

Kamet A.N.

International Information Technology University
Almaty, Kazakhstan

Scientific supervisor: Moldagulova A.N.

COMPARATIVE STUDY OF MACHINE LEARNING ALGORITHMS FOR CURRENCY RATE FORECASTING

Abstract. *In this article multiple Machine Learning algorithms have been analyzed in terms of currency rate forecasting, in order to identify the most accurate one. The scope of the study was limited to forecast USDKZT using market variables as a dataset. Comparative study shown that Extreme Random Forecast outperforms other suggested Machine Learning algorithms.*

Keywords: *comparative study, currency exchange rate, machine learning, currency rate forecasting.*

For decades currency forecasting have been crucial part of Foreign Exchange market, where representatives of various segments as government, banks, investors and traders take a part. In recent times, different advanced computing methods yield better results for time series forecasting [1]. Despite the increasing popularity of data-driven strategies, the literature that analyzes machine learning methods in financial forecasting is very limited, with most papers only focusing on stock return prediction [2].

Considering the lack of researches made on machine learning algorithms for currency rate forecasting, the comparative study of several machine learning algorithms were conducted. For dataset there were chosen monthly rate of USDKZT currency pair and M1, M2 money supply for dates from January 2001 to March 2020. As for machine learning algorithms there were chosen Linear Regression, Lasso Regression, ElasticNet, Random Forest and Extreme Random Forest.

Table 1. Results of comparative analysis

Algorithm	Score	Root mean squared error	Mean absolute error	R-squared
Linear Regression	0.89	28.47	19.37	0.90
Lasso Regression	0.89	28.46	18.66	0.90
ElasticNet	0.89	28.61	18.5	0.90
Random Forest	0.98	10.57	5.22	0.99
Extreme Random Forest	0.99	6.1	3.13	1

Comparative study were contained 3 steps: data preprocessing, implementation of algorithms and analysis of results.

For preprocessing part of comparative analysis the original data was modified by inputting previous time values known as lags. After several trials and errors 4 lags, which is equal to 4 months, were chosen. Another modification was an addition of year and month into columns in order to make data more specific. And the lastly added columns were the difference between the current and lag exchange rates.

As a main goal for comparative analysis it was decided to predict the USDKZT exchange rates for 3, 6 and 12 months. For this purpose 3 new columns with the future values of 3, 6, 12 months exchange rates from the original dataset were added to preprocessed data.

For implementation of algorithms there were used python as a programming language and jupyter notebook as visualization and testing tool. Besides the language itself there were used standard calculation and visualization libraries as pandas, matplotlib, numpy, seaborn and the library for machine learning – sklearn. Before beginning the implementation part the preprocessed data was split into train and test data by 80% and 20% respectively.

According to comparative analysis results in Table 1 Extreme Random Forecast shown the most accurate forecast results, with the Score of 0.99 and Mean Absolute error of 3.13. In order to assure the accuracy of Extreme Random Forest algorithm, the test set accuracy measurement was applied on it, results of which are depicted in Figure 2.

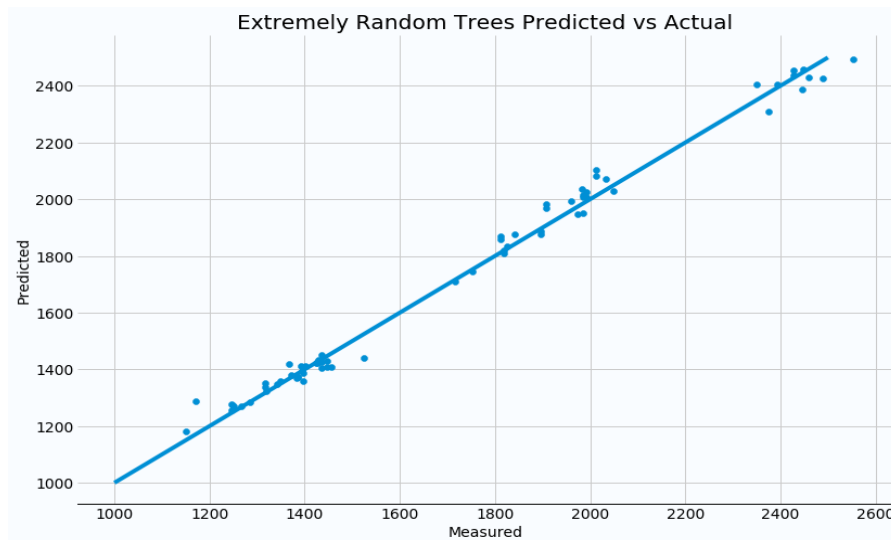


Figure 1. Comparison of predicted and actual values for Extreme Random Trees

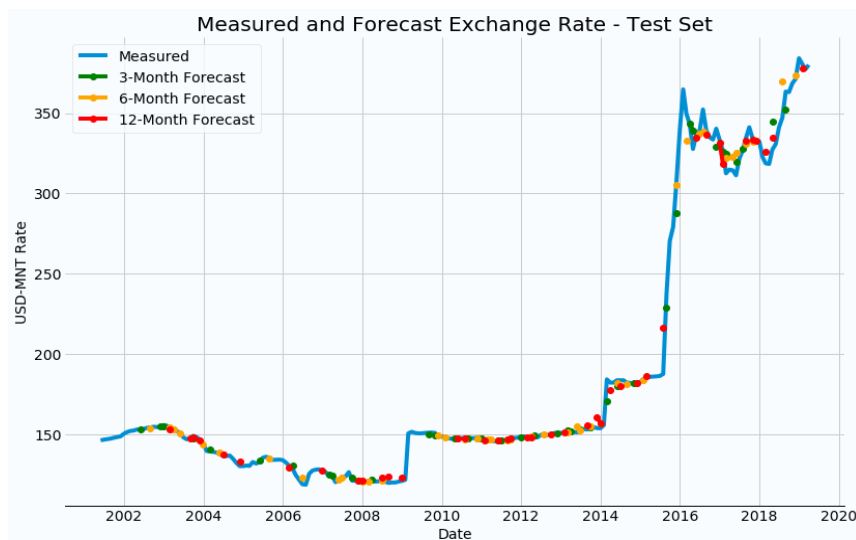


Figure 2. Accuracy measurement of Extreme Random Forest

Test set accuracy measurement was shown similar accurate results as the comparison of predicted and actual values for Extreme Random Trees depicted in Figure 1. Additionally, test set accuracy measurement shown the results on exchange rates for 3, 6 and 12 months.

In conclusion, for currency forecasting of chosen dataset the Extreme Random Forest algorithm shown the most accurate results. However, this results were not tested on real data, which would be implemented in future studies.

REFERENCES

1. Y. Muhammad, D. Mehr Yahya, A. Sitara, M. Muazzam, A. Farhan, M. Irfan and R. Seungmin, "An Intelligent Event-Sentiment-Based Daily Foreign Exchange Rate Forecasting System," *Applied Sciences*, vol. 9, no. 15, 2019
2. C. González Rojas and M. Herman, "Foreign Exchange Forecasting via Machine Learning," in *AIIPCC '19*, Sanya, 2019

Камет Ә.Н.

Ғылыми жетекші: Молдагулова А.Н.

Оффлайн және онлайн деректерді қаржылық көрсеткіштер бойынша талдау әдісін әзірлеу

Аңдатпа. Осы мақалада бірнеше машиналық оқыту алгоритмдері олардың неғұрлым дәліл анықтау үшін валюта бағамын болжау тұрғысынан талданды. Зерттеу аясы нарықтық айнымалыларды мәліметтер жиынтығы ретінде пайдалану арқылы USDKZT болжамымен шектелді. Салыстырмалы зерттеу көрсеткендей, Extreme Random Forecast машинаны оқытудың басқа алгоритмдерінен жоғары.

Кілт сөздер: салыстырмалы зерттеу, валютаның айырбас бағамы, машиналық оқыту, валюта бағамын болжау.

Камет А.Н.

Научный руководитель: Молдагулова А.Н.

Разработка метода анализа оффлайн и онлайн данных по финансовым параметрам

Аннотация. В этой статье были проанализированы несколько алгоритмов машинного обучения с точки зрения прогнозирования курса валюты, чтобы определить наиболее точный из них. Объем исследования был ограничен прогнозом USDKZT с использованием рыночных переменных в качестве набора данных. Сравнительное исследование показало, что Extreme Random Forecast превосходит другие предложенные алгоритмы машинного обучения.

Ключевые слова: сравнительный анализ, курс обмена валют, машинное обучение, прогноз курса валют.

About authors:

Aziza N. Kamet, Masters student, Information Systems Department, International Information Technology University.

Aiman N. Moldagulova, associate professor, Information Systems Department, International Information Technology University.