Foreign Exchange Prediction Using Machine Learning Approach: A Pilot Study

Conference Paper · August 2021

DOI: 10.1109/ICOIACT53268.2021.9563998

4 authors, including:

Sudimanto Sudimanto
STMIK LIKMI

Yaya Heryadi
Binus University

Lukas Lukas
Atma Jaya Catholic University of Indonesia

published on IEEE Xplore

Foreign Exchange Prediction Using Machine Learning Approach: A Pilot Study

Sudimanto
Computer Science Department, BINUS Graduate Program
Doctor of Computer Science, Bina Nusantara University
Jakarta, Indonesia
sudianen@yahoo.com

Yaya Heryadi
Computer Science Department, BINUS Graduate Program
Doctor of Computer Science, Bina Nusantara University
Jakarta, Indonesia
yayaheryadi@binus.edu

Lukas
Cognitive Engineering Research Group (CERG)
Faculty of Engineering, Universitas
Katolik Indonesia Atma Jaya
Jakarta, Indonesia
lukas@atmajaya.ac.id

Antoni Wibowo
Computer Science Department, BINUS Graduate Program
Doctor of Computer Science, Bina Nusantara University
Jakarta, Indonesia
anwibowo@binus.edu

***Abstract*— Foreign Exchange or FOREX trading is not only done on foreign currencies but, FOREX also can be done on several commodities such as Gold, Silver, Oil. Gold is one of the most valuable commodities in the world. Investors began to offer gold as a trading material against foreign currencies. Machine Learning (ML) in the FOREX trading world is usually used to predict future FOREX values. This pilot study aims to see a model from machine learning that has a fairly high level of accuracy in making FOREX predictions. This pilot study using historical data taken from the investing.com database where the FOREX data taken is FOREX XAU/USD data, with a period year from 2019 until 2021, and the indicator used is Moving Average Convergence/Divergence (MACD) technical analysis. The average accuracy obtained after training on the Tree model is 86.3%, the Support Vector Machine (SVM) model is 86.6% and the Ensemble model is 86.55%. Testing conducted using machine learning models for Tree, SVM and Ensemble models have the same level of accuracy, which is 88.3%.**

Keywords*— *Foreign Exchange, Machine Learning, MACD, Technical Analysis

I. INTRODUCTION

The stock market and foreign exchange or FOREX market according to Bisnis.com are the most popular financial markets in the investment world. Investments made in these two markets can have a high return but also have a high risk. The activities carried out in these two markets are trading [1]. Weerathunga and, Silva [2] suggested that FOREX market is one of the largest market in the world.

Foreign exchange is an event that occurs simultaneously between sales and purchases of foreign currency. The buying and selling of foreign currencies are carried out in pairs [3]. The FOREX market is the largest financial market in which the transaction value of the market is more than 2 trillion US$/day. And the FOREX market is different from financial markets which have a building or location from the place of transactions as FOREX transactions are carried out electronically through a network of banks, companies and, individual exchanges. In addition, the FOREX market operates 24 hours during weekday [3].

FOREX trading not only involves foreign currencies but, also involves several commodities such as Gold, Silver, Oil. Among those commodities, gold is one of the most valuable commodities in the world [4]. According to the Astonacci.com site, gold is one of the strong investment after the USD currency [5]. Investors began to offer gold as a trading material against foreign currencies, such as the euro, Swiss Franc, Australian dollar and, what is still well known is trading against the USD currency.

The high risk of FOREX trading has motivated many researcher to develop a robust model to predictive analytics. The prediction model based on Machine Learning (ML), for example, has been used in the FOREX trading world to predict future FOREX values. Some of the approaches in ML, like [6] using the Support Vector Machine (SVM) model in making predictions on FOREX, with an accuracy rate of 82.5% where the indicators used are Relative Strength Index, Moving Average, Bollinger Bands, some custom indicators. [7] using the Recurrent Neural Network method with the autoregressive integrated moving average (ARIMA) indicator, got an accuracy rate of 75.7% with a fairly small root mean square error (RMSE) value compared to other models compared.

This pilot study is divided into 5 sections. Section II will discuss a literature review about ML model and technical analyzes from previous research. Section III briefly discuss about framework ML model in Forex and method for prediction Forex. Section IV will discuss about result from this pilot study using some of ML model and MACD technical analyzes indicator. The last section will conclude this paper

II. LITERATURE REVIEW

Research on FOREX, not only uses a machine learning approach but also uses a model from Reinforcement learning (RL) as done by [8] using RL to predict FOREX, which uses a values-based approach and a policy-based approach. [9] uses the Deep learning method to improve prediction results on the stock market with several technical analyzes, namely Stochastic %K, Stochastic %D, Momentum, Rate of Change, William's %R, A/D Oscillator, and Disparity 5.

[10] used the Long-Short Term Memory (LSTM) model with the Forex Loss Function (FLF-LSTM) indicator in making predictions on FOREX. Based on the results of the research, the approach of the FLF-LSTM model was able to reduce the error by 13% when compared to using the ARIMA indicator.

Research on FOREX with commodities was also carried out by [11] using the Deep Reinforcement Learning approach, which succeeded in proving that Deep Reinforcement Learning can increase the ability to trade automatically. Other indicators were also introduced by [12] in predicting the movement of stock market prices using deep learning with engineering features. The method introduced by [12] is a Multi-Filter Neural Network (MFNN) with financial time series features and price movement prediction. This pilot study aims to see a model from machine learning that has a fairly high level of accuracy in making FOREX predictions. [13] comparing the moving average convergence/divergence (MACD) indicator with the faster take profit signal feature with the simple MACD indicator and the results obtained between both MACD does not have a significant difference.

III. Method

This pilot study uses historical data was taken from investing.com where the FOREX data taken is FOREX XAU/USD data, with a period year from 2019 until 2021. This data is divided into two parts for the year 2019-2020 was used to train the models that will be used to make FOREX predictions. While the data in 2021 is used as testing data. This experiment uses MATLAB applications like Classification Learner to generate machine learning models to predict FOREX.

The analysis used in this pilot study is a technical analysis where this analysis is only applied to price movements in the FOREX market [3]. This technical analysis is used by investors to see and find other technical indicators, to find new indicators using the help of computers. In this study, the indicator used is Moving Average Convergence/Divergence (MACD). The framework of this study can be seen in Fig 1.

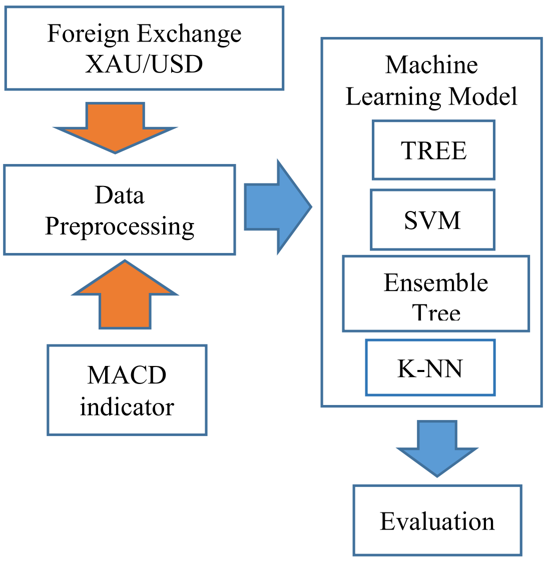

Fig. 1. Framework FOREX prediction.

IV. Result And Discussion

A. Data Preprocessing

The data preprocessing process carried out in this study is to prepare the data to be used as a dataset for training and testing. At this stage, all data will be checked for data types that will later be used in predicting. And at this stage, the dataset will be checked whether there is empty data or not.

Fig 2. is a graph from FOREX data were taken from 2019 until 2020 which is used as training data. At this stage the MACD indicator is also included to determine the conditions to be predicted, namely "buy" or "not buy".

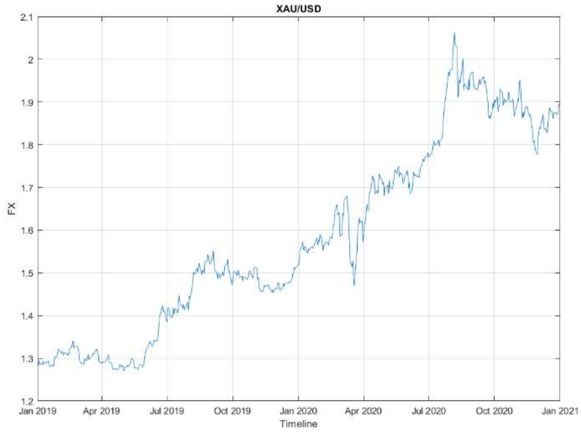

Fig. 2. FOREX chart from 2019 until 2020

B. Machine Learning Models

At this stage, training will be conducted on the prepared FOREX data using the MACD indicator. The training model of machine learning is carried out using the MATLAB program with the classification learner application which is provided by MATLAB. Fig 3. is the distribution of the classification of the training data set. Var1 in Fig 3 is a predictor of the opening price while Var2 is a predictor of the closing price. The class used in this study is "buy" class which is marked with a blue dot and "not buy" is marked with a red dot.

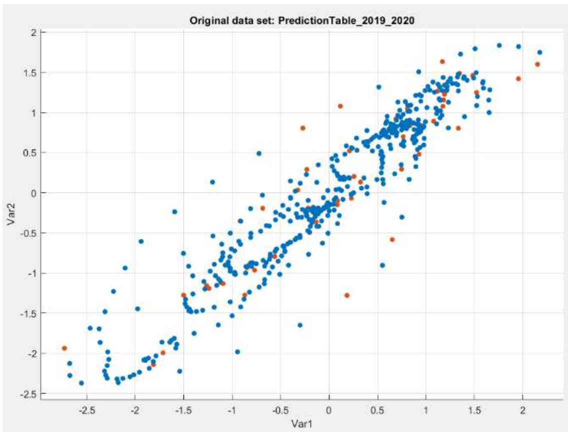

Fig. 3. Scatter plot from an original data set

The first model was the Tree model, which there are three models from the Tree model used to predict FOREX, such as Fine Tree, Medium Tree, and, Coarse Tree. The accuracy obtained from the training using the tree models can be seen in Table I.

TABLE I. TREE MODEL ACCURACY

Machine Learning model	Accuracy
Fine Tree	83.0%

Machine Learning model	Accuracy
Medium Tree	87.5%
Coarse Tree	88.3%

From Table I can be said that all Tree models used for training have an accuracy rate above 85%. And Coarse Tree model gets the highest accuracy is compared to the other models.

The second model in this pilot study is the SVM model. There are 6 SVM models used to predict FOREX, such as Linear SVM, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Coarse Gaussian SVM. The accuracy obtained from the training using the SVM models is shown in Table II.

TABLE II. SVM MODEL ACCURACY

Machine Learning model	Accuracy
SVM Linier	88.9%
Quadraric SVM	88.9%
Cubic SVM	75.6%
Fine Gaussian SVM	88.9%
Medium Gaussian SVM	88.9%
Coarse Gaussian SVM	88.9%

It can be seen that not all SVM models have a high enough accuracy in making predictions such as the Cubic SVM model, the accuracy obtained is 75.6%, while for other SVM models the accuracy rate is 88.9% based on table 2 of the SVM model which has an accuracy level above 85%.

The third model in this pilot study is the Ensemble model used in predicting the FOREX. In this model were used 2 models, namely Ensemble Boosted Tree and, Ensemble Bagged Tree. The accuracy of this model can be seen in Table III.

TABLE III. ENSEMBLE MODEL ACCURACY

Machine Learning model	Accuracy
Ensemble Boosted Tree	87.3%
Ensemble Bagged Tree	85.8%

From the prediction results using the Ensemble model, it can be seen that the Ensemble Boosted Tree model has an accuracy rate of 87.3%.

The Fourth model in this pilot study is K-Nearest Neighbor (K-NN). The KNN model used in this study uses different K values. The K values used in this study are 5 and 10. The difference in the level of accuracy obtained with K = 5 and K = 10 is 0.8%. Table IV shows the level of accuracy of the K-NN model.

TABLE IV. K-NN MODEL ACCURACY

Machine Learning model	Accuracy
5-NN	88.1%
10-NN	88.9%

C. Testing the Prediction Model

Several models of machine learning that have been trained will be tested using data testing. Testing data is FOREX data where period time the data is from 01/01/2021 until 06/30/2021. Fig 4 is a graph of FOREX taken from January 2021 until June 2021 which is used for data test. Machine learning models that will be used in testing are Coarse Tree, Coarse Gaussian SVM, Ensemble Boosted Tree and K-NN with K= 10, which has a high level of accuracy in each model.

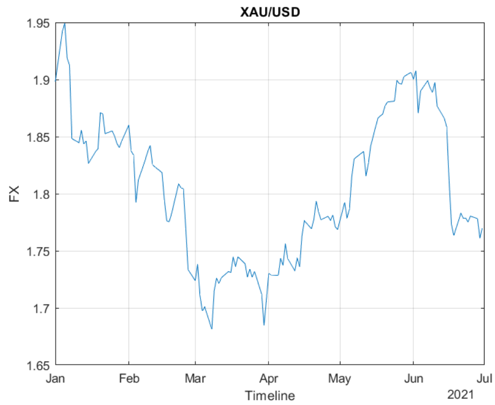

Fig. 4. FOREX chart from January 2021 until June 2021

The level of accuracy in the Coarse Tree, Coarse Gaussian SVM, Ensemble Boosted Tree, and K-NN models can be calculated using the accuracy formula obtained from the confusion matrix. Fig 5 is a confusion matrix from the prediction results from the coarse Tree model where the prediction and actual is "not buy" has a value of 83 while the predicted value for "not buy" and actual is "buy" is 11. Fig 5 shows the confusion matrix of the Coarse Tree.

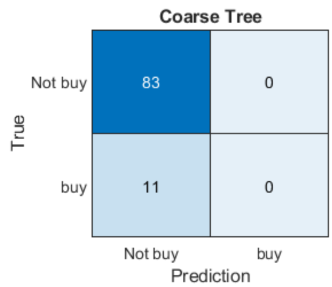

Fig. 5. Confusion Matrix for Coarse Tree

The confusion matrix Coarse Gaussian same as the confusion matrix coarse tree where the predicted and actual results have a value of 83 and a predicted value for "not buy" and the actual "buy" is 11. Fig 6 shows the confusion matrix of Coarse Gaussian SVM.

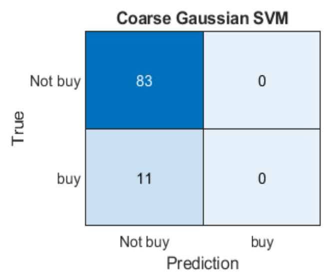

Fig. 6. Confusion Matrix for Coarse Gaussian SVM

The Confusion Matrix of Ensemble Boosted Tree is shown in Fig 7. The predicted and actual results from the boosted tree ensemble model have a value of 81 and the predicted value for "not buy" and the actual "buy" is 11. In this model, there is one condition where the prediction results for "buy" but actually "not buy" and the results obtained are 2.

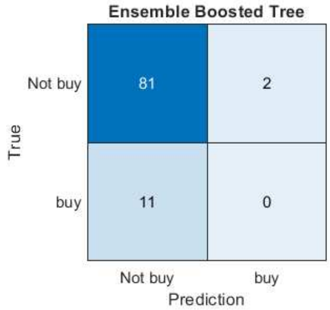

Fig. 7. Confusion Matrix for Ensemble Boosted Tree

The Confusion Matrix of the KNN model is shown in Fig 8. The number of K used in this model is 10, and the results of the confusion matrix from this model are the same as the confusion matrix found in the Tree and SVM models. The testing results of all the machine learning models can be seen in Table V.

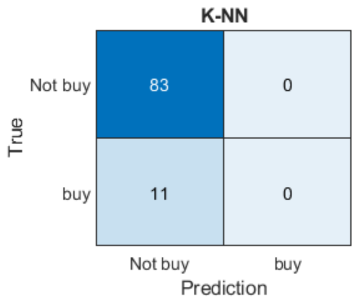

Fig. 8. Confusion Matrix for K-NN (K = 10)

TABLE V. RESULT MODEL ACCURACY

Machine Learning Model	**Accuracy**	
	Training	***Testing***
Coarse Tree	88.3%	88.3%
Coarse Gaussian SVM	88.9%	88.3%
Ensemble Boosted Tree	87.3%	88.3%
10-NN	88.9%	88.3%

From Table V it can be seen that the results of testing of several machine learning models there are differences such as the Coarse Gaussian SVM model when this model has trained the level of accuracy obtained is 88.9% but when testing on other datasets the accuracy rate decreases, as is the case with the Ensemble Boosted model. The Tree model accuracy obtained when the model is tested is the same when the model is training and the result is 88.3%.

V. CONCLUSION

The machine learning models used in this pilot study have different levels of accuracy in predicting FOREX. The average accuracy obtained in the Tree model is 86.3%, for the SVM model it has an average accuracy rate of 86.6% while for the ensemble model 86.55% among the models tested that the SVM model has the highest average level of accuracy. As further research, optimization of machine learning models will be carried out and using other technical indicators to get the accuracy higher than this pilot study.

REFERENCES

[1] D. R. Meilanova, "Mengenal Perbedaan antara Pasar Saham dan Forex - Finansial Bisnis.com."

[2] H. P. S. D. Weerathunga and A. T. P. Silva, "DRNN-ARIMA approach to short-term trend forecasting in forex market," *18th Int. Conf. Adv. ICT Emerg. Reg. ICTer 2018 - Proc.*, pp. 287–293, 2019, doi: 10.1109/ICTER.8615580.

[3] M. D. Archer, *Getting started in currency trading : winning in today's Forex market*. 2010.

[4] B. Lobel, "What is Gold? Understanding Gold as a Trader's Commodity," Nov. 07, 2020. https://www.dailyfx.com/education/commodities/what-is-gold.html (accessed May 27, 2021).

[5] H. Jumaidi, "MENGENAL TRADING GOLD/XAU/USD YUK," 2020. https://astronacci.com/blog/read/mengenal-trading-gold-xau-usd-yuk (accessed May 27, 2021).

[6] T. N. T. Thu and V. D. Xuan, "Using support vector machine in FoRex predicting," *2018 IEEE Int. Conf. Innov. Res. Dev. ICIRD 2018*, no. May, pp. 1–5, 2018, doi: 10.1109/ICIRD.2018.8376303.

[7] Z. Zeng and M. Khushi, "Wavelet Denoising and Attention-based RNN- ARIMA Model to Predict Forex Price," *Proc. Int. Jt. Conf. Neural Networks*, 2020, doi: 10.1109/IJCNN48605.2020.9206832.

[8] K. S. Zarkias, N. Passalis, A. Tsantekidis, and A. Tefas, "Deep Reinforcement Learning for Financial Trading Using Price Trailing," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2019-May, pp. 3067–3071, 2019, doi: 10.1109/ICASSP.2019.8683161.

[9] P. Oncharoen and P. Vateekul, "Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators," *ICAICTA 2018 - 5th Int. Conf. Adv. Informatics Concepts Theory Appl.*, pp. 19–24, 2018, doi: 10.1109/ICAICTA.2018.8541310.

[10] S. Ahmed, S. U. Hassan, N. R. Aljohani, and R. Nawaz, "FLF-LSTM: A novel prediction system using Forex Loss Function," *Appl. Soft Comput. J.*, vol. 97, p. 106780, 2020, doi: 10.1016/j.asoc.2020.106780.

[11] B. A. Usha, T. N. Manjunath, and T. Mudunuri, "Commodity and Forex trade automation using Deep Reinforcement Learning," *1st Int. Conf. Adv. Technol. Intell. Control. Environ. Comput. Commun. Eng. ICATIECE 2019*, pp. 27–31, 2019, doi: 10.1109/ICATIECE45860.2019.9063807.

[12] W. Long, Z. Lu, and L. Cui, "Deep learning-based feature engineering for stock price movement prediction," *Knowledge-Based Syst.*, vol. 164, pp. 163–173, 2019, doi: 10.1016/j.knosys.2018.10.034.

[13] D. Vezeris, T. Kyrgos, and C. Schinas, "Take Profit and Stop Loss Trading Strategies Comparison in Combination with an MACD Trading System," *J. Risk Financ. Manag.*, vol. 11, no. 3, p. 56, 2018, doi: 10.3390/jrfm11030056.