] C. Szegedy, W. Liu, Y. Jia, P. Sermanet and A. Rabinovich, Going deeper with convolutions, *IEEE Computer Society*, 2014.
[20] F. A. Gers and N. N. Schraudolph, Learning precise timing with LSTM recurrent networks, *Advances in Neural Information Processing Systems*, vol.15, pp.591-598, 2003.
[21] C. Yang, T. Kim, R. Wang, H. Peng and C. C. J. Kuo, Unsupervised image translation with self-regularization and attention, *IEEE Transactions on Image Processing*, vol.28, no.10, pp.4845-4856, 2019.

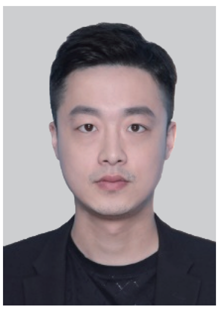
Author Biography

Jianming Li received the Bachelor's degree in Ship Engineering from Dalian University of Technology, China, 1999; the M.Sc. degree in Computer Application Technology from Dalian University of Technology, China, 2002; the Ph.D. degree in Computer Application Technology from Dalian University of Technology, China, 2007.

He is currently a full-time Associate Professor at the Dalian University of Technology, China. His main research interests include the machine learning, classification and prediction algorithms of deep learning, software automation, and quantitative analysis and strategy research in the financial field. He has published over 50 papers in journals and conferences.

Xiaojun Chen obtained Bachelor's degrees in Engineering and Literature, majoring in Computer Science and Technology, Dalian University of Technology, from September 2015 to June 2020.

She is currently pursuing a Master's degree at Dalian University of Technology. Her main research areas include machine learning, classification and prediction algorithms in deep learning, and quantitative analysis and strategy research in the financial field.



Suhao Guan received the Bachelor's degree in Management from Dongbei University of Finance and Economics, China, 2012; the Master's degree in Business Administration from Dongbei University of Finance and Economics, China, 2017.

Currently, he serves as the director of the Comprehensive Office in the Financial Department at Dalian University of Technology, China. He has presided over 2 provincial-level research projects and participated in 1 project sponsored by the Ministry of Finance, with his main research interests including budget management and performance management.

Xiangpei Hu received his B.S. (1983), M.S. (1987) and Ph.D. (1996) degrees from Harbin Institute of Technology, China, respectively. He is a Professor of Management Science at Dalian University of Technology, China, "Distinguished Young Scholars" of National Natural Science Foundation of China (NNSFC), "Chang-jiang Scholars Distinguished Professor" of Ministry of Education (MOE) of China.

His research and teaching interests are electronic commerce, supply chain and logistics management, intelligent operations research and the real-time optimization control for dynamic systems. He has published over 200 scholarly papers in refereed journals.