Stock Price Forecast: Comparison of LSTM, HMM, and Transformer

Qianzhun Wang[1] (✉) and Yingqing Yuan[2]

[1] Intelligent Manufacturing Engineering, School of Mechatronics Engineering and Automation, Shanghai University, Shanghai, China
nemowang@shu.edu.cn

[2] Finance, School of Economics, Ocean University of China, Qingdao, China

Abstract. With the development of deep learning, different kinds of neural network models are applied to the analysis and prediction of time series data. In the field of finance, deep learning models are widely used to forecast the stock market, which is an integration of technical data that can directly provide advice to investors. We chose three neural network models that have been very popular in the last decade: Long Short-Term Memory (LSTM), Hidden Markov model (HMM), and Transformer. We use the data of the new energy vehicles sector in the A-share market to establish and evaluate the model and compare the predictive performance of the three models. The result shows that Transformer performed the best-predicting capability of stocks of the new energy sector in the A-share market. The model's performance was quantified using the Mean Absolute Percentage Error (MAPE) and Matthews Correlation Coefficient (MCC).

Keywords: LSTM · HMM · Transformer · Stock Price Prediction · Time-series Forecasting

1 Introduction

The stock market is an important part of today's financial markets. The stock market has always been one of the most popular investing target because of its high returns. However, due to the unpredictability of the stock market, investing in the stock market carries a high level of risk. There is a lot of research in the academic community on stock price forecasting, and it is dedicated to finding a more suitable stock price prediction model for the stock market. In recent years, new energy vehicles have become a key industry in the world. In 2021, China/overseas electric vehicle sales will be 3.5 million/3 million units, up 158%/57% year-on-year. On the one hand, in 2021, the domestic new energy vehicle market will continue the high boom since the second half of the 20th year. The annual sales of new energy vehicles will reach 3.5 million, a year-on-year increase of nearly 160%, and the penetration rate will exceed 13%. The 22-year high boom continues, with sales of 2 million units from January to May, doubling year on year. In 2022, as the domestic epidemic recedes, the marginal impact will gradually weaken.

Q. Wang and Y. Yuan—These authors contributed equally to this work and should be considered co-first authors.

J. Yen et al. (Eds.): ICBIS 2023, AHCS 14, pp. 126–136, 2023.
https://doi.org/10.2991/978-94-6463-198-2_15

Many of the forecasting research has employed the statistical time series analysis techniques like HMM. Hassan and Nath [1]'s research showed that HMM is explainable and has a solid statistical foundation. In recent years, increasing number of stock prediction systems were based on AI techniques, including artificial neural networks (ANN) and Transformer [2] Chen et al. [3] used LSTM to highly improved the accuracy of stock prediction in the Chinese stock market. Nadeem Malibari1 and Iyad Katib [4] used the Transformer [2] to predict closing prices has a probability above 90% (which is at most 72% in other ways).

There are many stock price forecasting models based on financial time series analysis and artificial intelligence algorithms, and each has its advantages and disadvantages. LSTM is suitable for processing and predicting the important events of interval and long delays in time series. However, for LSTM, it's essentially critical to choose sequence learning features. There's a limit not to including the economic fundamentals, but also not the technical analysis data to avoid the co-founding pitfalls. HMM is proved to be more efficient in extracting information from the dataset. Transformer [2] employs a multi-head self-attention mechanism to learn the relationship among different positions globally, so that it can enhance the capacity of learning long-term dependencies. The specific differences in the prediction of stock prices between models have never been shown in previous studies, which has created difficulties for investors in the choice of models.

While the new energy vehicle industry maintains high growth, with the resumption of work and production and the introduction of land subsidy policies, the supply side and demand side of the industry It is expected to be gradually repaired. In the context of carbon neutrality, favorable overseas policies are frequently issued, the electrification process of mainstream car companies is accelerating, and high-quality supply is coming one after another. The demand for new energy vehicles is expected to continue to rise. On the other hand, the current development of electric vehicles in the world is still greatly affected by policies. If the follow-up stimulus policies do not meet expectations or the policy continuity is not strong, it will hurt the promotion of electric vehicles. In this context, the stock price changes of the new energy vehicle industry are ambiguous, and the research on its stock price forecast has great investment significance.

To solve the confusion about the future stock price of new energy vehicles and the hesitation between model selection, this paper is dedicated to using HMM, LSTM, and Transformer [2] models to study the stock price of China's new energy vehicle industry respectively, and then concludes that Transformer [2] is better than other models in the new energy vehicle industry, thus bringing some enlightenment to investors. Furthermore, to better show investors the differences in stock price forecasts of various models, this paper uses two-dimensional indicators to evaluate the prediction performance--MAPE and MCC.

2 Related Works

2.1 Stock Price Prediction

Since the invention of deep learning, quantitative finance starts to use the ANN models to predict stock prices. Normally, the researchers treat stock data as time-series data and build models to analyze its trend, periodicity, and volatility, so that the future price of the stock can be predicted. In the 1970s, Box and Jenkins [5] proposed the ARIMA model (Autoregressive Integrated Moving Average model), which is a simple linear-regression model according to historical data. ARIMA is suitable for data with high correlation and stability and has a good short-term prediction effect. In 2014, Adebiyi A. Ariyo and Adewumi O. Adewumi [6] used the ARIMA model to predict stock price. Their research presents the extensive process of building a stock price predictive model using the ARIMA model, and its result also indicated that the ARIMA model only has a potential for short-term prediction. Then with the rise of neural networks, more effective deep learning models were proposed. Another prediction system that is been widely used is artificial neural networks (ANN), so as its many extensions. The LSTM model was proposed in 1997 by Jürgen Schmidhuber [7], and used in predicting stocks by Chen Kai and Zhou Yi [3] in 2015. They use LSTM to predict Chinese stock returns, which improved the accuracy of stock returns prediction from 14.3% to 27.2%. Liu and Ma [8] introduced a quantum artificial neural network (QENN) to predict closing prices.

To solve the NP-complete problem of Recurrent neural networks (RNN), Hassan and Nath [1] proposed in 2005 a new approach for stock market forecasting – the Hidden Markov Model (HMM). They considered the opening price, closing price, highest price and the lowest price as 4 input features. It is shown that HMM has similar MAPE (mean absolute percentage errors) results as ANNs, but HMM is explainable and has a solid statistical foundation which is the weakness of ANNs. Li [9] also applied the normal hidden Markov model to Ping An Bank's stock price data. He divides stock states into bear markets and bull markets, corresponding to the two states in the hidden Markov model that can be transformed into each other. He estimated the parameters according to the data, then decoded it to find out the state hidden behind each data, and finally made the state prediction and the closing price distribution prediction respectively, and obtained a more reasonable result.

Although LSTM and HMM are two commonly used stock price prediction models, they suffer from long-range dependencies due to their lack of injection of attention mechanisms. In the follow-up research, scholars then proposed models related to the attention mechanism.

2.2 Transformer Deep Learning Model

In 2017, Transformer was proposed by Google Braint [2] is a deep learning model that adopts the mechanism of self-attention, differentially weighting the significance of each part of the input data, which is at first used in natural language processing and computer vision. But lately, researchers tried to use it in finance fields. Nadeem Malibari1 and Iyad Katib [4] used stock data from the Saudi Stock Exchange (Tadawul) to build the Transformer model [8], the result of which predicting closing prices has a probability above 90%(which is up to 72% in other ways). Researchers also compared Transformer [2] to other deep learning models for time-series data analysis. Li et al. [10] compared three kinds of time series forecasting models: the deep state space model(DSSM), the deep autoregressive model (DeepAR), and the Transformer [2] model. The result shows that all three methods are better than ARIMA, of which Transformer [2] is the best.

Traditional statistical models model time series individually, such as LSTM and HMM, with strict logic and can reflect the overall characteristics of the sequence. Artificial intelligence algorithms, such as the latest Transformer [2] method in academia, are a dynamic process with inductive reasoning as the core idea, which can better characterize potential influencing factors. As an emerging industry in the world, new energy vehicles have attracted much attention. Investors expect to make profits in this industry, but the development of this industry is greatly affected by policies. Therefore, the stock price forecast of the new energy vehicle industry has investment significance. This paper aims to use LSTM, HMM and Transformer [2] models to apply to the prediction analysis of stock data in the new energy vehicle industry market, to more intuitively observe the differences between the three models and give investors some enlightenment.

3 Method

This article aims to study stocks in the new energy vehicle market. We have used three different architectures, HMM, LSTM, Transformer [2] to predict the stock price. The three models are all realized by PyTorch environment and packages of python.

To better describe the stock price changes in the new energy vehicle market, this paper selects the data from 2019-6-17 to 2022-6-17 of NEW ENERGY VEHICLES (399976.SZ). The CSI New Energy Vehicle Index, which involves lithium batteries, charging piles, new energy vehicles, and other companies from the Shanghai and Shenzhen markets, are to reflect the overall performance of securities of listed companies related to new energy vehicles. We use python to realize the models based on PyTorch. For a fair comparison, we take the same data source of the index in the new energy section as inputs for each model, including the open point, close point, highest point, lowest point, and trading volume of the index each day.

The data varies in a range of 1000 to 7000. The first step is to standardize the data. When using prices and volume data, all the stock data must be within a typical value range. Generally, machine learning algorithms converge faster or perform better when they are close to normally distributed and/or on a similar scale. We use the MinMaxScaler function to scale the data to the range $[-1, 1]$ so that it can cause less error in the following steps.

$$\begin{aligned} X_{std} &= \frac{X - X.min(axis=0)}{X.max(axis=0) - X.min(axis=0)} \\ X_{scaled} &= X_{std} * (max - min) + min \end{aligned} \tag{1}$$

After the three trained models get the test results, we use Root Mean Squared Error (RMSE) to measure how well each model fits the actual data. RMSE is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are and a measure of how to spread out these residuals are. The formula of RMSE is as follows.

$$RMSE_{fo} = \sqrt{\frac{1}{N}\sum_{i=1}^{n}\left(z_{f_i} - z_{o_i}\right)^2} \tag{2}$$

where, $\begin{matrix} f = forcasts \\ o = observed\ values \end{matrix}.$

3.1 LSTM Model

LSTM introduces memory, one of the computational units that replace artificial neurons in the hidden layers of the network. Through the memory unit, the network can effectively associate and timely input remote data, to adapt to the dynamic structure of real-time prediction, and can display. An LSTM cell consists of an input gate, cell state, forget gate, output gate, a sigmoid layer, tanh layer, and point-wise multiplication operation. The equation of LSTM is as follows

$$C_t = g_{forget} \otimes C_{t-1} + g_{in} \otimes \tilde{C}_t \tag{3}$$

$$\tilde{C}_t = f(W \cdot x_t + V \cdot h_{t-1}) \tag{4}$$

$$h_t = g_{out} \otimes f(C_t) \tag{5}$$

$$y(t) = h(t) \tag{6}$$

$$g_{in}(t) = sigmoid(W \cdot x_t + V \cdot h_{t-1} + U \cdot C_{t-1}) \tag{7}$$

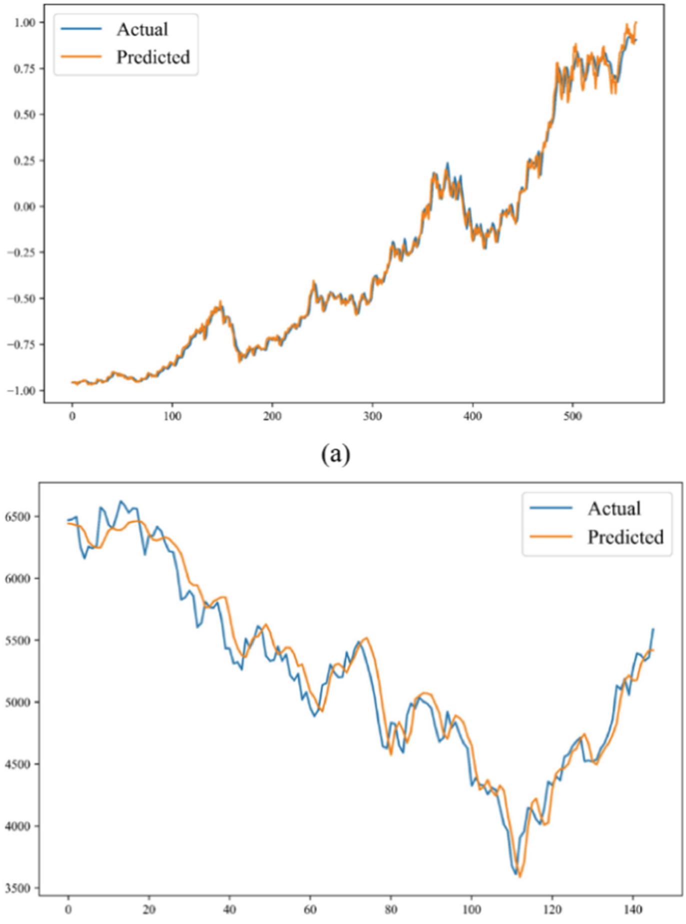
$$g_{forget}(t) = sigmoid(W \cdot x_t + V \cdot h_{t-1} + U \cdot C_{t-1}) \tag{8}$$

$$g_{out}(t) = sigmoid(W \cdot x_t + V \cdot h_{t-1} + U \cdot C_t) \tag{9}$$

As mentioned above, our model has four inputs. Therefore, the data feature number of the model is 4, the number of neurons in the hidden layer is 32, the number of layers of LSTM is 2, and the feature number of the predicted value is 1. After the model was instantiated, Adam was used to optimizing the algorithm and the mean square error was used as the loss function. The division ratio of the training set and test set is 8:2. Then Each group of time series contains 30 data, which are iterated 1000 times. The RMSE of the model on the testing set is 156.4924.

The predicting data of the training set and the real one are plotted in Fig. 1 (a). The predicted data fits well with the real data on the training set. The predicting data of the testing set and the real one are plotted in Fig. 1 (b). The predicted data fits poorly with the real data on the testing set.

It can be seen from the above results that in the sample of CSI new energy we selected, the stock price prediction performance of LSTM is average. The RMSE (Root Mean Squared Error) of the train is 0.01, and the RMSE of testing is 0.57.



(a)

(b)

Fig. 1. Good results of the training set and test set in LSTM

3.2 HMM Model

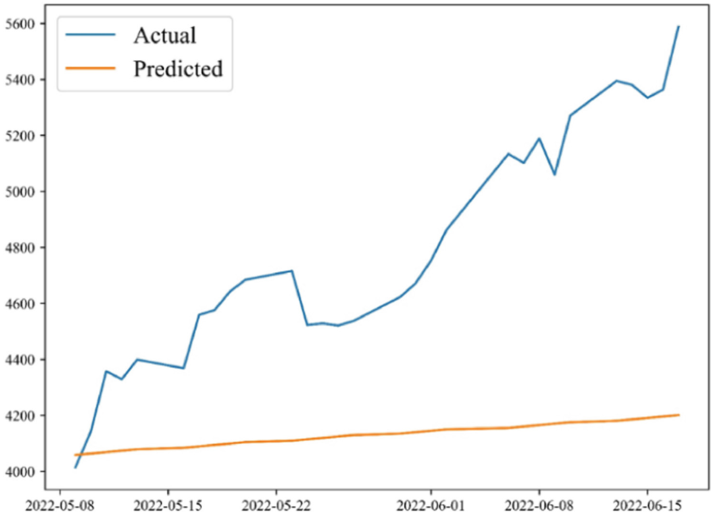
Hidden Markov models are based on a set of unobserved latent states, and each state is associated with a possible transition. Between base states are usually not obvious to investors. The transition of the basic state is based on the company's policies, decisions, economic conditions, etc.

An HMM (denoted by λ) model can be written as follows:

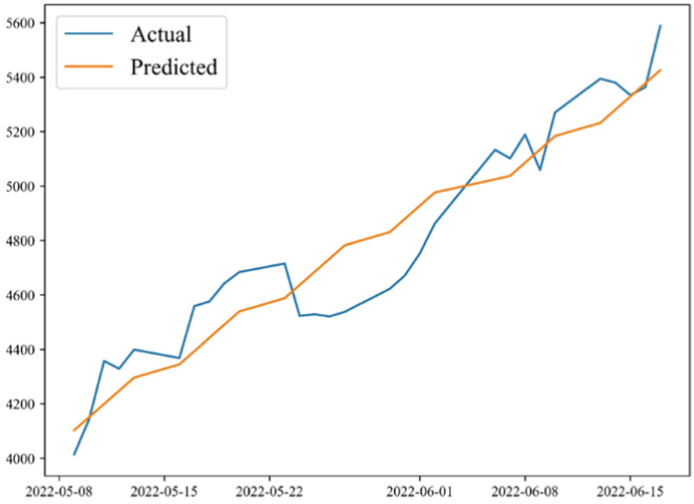
$$\lambda = (\pi, A, B) \tag{10}$$

where A is the transition matrix whose elements give the probability of a transition from one state to another, B is the emission matrix $b_j(O_t)$ giving the probability of observing O_t when in state j and π gives the initial probabilities of the states at $t = 1$.

The predicting data of the testing set and the real one using a 3-year training set are plotted in Fig. 2 (a). From the picture, we can see that the prediction effect of the HMM model trained with 3 years of data is not satisfactory, and it is very different from the actual data in terms of basic trend prediction and specific data prediction. This reflects that the HMM model is not suitable for long-term forecasting. To test the short-term prediction ability of the HMM model, we selected the CSI stock index data from



(a)



(b)

Fig. 2. The prediction of HMM using a 3-year training set and a 1-year training set that are not satisfying

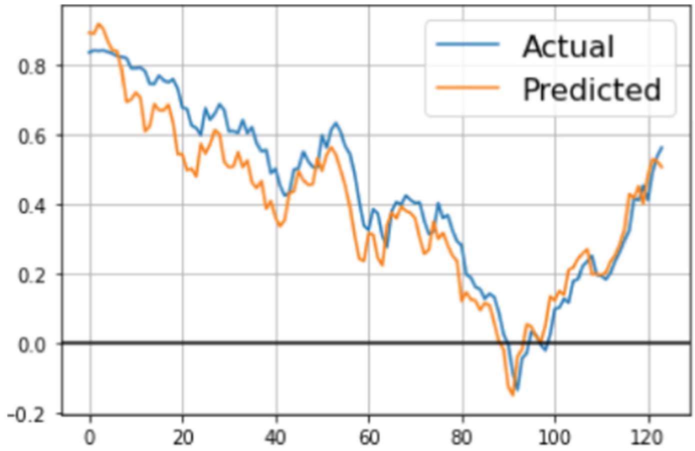


Fig. 3. The prediction of Transformer, which is the best result we get from the three methods

2021-1-25 to 2022-6-17 as the training set of the HMM short-term model, and obtained the stock price prediction results of the short-term HMM model, as shown in Fig. 2(b) shown.

By comparison, we can see that the prediction effect of the short-term HMM model has been greatly improved in stock price prediction compared with the HMM model trained with 3 years of data. The short-term HMM model can fit the trend of stock price changes, but there is still a gap between the specific data and the actual stock price.

3.3 Transformer

In 2017, Transformer [2], a well-known sequence-to-sequence model, achieved great success on natural machine translation tasks. Transformer [2] employs a multi-head self-attention mechanism to globally learn the relationship between different locations, enhancing the ability to learn long-term dependencies.

Transformer [2] proposes a self-attention mechanism, the core formula of which is as follows.

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{11}$$

The results of stock price prediction under the Transformer [2] method are shown in Fig. 3. From the figure, we can see that the prediction result of the stock price prediction model obtained by using the Transformer [2] has a high degree of fit with the actual stock price. So far, this is the best result we got from the three methods.

4 Discussion

The new energy vehicle market has become a hot topic in recent years, and it has also attracted the attention of the capital market. At the same time, the stock price of the new energy vehicle market is difficult to predict because the industry is seriously affected

Table 1. RMSE of three models

Model	HMM	LSTM	Transformer
RMSE	156.4924	132.0384	0.0730

Table 2. $\boldsymbol{R}^2$, MSE and MAE of three models

Accuracy	Model		
	HMM	LSTM	Transformer
$\boldsymbol{R}^2$	0.9553	0.8961	0.9275
MSE	24,489.8730	17,434.1378	0.0053
MAE	125.8156	118.8938	0.0603

by national policies and industry guidance, which makes people feel confusing. Under this circumstance, it is extremely urgent to apply stock price forecasting research in the new energy vehicle market. However, according to our observation, the existing stock price forecast research mainly focuses on the A-share market and has not been deeply cultivated in a certain industry. The stock price research in the field of new energy vehicles is relatively lacking.

We extracted the stock price data of CSI New energy in the past year to train and test HMM, LSTM and Transformer [2] models respectively. To compare the fit of the three model test results with the actual stock price, we use RMSE to measure the distribution of the residuals of each model. The RMSE of the test results obtained by each model is shown in Table 1.

From Table 1 we can see that the RMSE of the Transformer [2] model is much smaller than that of HMM and LSTM. This represents the best fit between the Transformer [2] test results and the actual data.

To further compare the three models, we again selected three indicators $\boldsymbol{R}^2$, MSE and MAE to measure the fitting effect, as shown in Table 2. We can see that Transformer [2] has the smallest error regardless of which fitting metric is used. Transformer [2] can almost capture the trends and gives more accurate predictions than HHM and LSTM.

5 Conclusion

This article aims to compare the stock price prediction models applied to the new energy vehicle market. We have selected three models of HMM, LSTM, and Transformer for comparison. It can be seen from our research that the short-term train model of HMM can predict the direction of stock price movement, but it is not suitable for long-term prediction. LSTMs work better in long-term predictions. Among the three models, the best stock price prediction effect is Transformer, which can closely track the trend of stock prices and make predictions in line with the direction of stock price changes.

The sample selected in this paper is the CSI new energy stock index from 2019–6-17 to 2022–6-17, and there may be differences in the prediction results based on different samples at different time points. We learned that the three models have been revised more carefully in the existing research so that the prediction results of these three models are more in line with expectations. This paper's efforts in this area are subpar. In conclusion, the Transformer has the best stock price prediction effect among the three models we selected. The analysis of these types of trends and cycles will give more profit to the investors.

References

1. M. R. Hassan and B. Nath, Stock market forecasting using hidden Markov model: a new approach, 5th International Conference on Intelligent Systems Design and Applications (ISDA 2005), 2005, pp. 192–196, https://doi.org/10.1109/ISDA.2005.85.
2. Vaswani, Ashish, et al., Attention is all you need., Advances in neural information processing systems 30 (2017).
3. Chen K, Zhou Y, Dai F., A LSTM-based method for stock returns prediction: A case study of China stock market[C]//2015 IEEE international conference on big data (big data). IEEE, 2015: 2823–2824.
4. Malibari N, Katib I, Mehmood R. Predicting Stock Closing Prices in Emerging Markets with Transformer Neural Networks: The Saudi Stock Exchange Case. International Journal of Advanced Computer Science and Applications, 2021, 12(12).
5. Box G E P, Jenkins G M, Reinsel G C, et al. Time series analysis: forecasting and control, John Wiley & Sons, 2015.
6. A. A. Ariyo, A. O. Adewumi and C. K. Ayo, Stock Price Prediction Using the ARIMA Model, 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, 2014, pp. 106–112, https://doi.org/10.1109/UKSim.2014.67
7. Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation, 1997, 9(8): 1735-1780.
8. Ge Liu, Wenping Ma, A quantum artificial neural network for stock closing price prediction, Information Sciences,Volume 598, 2022, Pages 75–85,ISSN 0020–0255
9. Renjun Li, The Applications of Hidden Markov Model in Quantitative Investment, Tsinghua University, 2019. https://doi.org/10.27266/d.cnki.gqhau.2019.000565.
10. Li Wen, Deng Sheng, Duan Yan, et al. Time Series Prediction and Deep Learning: Literature Review and Application examples. Computer Applications and Software, 2020, 37(10): 64-70.