Currency movement forecasting using time series analysis and long short-term memory

Kristina Sanjaya Putri[a*], Siana Halim[a]

[a]Department of Industrial Engineering, Universitas Kristen Petra, Surabaya, Indonesia.

*Corresponding Author: kristinaputri98@gmail.com

ARTICLE INFO

ABSTRACT

Foreign exchange is one type of investment, which its goal is to minimize losses that could occurred. Forecasting is a technique to minimize losses when investing. The purpose of this study is to make foreign exchange predictions using time series analysis called Auto Regressive Integrated Moving Average (ARIMA) and Long Short-term memory methods. This study uses the daily EUR / USD exchange rates from 2014 to March 2020. The data are used as the model to predict the value of the foreign exchange market in April 2020. The model obtained will be used for predictions in April 2020, where the RMSE values obtained from time series analysis (ARIMA) with a window size of 100 days and LSTM sequentially as follows 0.00527 and 0.00509. LSTM produces lower RMSE values than ARIMA. LSTM has better prediction results; this is because the LSTM has the ability to learn so that it can utilize a large amount of data while ARIMA cannot use it. ARIMA does not have the ability to learn even though given a large amount of data it gives poor forecasting results. The ARIMA prediction is the same as the values of the previous day.

1. Introduction

Foreign exchange (FOREX) is one type of trading activity that trades a country's currency to others for 24 hours continuously (Nagpure, 2019). FOREX market is the world's largest financial market. The daily trading volume has been increased six trillion dollars which it's 45% of the transaction volume comes from terminal retail customers (Ni et al., 2019). There are several techniques in FOREX trading, one of them is forecasting FOREX. Forecasting on FOREX can be done by the method of Statistical Learning (time series analysis), Technical analysis (candle stick), and deep learning (Recurrent Neural Network, LSTM). There are some research about forecasting FOREX with any method such using deep learning (Czekalski et al., 2015; Korczak & Hemes, 2017; Nagpure, 2019; Sezer et al., 2020), ARIMA (Reddy SK, 2015), fuzzy neuron (Reddy SK, 2015) and neuro-fuzzy system (Yong et al., 2018). Forecasting provides factors to be able to predict further whether there will be a bullish or bearish. Bullish symbolizes the optimism of the actors in market conditions whose prices are rising. Bearish symbolizes the pessimism of the actors in market conditions whose prices are falling (Ong, 2019). Forecasting is a technique that can help in minimizing losses on FOREX transactions. Forecasting techniques that have the lowest error rate are the most suitable techniques to use. Another consideration is for ease and

speed in using the method. Other paper have discussed about forecasting FOREX using deep learning (Czekalski et al., 2015; Korczak & Hemes, 2017; Nagpure, 2019; Sezer et al., 2020).

In this paper a comparison of forecasting performance between the two methods will be made where the method uses time series analysis specifically ARIMA and deep learning (LSTM) methods by observing the comparison of the Root Mean Squared Error (RMSE) values of the two methods and the speed of forecasting. ARIMA method is statistical learning which discard the trend of the data while LSTM is deep learning which able to learn the trend of data. The purpose of this paper is to find out which method is the best for making predictions in EUR/USD exchange rates.

2. Literature Review

Forecasting is done by making a model based on past data. The method used is time series analysis and long short-term memory. Broadly speaking, FOREX forecasting processes include data processing, finding optimal models and evaluating models.

2.1   Time Series Analysis

Time Series is a collection of observations taken sequentially in time (Palma, 2016). Time series analysis is important because it is used widely in the real world as an example of population growth in a country from the measure of its current population, it is useful to determine the future prospects of a population (Konar & Bhattacharya, 2017). This time series prediction believes events that have occurred in the past will happen again in the future based on the results of the plot of the time sequence (Palma, 2016).

2.2   Integrated Autoregressive Moving Average (ARIMA)

ARIMA, namely the Box-Jenkins model is the most common time series prediction model in the statistic model. ARIMA model has three basic types which are moving average (MA) model, autoregressive (AR) model, and autoregressive integrated moving average model (ARIMA). ARIMA generally denoted by ARIMA (p,d,q). p means the autoregressive parameters; q means the moving average order and d is the number of times the time series becomes stationary through difference. The value of d usually under 2. The p and q parameters are obtained from the partial auto correlation function (PACF) and auto correlation function (ACF). Model of ARIMA should be kept as small as possible to avoid overfitting to the sample data (Nielsen, 2019; Yang et al., 2020)

2.3   Recurrent Neural Network

Recurrent neural network (RNN) is a type of artificial neural network that is best suited to recognizing patterns in a sequence of data (Manaswi, 2018). The word recurrent comes from how this network works. This network applies the same method to each sequence of elements, gathering information about previous terms. The recurrent neural network is a very powerful algorithm that can classify, cluster, and make predictions about data, especially time series and text (Michelucci, 2018; Zheng et al., 2014).

2.4   Long Short-Term Memory

Long Short-term memory (LSTM) is a modification of RNN which deals with the problems of vanishing and exploding gradient also overcomes training problems in a very long sequence and retaining memory (Song et al., 2020). It is difficult to use RNN in solving problems that require long-term learning of temporal dependencies and long sequence relationships. The gradient of the loss function decay exponentially with time at which this causes difficulties in training the RNN (Manaswi, 2018). The problem of loss of gradient can be controlled better by providing additional gates. Gates LSTM consists of input gate, forget gate and output gate. This addition system is able to make a selection of information that needs to be stored and what information needs to be forgotten (Goyal et al., 2018; Lecun, et al., 2015).

LSTM has several types that are used for time series forecast, one of them is vanilla LSTM. Vanilla LSTM is the simplest LSTM model. This type has one hidden layer, and an output is used to make predictions. Figure 1 is a vanilla RNN. Vanilla LSTM has the same concept as vanilla RNN but different on the gate.
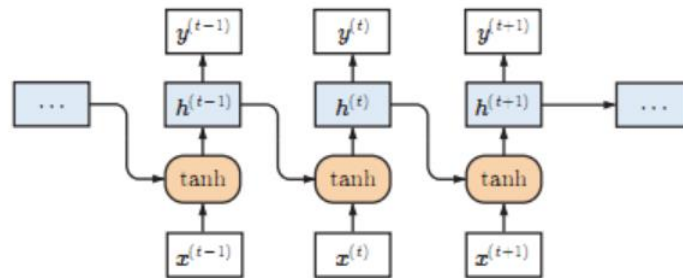

Figure 1. Vanilla RNN (Ming et al., 2017)

Where x_{t-1}, x_t, x_{t+1} is a sequential matrix input. h_{t-1}, h_t, h_{t+1} is the hidden layer that will be used in the next period. y_{t-1}, y_t, y_{t+1} is the output from the input x_{t-1}, x_t, x_{t+1}. When time t input x_t updates the value of h_{t-1} to h_t (Ming *et al.*, 2017).

3. Research Methodology

FOREX is the trading of currencies across foreign market. FOREX market is a place for people to buy, sell or exchange currencies (Rigters, 2019). Nowadays, FOREX trading can be done anywhere and easily accessed because it can be done online. Even now there are several applications that are used to do FOREX trading so that it can be accessed by using a gadget. With this easy access, people need a quick consideration to make a decision. Therefore, this research is needed to help traders not to get a lot of losses.

In a paper made by piotr (Czekalski et al., 2015) using ANN in making forex predictions. ANN was then developed into RNN and developed again into LSTM. This paper will compare the results of ARIMA and LSTM. The result will showed which method will produce accurate predictions that can help people related to this expertise and traders in making decisions about buying or selling in FOREX trading so that they can minimize the loss.

This research step starts with collecting historical data through http://www.investing.com. The historical data used is the daily closing exchange rates of the EUR/USD currency with a period

from January 2014 to April 2020. Data that has been collected will be used to make a model using 2 methods, namely ARIMA and LSTM. This process will produce the best model of each method.

The best model of ARIMA and LSTM will use to make prediction. Prediction results will be compared with the actual which will produce RMSE value that is used to find out which method is best in this case. Finally, we could find the conclusions about which method is better and has a lower RMSE value. RMSE value indicates the error that the prediction results have. In this case when the RMSE value is small, it means that the error or different of the prediction result is small than the actual price.

4. Results and Discussion

This section will discuss the result of the research. Each method will be explained in detail on how to get the model. In brief, this section will tell the best models of ARIMA and LSTM obtained in this study and the comparison of ARIMA and LSTM results.

4.1 Time Series Analysis

Time Series Analysis is a statistical method for prediction data that taken based on time series. Forecasting models used in this study are autoregressive (AR), moving average (MA), autoregressive moving average (ARMA) and integrated autoregressive moving average (ARIMA). The model is chosen based on the characteristic of the data processed (Hyndman et al., 2008; Wang et al., 2006).

AR, MA, and ARMA can only be used in time series that are stationary, while for time series that nonstationary using ARIMA. Globally steps in making predictions with ARIMA are input data, check stationary data, difference, identification of AR and MA, optimal model checking. Making ARIMA manually with the first step is to plot data that will be used to check stationarity of data.
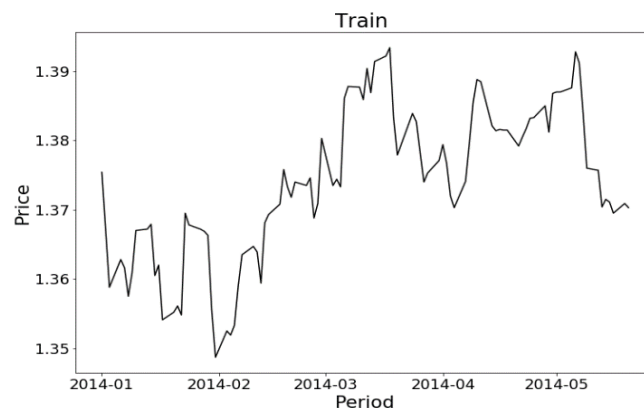

Figure 2. Plot 100 Data

Modelling uses a moving window where the way it works is there is a window which the size will be determined by the researcher in making the ARIMA and after that this window will move to predict the next day. The window size used is 4 types of amount of data that is 100 data using data from 2014 to March 2020, 100 data from 2014 to April 2020, 80% of total data (1304 data) using data from 2014 to 31 March 2020 and 80% of total data (1321 data) using data from 2014 to 31 April 2020. Stationarity of data is checked on each window size, for example 100 data.

Data is said to be stationary if it has a fixed mean and constant variance. The graph in Figure 2 visually can be seen the data has a mean and variance are not constant so that the data is not stationary. This is true since the forex market has asymmetric volatility (Baruník et al., 2017).

Stationarity test other than look at the plot data can also be able to use augmented dicky fuller (ADF) test (Diebold and Kilian, 2000). The hypothesis of this test are:

H_0: Data is not stationary

H_1: Data stationer

ADF Test results found that the ADF value is -2.04 and p-value = 0.27. This states that failed to reject the initial hypothesis, so the data is not stationary. The model used is ARIMA model (p,d,q). The value of d will be obtained from the number of times the difference in data until the data is stationary.

The data is stationary, so the next step is to look for the parameters of p and q using partial autocorrelation (PACF) and autocorrelation (ACF). The previous step is a manual step in making the ARIMA model. Modeling uses a moving window, therefore it is impossible to make it manually so the auto_ arima algorithm is used. The work of this algorithm is same as with manually process the difference is the algorithm will be searching model with the best AIC and BIC value from the guessed model optimal which get from plot the ACF and PACF. If the AIC and BIC values are minimal, the ARIMA (p,d,q) is obtained.

The results of this algorithm have obtained a model with the value of Akaike information criterion (AIC) and Bayesian information criterion (BIC). The model results from the first 100 data obtained the ARIMA (0,1,0). The prediction results show 2 window sizes namely 100 and 80% of the amount of data or 1304 data and 1321 data.

Table 1 shows the results of forecasting with ARIMA for the period March 23th 2020 to March 31st 2020 using historical data until March 2020. The forecasting result of ARIMA consists of 2 window sizes, 100 data and 1304 data.

Table 1. Prediction Using Historical Data until March 2020

DATE	ACTUAL	PREDICTION	
		100	1304
Mar 23th 2020	1.0723	1.0712	1.0695
Mar 24th 2020	1.0789	1.0739	1.0723
Mar 25th 2020	1.0882	1.0804	1.0789
Mar 26th 2020	1.1030	1.0893	1.0882
Mar 27th 2020	1.1142	1.1100	1.1030
Mar 30th 2020	1.1048	1.1196	1.1142
Mar 31st 2020	1.1031	1.1079	1.1048

Table 2 is the result of forecasting by ARIMA for the period of April 1st 2020 to April 10th 2020 using data until April 2020. The results of this forecasting consist of 2 window sizes namely 100 data and 1321 data. This prediction uses historical data from November 13th 2019 to April 30th 2020.

Table 2. Prediction Result for April

DATE	ACTUAL	PREDICTION	
		100	1304
APRIL 1st 2020	1.0964	1.0996	1.1031
APRIL 2nd 2020	1.0858	1.0960	1.0964
APRIL 3rd 2020	1.0809	1.0842	1.0858
APRIL 6th 2020	1.0793	1.0788	1.0809
APRIL 7th 2020	1.0891	1.0996	1.1031
APRIL 8th 2020	1.0858	1.0931	1.0891
APRIL 9th 2020	1.0929	1.0917	1.0858
APRIL 10th 2020	1.0936	1.0943	1.0929

4.2 Long Short-Term Memory

Long short-term memory (LSTM) is a modification of the Recurrent Neural Network (RNN) which with the addition of gates can overcome the problem of vanishing and exploding (Moore and Roche, 2015). The model used is adapted to the pattern of historical data obtained by trying various combinations such as the number of neurons, the number of hidden layers, the number of lookbacks, then number of epochs, activation functions, optimizer, the amount of output (step out), batch size, dropout. In brief, the process of finding a model in LSTM is started with the initialization of the parameters that will be used then load the data that will be used. The data needs to be prepared to make LSTM model consisting of supervised, divide data to train and test, scale. After the data is prepared it will enter the LSTM process which will repeat as many iterations. LSTM process contain of initialization LSTM, fitting and model evaluation. Fitting will repeat as many as the number of epochs. After the model search process, the model is stored then look for the lowest RMSE value. If the RMSE is not yet minimal, another parameter combination will be tried, and the process return to initialization. RMSE value when it is minimal then the model is selected. After model selected is doing prediction. The step in making this prediction is first load the data that will be used for the prediction and the selected model which have the lowest RMSE. Data needs to be prepared before entering the prediction process. Data preparation includes supervised and scale processes. The data that is ready will be used to make predictions. After that the results of the prediction will be inverted so that the value returns to before the scale or not worth between 0 to 1.

Supervised learning is done by shifting the data where it means the data will learn to do the forecasting for example if the value of t is x and the value of t+1 is y then the system will learn if the value of x then the value will be y. This study divides data into 80% of *train* and 20% of *test*. Train data is data used by the system for learning so that it can find the right weight for that data. Test data is used to test models with different data from those already studied.

Table 3. Fixed Parameters

PARAMETER	VALUE
OUTPUT (STEPOUT)	1
BATCH SIZE	1
OPTIMIZER	Adam
DROPOUT	0.0
ACTIVATION FUNCTION	Sigmoid
ITERATION	3

The data is scaled according to the activation function used. This research uses sigmoid, the data will be scaled to 0 to 1. Finding the lowest RMSE by trying a combination of parameters from the LSTM. The fixed parameters in this study can be seen in Table 3 while the parameters which are not fixed and the combination will be tried can be seen in Table 4.

Table 4. Non-fixed Parameters

PARAMETER	VALUE			
NEURON	8	10	15	27
EPOCH	2000	4000	6000	
HIDDEN LAYER	1	2		
LOOKBACK	1	2	3	

The first step of LSTM process is initialization the parameter that will use in LSTM. Afterward, compiling model which input the type of loss and optimizer will used. Type of loss will use mean square error (MSE) and the optimizer will use Adam. After that fitting model as the number of epochs. This process will result loss value which indicate the performance of the weight that machine get it. After that evaluating the model. RMSE value will result in evaluation model process. This process will loop as many as iteration. The combination of these parameters obtained the lowest RMSE value on the combination of epoch 4000, neuron 10, hidden layer 1 and lookback 1 with an RMSE test value of 0.0042 . This combination will be used to make forecasting.

4.3 LSTM and ARIMA Analysis

Analysis is done by comparing the predicted results of the two methods by looking at the RMSE value and the predicted value. This analysis is done using the same data and time period. ARIMA have 2 types of window size which 100 data and 80% of all data.

Table 5. RMSE Values of ARIMA and LSTM

DATA PERIOD	METHOD		
	ARIMA		LSTM
	100	80%	
JAN 2014 - MARCH 2020	0.0045	0.0044	0.0042
APRIL 2020 PREDICTION	0.0053	0.0055	0.0051

The RMSE of the ARIMA and LSTM methods can be seen in Table 5. The LSTM has the smallest RMSE value. This is because in LSTM there is a learning process in the past data so that it can produce better predictive results than ARIMA. The results of forecasting the ARIMA and LSTM models in the data from 2014 to March 2020 for the period 2 January 2019 to 15 January 2019 can be seen in Table 6.

Table 6. Prediction Result Using Historical Data until March 2020

DATE	ACTUAL	ARIMA PREDICTION	LSTM PREDICTION
JANUARY 2nd 2019	1.1343	1.1464	1.1458
JANUARY 3rd 2019	1.1394	1.1343	1.1348
JANUARY 4th 2019	1.1395	1.1394	1.1389

DATE	ACTUAL	ARIMA PREDICTION	LSTM PREDICTION
JANUARY 7th 2019	1.1476	1.1395	1.1397
JANUARY 8th 2019	1.1442	1.1476	1.1467
JANUARY 9th 2019	1.1543	1.1442	1.1440
JANUARY 10th 2019	1.1500	1.1543	1.1523
JANUARY 11th 2019	1.1469	1.1500	1.1491
JANUARY 14th 2019	1.1477	1.1469	1.1458
JANUARY 15th 2019	1.1417	1.1477	1.1465

Forecasting results from ARIMA are the same as the previous day while the LSTM is not the same. This is because ARIMA is not suitable when used on data with a long period of time which will produce an ARIMA (0,1,0) so that it is the same as the previous day's data. LSTM results do not follow previous data, which means that LSTM is better for a fairly long period of time. This is one of the advantages of LSTM. Another advantage of LSTM is LSTM able to learn the pattern of the data supplied, while the ARIMA trends and patterns in data is ignored. This is caused by data needs to fulfill the assumption of stationarity on ARIMA which causes loss of trends and patterns in data. LSTM can also look for maximum results (the lowest error value) while ARIMA can not find the maximum model but LSTM requires a long time for the process of finding a model compared to ARIMA.

Table 7. Prediction Result for April

DATE	ACTUAL	ARIMA PREDICTION	LSTM PREDICTION
APRIL 1st 2020	1.0964	1.0996	1.0964
APRIL 2nd 2020	1.0858	1.0960	1.0858
APRIL 3rd 2020	1.0809	1.0842	1.0809
APRIL 6th 2020	1.0793	1.0788	1.0793
APRIL 7th 2020	1.0891	1.0806	1.0891
APRIL 8th 2020	1.0858	1.0931	1.0858
APRIL 9th 2020	1.0929	1.0917	1.0929
APRIL 10th 2020	1.0936	1.0943	1.0936
APRIL 13th 2020	1.0915	1.0980	1.0915
APRIL 14th 2020	1.0981	1.0915	1.0981

Table 7 shows the results of forecasting in April 2020 for the period April 1st 2020 to April 14th 2020. LSTM can make predictions without having to fitting the model again where the process of fitting the model need time quite a long while in the ARIMA need to look for a model again. The time needed to find predictions using LSTM is 2 seconds while ARIMA takes 1 minute 4 seconds in forecasting this April where the data used totals 122 data. The specification of the computer being used have Intel Core i7-6500U CPU 2.50 GHz and the memory RAM of 8 GB. Forecasting time depends on the amount of data entered and the speed of the computer being used. The period of use of the selected LSTM model depends on the RMSE value generated from the prediction. The limits of this RMSE value are determined by each user (trader).

5. Conclusion

The results obtained by the RMSE values of ARIMA and LSTM while making model in sequence are 0.0044 and 0.0042 while when predicting in April 2020 the RMSE ARIMA and LSTM values are respectively 0.0053 and 0.0051. The RMSE used in ARIMA is the lowest RMSE value between 2 windows. LSTM produces lower RMSE values than ARIMA. LSTM has better prediction result,it is because the LSTM has the ability to learn so that it can utilize when given a growing amount of data while the ARIMA cannot use it. ARIMA does not have learning ability even though given a large amount of data for example 6 years provides the same forecasting results as the previous day compared to using the amount of data that is not too long for example 100 data.

References

Baruník, J., Kočenda, E., & Vácha, L. (2017). Asymmetric volatility connectedness on the forex market. *KIER Working Paper* 956, *Kyoto University Institute of Economic Research.*

Czekalski, P., Niezabitowski, M., & Styblinski, R. (2015). ANN for FOREX forecasting and trading. *Proceedings - 2015 20th International Conference on Control Systems and Computer Science, CSCS 2015*, 322–328. https://doi.org/10.1109/CSCS.2015.51

Diebold, F. X., & Kilian, L. (2000). Unit-root tests are useful for selecting forecasting models. *Journal of Business and Economic Statistics*, 18, 265-273.

Goyal, P., Pandey, S., & Jain, K. (2018). *Deep learning for natural language processing.* New York: Apress.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 26(3).

Konar, A., & Bhattacharya, D. (2017). Time-series prediction and applications. In *Intelligent systems reference library* (Vol. 127, pp. 1868–4394).

Korczak, J., & Hemes, M. (2017). Deep learning for financial time series forecasting in A-Trader system. *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, FedCSIS 2017*, 905–912. https://doi.org/10.15439/2017F449

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature* 521(7553):436-444

Manaswi, N. K. (2018). *Deep learning with applications using python.* New York: Apress.

Michelucci, U. (2018). *Applied deep learning: A case-based approach to understanding deep neural networks.* New York: Apress.

Ming, Y., Cao, S., Zhang, R., Li, Z., Chen, Y., Song, Y., & Qu, H. (2017). Understanding hidden memories of recurrent neural networks. *2017 IEEE Conference on Visual Analytics Science and Technology, VAST 2017*, (October 2017), 17. https://doi.org/10.1109/VAST.2017.8585721

Moore, M. J., & Roche, M. J. (2015). Less of puzzle: A new look at the forward forex market. *Journal of International Economics*, 58(2), 387–411.

Nagpure, A. R. (2019). Prediction of multi-currency exchange rates using deep learning.

International Journal of Innovative Technology and Exploring Engineering, *8*(6), 316–322.

Ni, L., Li, Y., Wang, X., Zhang, J., Yu, J., & Qi, C. (2019). Forecasting of Forex Time Series Data Based on Deep Learning. *Procedia Computer Science*, *147*, 647–652. https://doi.org/10.1016/j.procs.2019.01.189

Nielsen, A. (2019). *Practical time series analysis* (1st ed.). Sebastopol: O'Reilly media Inc.

Ong, E. (2019). *Technical analysis for mega profit*. Jakarta: PT Gramedia pustaka utama.

Palma, W. (2016). *Time series analysis*. New Jersey: John Wiley & Sons, Inc.

Reddy SK, B. A. (2015). Exchange rate forecasting using ARIMA, neural network and fuzzy neuron. *Journal of Stock & Forex Trading*, *04*(03). https://doi.org/10.4172/2168-9458.1000155

Rigters, G. (2019). *Forex trading for beginners & dummies*. Retrieved from https://play.google.com/books/reader?id=qFOxDwAAQBAJ&hl=id&pg=GBS.PP4.w.1.0.39

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing Journal*, *90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Song, X., Liu, Y., Xue, L., Wang, J., Zhang, J., Wang, J., … Cheng, Z. (2020). Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. *Journal of Petroleum Science and Engineering*, *186*(July 2019), 106682. https://doi.org/10.1016/j.petrol.2019.106682

Wang, X., Smith, K., & Hyndman, R. (2006). Characteristic-based clustering for time series data. *Data Mining and Knowledge Discovery*, 13(3), 335-364.

Yang, Q., Wang, J., Ma, H., & Wang, X. (2020). Research on COVID-19 Based on ARIMA ModelΔ—Taking Hubei, China as an example to see the epidemic in Italy. *Journal of Infection and Public Health*, 4–7. https://doi.org/10.1016/j.jiph.2020.06.019

Yong, Y. L., Lee, Y., Gu, X., Angelov, P. P., Ngo, D. C. L., & Shafipour, E. (2018). Foreign currency exchange rate prediction using neuro-fuzzy systems. *Procedia Computer Science*, *144*, 232–238. https://doi.org/10.1016/j.procs.2018.10.523

Zheng, Y., Liu, Q., and Chen, E., (2014), Time series classification using multi-channels deep convolutional neural networks. *Web-Age Information Management. Springer International Publising* 2014*:* 201-210