The Application of a Combination Model Based on Machine Learning in Offshore RMB Exchange Rate Forecasting

Yujia Liao*

Department of Economics and Management, South China Normal University, Guangzhou, China

*Corresponding author: 20230731023@m.scnu.edu.cn

Abstract. With the vigorous development of China's economy, the internationalization level of Renminbi (RMB) is still improving, and the offshore RMB market is also expanding. Compared with onshore RMB, the price of offshore RMB is formed by free bidding of buyers and sellers. Market supply and demand are two extremely important factors affecting the exchange rate of offshore RMB, and the fluctuation of exchange rate is relatively free, so the offshore RMB exchange rate is relatively volatile. Under such circumstances, it is of great significance to forecast the offshore RMB exchange rate. This review focuses on the application of combinatorial models related to machine learning in offshore RMB exchange rate forecasting, and systematically combs the data types, modeling methods, and related forecasting performance involved in recent research. This paper evaluates the application effectiveness of the machine learning fusion Autoregressive Integrated Moving Average (ARIMA) model, supervised learning based on neural network and random forest, and the Long Short-Term Memory-Temporal Convolutional Network-Convolutional Neural Network (LSTM-TCN-CNN) hybrid model in offshore RMB exchange rate forecasting, compares and analyzes the advantages of each model, and summarizes the applicable scenarios of different models. The results show that the machine learning fusion ARIMA model is suitable for the prediction of time series with obvious rules and a small amount of data, the combination model of random forest and neural network is suitable for the prediction of structured data, and the LSTM-TCN-CNN model is suitable for the prediction of multivariate time series with a large amount of data and high dimension.

Keywords: machine learning, portfolio model, offshore RMB, exchange rate forecasting.

1. Introduction

With the continuous improvement of the level of RMB globalization, the status of offshore RMB in the global financial system is constantly improving. The offshore RMB exchange rate is affected by international market, monetary policy, geopolitics, market sentiment, and other factors, showing significantly high volatility characteristics, so the research on the offshore RMB exchange rate forecast has important practical significance. From the macro level, the study is helpful for regulators and decision-makers to judge the trend of the offshore RMB market and grasp the trend of the exchange rate market, maintaining national financial stability. From the micro level, for many offshore RMB users, whether enterprises or individuals, more accurate exchange rate forecasts provide a more reliable basis for decision-making, help them to better manage risk management, avoid losses, and obtain greater profits.

In the existing research, many scholars have carried out research work of offshore RMB exchange rate prediction: Sun et al. Used the multifractal method to predict the onshore and offshore RMB exchange rates, and analyzed the fluctuation relationship between them [1]. Wang added NDF to the NARX network for model training and performance promotion; the error is less than 0. 005, and achieved good results [2]. Zhan selected the spot exchange rate of RMB against the US dollar in the offshore market of Hong Kong from January 2, 2017 to March 30, 2018, as a data set, and used the BP neural network model to construct and train it. The final training mean square error was 0.0006, and the prediction effect was good [3]. With the development and application of machine learning methods, especially supervised learning and deep learning methods, new tools are provided for the research of offshore RMB prediction. At the same time, with the rise of the application of the portfolio model in forecasting research, it also provides new possibilities for offshore RMB forecasting.

This paper focuses on the application of the combination model of machine learning in the prediction of offshore RMB exchange rate, compares the advantages and disadvantages of different models, and fills in the blank of the comparative study of the combination model of machine learning.

2. Theoretical Background

2.1. Traditional Offshore RMB Exchange Rate Forecast

The traditional forecasting methods of offshore RMB exchange rate mainly include time series models (typically, Autoregressive Integrated Moving Average (ARIMA) models) and econometric models, which improve the accuracy of the model to a certain extent, but due to the complexity of the fluctuation of offshore RMB exchange rate, they still have limitations in data adaptability and generalization ability of the results.

In the early days, some scholars also applied neural networks to the prediction of offshore RMB exchange rate, for example, using NARX networks and adding NDF to predict the offshore RMB exchange rate, but the effectiveness of NDF and the promotion of the network still need to be further verified [2]. In addition, in order to deal with the nonlinear characteristics of exchange rate data, the combination model has also been applied in the prediction of the offshore RMB exchange rate.

2.2. Combinatorial models of machine learning in exchange rate forecasting applications

As the data dimensions increase, machine learning and more accurate combination models are gradually applied to predicting the offshore RMB exchange rate. Supervised learning, with its natural advantages in dealing with complex nonlinear relationships, has gradually become a new tool and trend for exchange rate forecasting. In the supervised learning model, all kinds of data can be easily input as input features (X), which just caters to the extensive characteristics of offshore RMB exchange rate factors. Gao et al. The research shows that the combination of single time series neural network model and random forest model can improve the accuracy of offshore RMB exchange rate prediction, and the time of the algorithm is less than 60 seconds, which is less than exchange rate time interval of the input characteristics of the exchange rate prediction model, and can meet the needs of short-term exchange rate prediction [4]. It has certain significance for the short-term prediction of the offshore RMB exchange rate.

At the same time, the combination model also has a good effect in the application of exchange rate forecasting because of its "complementary" characteristics. In the research of Zhou, it uses ARIMA fusion machine learning model to fit, uses ARIMA model to capture linear components in time series, and then uses machine learning model to capture nonlinear residual patterns that ARIMA model can not explain. Experiments show that the ARIMA model has the best performance when making a single model prediction, and the fusion model can significantly improve the prediction performance of a single model and achieve complementary results no matter what machine learning model is used [5].

To sum up, the combination model of machine learning can combine the advantages of each model, and improve the accuracy of the model after combining a single model with other models, so as to achieve a complementary effect.

3. Case studies

3.1. Case study 1: Application of Machine Learning Fusion ARIMA Model in Offshore RMB Exchange Rate

With the wide application of machine learning methods, it has also become an important tool for forecasting the offshore RMB exchange rate. Because the exchange rate series has the basic characteristics of financial time series, such as a certain degree of autocorrelation, the time series method is also applied to predicting the offshore RMB exchange rate. Zhou used the method of neural network and random forest fusion ARIMA model in his research, and selected all the daily data of offshore RMB exchange rate from May 2012 to September 2019 as the data set by virtue of the advantage of ARIMA model in dealing with linear characteristics and the nonlinear modeling method of machine learning. RMSE, MAE, MAPE, Theil-U and DAR are used to comprehensively evaluate the Predictive performance of machine learning, ARIMA and their fusion model.

The experimental results show that the RMSE, MAE and Theil-U index of ARIMA model are all less than 0. 03, and the MAPE is 0. 2622, which has the smallest prediction error compared with other models used in the study. Among the fusion models, the best one is the ARIMA fusion random forest model, its RMSE is only 0.0193, Theil-U index of only 0.0014, and DAR of 0.8218 [5]. The robustness tests on the number of units in the input layer, the number of hidden layers and the number of neurons show that the ARIMA model has the best performance when making a single model prediction, and the prediction performance of the model is improved no matter which machine learning model is combined. In addition, the integration of machine learning and ARIMA model also shows higher flexibility in dealing with exogenous variables.

3.2. Case study 2: Research on Forecasting of Offshore RMB Exchange Rate Based on Neural Network and Random Forest

Neural networks and random forests are important methods of supervised learning in machine learning. Gao proposed NN-RF exchange rate forecasting model, GRU-RF exchange rate forecasting model and LSTM-RF model based on the combination of random forest (RF) and recurrent neural network (RNN), gated recurrent neural network (GRU), and long-term and short-term memory model (LSTM). The researchers selected the foreign exchange trading price data of offshore RMB against six currency pairs, including the US dollar, the euro, the pound sterling, the Singapore dollar, the Australian dollar, and the Hong Kong dollar, for the training and testing of the exchange rate prediction model, and combined the Huber function to ensure the error and accuracy of the model. Finally, the prediction accuracy of the model is evaluated by mean square error (MSE) in different time intervals.

The experimental results show that the three exchange rate prediction models have high algorithm complexity and long short-term exchange rate prediction time (but less than 60 seconds), and they all improve the accuracy of exchange rate prediction by combining with the random forest model [4]. And the prediction accuracy when the time interval of the input historical exchange rate data is at the minute level is greater than the prediction accuracy when the time interval of the input historical exchange rate data is at the day level.

3.3. Case study 3: Prediction of offshore RMB exchange rate based on LSTM-TCN-CNN hybrid model

The nonlinear characteristic of exchange rate fluctuation is an important problem to be solved in forecasting the offshore RMB exchange rate. To deal with this problem, Yue et. Al proposed a hybrid model of LSTM-TCN-CNN, which aims to combine the ability of LSTM to capture long-term dependencies, the ability of TCN to process long-term historical information, and the ability of CNN to extract local features, so as to achieve complementary effects. The researchers selected the offshore RMB exchange rate data from April 30, 2012, to February 27, 2025, and added the EPU and VIX indices to improve the prediction accuracy of the model. After data preprocessing, the model was divided into a training set and a test set, and three-level modules were used for modeling and training. Finally, the determination coefficient (R *), MAE, MSE, and MAPE are selected as the evaluation indices to comprehensively measure the prediction effect of the model.

The experimental results show that the MAPE values of the LSTM model, the LSTM-TCN hybrid model, and the LSTM-TCN-CNN hybrid model are 0. 4528%, 0.2889% and 0. 2349%, respectively [3]. It shows that the prediction effect is significantly better than that of a single LSTM model in the prediction of offshore RMB exchange rate by integrating the advantages of LSTM, TCN, and CNN.

At the same time, it also shows that the mixed model has better prediction ability than a single model in the prediction of offshore RMB exchange rate.

3.4. Comparative Analysis of Different Models

There are differences in feature extraction among the models. The machine learning fusion ARIMA model mainly uses the HP filtering method to decompose the exchange rate data series into trend series and periodic series, the combination model of random forest and neural network extracts features based on neural network units, and the LSTM-TCN-CNN model outputs features through the TCN layer.

There are differences in the time levels of exchange rate data used by the models. The combined model of random forest and neural network uses minute-level historical exchange rate data, while the machine learning fusion ARIMA model and LSTM-TCN-CNN model use day-level historical exchange rate data.

Different models have different prediction accuracy. The LSTM-TCN-CNN model shows the lowest MSE in daily exchange rate forecasting, the combined model of random forest and neural network shows the lowest MSE in minute exchange rate forecasting, and the machine learning fusion ARIMA model performs well in the fitting of exchange rate series.

Different models have their own advantages and applicability. The machine learning fusion ARIMA model is suitable for the prediction of time series with obvious rules and a small amount of data, the combination model of random forest and neural network is suitable for the prediction of structured data, and the LSTM-TCN-CNN model is suitable for the prediction of multivariate time series with a large amount of data and high dimension.

4. Discussion

4.1. Limitation

At present, the combination model based on machine learning still faces challenges in the prediction of offshore RMB exchange rate, mainly including the following three aspects:

4.1.1 Data issues

Offshore RMB exchange rate forecasting requires high quality and breadth of basic exchange rate data, and model training also depends on the support of information. However, nowadays, market sentiment and exchange rate fluctuations still lag behind in updating multi-dimensional data, and information can not be disclosed in real time. The delay of data release affects the ability of the model to capture market mutations and micro-dynamics, thus affecting the prediction performance of the model and reducing the accuracy and reliability of the prediction.

4.1.2 Model problem

The combined model itself is complex, although it is conducive to capturing complex patterns and features, but it is also very easy to memorize the noise and details in the training data, rather than learning general rules, leading to over-fitting. And this kind of complex combination model needs to consume more training time and computing resources than a single model. In the combination research of neural network and random forest, the exchange rate prediction model itself has high complexity and many parameters, which requires longer training time and higher computing resources, especially in long time series and large-scale data sets. May face high computational overhead and memory requirements. This may lead to fatal defects in real-time prediction.

4.1.3 Practical issues

Although the combination model based on machine learning has a good application prospect in the prediction of offshore RMB exchange rate, the foreign exchange market is highly complex, the economic situation at home and abroad changes rapidly, and the exchange rate changes are affected by many factors, such as the international political situation, the uncertainty of fiscal policy, and the difference of inflation, so the accuracy of the model in simulating the complex market environment still needs to be improved.

4.2. Inspiration

In terms of data, we can increase the information disclosure of offshore RMB exchange rate data. In view of the real-time fluctuation of the exchange rate market, by updating the exchange rate data in real time, more comprehensive data support is provided for the training of the model, thereby improving the prediction effect of the model, improving the accuracy and robustness, and further enhancing the depth and breadth of the input features by enriching the data sources, so that the model can capture more complex market dynamics.

For the model and associated costs, model compression techniques can be used, such as knowledge distillation: transferring knowledge from a large, highly accurate teacher model to a lightweight student model, reducing inference time and resource consumption while maintaining no significant decline in predictive power. Similarly, it is also possible to simplify the overall structure, reduce the computational cost and improve the operational efficiency by deleting the sub-models, features or neurons with lower importance in the neural network, so as to deal with the problem of high computational load usually brought by complex combined models. Furthermore, it can strengthen the construction of hardware infrastructure and provide more powerful technical support for model prediction.

For practical applications, we can explore more new combination models and enrich feature extraction to simulate the real exchange rate market environment more accurately, so as to improve the accuracy of model prediction.

5. Conclusion

This paper reviews the progress in the application of combinatorial models based on machine learning in offshore RMB forecasting. This paper focuses on the analysis of data selection, feature extraction, model performance and prediction effect, and discusses the application of different machine learning combination models in the prediction of offshore RMB exchange rate. The machine learning fusion ARIMA model is suitable for the prediction of time series with obvious rules and small amount of data, the combination model of random forest and neural network is suitable for the prediction of structured data, and the LSTM-TCN-CNN model is suitable for the prediction of multivariate time series with large amount of data and high dimension. On the whole, the prediction effect of the combined model is better than that of the single model, which has a good application prospect in the prediction of offshore RMB exchange rate and has a certain reference value for market participants. At the same time, the current research also exposed some challenges, such as data lag, lack of model interpretability and lack of practical simulation, which need to be further improved in the future. In the future, we can further improve the prediction accuracy of the model from the aspects of information disclosure, model optimization and hardware infrastructure construction. With the advancement of the era of big data and the continuous updating of algorithms, the combination model of machine learning is expected to play a greater role in offshore RMB forecasting, which will help to provide forward-looking reference and guidance for investors and other relevant parties.

References

[1] Sun L, Xiang M, Marquez L. Forecasting the volatility of onshore and offshore USD/RMB exchange rates using a multifractal approach. Physica A: Statistical Mechanics and its Applications, 2019, 532: 121787.

[2] Wang N. Research on RMB exchange rate forecasting based on NDF and NARX network [D]. Dalian University of Technology, 2013.

[3] Gao J. Research on the forecasting model of offshore RMB exchange rate based on neural network and random forest. East China Normal University, 2024.

[4] Zhou L. Prediction of offshore RMB exchange rate based on machine learning and ARIMA model. Journal of Statistics, 2020, 1(2): 48-56.

[5] Yue C, Zhuang H. Offshore RMB exchange rate forecasting based on LSTM-TCN-CNN hybrid model. China Price, 2025, (6): 79-85.