Implementations of Hybrid Prediction Models for Stock Price Forecasting

Jiaji Zeng
Mathematical Institute, Southwestern University of Finance and Economics, Chengdu, China

Keywords: Prediction of Stock Price, Hybrid Prediction Model, Financial Markets.

Abstract: With the rapid advance of the world economy since the 20th century, many people are devoting their energy to studying the stock market for profit. Therefore, various research methods for stock price prediction have emerged, among which a single prediction model is an important part of these research methods. However, with the progress of technology, the limitations of single model prediction are gradually amplified. It is difficult to fully capture the dynamic changes and uncertainties of complex financial systems. So, aiming to augment the accuracy and reliability of predictions, scientists began exploring the possibility of combining multiple prediction models, namely hybrid prediction models. This research will start with stock price prediction, based on variable analysis, representative configuration, model application results and performance, as well as its limitations and prospects, to explore the implementation of hybrid prediction models for the prediction of stock price. These results are of great significance in exploring the development prospects, risk assessment, optimization and adjustment of financial markets.

1 INTRODUCTION

With the rapid progress of the times and economy, plenty of people are participating in the stock market with the expectation of getting rich. Therefore, the prediction of stock price has become a crucial concept in modern financial markets. There have been a lot of methods for predicting stock price in the past, which can be generally separated into several types: technical analysis, basic analysis, macroeconomic analysis, market sentiment analysis, and quantitative analysis. In recent years, methods for predicting stock prices have been constantly developing and innovating. Stock price prediction is a rather hazardous operation. A good analyst is therefore not someone who is always right, someone who has a higher efficiency than his colleagues (Op't Landt, 1997). In recent years, with the obvious progress of new technologies, researchers have adopted various cutting-edge hybrid models in the field of predicting stock price to address the complexity and nonlinear characteristics of the stock market. These hybrid models combine several advantages of dissimilar algorithms and techniques to promote the accuracy and stability of these predictions (Ji et al., 2021).

The first new method is the combination of deep learning and traditional statistical models, including two types of models. Long Short Term Memory Network (LSTM), as a variant of Recurrent Neural Network (RNN) in deep learning, excels at handling long-term dependency problems in sequential data. The autoregressive integrated moving average model (ARIMA) is a classic model in traditional time series analysis. Some studies combine LSTM with ARIMA to capture nonlinear trends and fluctuations, while using ARIMA to process linear parts, thereby improving prediction accuracy. Stock price prediction is a challenging problem due to its random movement. This hybrid model is a combination of two well-known networks, LSTM and Gated Recurrent Unit (GRU) (Hossain et al., 2018). Transformer model has achieved significant results in the area of natural language development with powerful parallel computing capability and efficient information extraction ability. In recent years, researchers have begun to combine Transformer with LSTM for stock prediction. This hybrid model can simultaneously focus on global and local information, further improving prediction performance (Jakkula, 2006).

The second method is the fusion of machine learning algorithms. Support Vector Machine (SVM)

Zeng, J.
Implementations of Hybrid Prediction Models for Stock Price Forecasting.
DOI: 10.5220/0013270700004568
In *Proceedings of the 1st International Conference on E-commerce and Artificial Intelligence (ECAI 2024)*, pages 568-573
ISBN: 978-989-758-726-9

is based on the principle of minimizing structural risk, striving to minimize empirical risk while maximizing the generalizability of the learning model (Vishwanathan et al., 2002). However, a single SVM model may not be capable of handling complex stock markets. Therefore, researchers combined SVM with self-organizing map neural networks (SOM), particle swarm optimization algorithms (PSO), and other methods to construct a hybrid model to cope with the variability of the stock market (Biswas et al., 2021). Besides, some other machine learning algorithms like random forest, gradient boosting tree, etc. are also used for stock prediction. Research work uses the frequently used algorithms LSTM, Extreme Gradient Boosting (XGBoost), Linear Regression, Moving Average, and Last Value model, on more than twelve months of past stock data to erect a prediction model to forecast stock price. Ensemble learning method can effectively improve the general prediction performance by combining prediction outcomes of many single models.

Thirdly, the combination of Big Data and Artificial Intelligence (AI). Along with the advancement of Big Data technology, researchers can obtain more dimensions of stock market data. These data not only include traditional indicators like past stock prices and trade volumes, but also non-traditional data like social media sentiment and macroeconomic indicators. A hybrid model based on big data could completely utilize this key information to promote the accuracy and timeliness of predictions. Some cutting-edge research has begun to apply hybrid models to automated trading systems. By monitoring market dynamics in real-time and predicting stock price trends, automated trading systems can automatically execute buy or sell operations at the appropriate time, reducing the risk of human intervention and improving trading efficiency. In recent years, the number of academic papers and patents on hybrid model prediction of stock prices has been increasing. These research results not only promote the development and improvement of relevant theories, but also provide strong support for practical applications.

Besides, some financial institutions and technology companies have begun to apply hybrid models to stock prediction and trading strategies. By continuously optimizing models and algorithms, these institutions are able to provide customers with more accurate and personalized investment advice and services. In summary, based on previous research, this article will explore the implementation of a hybrid prediction model for predicting of stock price through specific examples or methods. The framework would base on these topics: Descriptions of stock price prediction, Configurations for hybrid model, Implementation results and Limitations and Prospects.

2 DESCRIPTIONS OF STOCK PRICE PREDICTION

In stock price prediction, the dependent variable usually refers to the target variable that one attempts to predict or explain, namely the stock price itself or its related indicators (Pinto & Asnani, 2011). Specifically, the dependent variables for stock price prediction can include these aspects. For prices:

- Opening price: First trading price of 1 stock at the beginning of a market day.
- Closing price: Last trading price of 1 stock at the end of a trading day.
- Highest and lowest prices: The highest and lowest trading prices of a stock during a trading day.

These prices directly reflect the market performance of stocks and are one of the most important indicators for investors to pay attention to. For stock return rate

- Daily return: Percentage change in the price of 1 stock within a single market day.
- Weekly return, monthly return, and annual return: represent percentage change of stock prices over a week, month, and year, respectively.

Stock return plays an important role in measuring stock investment returns and a common dependent variable in stock price prediction. Regarding to other relevant indicators

- Trading volume: The number of trades in a stock during a specific period of time, reflecting the level of market activity and investor participation (Cakra & Trisedya, 2015).

Some financial indicators, such as price to earnings ratio and price to book ratio: These factors are used to evaluate the investment value of stocks. Although they are not direct price indicators, they are often used as references in stock price forecasting (Selvin et al., 2017).

In stock price prediction, the dependent variable is usually the future price or price trend of the stock. To predict this dependent variable, a series of independent variables (also known as features, factors, or input features) need to be used, as well as selecting an appropriate prediction model (Ji et al., 2021). The independent variables for stock price prediction can include various types of data, which typically reflect multiple aspects such as the historical performance of the stock market, company fundamentals, macroeconomic environment, and market sentiment (Song & Lee, 2020). There are several common independent variables:

- Historical price data: including past stock prices, opening and closing prices, highest and lowest prices, etc. These data can reflect the historical fluctuations of stock prices.
- Trading volume data: The trading volume of stocks is also an important predictor, which can reflect the level of market activity and investor participation.
- Technical indicators: like moving averages, RSI, stochastic oscillators can help analyze the overbought and oversold situation of stock prices, trend strength, etc.
- Fundamental data: including the company's financial statement data (such as revenue, profit, balance sheet, etc.), price to earnings ratio or book ratio, dividend yield, etc. Those data can reflect the company's profitability, the financial condition, and the market valuation.
- Macroeconomic data: Rate of GDP growth, rate of inflation, interest rate level, can reflect the impulse of overall economic environment on stock market.
- Market sentiment data: including news sentiment analysis, social media sentiment index, etc., can reflect investors' overall views and emotional changes towards the stock market.

In stock price prediction, there are multiple models to choose from, each with its own characteristics and scope of application. Here are some common prediction models:

- Linear regression model: Predicting stock prices by fitting a linear relation between the independent variables and dependent variables. Although simple, it may be effective in certain situations.
- The time series analysis models, like ARIMA models, seasonal decomposition models are particularly suitable for processing data with temporal dependencies.
- Machine Learning (ML) models: including decision trees, random forests, SVM, neural networks, etc. These models are capable of handling complex nonlinear relationships and perform well on large amounts of data.
- Deep learning models: RNN, LSTM, Transformers, etc. These models perform well in processing time series data and capturing long-term dependencies, making them particularly suitable for stock price prediction.

3 CONFIGURATIONS FOR HYBRID MODEL

Regarding the LSTM-SDE configuration of hybrid models, there is no widely recognized hybrid model directly named LSTM-SDE in standard deep learning architectures. However, the LSTM-SDE is an attempt to combine LSTM with a method based on Stochastic Differential Equations (SDE) to integrate the advantages of both in processing time series data and dynamic system modeling (Melo et al., 2022). Base on both advantages of LSTM model and advantages of SDE, here are some potential characteristics of LSTM-SDE configuration. Firstly, LSTM is a special type of recurrent neural network (RNN) that excels at processing sequential data and capturing long-term dependencies. Through structures such as forget gates, input gates, and output gates, LSTM can effectively control the flow of information and reduce the problem of gradient vanishing or exploding. Secondly, SDE (stochastic differential equation) is a mathematical model that describes the dynamic changes of stochastic processes and is widely used in fields such as finance, physics, and biology. SDE can capture the uncertainty and randomness of systems, and has unique advantages for modeling complex dynamic systems (Araújo et al., 2015). As a result, there is a combination method. In the LSTM-SDE configuration, LSTM may be used for feature extraction and temporal modeling of time series data, while SDE is used to describe the random changes in system state. The two can be combined in some way, such as using the output of LSTM as the input of SDE, or associating certain parameters of SDE with the hidden state of LSTM. Because of the high nonlinearity and randomness Financial market data always has, LSTM can capture the long-term trends of the market, while SDE can describe the random fluctuations of the market. By combining LSTM and SDE, more accurate financial market prediction models can be constructed as well.

Before studying the CNN-LSTM-AM model, it is necessary to first introduce the CNN-LSTM model. The CNN-LSTM model has combined both the advantages of CNN and LSTM. CNN excels at processing data with grid structures, such as images and videos, and automatically extracts features from input data through convolutional layers. LSTM is good at processing sequential data and can capture long-term dependencies. In the CNN-LSTM model, CNN is first used to extract spatial features of input data, and then these feature sequences are used as inputs to LSTM to learn the temporal dependencies between features (Sun et al., 2024). In that case, the CNN-LSTM-AM model introduces an attention mechanism based on CNN-LSTM, allowing the model to focus more on important features or time points. The attention mechanism allows the model to dynamically adjust the importance weights of different time steps or features when processing sequential data. Therefore, the application of CNN-LSTM-AM (Convolutional Neural Network Long Short Term Memory Network Attention Mechanism) in stock price prediction is a comprehensive method that combines multiple deep learning techniques, aiming to improve the accuracy and efficiency of prediction.

4 IMPLEMENTATION RESULTS

In prediction of stock price, hybrid prediction models have been widely applied and researched due to their ability to combine multiple methods and techniques to promote prediction accuracy and stability. There are some specific results and performance of implementing hybrid prediction model for predicting stock prices. This model is constructed by combining self attention mechanism, LSTM, and GRU. Compared with models such as LSTM, GRU, RNN-LSTM, and RNN-GRU, the ATLG model has higher accuracy. Introducing self attention mechanism makes the model to better focus on stock feature information at important time points. Through backtesting using the MACD (Moving Average Convergence and Divergence) indicator, a return of 53% was obtained, which is higher than the return of the Shanghai and Shenzhen 300 during the same period, proving the effectiveness and practicality of the model (Xie et al., 2021). Some researchers use SVR. This model has combined SVR (Support Vector Regression) with various methods such as wavelet transform, sliding window, etc. The original data is transformed using the difference method to highlight the trend of stock price fluctuations. Wavelet transform is introduced to separate the main and detailed components of stock prices, and multi-scale sliding windows are used to smooth stock prices, especially those with strong noise (Liu et al., 2022). Compared with the SVR model based on a single feature, this hybrid model has improved prediction accuracy. However, due to the lack of detailed experimental data and results in the reference article, specific accuracy improvement values cannot be provided.

This model is constructed by combining Variational Mode Decomposition (VMD) with two different neural networks (such as GRU and Echo State Network ESN) for mixed prediction. VMD decomposition effect: VMD decomposition effectively reduces the complexity of raw data and helps extract different frequency features of stock prices. By separately predicting the sub sequences decomposed by VMD and combining them with ARIMA model for error correction, the overall prediction accuracy and stability of the model have been significantly improved. The experimental results have showed that this model has a advance accuracy than the traditional single model and other hybrid models when predicting the historical data of stocks such as Shanghai Shenzhen 300 and the US Standard&Poor's 500.

To sum up, these hybrid forecasting models show high accuracy and stability in stock price forecasting, and effectively improve the forecasting ability of a single model by combining different technologies and methods. However, the specific prediction performance will be affected by various factors such as model design, data quality, and market environment, so adjustments and optimizations need to be made according to specific situations in practical applications.

5 LIMITATIONS AND PROSPECTS

When using mixed forecasting models for stock price prediction, although these models combine the superiority of various techniques and algorithms to promote the accuracy and stability of predictions, there are still some limitations. Firstly, hybrid prediction models typically combine multiple algorithms and techniques, making the model structure more complex. This complexity may increase the training difficulty and computational cost of the model, while also potentially reducing its interpretability. Secondly, due to the combination of multiple algorithms in the hybrid model, its prediction results may be difficult to explain directly. This may be a problem for investors as they need to understand the logic and basis behind their predictions. Hybrid prediction models are highly depending on quality and completeness of input data. Whether there are deviations, missing or outliers in the data, it will significantly affect the training effectiveness and prediction accuracy of the model. Besides, stock market is influenced by various factors, including macroeconomics, industry policies, market sentiment, etc. However, data on these factors may be incomplete or difficult to obtain, resulting in the model being unable to fully consider all relevant factors when making predictions. Hybrid prediction models are often built based on a series of assumptions. These assumptions may not fully conform to actual situation of the stock market, thereby affecting the accuracy of the model's predictions. Also, the parameter settings in the hybrid model have a obvious impulse on the prediction results. Whereas, determining the optimal parameter combination is a challenging problem. Inappropriate parameter settings may lead to overfitting or underfitting of the model, thereby reducing predictive performance. On one hand, current technologies and algorithms still have certain limitations when dealing with complex data and relationships. These limitations may limit the effectiveness of hybrid prediction models in predicting stock prices. On the other hand, different hybrid prediction models may use different combinations of algorithms. However, not all algorithms are suitable for stock price prediction. Improper selection of algorithms may lead to a decrease in model performance. Under the influence of numerous factors, the stock market has a great degree of uncertainty and is influenced by various unpredictable factors. These factors may include policy changes, natural disasters, international situations, etc. Although hybrid prediction models can handle complex data and relationships, their predictive ability may be limited when faced with these sudden unpredictable events. What’s more, the volatility of the stock market is also a crucial factor affecting the accuracy of predictions. In that case of significant market volatility, the predicted results of the hybrid model may have significant deviations from the actual trend.

Although there are still many limitations to prediction, one can still optimize it through various methods. For instance, one can improve data quality and integrity to make sure enough accuracy and reliability of input data, and simplify the model structure and improve its interpretability. One needs to be cautious in setting model parameters to avoid overfitting or underfitting. Besides, both paying attention to market dynamics and changes and adjusting models in a timely manner are significant for adapting to market changes. At last, one should continuously explore and try new technologies and algorithms to improve the performance and application effectiveness of hybrid prediction models.

All in all, the development prospects of using mixed forecasting models for stock price prediction are quite broad, mainly due to several factors. With the continuous advancement of deep learning technology, models such as LSTM and CNN have demonstrated forceful capabilities in processing time series and spatial feature data. These technologies provide more accurate and efficient algorithmic support for hybrid prediction models. Additionally, hybrid prediction models can more comprehensively capture the complexity and nonlinear characteristics of stock price fluctuations by combining the advantages of different models. For example, LSTM models are good at processing time series data, while CNN is good at extracting spatial features. The combination of the two can significantly improve the accuracy and stability of predictions. In the future, with the continuous development of model fusion technology, hybrid prediction models will become more intelligent and efficient. The widespread use of big data technology has led to a sharp increasing in the amount of data in stock market, including historical trading data, market sentiment data, macroeconomic data, etc. These data provide a rich source of information for hybrid forecasting models, helping them to more accurately capture market trends and changes. At the same time, with the continuous improvement of data cleaning and preprocessing techniques, hybrid prediction models can more effectively handle noisy data and outliers, improving the accuracy and reliability of predictions. Along with the continuous advancement of financial markets and the multiplicative demand for risk control among investors, stock price forecasting, as an important investment decision-making tool, has seen a continuous increase in market demand. The hybrid prediction model, with its high accuracy and stability, is expected to become the mainstream tool in the field of stock price prediction in the future. In that case, spontaneously, governments and regulatory agencies of various countries have attached great importance to and supported the development of financial technology. The promotion of policies will facilitate the application and dissemination of hybrid forecasting models in the financial sector, providing a broader space for their development. In addition to deep learning technology, many emerging technologies such as natural language processing (NLP), knowledge graphs, etc. are also constantly emerging and developing. These innovative technologies will provide richer information sources and more efficient algorithm support for hybrid prediction models, promoting their application and development in the area of stock price prediction.

To sum up, the development prospects of using mixed forecasting models for stock price prediction are very broad. With the optimization of technology, the improvement of data quality, market demands and policy supports, and the continuous promotion of interdisciplinary integration and innovation, hybrid forecasting models will play a more crucial role in financial field, constantly providing the investors with the most accurate and reliable decision support.

6 CONCLUSIONS

To sum up, the hybrid forecasting model is not only playing a significant role in the field of stock prices, but also directly or indirectly affecting the entire financial market and other related fields. In summary, the research on using hybrid prediction models to predict stock prices is of great significance and has broad prospects. In the future, string along with both constantly innovation of technology and continuous development of the market, hybrid forecasting models will play a big role in the financial market.

REFERENCES

Araújo, R. D. A., Oliveira, A. L., Meira, S., 2015. *A hybrid model for high-frequency stock market forecasting.* Expert Systems with Applications, 42(8), 4081-4096.

Biswas, M., Shome, A., Islam, M. A., Nova, A. J., Ahmed, S., 2021. *Predicting stock market price: A logical strategy using deep learning.* 2021 IEEE 11th IEEE symposium on computer applications & industrial electronics (ISCAIE), 218-223.

Cakra, Y. E., Trisedya, B. D., 2015. *Stock price prediction using linear regression based on sentiment analysis.* 2015 international conference on advanced computer science and information systems (ICACSIS), 147-154.

Hossain, M. A., Karim, R., Thulasiram, R., Bruce, N. D., Wang, Y., 2018. *Hybrid deep learning model for stock price prediction.* 2018 IEEE symposium series on computational intelligence (ssci), 1837-1844.

Jakkula, V., 2006. *Tutorial on support vector machine svm*. School of EECS, Washington State University, 37(2.5), 3.

Ji, X., Wang, J., Yan, Z., 2021. *A stock price prediction method based on deep learning technology*. International Journal of Crowd Science, 5(1), 55-72.

Liu, T., Ma, X., Li, S., Li, X., Zhang, C., 2022. *A stock price prediction method based on meta-learning and variational mode decomposition.* Knowledge-Based Systems, 252, 109324.

Ma, C., Ning, Y., Jin, H., Wu, J., 2019. *The hybrid dynamic stock forecasting model based on ANN and SVR*. 2019 International Conference on Intelligent Computing, Automation and Systems (ICICAS), 715-718.

Melo, F. M., Magnani, F. S., Carvalho, M., 2022. *A decision-making method to choose optimal systems considering financial and environmental aspects: Application in hybrid CCHP systems.* Energy, 250, 123816.

Op't Landt, F. W., 1997. *Stock price prediction using neural networks.* Maters Thesis, Leiden University.

Pinto, M. V., Asnani, K., 2011. *Stock price prediction using quotes and financial news.* International Journal of Soft Computing and Engineering (IJSCE), 1(5), 266-269.

Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., Soman, K. P., 2017. *Stock price prediction using LSTM, RNN and CNN-sliding window model.* 2017 international conference on advances in computing, communications and informatics (icacci), 1643-1647.

Song, Y., Lee, J., 2020. *Importance of event binary features in stock price prediction.* Applied Sciences, 10(5), 1597.

Sun, Y., Zhou, Q., Sun, L., Sun, L., Kang, J., Li, H., 2024. *CNN–LSTM–AM: A power prediction model for offshore wind turbines.* Ocean Engineering, 301, 117598.

Vishwanathan, S. V. M., Murty, M. N., 2002. S*SVM: a simple SVM algorithm.* Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290), 3, 2393-2398.

Xie, Y., Gui, F. X., Wang, W. J., Chien, C. F., 2021. *A two-stage multi-population genetic algorithm with heuristics for workflow scheduling in heterogeneous distributed computing environments*. IEEE Transactions on Cloud Computing, 11(2), 1446-1460.