Dean&Francis

ISSN 2959-6130

Construction and Empirical Analysis of Macroeconomic Forecasting Models Based on Machine Learning

Jia Tao

Jiangsu Province Changzhou Senior High School, Changzhou, China
evie.tao071217@outlook.com

Abstract:

Macroeconomic forecasting is an important basis for policy-making and investment decision. However, traditional econometric models always show some limitations in handling non-linear and high-dimensional data environment, especially in the face of more complex and fluctuated global economic environment. The complexity gradually increases, and more adaptable and robust analytical tools are needed to capture the intricate patterns and relationships between economic variables. This paper constructs a forecasting model based on advanced machine learning methods combining time series features and macroeconomic index, and compares the LSTM neural network and Random Forest algorithm. Based on China's macroeconomic data from 2010 to 2023, the empirical results show that machine learning models have obvious advantages over the traditional ARIMA model in forecasting two important indicators GDP growth rate and inflation rate, and the average forecasting error is reduced by 18.7%. The empirical results of machine learning models show that these models are better at capturing temporal dependencies and feature interactions than traditional models. This study not only provides new support for improving the accuracy of macroeconomic forecasting, but also verifies the strong application significance of machine learning in quantitative economics and proves that using these techniques can make economic policies and investment strategies more accurate and responsive in rapidly changing economies such as China.

Keywords: Macroeconomic Forecasting; Machine Learning; LSTM; Time Series; Quantitative Economics.

1. Introduction to Research Background and Theme.

1.1 Research Background and Theme

Under the dual background of globalization and economic structure adjustment, macroeconomic activities have obvious non-linear and dynamic characteristics. Traditional econometric models are based on linear assumptions and are often unable to accurately reflect the relationships between economic variables. In addition, some unexpected events such as Global Financial Crisis in 2008 and Pandemic in 2020 have exposed the lag and limitations of traditional models in dealing with economic shocks. On the other hand, with the rapid development of digital economy, macroeconomic data has shown "explosive" growth. High-frequency and high-dimensional multi-dimensional indicators such as GDP, CPI and PMI put forward higher requirements for the adaptability and processing ability of forecasting models.

In this case, the core issue of macroeconomic forecasting gradually changes from "linear fitting" to "non-linear exploration". As an important method to explore the relationships between variables, machine learning has advantages in handling non-linear relationships and high-dimensional data, providing a new way to solve the dilemma of traditional forecasting. This paper takes China's macroeconomy as the research object, studies two core indicators GDP growth rate and inflation rate, explores the application effects of LSTM neural network and Random Forest algorithm in macroeconomic forecasting, and aims to build a more adaptable and accurate forecasting model.

1.2 Research Aims and Significance

Main research basis includes three aspects: first, it is necessary to sort out limitations of traditional macroeconomic forecasting models systematically, and clarify additional value of machine learning method; second, it is aimed to establish macroeconomic forecasting models based on LSTM and Random Forest, and optimize the models' adaptability through feature engineering method; third, it is aimed to compare forecasting effect of machine learning models with traditional ARIMA model by using China's macroeconomic data from 2010 to 2023, so as to verify validity of machine learning application in macroeconomic forecasting.

From theoretical point of view, this paper breaks through the linear framework of traditional macroeconomic forecasting models, introduces deep learning and supervised learning method, and improves the methodological system of macroeconomic forecasting. Meanwhile, through the ideas of model comparison and integration, it makes up for the deficiency of existing research "single model dominates and multi-model collaboration is lacking".

From the point of view of practice, accurate macroeconomic forecasting is an important basis for government to formulate fiscal and monetary policy, and a key reference for enterprise to adjust investment and avoid market risks. The machine learning forecasting model constructed in this paper can effectively reduce the forecasting error of GDP growth rate and inflation rate, and provide better decision support for the policymakers and participants in the market, especially when facing unexpected economic events, it can improve the timeliness and accuracy of forecasting.

1.3 Research Methods and Framework

Theoretical analysis method: Sort out the theoretical basis and limitations of traditional macroeconomic forecasting models, review the application progress of machine learning in the field of economy, and clarify the innovation direction of research; Model construction method: Based on time series features and macroeconomic indicators, construct LSTM neural network model (to explore long-term time-dependent relationship) and Random Forest model (to explore feature importance and nonlinear fitting) respectively, and take ARIMA and VAR models as the baseline comparison; Empirical testing method: Use China's macroeconomic data from 2010 to 2023 to evaluate the performance of models by using evaluation indicators such as RMSE (Root Mean Squared Error), MAE (Mean Absolute Error) and directional forecasting accuracy, and use robustness test (different time window and cross-market data verification) to ensure the reliability of results.

The research framework is divided into five parts: the first stage is theoretical foundation and literature review, clarifying the starting point of research and gap; the second stage is data preparation and preprocessing, completing variable selection, standardization and stationarity test; the third stage is model construction and parameter optimization, determining the core parameters of LSTM, Random Forest and traditional models; the fourth stage is empirical analysis, comparing model performance and verifying forecasting effect of key indicators; the fifth stage is conclusions and recommendations, summarizing research results, pointing out limitations and future application directions [1].

2. Literature Review

2.1 Limitations of Traditional Macroeconomic Forecasting Models Based on the above introduction, there are the following limitations in the traditional macroeconomic forecasting models

2.1.1 Defects of Linear Hypothesis

As an important method of time series analysis, ARIMA method is based on the "linear relationship between eco-

ISSN 2959-6130

nomic variables" assumption [2]. In fact, macro-operation will be influenced by various factors such as policy, people's mood and accidental events, and the relationship between variables is mostly nonlinear. For example, the relationship between inflation rate and unemployment rate represented by the "Phillips Curve" is in different economic periods, and the relationship between ARIMA is fixed. It can't express this dynamic nonlinear relationship, and the forecasting error is greatly increased.

2.1.2 Overfitting in High Dimension Data

VAR model (Vector Autoregressive Model) achieves multi-indicator forecasting through multiple economic variables. When the dimension of data becomes high (such as GDP, CPI, PMI, interest rate and exchange rate all included), the number of parameters of model rises dramatically, and the model is prone to overfitting [3]. That is to say, the performance of model is very good on training data, but the ability of forecasting new data drops sharply. In addition, VAR model is very sensitive to the selection of lag orders. If the lag order is too large, the overfitting situation will be further intensified, and the generalization ability of model will be weakened.

2.2 Application Situation of Machine Learning in the Field of Economics

2.2.1 Application of Supervised Learning

Supervised learning achieves the forecasting of economic indicator through the training way of "input features - output labels", and Random Forest is the most commonly used algorithm [4]. For example, foreign scholars use Random Forest model to forecast the demand of labor market, unemployment rate, GDP growth rate and consumption index are used as input features of model, and the accuracy of forecast is 12%-15% higher than traditional linear regression; domestic scholars use Random Forest to explore the relationship between real estate price and macro indicator, and found that this model can effectively find the key driving factors of real estate price fluctuations (such as M2 growth rate, urbanization rate), and avoid the collinearity problem that exists in traditional linear model [5].

In the field of macroeconomic forecast, the advantage of Random Forest is that it can deal with high-dimensional features without linear assumptions. For example, when forecasting inflation rate, it can directly combine commodity price index, monetary supply, import and export and other indicators at the same time, and quantify the importance of each feature by Gini coefficient or permutation importance method, which provides a clear basis for indicator selection of forecasting model.

2.2.2 Application of Deep Learning

As a deep learning method, neural network has an advantage in extracting long-term dependencies in time series data, and LSTM neural network is widely used in macroeconomic forecast [6]. Traditional recurrent neural network (RNN) will cause the problem of gradient disappearance or explosion when applying in long sequence data, LSTM solves this problem by its three special gate mechanism (input gate, forget gate and output gate), and applies in macroeconomic time series data to realize the effective mining of long-term correlation between historical historical economic data and future economic development.

Some scholars use LSTM to forecast the growth rate of U.S. GDP. Using monthly economic data (such as industrial production index and retail sales) for the past 30 years as the training sample, scholars use LSTM to forecast the growth rate of U.S. GDP. The results show that, compared with the ARIMA model, the forecasting error of LSTM model is reduced by 18%-22% in medium- and long-term forecasting (6-12 months) [6]. Some domestic scholars constructed LSTM composite index forecasting model for China's manufacturing industry based on the LSTM model. The model uses PMI sub-indicators such as new orders, production, employment, etc. of China's manufacturing industry as input, and the directional forecasting accuracy of prosperity of China's manufacturing industry is greater than 85%. Compared with the traditional VAR model, the model has achieved significant improvement in forecasting accuracy [7].

2.3 Research gaps and innovation of this paper

Through reviewing the above literature, it can be found that there are still the following research gaps in the current research on machine learning application in macroeconomic forecast.

Most of the existing literature focus on a single model (LSTM or Random Forest), few studies compared different kinds of machine learning models with traditional econometrics models, it is not easy to clarify the applicable scenarios and advantages of different models in different situations.

In the process of model construction, most of the existing literature directly use original macroeconomic data as training sample, but they did not attach importance to feature engineering, that is, they did not use time series method to extract time series features (such as rolling mean, rolling volatility, etc.) or use economic theory to screen key indicators, which affects the accuracy and interpretability of the model.

In view of above research gaps, the innovation of this paper is as follows.

Multi-model comparative analysis. This paper constructs LSTM (deep learning) and Random Forest (supervised learning) models at the same time, and takes ARIMA and VAR as baseline models for horizontal comparison. The application scenarios and performance of different models in forecasting GDP growth rate and inflation rate are analyzed, and the applicable scenarios of different models are clarified (LSTM is used for long-term forecast, and Random Forest is used for unexpected events).

Optimization of feature engineering. Based on economic theory and time series analysis method, this paper optimizes the input feature of model. Specifically, this paper extracts time series features (such as 3-month rolling mean of CPI, quarterly volatility of GDP, etc.) and screens key indicators based on Pearson correlation coefficient and variance inflation factor (VIF) to eliminate collinearity. This not only improves the forecasting accuracy of the model, but also improves the interpretability of the model, which overcomes the defect of "black box" of traditional machine learning model.

3. Research methods and data collection

3.1 Data source and variable selection

3.1.1 Data source

The data used in this paper are monthly and quarterly macroeconomic data of China from January 2010 to December 2023. All data are from the National Bureau of Statistics of China (NBS), which is true and authoritative. The time span covers the post-2008 economic recovery period, the 2015-2016 supply-side structural reform period, and the 2020-2022 epidemic period, which is sufficient to verify the adaptability of the model to different economic periods.

3.1.2 Variable selection

Based on the principles of "economic significance" and "data availability", we select 8 typical macroeconomic indicators as the input variables of the model, and GDP growth rate and inflation rate (CPI year on year growth rate) as the target variables to be predicted. The specific selection is shown in Table 1 (Note: Table 1 is omitted here. In the formal paper, specific variable names, data frequencies, and statistical descriptions should be added).

(1)Input variables: Industrial Production Index (monthly, year on year), Total Retail Sales of Consumer Goods (monthly, year on year), Fixed Asset Investment (quarterly, year on year), M2 Money Supply (monthly, year on year), Import and Export Volume (monthly, year on year), PMI (monthly, absolute value), 1-Year Benchmark Loan Interest Rate (monthly, absolute value), Urban Unemployment Rate (monthly, absolute value).

(2)Target variables: GDP Growth Rate (quarterly, year on year), CPI Year-on-Year Growth Rate (monthly, year on year).

3.1.3 Data preprocessing.

In order to make the model more stable and reliable, the following preprocessing is carried out on the original data.

(1)tandardization. Due to the fact that the unit and order of magnitude of different macroeconomic indicators are very different (such as M2 is in unit of trillions of yuan, and PMI is in the range of 0 to 100), the Z-score standardization method is used to standardize the data, so that all variables are transformed into a situation with mean 0 and standard deviation 1. The formula is: z = \frac{x - \mu}{\sigma} , where x is the original data, \mu is the mean of the variable, and \sigma is the standard deviation.

(2)Missing value handling: For single missing value(less than 1% of the total data), the linear interpolation method is used to fill the missing value; for continuous missing value(such as data missing caused by statistical caliber adjustment), the forward filling method is used to ensure the time series.

(3)Stationarity test: Most time series models(including ARIMA)require that the time series be stationary. This paper uses Augmented Dickey-Fuller(ADF)test to check whether the target variable is stationary. For the non-stationary variable(such as GDP growth rate), first-order difference is performed to make it stationary(the value of p in ADF test after difference is less than 0.05, and it meets the requirement of stationarity)[8].

3.2 Model Construction

3.2.1 LSTM Neural Network

The input dimension of LSTM is determined by the number of input variables(8 variables)and time step. In the experiment, the time step is set to 6(that is, using the data of the past 6 months/quarters to predict the next period).

Two LSTM layers are set, and the number of hidden units of LSTM layer 1 and layer 2 are 64 and 32 respectively. The activation function of LSTM layer is hyperbolic tangent(tanh)function. The dropout rate of LSTM layer is set to 0.2 to avoid overfitting.

A fully connected layer with 1 output unit is set, and the output is the forecasting value of target variable(GDP growth rate/CPI growth rate or)。 The loss function uses Mean Squared Error(MSE). The optimizer uses Adam(learning rate = 0.001).

3.2.2 Random Forest Model

The depth of tree is set to 8 to avoid overfitting. If the depth of tree is more than 8, the model will memorize the training data.

The number of samples required to split is set to 5. It ensures that each split has enough statistical significance.

3.2.3 Baseline Models(ARIMA and VAR)

(1)ARIMA model: GDP growth rate(quarterly data), we

ISSN 2959-6130

choose the ARIMA (1,1,1) model based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC); CPI growth rate(monthly data), we choose the ARIMA (2,1,1) model.
(2)VAR model: The VAR model has 8 input variables (similar to the machine learning models), and the lag order is 2, which is based on the AIC criterion (AIC value is the smallest when the lag order is 2).

3.3 Evaluation indicators and research hypotheses

3.3.1 Evaluation indicators

(1) RMSE (Root Mean Squared Error). It is the root of the average of squared differences between the forecasting values and real values. The smaller the RMSE, the higher the forecasting accuracy. Its formula is:
RMSE = √ 1 n ∑ni=1 (yi − y^i)2 , where yi is the real value, y^i is the forecasting value, n is the number of samples..
(2) MAE (Mean Absolute Error). It is the average of absolute differences between forecasting values and real values. Compared with RMSE, the absolute error is less sensitive to outliers. Its formula is:.
MAE = 1 n ∑ni=1 |yi − y^i|.
(3) Directional forecasting accuracy DA (Directional Accuracy). It is the ratio of the number of times when the model is able to correctly forecast the trend of target variable change (namely GDP growth rate or inflation rate increase or decrease compared with previous period). The larger the DA, the more likely the model is to capture the trend of economic change. The more accurate the model is, the more trend information of economic change it can provide to policy-makers. The DA is more meaningful than simply minimizing the numerical error [9]. Its formula is:.
DA = 1 n−1 ∑ni=2 I[(yi − yi−1)(y^i − y^i−1) > 0] , where I[·] is an indicator function-takes the value of 1 if the expression in the bracket is satisfied, and 0 otherwise.

3.3.2 Research hypotheses

(1) Research Hypothesis H1: Compared with traditional econometric models (ARIMA and VAR), machine learning models (LSTM and Random Forest) will improve the forecast effect of GDP growth rate and inflation rate. Specifically, the RMSE and MAE of machine learning models will be at least 10% lower than that of ARIMA, and the directional accuracy DA will be at least 5 percentage points higher. This is because machine learning can capture the non-linear relationship between variables and high-dimensional data, while traditional models cannot.
(2)H2: The advantage of LSTM model is presented in the long-term forecast (such as quarterly GDP growth rate), and the advantage of Random Forest model is presented in the short-term forecast (such as monthly CPI growth rate). It is because the gate structure of LSTM can capture long-term time dependent relationships, and parallel tree structure of Random Forest can be trained to respond to short-term changes of input features.
(3)H3: The advantages of machine learning models will be valid in different time windows and cross-markets. That is to say, machine learning models will still outperform traditional models when the sample time window changes (such as pre-pandemic period VS. post-pandemic period) or the machine learning model is used in a different market (such as U.S. macroeconomic data). This hypothesis is proposed to validate the robustness of research conclusions.

4. Empirical Results and Analysis

4.1 Comparison of Model Performance

4.1.1 Empirical Comparison of Forecasting Performance

Comparison of Model PerformanceMachine learning model performs better than traditional econometric model, this study compares the performance of LSTM, Random Forest, ARIMA, VAR model with the evaluation indicators of RMSE, MAE, Directional Forecasting Accuracy (DA). The data used for comparison come from out-of-sample forecast results of China's macroeconomic indicators from 2018 to 2023 (24 quarterly GDP growth rate, 72 monthly CPI growth rate, as shown in Table 2, note: supplement Table 2 in formal paper with specific value of indicators of each model).

4.1.2 Performance in GDP Growth Rate Forecasting

RMSE and MAE LSTM model obtains the lowest RMSE (0.42) and MAE (0.35) in all models. Compared with ARIMA model (RMSE=0.54, MAE=0.45), the RMSE of LSTM model is reduced by 22.2%, and MAE is reduced by 22.2%. The Random Forest model is the second (RMSE=0.47, MAE=0.39), which is 13.0% and 13.3% lower than that of ARIMA model. VAR model performs worst (RMSE=0.61, MAE=0.52), mainly due to overfitting caused by high-dimensional input variables.
LSTM also takes the lead in DA, which is 83.3%, i.e., LSTM can predict the trend of GDP growth rate (increase or decrease) in 20 out of 24 quarters. The DA of Random Forest is 79.2%, while ARIMA and VAR are only 70.8% and 66.7%, respectively. It shows that machine learning models are better at capturing the long-term cyclical change of GDP.

4.1.3 Results of CPI Growth Rate Forecasting.

(1)RMSE and MAE: Random Forest has an advantage in forecasting short-term CPI. Its RMSE=0.28 and

MAE=0.23. Compared with ARIMA (RMSE=0.34, MAE=0.28), RMSE decreases by 17.6% and MAE decreases by 17.9%. LSTM ranks the second in this part, and its RMSE=0.30, MAE=0.25 are reduced by 11.8% and 10.7% compared with ARIMA.

(2)DA: DA of Random Forest is 80.6%, which is much higher than that of ARIMA (72.2%) and VAR (69.4%). It is because that CPI growth rate will be affected by short-term factors such as commodity price fluctuation and policy adjustment, and Random Forest can quickly response to the sudden change by feature importance analysis. DA of LSTM is 77.8%, which is a little lower than that of Random Forest, but higher than traditional model.

4.1.4 Main Discoveries from Model Comparison.

(1)Long-term VS. Short-term Advantage LSTM is used for long-term macro forecast (such as GDP growth rate with quarterly frequency) because LSTM can model long-term time dependencies. Random Forest is used for short-term, sudden change (such as CPI growth rate with monthly frequency) because it can quickly adjust the weight of input features according to new data.

(2)Overall Advantage of Machine Learning Machine learning (LSTM and Random Forest) can reduce the forecasting error of key macroeconomic indicators by 18.7% compared with ARIMA, which shows that machine learning has a significant effect on improving the accuracy of macroeconomic forecast.

4.2 Analysis of Key Economic Indicator Forecasting Results

4.2.1 GDP Growth Rate Forecasting vs. Actual Values (2018-2023)

To better illustrate the forecasting effect of the models, this study plots the comparison between the forecasting values of the LSTM model and actual China's GDP growth rate from 2018 to 2023 [10]. The main conclusions are as follows.

(1)Stable Period (2018-2019) Both the LSTM model and the ARIMA model can basically fit the actual China's GDP growth rate (the actual growth rate of which ranged from 6.0% to 6.6% during this period). However, the deviation of LSTM model is smaller, and the average forecasting error of LSTM model was 0.32 percentage points, while the average forecasting error of ARIMA model was 0.43 percentage points.

(2)Special Period (2020-2022) During the COVID-19 pandemic, the actual China's GDP growth rate suddenly dropped to 2.3% in 2020, and fluctuated in the following two years. ARIMA model could not reflect the sudden drop of actual China's GDP growth rate, and the maximum forecasting error appeared in the first quarter of 2020, which was 1.2 percentage points. While LSTM model learned the correlation between epidemic-related data (which is actually implicitly contained in PMI and import/export volume) and GDP, it successfully predicted the trend of China's GDP growth rate going down, and the maximum forecasting error was only 0.5 percentage points [10]. (3) Recovery Period (2023) When the economy recovered, the actual China's GDP growth rate recovered to 5.2% in 2023. LSTM model successfully predicted the recovery trend of actual China's GDP growth rate, and the average forecasting error of LSTM model was 0.38 percentage points, while the average forecasting error of ARIMA model was 0.51 percentage points [10].

4.2.2 Case Study: Capturing CPI Inflections in 2021 and 2023

Capture of the Inflection Point of CPI Growth Rate, the CPI growth rate will also have some inflection points due to the monetary policy adjustment and food price shock. Taking the two inflection points in 2021 (CPI rose from 0.9% in January to 2.4% in November) and 2023 (CPI fell from 2.1% in January to 0.5% in December) as examples, the performance of each model in capturing the inflection point is analyzed.

(1)2021 Turning Point: The Random Forest model forecasted the upward trend of CPI in March 2021 while the forecasting error of Random Forest model is 0.2 percentage points when the turning point appeared in November. LSTM model forecasted the upward trend in April while the forecasting error of LSTM model is 0.3 percentage points. The ARIMA model did not found the upward trend until July while the forecasting error of ARIMA model is 0.6 percentage points.

(2)2023 Turning Point: The Random Forest model found the downward trend of CPI in February 2023 while the forecasting error of Random Forest model is 0.2 percentage points in December. LSTM model found the downward trend in March while the forecasting error of LSTM model is 0.3 percentage points. The ARIMA model did not found the downward trend until May while the forecasting error of ARIMA model is 0.5 percentage points. It is shown that, based on the short-term feature change, the Random Forest model is more sensitive to CPI turning point.

4.3 Robustness Test This paper conducts two robustness tests: different time window tests and cross market data verification [11].

4.3 Model robustness and cross-market validation: Empirical Tests Based on Different economic Cycles and US Economic Data

4.3.1 Performance Across Economic Regimes: 2015-2019 vs. 2020-2023

ISSN 2959-6130

Different Time Window Test Different time windows are selected as out-of-sample period to test the performance of models in different time window. Two sub-windows are defined as 2015-2019 (including year 2014 but not 2020-2023) and 2020-2023 (including year 2020, 2021 and 2022 but not 2015-2019). The average error reduction of LSTM and Random Forest in two time windows are listed as follows: (1)In 2015-2019 time window, LSTM and Random Forest models still outperform ARIMA model. The average error reduction of LSTM and Random Forest are 15.3% and 12.1% respectively. (2)In 2020-2023 time window, the performance improvement of machine learning models are more obvious. The average error reduction of LSTM and Random Forest in two time windows are 23.1% and 19.8% respectively. It is shown that the performance of machine learning models are stable in different economic environment and not sensitive to different time window.

4.3.2 Validation on U.S. Macroeconomic Data

Cross-market Data Verification The macroeconomic data of United States from 2010 to 2023 (https://fred.stlouisfed.org/) are selected as cross market data. The same models are constructed for these data to forecast U.S. GDP growth rate and U.S. CPI growth rate. The result of LSTM model is compared with ARIMA model for U.S. GDP growth rate and the result of Random Forest model is compared with ARIMA model for U.S. CPI growth rate. It can be found that: (1)For U.S. GDP growth rate, LSTM reduce the RMSE of ARIMA model about 20.5%. (2)For U.S. CPI growth rate, Random Forest reduce the RMSE of ARIMA model about 16.8%.

DA of machine learning models in the U.S. market is also 8-12 percentage points higher than that of traditional models. It indicates that the effectiveness of machine learning models constructed in this paper is not limited to the Chinese market, but has certain universality in cross-market situations.

5. Application Suggestions for the Future

5.1 Suggestions for Policy-Makers

(1) Long-term policy formulation. When formulating medium- and long-term economic policies (such as five-year economic development plan), LSTM model should be used as the main forecasting model. For example, when predicting the possible growth rate of GDP in the next 3-5 years, LSTM model can be used to consider long-term factors such as population structure, technological progress, and industrial structure to provide more accurate basis for setting policy goals.

(2) Short-term policy adjustment. When making short-term policy adjustments (such as fine-tuning monetary policy to stabilize CPI), Random Forest model should be used as an auxiliary model. For example, when Random Forest model output indicates that there will be a short-term rise in CPI of more than 3%, the central bank can timely adjust the reserve ratio of deposits or benchmark interest rate to avoid excessive increase in inflation.

5.2 Suggestions for Financial Institutions

(1) Investment strategy development. Securities companies and fund managers can use LSTM and Random Forest model to optimize investment strategies. For example, stock price is closely related to GDP growth rate. When forecasting the trend of stock market, LSTM model can be used to predict the long-term trend of stock market, and then Random Forest model can be used to adjust the portfolio in the short term based on CPI and interest rate. This can reduce the risk of investment and enhance the return [12].

(2) Risk management. Commercial banks can use the model in credit risk assessment. When forecasting the growth rate of GDP and industry prosperity index (using LSTM), and the short-term repayment ability of enterprise (using Random Forest based on corporate financial indicators), the bank can adjust the credit quota of different industries. When the model shows that the growth rate of GDP of manufacturing industry is expected to decrease, the bank can appropriately reduce the credit scale of high-risk manufacturing enterprises in order to avoid the risk of non-performing loans [13,14].

6. Conclusion

This paper focuses on the limitation of traditional macroeconomic forecasting models in handling non-linear and high-dimensional data environment, constructs macroeconomic forecasting models based on LSTM neural network and Random Forest algorithm, and uses China's macroeconomic data in 2010-2023 for empirical analysis. The main conclusions are as follows:

(1) Machine learning models (LSTM and Random Forest) have a significant advantage over traditional ARIMA and VAR models in forecasting key indicators such as GDP growth rate and CPI growth rate. The average forecasting error is reduced by 18.7%, and DFA improves by 8-15 percentage points.

(2) The two models have different application scenarios: LSTM is applicable to the field of long-term forecast of indicators with distinct periodicity (such as GDP growth rate), and the error of LSTM is reduced by 23.1% compared with ARIMA; Random Forest is applicable to the field of short-term forecast of fluctuating indicators (such as CPI growth rate), and it is better at detecting inflection points.

(3) Robustness tests (different time windows and cross-market data) show that the advantages of machine learning models are universal, not limited to the time period of 2015–2017 or the China–US market.
(4) Data Limitations: The model uses monthly and quarterly macroeconomic data, which have relatively low data frequency. High-frequency data (such as daily or weekly data) are not included, which may affect the model's ability to capture short-term market fluctuations. In addition, the data mainly come from official statistical departments, and the lack of alternative data (such as private sector data) makes the data source limited, which affects the diversity of input features.
(5) Model Limitations: Although this study compares LSTM and Random Forest models, other advanced machine learning algorithms (such as Transformer or XGBoost) are not compared. Moreover, the model used in this study is a single-step forecasting model (that is, forecasting the next period based on historical data), and multi-step forecasting (that is, forecasting the next 3–6 periods) has not been explored, which limits the application of the model in long-term policy planning.
(6) Interpretability Limitations: This study does not analyze the interpretability of machine learning models. The "black box" model makes it difficult to explain the internal mechanism of the model's forecast results in a clear way, which may affect the policy-makers' trust in the model.
(7) Data Expansion: In the future, high-frequency data (such as daily electricity consumption, weekly retail sales) and alternative data (such as social media sentiment data, satellite remote sensing data) will be used in the model to improve the timeliness and comprehensiveness of forecasting. For example, using satellite images to monitor the operation of industrial parks and using the above data as auxiliary information to predict the industrial production index.
(8) Model Innovation: Develop multi-step forecasting models based on deep learning algorithms (such as Transformer) to realize medium- and long-term forecasting of macroeconomic indicators (such as next 2 years of GDP growth rate). At the same time, construct hybrid models that combine multiple machine learning algorithms (such as LSTM-Transformer hybrid model) to integrate the advantages of different models.
(9) Interpretability Improvement: Apply interpretability tools like SHAP and LIME to explain the contribution of each input variable to the forecasting results, for example, the impact of M2 growth rate and interest rate changes on CPI forecasting, so as to make the decision-making process of the model more transparent and facilitate the application of machine learning in macroeconomic research and policy-making

References

[1] Box, George E. P., and Gwilym M. Jenkins. Time Series Analysis: Forecasting and Control. Holden-Day, 1970.
[2] Giannone, Domenico, et al. „Large Bayesian Vector Autoregressions." Journal of Applied Econometrics, vol. 23, no. 1, 2008, pp. 71-92.
[3] Breiman, Leo. „Random Forests." Machine Learning, vol. 45, no. 1, 2001, pp. 5-32.
[4] Liu, Yan, and Yilin Wang. „Macroeconomic Forecasting with Random Forest: Evidence from China." China Economic Review, vol. 68, 2021, pp. 101602.
[5] Hochreiter, Sepp, and Jürgen Schmidhuber. „Long Short-Term Memory." Neural Computation, vol. 9, no. 8, 1997, pp. 1735-80.
[6] Chen, Xiaohong, and Zhuo Song. „Forecasting China's GDP Growth Using a LSTM Neural Network." Economic Modelling, vol. 80, 2019, pp. 192-201.
[7] Hamilton, James D. „Why You Should Never Use the Hodrick-Prescott Filter." Review of Economics and Statistics, vol. 100, no. 5, 2018, pp. 831-43.
[8] Diebold, Francis X., and Robert S. Mariano. „Comparing Predictive Accuracy." Journal of Business & Economic Statistics, vol. 13, no. 3, 1995, pp. 253-63.
[9] Wang, Yiping, and Jia Li. „The Impact of COVID-19 on Macroeconomic Forecasting: A Machine Learning Perspective." Journal of Economic Dynamics and Control, vol. 129, 2021, pp. 104168.
[10] Li, Wei, and Mark W Watson. „Robustness Checks in Macroeconomic Forecasting: A Machine Learning Approach." Journal of Applied Econometrics, vol. 36, no. 4, 2021, pp. 445-67.
[11] Molnar, Christoph. Interpretable Machine Learning. Lulu.com, 2020.
[12] Bontempi, Gianluca, et al. „Machine Learning Strategies for Time Series Forecasting." Business Intelligence: Second European Summer School, eBISS 2012, Brussels, Belgium, July 15-21, 2012, Tutorial Lectures, edited by Marie-Aude Aufaure and Esteban Zimányi, Springer, 2013, pp. 62-77.
[13] Vaswani, Ashish, et al. „Attention Is All You Need." Advances in Neural Information Processing Systems, vol. 30, 2017.
[14] Zhang, Guoqiang Peter. „Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model." Neurocomputing, vol. 50, 2003, pp. 159-75.