Principe and Applications of Hybrid Prediction Models for Stock Price Forecasting

Yikun Liu
Department of Information Science and Technology, Beijing University of Technology, Beijing, China

Keywords: Stock Prediction, Hybrid Models, Machine Learning.

Abstract: The stock market, one of the key elements of the financial industry, is a risky and lucrative arena that draws in a large number of traders. Contemporarily, there has been a surge in interest in research concerning stock price prediction. This research focuses on the concept and utilization of hybrid prediction models in predicting stock prices. This study first introduces some traditional and deep learning-based single models and the relevant background of stock forecasting, and then introduces some cutting-edge hybrid model configurations. The prediction results of these models were compared. By analysing mean average error (MAE), root mean square error (RMSE) and other performance metrics, it can be found that these hybrid models have a great improvement compared with the single model, and different models have different advantages. The research on hybrid model stock forecasting is helpful to understand its application in the stock market, better forecast stocks, and lay the foundation for the establishment of more diversified and effective models in the future.

1 INTRODUCTION

The stock market, one of the key elements of the financial industry, is a risky and lucrative arena that draws in a large number of traders. The correct prediction of stock prices can enable investors to reduce risks and improve returns when making decisions. Research on stock price prediction has become very popular in recent years. According to different theories of model construction, there are two primary groups for stock price single prediction models (Zhang, 2020). Classic statistical models, like the time sequence and hidden Markov model, are founded on statistical theory. Neural networks, support vector machines, decision trees, and other cutting-edge stock prediction techniques are examples of machine learning models. Each of these approaches has advantages and disadvantages, and several typical models are described below.

In time series analysis, the Auto-Regressive Integrated Moving Average Model (ARIMA) and the Auto-Regressive Moving Average Model (ARMA) are frequently used to forecast and evaluate data with time changes, and ARIMA is an extension of ARMA (Zhao, 2021). Compared with ARMA, ARIMA has one more difference step in data processing and performs better in processing non-stationary time series data. However, this model is not suitable for long-term prediction (Huang, 2023).

It was during the latter part of the 1960s that the Hidden Markov Model (HMM) was developed, and Baum et al. gave the original prototype of the model in a series of statistical papers. It also has its application in the financial field. In 2005, Hassan and Nath introduced a new technique for predicting stock prices by using HMM to the task. The method takes the opening price, closing price, highest price and lowest price as the model input, and predicts the stock price by parameter estimation and state decoding (Hassan, 2005). HMM has a good effect on process state prediction, and can be used in state prediction where state classification is more obvious. HMM has a good effect on global (the whole) prediction, but it also has the disadvantages of not suitable for local prediction and poor prediction accuracy in the medium and long term.

The vast amount of data in the stock market has drawn the interest of numerous academics since the big data era began. The subject of stock prediction makes extensive use of machine learning techniques including support vector machines, neural networks, decision trees, etc. Many of the drawbacks of conventional approaches are offset by their benefits in processing complex and massive amounts of data. After combining machine learning algorithms with a

Liu, Y.
Principe and Applications of Hybrid Prediction Models for Stock Price Forecasting.
DOI: 10.5220/0013270600004568
In *Proceedings of the 1st International Conference on E-commerce and Artificial Intelligence (ECAI 2024)*, pages 561-567
ISBN: 978-989-758-726-9

vast amount of historical stock market data, researchers are able to develop and train their model, which then helps them forecast future stock market trends. When compared to conventional methods, machine learning techniques greatly improve precision in forecasting and have significant implications in both theory and practice.

Regardless of the forecasting technique employed, stock forecasting has specific forecasting constraints and uses a finite amount of data and information. Consequently, Bates and Granger combined forecasting method highlights the benefits of various forecasting techniques while avoiding drawbacks, making full use of the data in various forecasting models to forecast. The stock market prediction is progressively made using the integrated approach (Bates, 1969). In order to forecast the closing price of the stocks for the following day, Lu proposed the convolutional neural networks-bidirectional long short-term Memory-attention Mechanism(CNN-BiLSTM-AM) approach. The attention mechanism (AM), bi-directional long short-term memory (BiLSTM), and convolutional neural networks (CNN) make up this technique. Its performance is superior to that of the other models (Lu, 2020). By combining deep neural network (DNN) and predictive rule integration (PRE) technology, a hybrid stock prediction model was proposed by Srivinay. RMSE of this model was improved about 6% compared to the sinlge prediction model (Srivinay, 2022). A novel method of prediction based on generative adversarial networks (GANs) is called the Hybrid Prediction Algorithm (HPA). Multi-Model based Hybrid Prediction Algorithm (MM-HPA) and GAN-HPA were coupled by Nagagopiraju to create a new hybrid model called MMGAN-HPA (Nagagopiraju, 2023).

This paper mainly focuses on the research history and research development of stock forecasting, and introduces some current cutting-edge stock forecasting methods based on mixed models, aiming to provide theoretical basis for further research of stock forecasting, hoping to find more accurate and efficient forecasting methods. In the second section, the factors and some common models often used in stock forecasting are introduced; in the third section, the structure of mixed models used in stock forecasting is introduced; in the fourth section, the performance of mixed models in stock forecasting is introduced; in the fifth chapter, the limitations of these models in application and the future outlook of stock forecasting based on mixed models are introduced.

2 DESCRIPTIONS OF STOCK PRICE PREDICTION

The fundamental concept behind stock prediction is to develop a model that, using historical data from the past, projects future stock prices. At first technical analysis were mostly used for subjective projections for shares.. The closing values of stocks were later enumerated in chronological sequence to create a new model. Based on the stock's past price movement, forecast the short-term change trend for the future. At present, people use the large amount of historical data generated by the stock market, combined with machine learning algorithms for modeling and training. They trained their models to predict future movements of stocks.

There are some basic data types in stock prediction (Lu, 2021). The initial transaction price per share of a particular securities following the stock exchange's opening each trading day is referred to as the opening price. One minute before a stock deals for its final session of the day, its volume-weighted average price equals its closing price. The total number of equities exchanged during the day is referred to as volume. The highest price equals the maximum price a stock can produce during the course of trading everyday. The lowest price is the minimum price a stock can produce during the course of trading everyday. Turnover quantity is the total number of shares of all stocks that were traded in that particular day. Some technical indicators are also used in stock prediction.Technical indicators refer to the collection of raw trading data calculated by different mathematical formulas. The internal information of different aspects of the stock market can be directly reflected by making the calculation results of these indicators into charts. Assessing and forecasting the stock market's movement and behavior is advantageous. Stochastic Indicator (KDJ-K, KDJD, KDJ-J), Relative Strength Index (RSI-6, RSI-10), Boll, Boll upper bound, Boll lower bound,Moving Average Convergence Divergence (MACD), MACD signal line, MACD histogram (Shao, 2022). In addition to technical indicators, some macroeconomic factors are also commonly used in stock forecasting, such as China's 2-year treasury yields, US 10-year debt (Liu, 2022).

3 CONFIGURATIONS FOR HYBRID MODEL

CNN is a frequently utilized model. The fundamental building blocks of the CNN model, the convolution layer and pooling layer, are capable of automatically extracting and reducing the dimension of the input features. This reduces the adverse impact of the traditional model and works better for extracting characteristics of stock information. The convolution layer extracts some local features of the input stock price sequence through the convolution kernel, which is equivalent to the feature extractor. The stock price series is reduced and the secondary features are extracted during the pooling layer, so as to further enhance the model's capacity for generalization. Long Short-Term Memory(LSTM) is generally limited to transmitting data in a single direction and only accepting input from the past; it cannot process input from the future. Data from the past and the future can be taken into account by BiLSTM simultaneously. Its principle is: Compute LSTM's path starting from both ends respectively, and then merge the LSTMs of the two directions(Althelaya, 2018). Past data information of the input sequence is stored in the forward LSTM. Information regarding the input sequence's future is contained in the backward LSTM. The human brain's capacity to focus on items is simulated in the Attention Mechanism (AM). Giving more weight to relevant information and less weight to unimportant information is the fundamental tenet of AM. The main structure of CNN-BiLSTM-AM model is CNN, BiLSTM, and AM, including input layer, CNN layer, BiLSTM layer, AM layer, and output layer(Lu, 2021).In the training process, the standardization process is applied to reduce the difference between the data and better adapt to the model training.Separately, every network layer extracts and processes characteristics of data. Lastly, the output layer stores the model's forecasting findings. To continually improve the model's predictive power, every discrepancy between the actual value and the anticipated results is computed, and backpropagation is used to update the model's weight and deviation. The training process continues until a set termination condition is met (completion of a predetermined number of cycles or an error threshold is reached). After the training process, the model can be used for prediction. First, the input data is standardized, and then the trained CNN-BiLSTM-AM model is used to generate prediction results. And then restore the output results to the original data format, and finally output them.

An enhanced neural network model built on the foundation of a recurrent neural network(RNN) is the LSTM model, which resolves the gradient explosion and disappearance issues with RNNs and offers a greater capacity for generalization. The input, output, and forgetting gates are added to the LSTM model in comparison to the RNN. These gate units remove or add data information, so that it can retain important information as much as possible and remove interference information. In these door units, first of all, the forgetting door is responsible for forgetting the useless historical stock information, then, based on input stock data and historical information, the input door modifies the unit status. The current stock information is finally output by the output door based on the status of the unit. Bidirectional Encoder Representations from Transformers(BERT)'s design is inspired by bidirectional and Transformer, which, unlike traditional one-way language models, takes into account both left and right contexts to more fully capture the context of text. BERT model uses self-attention mechanism to construct deep neural network, and Transformer is the core to implement bidirectional text coding (Zhang, 2024). BERT model and BiLSTM model are used to extract the emotional features of some financial news, and BERT self-supervision function is used to predict the emotional polarity of the remaining financial news. The forecasting step utilizing LSTM can be continued After integrating the obtained emotional features with stock information.

4 IMPLEMENTATION RESULTS

The Shanghai Composite Index (000,001) stock is chosen as the experimental data in an experiment (Lu, 2021). Models are trained using the training data set that has been analyzed. MAE, RMSE, and R-square (R^2) are employed as the methodologies' evaluation criteria to assess each model's ability to predict outcomes. Higher prediction effect is associated with lesser MAE and RMSE. The model's predictive power increases with the proximity of its R^2 value to zero, which runs from zero to one. Of the nine methods in the Table 1, CNN-BiLSTM-AM performs the best since its MAE and RMSE are the lowest and its R^2 is closest to 1. Comparing BiLSTM with LSTM, MAE decreased 4%, RMSE decreased 2%, indicating that BiLSTM has better effect. BiLSTM and LSTM are combined with CNN to form BiLSTM-CNN and LSTM-CNN, respectively. The results show that CNN-bilSTM has a higher R^2 and smaller MAE and RMSE than CNN-LSTM. It shows that CNN-

BiLSTM performs better than CNN-LSTM. The combination of BiLSTM and CNN is changed to BiLSTM and AM to form BiLSTM-AM, MAE and RSME further decrease, and R2 improves slightly, which indicates that bilSTM-AM performs better than CNN-BiLSTM.

Table 1: Comparison of evaluation error indexes of the five methods.

Method	MAE	RMSE	R^2
MLP	31.496	39.260	0.9699
CNN	25.665	36.878	0.9735
RNN	26.822	35.801	0.9751
LSTM	24.361	34.331	0.9770
BiLSTM	23.409	33.579	0.9780
CNN-LSTM	23.195	32.640	0.9792
CNN-BiLSTM	22.715	32.065	0.9800
BiLSTM-AM	22.337	31.955	0.9801
CNN-BiLSTM-AM	21.952	31.694	0.9804

A study uses the DJIA stock dataset from 2018 to 2023 to examine a number of algorithms, including ANN, LSTM-GA, LSTM1D, LSTM2D, LSTM3D, and optimized LSTM with ARO (LSTM-ARO) (Gülmez, 2023). An indicator used to assess the precision of forecasting models is MAE. As can be seen from the summary of the table, for most of the stocks used in the experiment, the MAE of LSTM-ARO is the lowest among the several methods, indicating that LSTM-ARO performs most effectively out of all the experiment's approaches. However, another crucial point to remember is that different stocks have varying degree of precision of prediction using LSTM-ARO. The precision of a model's predictions is also assessed by its R^2 value. The model's predictive power increases with the proximity of its R^2 value to zero, which runs from zero to one. As can be seen from the summary of the Table 2 and Table 3, for most of the stocks used in the experiment, the R^2 of LSTM-ARO is the lowest among the several methods, indicating that LSTM-ARO performs most effectively out of all the experiment's approaches. A further point to consider is that certain models, like LSTM1D and LSTM2D, perform worse than a straightforward data average when it comes to specific indicators, as shown by their negative R^2 ratings.

Table 2: Comparison of the models for MAE criteria.

MAE Ticker	LSTM-ARO	LSTM-GA	LSTM1D	LSTM2D	LSTM3D	ANN
AXP	3.804	3.848	5.159	5.415	5.082	4.704
AMGN	3.318	3.425	9.614	7.867	11.513	6.193
AAPL	3.846	3.955	5.851	4.878	6.632	7.673
BA	4.425	4.805	15.994	14.042	15.369	13.813
CAT	4.947	4.748	8.875	10.310	8.661	9.235
CSC0	0.744	0.813	0.828	1.176	1.430	0.988
CVX	3.830	3.531	12.940	5.969	16.006	10.518
GS	7.060	7.272	9.653	10.129	8.844	12.525
HD	6.206	6.026	6.046	14.317	7.487	10.753
HON	2.876	3.068	5.851	5.413	3.854	5.769
IBM	2.113	2.096	2.652	3.117	3.755	2.836
INTC	1.092	1.072	4.107	4.486	5.064	3.059
JNJ	1.912	1.795	3.436	3.860	4.525	2.139
KO	0.738	0.751	2.584	2.624	3.217	1.439
JPM	2.523	2.894	4.570	3.415	3.380	4.569
MCD	3.630	3.646	5.864	10.878	8.596	8.233
MMM	2.602	2.532	6.410	5.845	6.210	5.538
MRK	1.551	1.212	3.810	5.563	5.906	3.165
MSFT	6.584	7.583	8.889	8.659	8.784	9.363
NKE	2.907	3.050	3.371	4.501	3.939	4.563
PG	2.180	2.105	5.392	5.186	5.357	5.968
TRV	2.853	2.808	2.796	7.112	7.962	7.323
UNH	9.016	7.852	40.169	45.718	29.796	13.116
CRM	5.024	5.879	7.527	6.616	6.419	6.931
VZ	0.618	0.609	1.274	1.274	1.258	1.303
V	3.724	3.849	5.347	4.752	5.583	6.540
WBA	0.667	0.693	1.124	1.270	1.313	1.153
WMT	2.291	2.309	3.314	3.464	3.648	4.770
DIS	2.923	2.935	4.472	5.300	5.338	3.431
DOW	1.059	1.114	1.178	1.203	1.489	2.337

Table 3: Comparison of the models for R^2 criteria.

R^2 Ticker	LSTM-ARO	LSTM-GA	LSTM1D	LSTM2D	LSTM3D	ANN
AXP	0.907	0.900	0.836	0.815	0.836	0.862
AMGN	0.943	0.942	0.606	0.696	0.354	0.834
AAPL	0.857	0.848	0.675	0.764	0.602	0.458
BA	0.954	0.945	0.527	0.625	0.540	0.661
CAT	0.909	0.914	0.724	0.649	0.732	0.705
CSC0	0.961	0.954	0.952	0.913	0.870	0.933
CVX	0.911	0.919	0.227	0.802	-0.214	0.467
GS	0.886	0.885	0.806	0.781	0.826	0.651
HD	0.902	0.906	0.909	0.587	0.868	0.708
HON	0.928	0.921	0.725	0.766	0.880	0.728
IBM	0.884	0.885	0.826	0.772	0.650	0.798
INTC	0.972	0.970	0.593	0.498	0.338	0.803
JNJ	0.834	0.854	0.515	0.407	0.192	0.786
KO	0.841	0.845	—0.246	-0.339	—0.943	0.537
JPM	0.940	0.924	0.821	0.890	0.892	0.807
MCD	0.875	0.869	0.689	0.133	0.401	0.442
MMM	0.947	0.952	0.700	0.728	0.698	0.789
MRK	0.962	0.976	0.756	0.520	0.446	0.858
MSFT	0.892	0.865	0.797	0.820	0.820	0.773
NKE	0.949	0.946	0.935	0.887	0.908	0.874
PG	0.900	0.903	0.568	0.589	0.574	0.424
TRV	0.866	0.866	0.869	0.359	0.121	0.314
UNH	0.823	0.865	-1.517	-2.239	-0.512	0.664
CRM	0.947	0.930	0.892	0.913	0.917	0.902
VZ	0.975	0.975	0.917	0.910	0.922	0.918
V	0.814	0.805	0.652	0.708	0.611	0.465
WBA	0.967	0.965	0.913	0.890	0.879	0.910
WMT	0.888	0.885	0.808	0.760	0.762	0.612
DIS	0.965	0.963	0.924	0.899	0.895	0.954
DOW	0.956	0.953	0.944	0.937	0.917	0.806

Table 4: Performance of MMGAN-HPA.

Stock ticker	MAE	MSE	CORRELATION
TCS	0.00263397	0.00003490	0.99586647
BHEL	0.00267097	0.00002290	0.99686559
WIPRO	0.00239697	0.00002400	0.99716163
AXISBANK	0.00280396	0.00003020	0.99837178
MARUTI	0.00221797	0.00002160	0.99721564
TATASTEEL	0.00463494	0.00006770	0.99736265

MMGAN-HPA is also an efficient hybrid model. Seen from Table 4, MAE, MSE of stocks used throughout the experiment are very low, the correlation prediction performance is very high, demonstrating the algorithm's effectiveness.

A FASTRNN_CNN_BiLSTM model was proposed in a recent study (Yadav, 2022). The RMSE and calculation time are used to evaluate the models' performance. Out of the 500 stocks in the trial, 86% were used for training and 14% were used for testing. The four businesses' stock values, i.e., Apple, Facebook, Nike, and Uber, have been used to test the models. Nine additional cutting-edge models are contrasted with these suggested models. It is evident that the suggested models outperform other models in terms of computing time and RMSE. Though the computation time of FBProphet, it has a higher RMSE. The suggested models perform optimally across a variety of stocks overall. A typical results for Apple is given in Table 5.

Table 5: RMSE and computation time calculated for the state-of-the-art and the proposed models for Apple Inc. stock values.

Model name	RMSE	Time (in s)
ARIMA	0.796109	1.63
BiLSTM_Attention_CNN_BiLSTM	0.234644	25.72292113
CNN_LSTM_Attention_LSTM	0.214821	16.00164294
FBProphet	0.935556	0.659962893
LSTM	0.228731	13.28157353
LSTM_Attention_CNN_BiLSTM	0.263613	19.96186757
LSTM_Attention_CNN_LSTM	0.27994	17.32732081
LSTM_Attention_LSTM	0.299334	19.28274226
LSTM_CNN_BiLSTM	0.23489	16.76800251
FastRNN (proposed)	0.202456	3.337492943
FASTRNN_CNN_BiLSTM(proposed)	0.205647	13.49208355
Abbreviations: ARIMA, auto regressive integrated moving average; RMSE,root mean squared error		

5 LIMITATIONS AND PROSPETCS

The stock data from the Chinese stock market is used to train certain mixed models. When utilizing the model on foreign stock markets, there will be an obvious lag, which might be connected to how Chinese and international stock markets trade. Such difference can lead to situations where the trained model cannot be applied to multiple different stock markets. The data in the real stock market is noisy, which may lead to poor results if used directly for training. Additionally, overfitting phenomena might exist while training. To some model, the result is good while training, however, they perform poorly when using other data.

Numerous causes, including abrupt political events, changes in economic policy, and major global events, frequently have an impact on the stock market. These factors are often unpredictable and can lead to errors in model predictions. In the subsequent development of stock prediction, more consideration can be given to how market sentiment, financial news and economic policies affect the fluctuating pattern of shares, so that the forecasting model will not only rely on stock historical data, but become more comprehensive.

6 CONCLUSIONS

To sum up, this research focuses on the concept and utilization of hybrid prediction models in predicting stock prices. This article first introduces some traditional and deep learning-based single models and the relevant background of stock forecasting, and then introduces some cutting-edge hybrid model configurations. The prediction results of these models were compared. By analysing MAE, RMSE and other performance analysis indicators, it can be found that these hybrid models have a great improvement compared with the single model, and different models have different advantages. In the subsequent development of stock prediction, more consideration can be given to how market sentiment, financial news and economic policies affect the fluctuating pattern of shares, so that the forecasting model will not only rely on stock historical data, but become more comprehensive. The research on hybrid model stock forecasting is helpful to understand its application in the stock market, better forecast stocks, and lay the foundation for the establishment of more diversified and effective models in the future.

REFERENCES

Althelaya, K. A., El-Alfy, E. S. M., Mohammed, S., 2018. *Evaluation of bidirectional LSTM for short-and long-term stock market prediction*. 2018 9th international conference on information and communication systems (ICICS), 151-156.

Bates, J. M., Granger, C. W., 1969. *The combination of forecasts*. Journal of the operational research society, 20(4), 451-468.

Gülmez, B., 2023. *Stock price prediction with optimized deep LSTM network with artificial rabbits optimization* *algorithm.* Expert Systems with Applications, 227, 120346.

Hassan, M. R., Nath, B., 2005. *Stock market forecasting using hidden Markov model: a new approach*. 5th international conference on intelligent systems design and applications (ISDA'05), 192-196.

Huang M., 2023. *Empirical research on stock price prediction based on ARIMA model.* Neijiang Science and Technology, 3, 61-62.

Liu, X., 2022. *Research on Multi-factor Stock prediction combined with quantitative trading*. Master's Thesis of Soochow University.

Lu, W., Li, J., Wang, J., Qin, L., 2021. *A CNN-BiLSTM-AM method for stock price prediction*. Neural Computing and Applications, 33(10), 4741-4753.

Shao, G., 2022. *Stock price prediction based on multifactorial linear models and machine learning approaches.* 2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), 319-324.

Srivinay, Manujakshi, B. C., Kabadi, M. G., Naik, N., 2022. *A hybrid stock price prediction model based on PRE and deep neural network.* Data, 7(5), 51.

Vullam, N., Yakubreddy, K., Vellela, S. S., Sk, K. B., Reddy, V., Priya, S. S., 2023. *Prediction And Analysis Using A Hybrid Model For Stock Market.* 2023 3rd International Conference on Intelligent Technologies (CONIT) (pp. 1-5. IEEE.

Yadav, K., Yadav, M., Saini, S.: *Stock values predictions using deep learning based hybrid models.* CAAI Trans. Intell. Technol. 7(1), 107–116.

Zhang, Q., Lin, T., Qi, X., Zhao, X., 2020. *Overview of stock prediction based on Machine learning.* Journal of hebei academy of sciences, 4), 15-21.

Zhang, S., Su, C., 2017. R*esearch on stock index prediction based on sentiment dictionary and BERT-BiLSTM.* Computer Engineering and Applications 1-16.

Zhao, T., Han, Y., Yang, M., et al., 2021. *A review of time series data prediction methods based on machine learning*. Journal of Tianjin university of science and technology, 5, 1-9.