UNIVERSITY OF INFORMATION TECHNOLOGY

FACULTY OF INFORMATION SYSTEMS

Gold prices forecasting: A comparison of various forecasting models

- PHẠM THÀNH ĐẠT - 20521175
- THIỀU HUY HOÀNG - 20521350
- NGUYỄN QUANG VY - 20522181

Gold prices forecasting:

A comparison of various forecasting models

1st Pham Thanh Dat
University of Information Technology
20521175@gm.uit.edu.vn

2nd Thieu Huy Hoang
University of Information Technology
20521175@gm.uit.edu.vn

3rd Nguyen Quang Vy
University of Information Technology
20521175@gm.uit.edu.vn

***Abstract*— The rarity and high value of gold are increasingly being recognized on a global scale. The price of gold fluctuates a lot almost daily and the government, financial institutions, and business leaders pay close attention to it. Gold is one of the best investment options because of its steadily rising price over time. Given that gold is valued as an investment and is used as a raw material in many industrial products, anticipating the price of gold is a significant financial issue for many people. Therefore, predicting the price of gold is crucial for every nation's financial economy. We can only earn a large profit by carefully forecasting the price of gold. It is essential to determine the gold price using a reliable forecasting method. To identify which forecasting model performs best on the available data set and produces superior investment metrics, this study uses six forecasting models: LSTM, Bi-LSTM, ARIMA, PROPHET, Linear Regression, and SVR.**

***Keywords* — Tine-series analysis, forecasting, LSTM, Bi-LSTM, ARIMA, LSTM, SVR, PROPHET**

I. Introduction

In any unstable political and economic situation or crisis, gold is one of the most effective financial currencies for maintaining price and avoiding risk. From ancient times until now, gold has been an important currency for the Vietnamese people in general, so every year the Vietnamese people have a god of fortune, and on that day, everyone buys gold with the hope of luck. Most well-to-do families in Vietnam own some gold, and they are particularly interested in fluctuations in gold prices.

As a result, in this project, our group has chosen to predict the gold price using machine learning and deep learning models in the hope of assisting the people of Vietnam.

Although there are numerous time-series forecasting models available, this project presents an empirical evaluation of seven popular time-series forecasting models for forecasting gold prices. In particular, six forecasting models are: Linear regression, Support vector regression, Autoregressive integrated moving average (ARIMA), Prophet, LSTM and Bi-LSTM CNN model.

Our project will be implementing multiple forecasting models and comparing their performance using error measures namely MAPE and RMSE for each model to determine which one is best for estimating the price.

II. Related Research

Several time-series prediction methods have been proposed over the years. Our research will compare the accuracy of six gold price prediction models: LSTM, Bi-LSTM, SVR, Linear, and ARIMA. In terms of work, Pandey et al. (2019) **[1]** used linear regression and random forest to analyze and examine the patterns of previous close gold prices, and the results showed that the models were efficient and produced better results. Another case in point is Vidya and Hari (2020) **[2],** who forecast gold prices using long-short-term memory networks (LSTM), which outperformed traditional forecasting models. Aside from that, Deng, Yu-Feng, Xing Jin, and Yi-Xin Zhong (2005) **[3]** used Ensemble SVR for time series prediction. In addition, Yang and Xiaohui **[4]** used ARIMA to forecast gold prices. Finally, it appears that ARIMA (3, 1, 2) is the best model for predicting gold prices. Others, such as Yurtsever, M. (2021) **[5]** forecast gold prices using LSTM, Bi-LSTM, and GRU. Kishanna, Haari, L. RamaParvathyb, and Chennai are the last but not least. SIMATS (2022) **[6]** employs FBProphet and linear regression to forecast gold prices. After all, the FBProphet model is said to be superior to the linear model.

III. Materials

1) *Data representation*

The data set was gathered on Kaggle Web between January 2, 2012, and December 30, 2022 **[7].** It has 2870 rows and 2 attribute columns in total. One column describes the timeline of data representation, while the others show the gold price in VND over time.

Date	US dollar (	Euro (EUR	Japanese y	Canadian c	Chinese re	Indonesiar	Thai baht T	Vietnames	Korean wc	Russian ru	South Afric	Australian
1/2/2012	1531	1179.37	117795.1	1558.94	9636.11	13882343	48303.03	32202289	1763712	49180.29	12360.37	1493.37
1/3/2012	1598	1224.24	122646.5	1612.54	10057.81	14597730	50416.88	33607538	1838899	50657.38	12838.89	1539.57
1/4/2012	1613	1249.52	123781.6	1636.31	10153.19	14750885	50753.04	33923003	1853014	51401.13	13175.06	1561.25
1/5/2012	1599	1249.9	123298.9	1632.98	10076.42	14590875	50616.34	33628569	1843167	51220.92	13082.94	1560.61
1/6/2012	1616.5	1271.43	124656.4	1655.7	10199.31	14702068	51121.81	34000653	1879747	51664.95	13167.44	1580.93
1/9/2012	1615	1267.91	124112.7	1657.88	10198.08	14777250	51292.39	33970718	1879295	51549.66	13180.01	1579.39
1/10/2012	1637	1281.11	125713.4	1667.53	10337.65	14986735	51835.59	34406466	1893518	51672.71	13286.38	1584.78
1/11/2012	1634.5	1288.38	125725.7	1668.74	10322.68	14972020	51895.38	34374352	1893977	51926.9	13289.14	1589.29
1/12/2012	1661	1297.96	127490	1695.22	10493.87	15214760	52869.61	34939135	1923604	52517.82	13411.33	1611.06
1/13/2012	1635.5	1291.2	125900.8	1675.81	10314.44	14850340	52017.07	34400289	1877881	52203.76	13339.46	1590.26
1/16/2012	1641	1294.93	125881.1	1670.78	10365.38	14998740	52339.68	34490538	1895026	51984.89	13258.38	1588.58
1/17/2012	1656	1300.15	127197.4	1679.51	10457.64	15044760	52619.39	34805808	1897030	52378.1	13358.12	1593
1/18/2012	1647	1285.06	126489.6	1669.81	10395.86	14901233	52308.71	34471710	1880462	51908.17	13203.92	1585.87
1/19/2012	1655	1283.69	127550.8	1669.65	10454.14	14895000	52422.1	34589500	1881900	51863.55	13131.43	1588.83

Figure 1. Data representation

IV. METHODS

1) *Performance measure*

To assess the predictive power of our proposed models, we use two performance measures: the root means square error (RMSE) and the MAPE. When we train models, we use RMSE as a loss function, and MAPE is a statistical measure of prediction accuracy. The following are the equations:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_{1,i} - x_{2,i}\right)^2} \quad [8]$$

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{x_{2,i} - x_{1,i}}{x_{1,i}}\right| \quad [9]$$

Where N is the number of data points, $x_{1,i}$ is a predicted value and $x_{2,i}$ is a real value.

A. *Linear Regression*

Linear Regression is an algorithm of machine learning, based on supervised learning. This method is frequently used to forecast and determine the relationship between variables that cause and effect. The number of independent variables and the type of relationship between the independent and dependent variables are the primary differences between regression methods. The regression procedure enables you to confidently establish which elements are most important, which can be ignored, and how those factors interact with one another.

In order to do a regression analysis, we must first define a dependent variable that is determined by one or more independent variables that you hypothesize. Otherwise, we'd need a big dataset to work with.

Figure 2.Result of Linear Regression model

The following table are the experimental results of three data division methods for training and testing:

Model	Train-Test	RMSE	MAPE
LR	7-3	13634173.54	32.31%
	8-2	8380420.58	19.57%
	9-1	4456279.83	9.33%

Table 1. Result of three data division methods

Conclusion: After measuring the Linear Regression model, the model with 90% data for training and 10% data for testing is the most optimal in all three cases. We can see that the results are: RMSE (4456279.83) and MAPE (9.33%).

B. *Support vector regression*

In the fields of machine learning and data mining, Support Vector Machines have received a lot of attention. Support Vector Machines were created in 1992 by Cortes and Vapnik **[10]** for the categorization of supervised learning frameworks. The SVMs have been extended for regression and rank learning **[11], [12].** While first utilized as a binary kernel technique, these two functions may be successfully substituted by kernel functions that transfer the input data into a higher dimension space whose linear regression is equivalent to the non-linear regression in the original space. The transformation of the function from low dimension to high dimension requires no knowledge of ϕ, making the transition easier. The kernel function used in this study is the Sigmoid. Some of the common kernels are as below:

Kernel	Function
Polynomial	$[(x * x_i) + 1]^d$
RBF	$Exp\{-\gamma\lvert x - x_i\rvert^2\}$
Sigmoid	$\tanh(\gamma x^T z + r)$

Table 2. Kernels used in the Support Vector Machine

Figure 3. Result of SVR model

The following table are the experimental results of three data division methods for training and testing:

Model	Train-Test	RMSE	MAPE
SVR	7-3	3730356.55	7.69%
	8-2	3011772.56	5.80%
	9-1	1694564.28	3.42%

Table 3. Result of three data division methods

Conclusion: After measuring the SVR model, the model with 90% data for training and 10% data for testing is the most optimal in all three cases. We can see that the results are: RMSE (1694564.28) and MAPE (3.42%).

C. LSTM (Long short-term memory.)

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network that can learn order dependence in sequence prediction problems. This is a necessary characteristic in complex problem domains such as machine translation, speech recognition, and others. **[13]**

An LSTM layer is made up of a collection of recurrently connected memory blocks. Each block contains one or more recurrently connected memory cells through three multiplicative units - the input, output, and forget gates. These provide continuous analogs of the cells' write, read, and reset operations.

The advent of LSTM networks minimizes the drawback of gradient vanishing in part by allowing information to propagate more directly through the cell state.

LSTM cell:

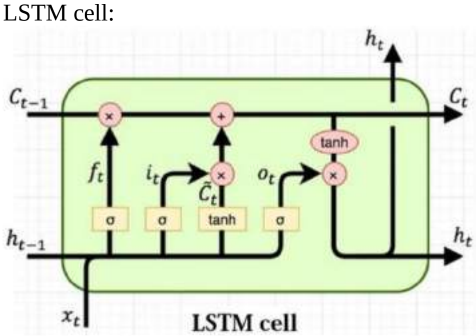

Figure 4. LSTM Architechture Flow Diagram **[14]**

Calculate in LSTM cell:

$$\textit{Forget gate}: f_t = \sigma(\mathrm{W}_f x_t + U_f h_{t-1})$$

$$\textit{Input gate}: i_t = \sigma(\mathrm{W}_i x_t + U_i h_{t-1})$$

$$\textit{Cell gate}: c_t = \tanh(\mathrm{W}_c x_t + U_c h_{t-1}) \quad \textbf{[15]}$$

$$\textit{Output gate}: o_t = \phi h(\mathrm{W}_o x_t + U_o h_{t-1})$$

$$\textit{Cell state}: c_t = f_t \times c_{t-1} + i_t \times c_t$$

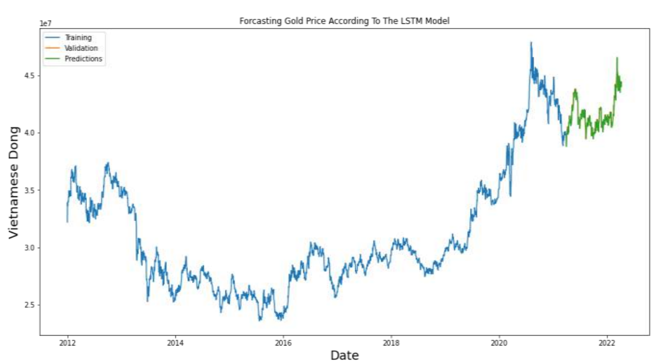

Figure 5. Result of LSTM model

The following table are the experimental results of three data division methods for training and testing:

Model	Train-Test	RMSE	MAPE
LSTM	7-3	497248	0.95%
	8-2	456035	0.78%
	9-1	391038	0.68%

Table 4. Result of three data division methods

Conclusion: After measuring the LSTM model, the model with 90% data for training and 10% data for testing is the most optimal in all three cases. We see that the results are very good, with very low RMSE (391038) and MAPE (0.68%), because there has been no strong fluctuation or sudden change over the years due to the characteristics of the gold data set.

D. Bi-LSTM

Bidirectional long-short term memory (Bi-LSTM) is a technique that allows any neural network to store sequence information both forward and backward. Bi-LSTM allows input flow in both directions, whereas normal LSTM only allows input flow in one direction. **[16]**

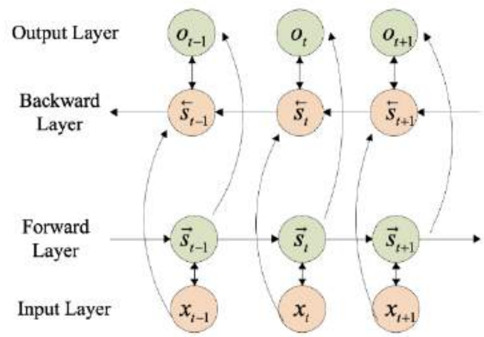

Figure 6. the basic structure of bidirectioal LSTM **[17]**

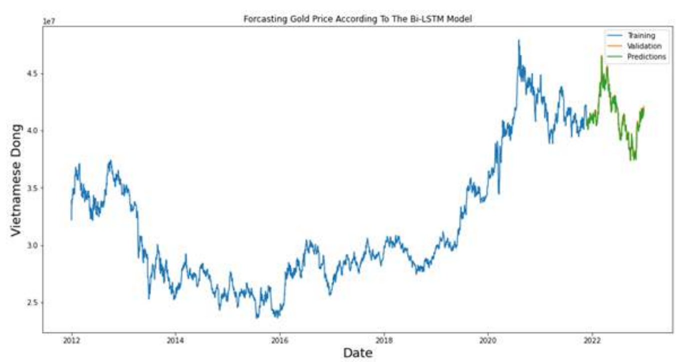

Figure 7. Result of Bi-LSTM Model

Model	Train-Test	RMSE	MAPE
Bi-LSTM	7-3	493253	0.96%
	8-2	511867	0.97%
	9-1	397070	0.73%

Table 5. Result of three data division methods

Conclusion: After measuring the Bi-LSTM model, the model with 90% data for training and 10% data for testing is the most optimal in all three cases. We see that the results are very good, with very low RMSE (397070) and MAPE (0.73%), because there has been no strong fluctuation or sudden change over the years due to the characteristics of the gold data set.

E. ARIMA

An autoregressive integrated moving average, or ARIMA, is a statistical analysis model that uses time series data to either better understand the data set or to predict future trends.

A statistical model is autoregressive if it predicts future values based on past values. For example, an ARIMA model might seek to predict a gold future price based on its past performance or forecast a company's earnings based on past periods. **[18]**

The ARIMA model stands for Auto Regression (AR), Moving Average (MA) and Differential Integration Integrated - I.

Important point: The ARIMA model is not a perfect predictive model for any time series data.

The ARIMA model only works best if the data is highly time dependent. Randomized data usually do not work for ARIMA models.

The ARIMA model is only good at predicting time points.

- **Types of Models ARIMA:**

The ARIMA model is not seasonal

Seasonal ARIMA model (Seasonal ARIMA – SARIMA)

- **Stationary**

A stationary time series is a series of mean, variance, and autocorrelation values that do not change over time and it does not include the trend factor. With most statistical predictive methods, the calculation must be ensured. stationarity of the data series, so checking for stationarity is very important. To test the stationarity of data, we have two popular testing methods: Dickey (DF) test and Improved Dickey Fuller (ADF4). **[19]**

- **ARIMA (*p, d, q*)**

The parameter p is the number of autoregressive terms or the number of "lag observations." It is also called the "lag order," and it determines the outcome of the model by providing lagged data points.

The parameter d is known as the degree of differencing. it indicates the number of times the lagged indicators have been subtracted to make the data stationary.

The parameter q is the number of forecast errors in the model and is also referred to as the size of the moving average window.

The parameters take the value of integers and must be defined for the model to work. They can also take a value of 0, implying that they will not be used in the model. In such a way, the ARIMA model can be turned into:

ARMA model (no stationary data, ***d* = 0**)

AR model (no moving averages or stationary data, just an autoregression on past values, ***d* = 0, *q* = 0**)

MA model (a moving average model with no autoregression or stationary data, ***p* = 0, *d* = 0)**

Therefore, ARIMA models may be defined as:

1. ARIMA (1, 0, 0) – known as the **first-order autoregressive model**
2. ARIMA (0, 1, 0) – known as the **random walk model**
3. ARIMA (1, 1, 0) – known as the **differenced first-order autoregressive model**, and so on.

Once the parameters (*p, d, q*) have been defined, the ARIMA model aims to estimate the coefficients **α** and **θ**, which is the result of using previous data points to forecast values. **[20]**

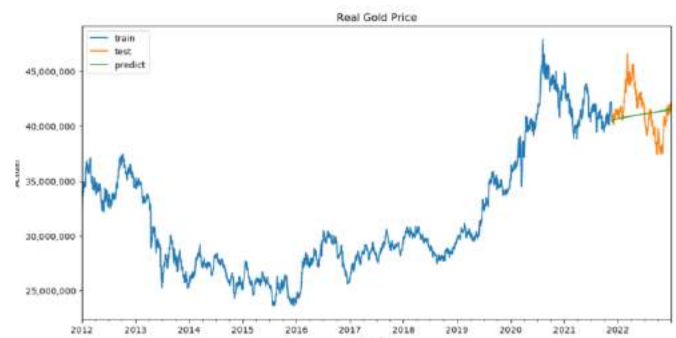

Figure 8. Result of ARIMA

Table 7. Result of three data division methods

model

The following table are the experimental results of three data division methods for training and testing:

Model	Train-Test	RMSE	MAPE
ARIMA	7-3	5578826.49	11.85%
	8-2	4697358.81	10.32%
	9-1	2074521.9	3.94%

Table 6. Result of three data division methods

Conclusion: After measuring the ARIMA model, the model with 90% data for training and 10% data for testing is the most optimal in all three cases. We can see that the results are: RMSE (2074521.9) and MAPE (3.94%).

F. Prophet

Prophet is a free, open-source application developed by Facebook for forecasting time series data, which aids in understanding and potential market forecasts for organizations. It is based on a decomposable additive model, which also accounts for the effects of vacations, and fits non-linear trends with seasonality.

Trend:

The trend shows the tendency of the data to increase or decrease over a long period of time and it filters out the seasonal variations.

Seasonality:

Seasonality is the variations that occur over a short period of time and is not prominent enough to be called a "trend".

Understanding the Prophet Model

The general idea of the model is similar to a generalized additive model. The "Prophet Equation" fits, as mentioned above, trend, seasonality and holidays. This is given by,

$$\mathbf{y(t) = g(t) + s(t) + h(t) + e(t)}$$

where:

g(t) refers to trend (changes over a long period of time)

s(t) refers to seasonality (periodic or short-term changes)

h(t) refers to effects of holidays to the forecast

e(t) refers to the unconditional changes that is specific to a business or a person or a circumstance. It is also called the error term.

y(t) is the forecast.

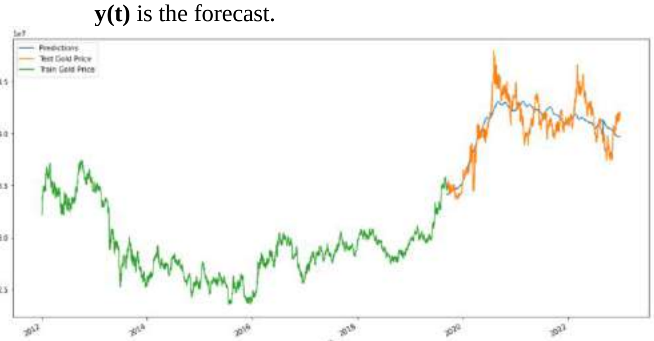

Figure 9. Result of PROPHET model

Model	Train-Test	RMSE	MAPE
PROPHET	7-3	1573796.94	2.99%
	8-2	1664335.79	3.28%
	9-1	1753647.79	3.39%

Table 7. Result of three data division methods

Conclusion: After measuring the PROPHET model, the model with 70% data for training and 30% data for testing is the most optimal in all three cases. We can see that the results are: RMSE (1573796.94) and MAPE (2.99%).

V. CONCLUSION

The following table are the experimental results of six models for training and testing we got it:

Model	RMSE	MAPE
LR	4456279.83	9.33%
SVR	1694564.28	3.42%
ARIMA	2074521.9	3.94%
PROPHET	1573796.94	2.99%
LSTM	404883	0.74%
BI-LSTM	397070	0.73%

Table 8. Measuring the six models according to the past values

After analysis of six models of time series forecasting and its corresponding RMSE and MAPE. This can be easily verified that the Recurrent Neural Network (RNN) based Bidirectional - Long Short-Term Memory (Bi - LSTM) Model-design gives the best forecast on test data with RMSE error of 397070 and MAPE Error of 0.73%.

ACKNOWLEDGMENT

First and foremost, I would like to express my deepest gratitude to ***Assoc. Prof. Dr. Nguyễn Đình Thuân*** and ***Mr. Nguyễn Minh Nhựt*** who always cared for our team during this course. I appreciate your assistance and the sharing of your valuable experience, which will help us accomplish our project. I would sincerely like to thank ***Mr. Nhựt***, who teaches and supports our team throughout our course.

I would like to take this opportunity to thank ***Assoc. Prof. Dr. Nguyễn Đình Thuân*** for permitting us to carry out our project. Finally, I would like to express how honored when our team was able to learn from and attend your class.

Many thanks to ***Assoc. Prof. Dr. Nguyễn Đình Thuân*** and ***Mr. Nguyễn Minh Nhựt*** for your tireless efforts in guiding the team to success and encouraging our team to keep moving forward. My heartfelt gratitude goes to all of my classmates, especially my friends, for devoting their time to assisting and supporting our team in the fabrication of our project.

REFERENCES

[1] Pandey, A.C., Misra, S. and Saxena, M., (2019) Gold and Diamond Price Prediction Using Enhanced Ensemble Learning. 2019 12th International Conference on Contemporary Computing, IC3 2019, pp.2019–2022.

[2] Vidya, G.S. and Hari, V.S., (2020) Gold Price Prediction and Modelling using Deep Learning Techniques. 2020 IEEE Recent Advances in Intelligent Computational Systems, RAICS 2020, pp.28–31.

[3] Deng, Yu-Feng, Xing Jin, and Yi-Xin Zhong. "Ensemble SVR for prediction of time series." In 2005 International Conference on Machine Learning and Cybernetics, vol. 6, pp. 3528-3534. IEEE, 2005.

[4] Yang, X., 2019, January. The prediction of gold price using ARIMA model. In 2nd International Conference on Social Science, Public Health and Education (SSPHE 2018) (pp. 273-276). Atlantis Press.

[5] YURTSEVER, M., 2021. Gold price forecasting using LSTM, Bi-LSTM and GRU. Avrupa Bilim ve Teknoloji Dergisi, (31), pp.341-347.

[6] Kishanna, H., RamaParvathyb, L. and SIMATS, C., 2022. A Novel Approach for Correlation Analysis on FBProphet to Forecast Market Gold Rates with Linear Regression.

[7] JISHNU, Global Gold price - Historical Data (1979-Present)

[8] Chai, T. and Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE). Geoscientific Model Development Discussions, 7(1), pp.1525-1534.

[9] Chicco, D., Warrens, M.J. and Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, 7, p.e623.

[10] C. Cortes and V. Vapnik, Machine Learning, vol. 20, no. 3, pp. 273-297, 1995.

[11] Herbrich, R., Graepel, T., Obermayer, K. (eds.): Large margin rank boundaries for ordinal regression. MIT-Press. 2000.

[12] H. Yu, 'Selective sampling techniques for feedback-based data retrieval', Data Mining and Knowledge Discovery, vol. 22,no. 1-2, pp

[13] Graves, A. and Schmidhuber, J., 2005. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural networks, 18(5-6), pp.602-610.

[14] Varsamopoulos, Savvas & Bertels, Koen & Almudever, Carmen. (2018). Designing neural network based decoders for surface codes.

[15] Cai, Changchun, Yuan Tao, Tianqi Zhu, and Zhixiang Deng. 2021. "Short-Term Load Forecasting Based on Deep Learning Bidirectional LSTM Neural Network" Applied Sciences 11, no. 17: 8129. https://doi.org/10.3390/app11178129

[16] Towardsai - Univariate Time Series With Stacked LSTM, BiLSTM, and NeuralProphet

[17] Cai C, Tao Y, Zhu T, Deng Z. Short-Term Load Forecasting Based on Deep Learning Bidirectional LSTM Neural Network. Applied Sciences. 2021; 11(17):8129. https://doi.org/10.3390/app11178129

[18] ADAM HAYES,Autoregressive Integrated Moving Average,December 18, 2022.

[19] Nguyễn Minh Nhựt – LAB4_Phan Tich Du Lieu Chuoi Thoi Gian.

[20] CFI Team, Autoregressive Integrated Moving Average (ARIMA), December 28, 2022.

[21] geeksforgeeks, Time Series Analysis using Facebook Prophet, 30 Jun, 2020.