CNY EX Rate Prediction Based on LSTM and Machine Learning Methods

Jiaqi Lu
School of Economics and Management, Tongji University, 1239 Siping Road, Shanghai, China

Keywords: Foreign Exchange Rate, LSTM, Machine Learning, Commodity Features, Technical Features.

Abstract: The foreign exchange market is volatile and unpredictable and the foreign exchange rate is challenging to forecast in almost all the regions. With the maturity of the foreign exchange market, more and more traders make transactions on foreign exchange products. The ability to estimate this foreign exchange rate has therefore become crucial in the financial market. In this study, machine learning methods are used to predict the exchange rate of the Chinese yuan (CNY). The feature inputs include three categories, which is respectively technical features, commodity features, and forex features. The technical features include some powerful technical factors. The commodity features include gold price, oil price, and stock index. The forex features include some frequently traded currency. The models include Linear Regression, Lasso Regression, Ridge Regression, long short-term memory (LSTM), Random Forest, and XG-Boost. In conclusion, this study finds that the Long Short-Term Memory model has the best performance and the tech features are the best inputs for predicting the CNY exchange rate.

1 INTRODUCTION

With the maturity of the financial system, foreign exchange plays a more important role in global trading and it becomes more urgent to have a forecast of the trend of the exchange rate. However, the exchange rate prediction has been one of the most challenging tasks for long. It is necessary to comprehend the intricacies of global political economy, sociological and economic infrastructures, and occasional political and social events since they have a comprehensive impact on the exchange rate. It means too many complex factors need to be taken into consideration.

In the past, emphasis was placed on employing macroeconomic indicators such as spot rates, unemployment rates, or inflation rates to discern long-term trends in exchange rates. However, these approaches offered only broad predictions based on empirical observations and were insufficient for providing concise, short-term investment or business advice. Statistical models like integrated moving averages and auto-linear regression were also utilized for financial time series predictions, but they were constrained by their inability to transcend historical data. With advancements in computational capabilities, machine learning algorithms have emerged as transformative tools for financial forecasting (Singh et al, 2009). Going beyond traditional qualitative analysis, this paper uses machine learning methods to predict the Chinese Yuan (CNY) exchange rate. Notably, different from traditional macro features, this paper introduces a series of tech, commodity, and forex feature inputs, thereby enhancing the model's capacity to capture nuanced market dynamics.

Section II of this paper discusses related works. Section III discusses data analysis, feature engineering, and modeling. Section IV discusses the results and analysis.

2 LITERATURE REVIEW

Conventional econometric models predict exchange rates using underlying economic circumstances, assuming that long-term patterns are determined by economic fundamentals. However, Meese and Rogoff demonstrate the failure of econometric models to anticipate short-term exchange rates (Meese et al, 1983). Two popular time series models for predicting currency rates are exponential

Lu, J.
CNY EX Rate Prediction Based on LSTM and Machine Learning Methods.
DOI: 10.5220/0012818800004547

In *Proceedings of the 1st International Conference on Data Science and Engineering (ICDSE 2024)*, pages 453-458
ISBN: 978-989-758-690-3

smoothing (ETS) models and autoregressive integrated moving average (ARIMA) models. ARIMA models can handle nonstationary data by differences. ETS models can take seasonality and trends into account (Galeshchuk et al, 2017). The more powerful methods are machine learning approaches which have developed over the years. Qian and Rasheed use several inductive machine-learning classifiers to get a prediction accuracy of up to 67% (Amat et al, 2018). Amat et al. constructs sequential ridge regression and the exponentially weighted average strategy, both with discount factors. They do not estimate an underlying model but combine the fundamentals to directly output forecasts (Amat et al, 2018). Pradeepkumar and Ravi update the Quantile Regression Neural Network to predict the volatility of financial time series (Pradeepkumar & Ravi, 2017). Fischer and Krauss demonstrate that the long short-term memory (LSTM) network can extract relevant information from financial time series data that is noisy (Fischer et al, 2018). Gyamerah and Moyo capture the exchange rate uncertainty using probability density forecasting functions (Gyamerah, 2020). Wang and Guo propose a hybrid model which has good approximation and generalization ability, greatly improving the performance (Wang & Guo, 2020). Cao et al. developed a new deep-coupled LSTM method, to capture the complex couplings for exchange rate prediction (Cao et al, 2020).

In this study, several popular machine learning methods are constructed to compare their performance on three categories of feature inputs, trying to find the best model and feature combination to predict the CNY exchange rate.

3 METHODOLOGY

3.1 Data Analysis

Both the autocorrelation function and partial autocorrelation function of returns exhibit a cutoff at the one-day lag, indicating weak correlation and stochastic behaviour in the daily returns of CNY. This leads to the conclusion that using CNY daily returns alone may not be sufficient for accurate predictions. Further feature construction is imperative for enhancing predictive capabilities in the model (Fig. 1).

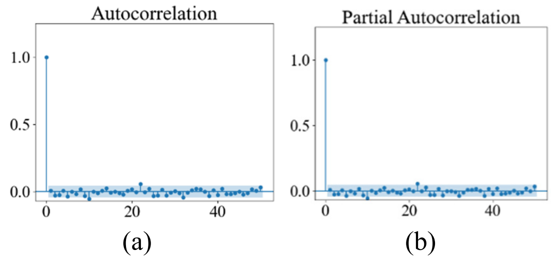

Figure 1: ACF and PACF of return (Picture credit: Original).

Time series decomposition serves as a valuable technique for disentangling various components within a dataset. Time series data typically exhibits three fundamental components: seasonality, trend, and noise. In Fig. 2. and Fig. 3., the decomposition of CNY reveals a notable stochastic nature. The observed fluctuations can be predominantly attributed to the trend factor, with minimal discernible influence from seasonal components. The absence of a significant seasonal factor suggests that regular and predictable patterns over specific time intervals are not evident in the exchange rate dynamics. In conclusion, the CNY exchange rate changes are predominantly driven by external factors that can’t be explained by time series models. In this sense, the traditionally empirical rule is usually out of action and researchers should resort to machine learning methods.

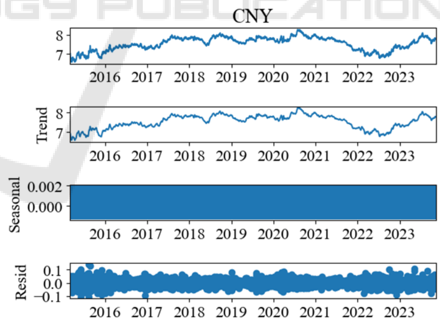

Figure 2: Seasonal decomposition (Picture credit: Original).

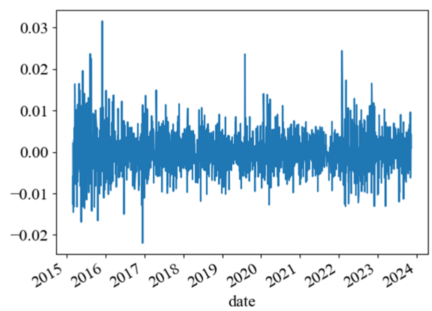

Figure 3: First-order difference (Picture credit: Original).

The study finds that foreign exchange rates have a strong correlation. Fig. 4. shows that CNY has a very strong correlation with the Canadian dollar (CAD) 0.75 and the Great Britain Pound (GBP) 0.73 as well as the Australian dollar (AUD) 0.63 and USD 0.61, which is because of the implementation of a fixed exchange rate policy. It indicates that the price of such currency may grow with the increase of the corresponding currency. It can be used for prediction empirically. However, the JPY seems to have little correlation with the other five currencies. This may be partly due to the macroeconomic trilemma. The Japanese government clings to monetary independence and the free movement of capital and gives up the fixed exchange rate. The result is that the Japan Yen (JPY) changes stochastically. In the feature engineering session, JPY is abandoned for its weak correlation with CNY, and the other four currencies which are USD, CAD, GBP, and AUD are selected as the forex features.

Figure 4: Correlation matrix (Picture credit: Original).

3.2 Feature Engineering

This section introduces the feature inputs of the models. In general, there are three categories of inputs in this study which are respectively the tech features, the commodity features, and the forex features. The raw data are the daily price for CYN, USD, JPY, CAD, GBP, and AUD (direct quotation by Euro), Crude Oil Prices (West Texas Intermediate), Global Gold Price, and Shanghai Composite Index ranging from 2015.3 till 2023.11. The CNY daily price data are used to construct technical indicators such as Exponential Moving Average, Relative Strength Index, Momentum, Commodity Channel Index, Bollinger Bands, and Moving Average Convergence Divergence. Crude Oil Prices WTI, Global Gold Prices, and the Shanghai Composite Index constitute the commodity features. The last category of forex features consists of USD, CAD, GBP, and AUD's daily returns. The label is the daily return of CNY. The dataset is then split in the time series order so that the former 80% of the data was used for training the models and the latter 20% was used for testing the predictions.

The Exponential Moving Average (EMA) gives the most weight to the recent values in a period. Thus, past values have a decreasing contribution, while more recent values dominate. This technique allows the moving average to be more responsive to variations. The 12-day and 26-day EMA are selected in the tech features to consider both the long and short factors.

$$K = \frac{2}{(n+1)} \tag{1}$$

$$EMA = K \times close + (1-K) \times EMA_{-1} \tag{2}$$

The Relative Strength Index (RSI) determines a ratio of the upward price changes to the absolute price changes in a period. The value ranges from 0 to 100. The most popularly used 14-day RSI is selected in the tech features.

$$\begin{cases} up_i = \max\left(close - close_{-1}, 0\right) \\ dn_i = \max\left(close_{-1} - close, 0\right) \end{cases} \tag{3}$$

$$upavg = \frac{\sum_{i=1}^{n} up_i}{n} \tag{4}$$

$$dnavg = \frac{\sum_{i=1}^{n} dn_i}{n} \tag{5}$$

$$RSI = 100 \times \frac{upavg}{upavg + dnavg} \tag{6}$$

The Commodity Channel Index (CCI) is intended to identify initial and final market trends. The value normally ranges from -100 to 100. The 14-day CCI is selected in the tech features.

$$CCI = \frac{\left(TP - SMA(TP, n)\right)}{(0.015 \times MD)} \tag{7}$$

$$TP = \frac{high_n + low_n + close}{3} \tag{8}$$

$$MD = \frac{\sum_{i=1}^{n} TP_i - SMA(TP, n)}{n} \tag{9}$$

The Momentum (MOM) is a gauge of the acceleration and deceleration of prices. 1-day, 5-day, and 14-day MOM are selected in the tech features.

$$Momentum = close - close_{-n} \tag{10}$$

Bollinger Band contains three lines. A straightforward moving average of the average price makes up the middle band. F standard deviations above and below the middle band correspond to the upper and lower bands. The distance between the upper and lower Bollinger Bands measures volatility which is known as the Bollinger Band Width indicator. When volatility is high, the Band Width value is higher; when volatility is low, it is lower. The Band Percent value indicates the location of the close price. 20-day Band Width and Band Percent are selected in the tech features.

$$TP = \frac{high + low + close}{3} \tag{11}$$

$$MidBand = \text{SimpleMovingAverage}(TP) \tag{12}$$

$$UpperBand = MidBand + F \times \sigma(TP) \tag{13}$$

$$LowerBand = MidBand - F \times \sigma(TP) \tag{14}$$

$$BandWidth = 2 \times F \times \sigma(TP) \tag{15}$$

$$BandPercent = \frac{close - LowerBand}{UpperBand - LowerBand} \tag{16}$$

The Moving Average Convergence Divergence (MACD) is the difference between two Exponential Moving Averages. It forecasts trend shifts and the beginning of a new trend direction. 12-26 days Difference (DIF) and Difference Exponential Average (DEA), 9 days MACD are selected in the tech features.

$$DIF = EMA_{12}(close) - EMA_{26}(close) \tag{17}$$

$$DEA = EMA_{9}(DIF) \tag{18}$$

$$MACD = DIF - DEA \tag{19}$$

Crude Oil Prices WTI refers to a specific grade of crude oil that serves as a key benchmark for oil pricing globally. Changes in Crude Oil Prices WTI have broad macroeconomic implications. They affect inflation rates, transportation costs, and the overall economic health of oil-producing and oil-consuming nations. The daily return of Crude Oil Prices WTI is selected in the commodity features.

Gold serves as both a commodity and a financial asset, and its price is a key indicator in the global financial markets. The price of gold is of paramount significance due to its role as a safe-haven asset against inflation and economic uncertainty. Investors and central banks closely monitor gold prices as part of their risk management and wealth preservation strategies. The daily return of Global Gold Price is selected in the commodity features.

The Shanghai Composite Index (SCI) is a key stock market benchmark in China, representing the performance of a diverse range of equities listed on the Shanghai Stock Exchange (SSE). The SCI covers companies across various sectors, providing a comprehensive snapshot of the performance of the Chinese stock market. The daily return of the Shanghai Composite Index is selected in the commodity features.

3.3 Models

This study uses six different machine learning models to train the dataset including Linear Regression, Lasso Regression, Ridge Regression, Long Short-Term Memory, Random Forest, and XG-Boost.

Linear regression is a fundamental statistical method. Lasso regression and Ridge regression are two extensions of linear regression with Lasso introducing a penalty term of the absolute values of the coefficients and Ridge introducing a penalty term of the squared values of the coefficients. It is effective for feature selection and can lead to sparse models.

LSTM is a type of recurrent neural network, particularly effective in modelling sequences and time-series data due to its ability to capture long-term dependencies. The LSTM architecture includes memory cells with gating mechanisms, allowing it to selectively remember or forget information. To be specific, the LSTM structure is composed of the cell state, hidden state, input gate, forget gate, and output gate. Cell state runs along the entire sequence, allowing them to absorb and retain information over time. Input Gate updates the cell state with new data. Forget gate determines what information from the cell state should be retained or deleted. The output gate generates the final output depending on the current cell state. The hidden state transmits information from the past to the present, allowing the model to consider historical data. This architecture enables the model to capture and utilize information over time, making it suitable for time-series prediction. In this study, the architecture starts with an LSTM layer with 100 neurons, using the 'ReLU' activation function. It is followed by three fully connected dense layers with respectively 50 neurons, 10 neurons, and one neuron output.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{20}$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{21}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{22}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_{t-1}] + b_o) \tag{23}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{24}$$

Random Forest is an ensemble learning method that constructs numerous decision trees during training and aggregates their predictions through averaging or voting. Each tree is trained on a random subset of the data and features, introducing diversity and reducing the risk of overfitting. In this study, the architecture consists of 500 decision trees, with a maximum depth of 10 and minimum node samples of 15.

XG-Boost is a gradient-boosting algorithm that builds an ensemble of weak learners and sequentially refines their predictions. XG-Boost minimizes a loss function by iteratively adding weak learners, each compensating for the errors of the existing ensemble. It uses gradient descent optimization to find the best parameters for weak learners. In this study, the architecture consists of 500 decision trees, with a learning rate of 0.001.

4 RESULTS

4.1 Model Performance

The predictions for the CNY exchange rate generated by the six machine learning models are visually presented in Figure 5. Notably, the predicted price trajectories across all six models align closely with the actual price movements, indicating a robust fit between the predicted and observed values.

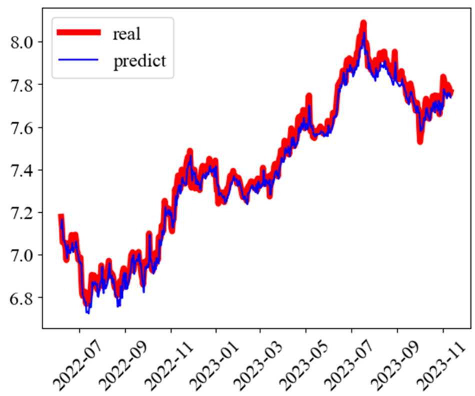

Figure 5: CNY prediction by LSTM (Picture credit: Original).

Table 1 presents the average performance scores of six distinct models employed in the study. The LSTM model demonstrates superior performance with a remarkably low RMSE of 0.063. This outstanding result underscores the effectiveness of LSTM in capturing long-term dependencies. The exceptional performance of the LSTM model can be attributed to its utilization of memory cells, which enable the model to selectively retain or discard information over extended temporal sequences. Following the LSTM model, the Ridge regression, linear regression, and random forest models exhibited commendable performance, though with slightly higher RMSE values. These results highlight the competitive nature of these traditional regression and ensemble models in the context of the conducted time-series predictions.

Table 1: Regression metrics of different models.

Model	MAE	MSE	RMSE
Linear Regression	0.062783	0.005129	0.071596
Lasso	0.068414	0.005955	0.077168
Ridge	0.062589	0.005095	0.071346
LSTM	0.049589	0.004117	0.063395
Random Forest	0.062491	0.005296	0.071757
XG-Boost	0.067127	0.005844	0.075878

Table 2 provides an overview of the average performance scores across different feature inputs utilized in the study. The analysis reveals that the forex features exhibit the most favourable performance, achieving the lowest error with an RMSE of 0.067. The preference for forex features in training predictive models can be attributed to their ability to provide nuanced and comprehensive information about market dynamics. The forex market's inherent characteristics, including liquidity and global interconnectivity, contribute to the robustness of features derived from this domain. The observed success of utilizing forex features is further substantiated by the strong intercorrelation inherent in foreign exchange variables. The inherent relationships between these features contribute to a more coherent and representative model, leading to more accurate predictions.

Table 2: Regression metrics of different features.

Category	MAE	MSE	RMSE
Tech	0.065612	0.00585	0.076233
Commodity	0.063677	0.005335	0.072361
Forex	0.057208	0.004533	0.066976

5 CONCLUSION

In this research, a synergy of six distinct models and three feature categories is designed to formulate a predictive framework for forecasting the daily return of the CNY exchange rate. The outcomes unequivocally affirm the efficacy of each model in forecasting the CNY exchange rate, thereby underscoring their collective predictive capabilities. The optimal model is LSTM which showcases its prowess in handling intricate temporal patterns and reinforces its role as a preferred choice for time-series prediction tasks. As for feature engineering, forex features minimize prediction errors due to their reflection of market dynamics and strong correlation with CNY. In a word, the findings give a more strategic approach to model training, contributing to the advancement of predicting exchange rates.

REFERENCES

Y. Singh, A. S. Chauhan. J. Theor. Appl. Inf. Technol. 5(1) (2009)
R. Meese, K. Rogoff. J. Frenkel. Chi. U - Chi. P (1983)
S. Galeshchuk, S. Mukherjee. In. Sys. in Acc. Fin. Mgmt, 24(4), 100-110 (2017)
B. Qian, K. Rasheed. Journal of Forecasting, 29(3), 271-284 (2010)
C. Amat, T. Michalski, G. Stoltz. Journal of International Money and Finance, 88, 1-24 (2018)
D. Pradeepkumar, V. Ravi. App. Sof. Comp., 58, 35-52 (2017)
T. Fischer, C. Krauss. Eur. J. Oper. Res. 270(2), 654-669 (2018)
S. A. Gyamerah, E. Moyo. Comp. 1-11 (2020)
Y. Wang, Y. Guo. Chn. Comm. 17(3), 205-221 (2020)
W. Cao, W. Zhu, W. Wang, Y. Demazeau, C. Zhang. IEEE Intel. Sys. 35(2), 43-53 (2020)