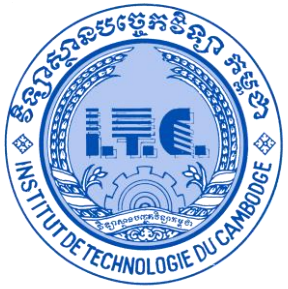
Institute of Technology of Cambodia

Graduate School of ITC

Master of Data Science



Prediction on Ribbed Smoked Sheet No.3 (RSS3)'s Price Using Time Series Forecasting: ARIMA and LSTM Model

Presented by: Han Chandeth

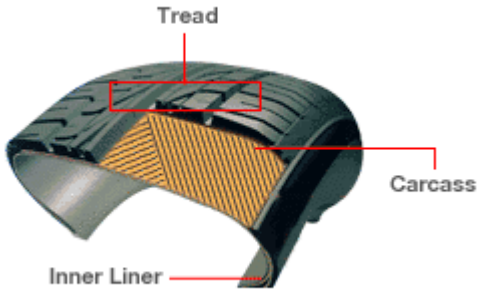
CONTENT

- INTRODUCTION
- LITERATURE REVIEW
- DATA COLLECTION
- METHODOGY
- EXPECTED RESULT
- CONCLUSION
- REFERENCES

INTRODUCTION

Ribbed Smoke Sheet No. 3 (RSS3) is a significant grade of natural rubber that holds immense importance in the rubber industry. RSS3 rubber sheets are used in the production of tires, Tread Carcass, Off Road Tires & ADV tires, Extruded hoses, and Footwear etc. The price fluctuations of RSS3 significantly impact rubber producers, traders, and manufacturers, making accurate price forecasting crucial for effective decision-making, risk management, and inventory planning.

Time series analysis techniques have emerged as powerful tools for predicting commodity prices, and this research aims to explore the application of two widely used techniques, Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM), in predicting the price of Ribbed Smoke Sheet No. 3 natural rubber.



INTRODUCTION (Cont.)

Research Objectives

- Develop two time series forecasting models for ribbed smoked sheet natural rubber's price: an ARIMA model and an LSTM model.
- Evaluate the performance of the two models using different evaluation metrics (RMSE, MAE, MAPE).
- Compare the performance of the two models and identify the best model for predicting RSS3's price.
- Using the best model to forecast the price of RSS3 in the next 10 months.

LITERATURE REVIEW

Research Title and Results	Data	Articles
According to the empirical investigations done and described in this article "A Comparison of ARIMA and LSTM in Forecasting Time Series", deep learning-based algorithms such as **LSTM outperform traditional-based algorithms such as the ARIMA model**. • The data related to the financial time series or stock market :Rooted Mean Squared Error (RMSE) using Rolling ARIMA and Rolling LSTM models are 511.481 and 64.213, respectively, • The economic related data: average RMSE values for Rolling ARIMA and Rolling LSTM are computed as 5.999 and 0.936,	• monthly financial time series from Jan 1985 to Aug 2018 from the Yahoo finance Website • monthly economics time series for different time periods from the Federal Reserve Bank of St. Louis, and the International Monetary Fund (IMF) Website	Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2018)
In this study "Possible Method for Monthly Natural Rubber Price Forecasting", the authors found that **ARIMA is the possible methods for world rubber price forecasting while Double Exponential Smoothing should be applied for predicting domestic rubber prices since it allows for better predictive performance.** • For world rubber price, ARIMA generates 5.31 % forecasting error, the lowest value MAPE. • For domestic rubber prices, the best forecasting model is the Double Exponential Smoothing Model with the lowest value of MAPE: 3.00 and MSD: 142.285.	• monthly natural rubber prices in Indonesia and World Markets (from 2012:1 – 2016:12)	Sukiyono, K., Nusril, ., Sumartono, E., Cahyadinata, I., Yuliarso, M., Mulyasari, G., Nabiu, M., & Annisa, F. (2020).

LITERATURE REVIEW

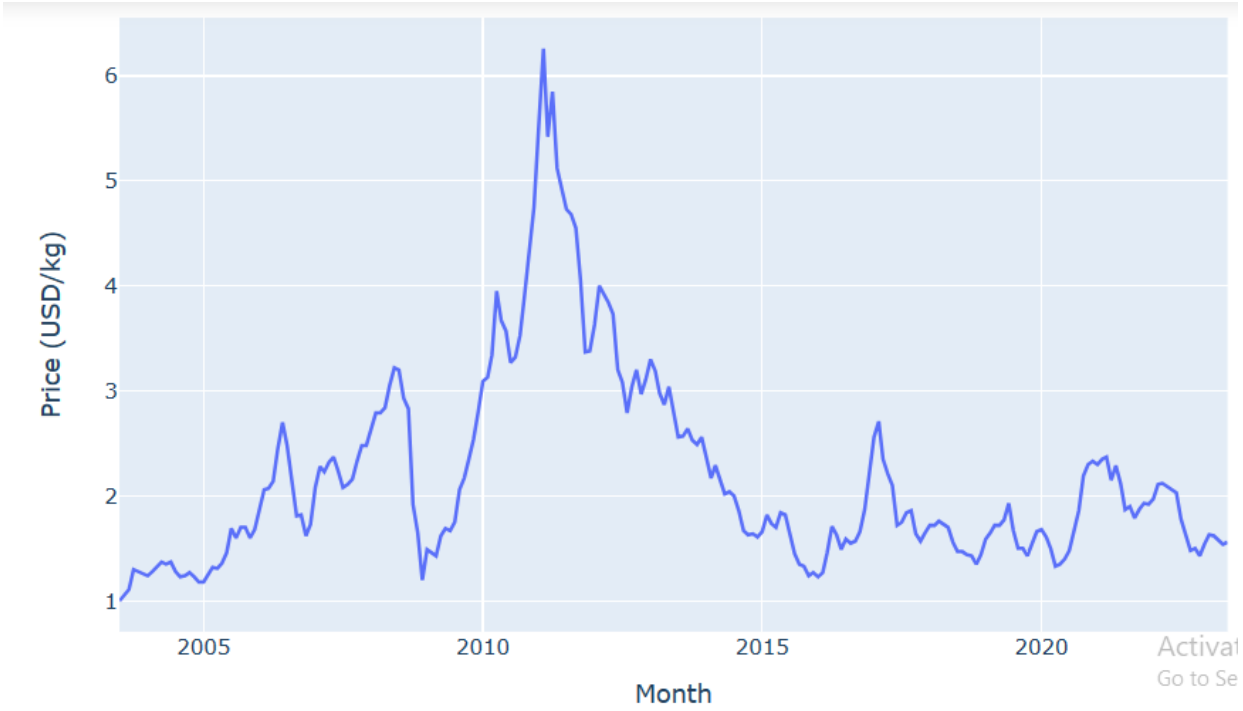
Research Title and Results	Data	Articles
In “Forecasting smoked rubber sheets price based on a deep learning model with long short-term memory”, the researchers found that a **9-layer LSTM model with 3 primary LSTMs** is the best forecasting model for the price of rubber smoked sheets (RSS3). It had a root mean square error (**RMSE) of 2.4121, a mean absolute percentage error (MAPE) of 0.0413, and an accuracy rate of 95.88%.**	• 2,631 data points from the Rubber Authority of Thailand for the past 10 years.	Phoksawat, K., Phoksawat, E., & Chanakot, B. (2023)

DATA COLLECTION

- Data: Monthly Price (US Dollars per Kilogram), No. 3 Smoked Sheet (RSS3)
- Date: Jul.2003 - May.2023 (239 data points).
- Source: Singapore Commodity Exchange (SICOM)

Monthly RSS3 price from 2013-2023

Months	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Jan		1.24	1.18	1.88	2.08	2.62	1.49	3.09	5.52	3.63	3.3	2.34	1.66	1.23	2.56	1.72	1.59	1.68	2.3	1.97	1.63
Feb		1.28	1.26	2.06	2.28	2.79	1.46	3.13	6.26	4	3.19	2.17	1.82	1.27	2.71	1.72	1.65	1.61	2.35	2.11	1.62
Mar		1.33	1.32	2.07	2.23	2.79	1.43	3.34	5.42	3.93	2.98	2.29	1.74	1.46	2.35	1.76	1.72	1.5	2.37	2.12	1.58
Apr		1.37	1.31	2.14	2.32	2.84	1.62	3.95	5.85	3.84	2.87	2.15	1.7	1.71	2.21	1.73	1.72	1.33	2.15	2.09	1.54
May		1.35	1.36	2.45	2.37	3.05	1.69	3.67	5.12	3.73	3.04	2.02	1.84	1.63	2.1	1.7	1.77	1.35	2.29	2.06	1.56
Jun		1.37	1.46	2.7	2.23	3.22	1.67	3.57	4.93	3.2	2.81	2.04	1.82	1.49	1.72	1.56	1.93	1.4	2.12	2.03	
Jul	1	1.28	1.69	2.48	2.08	3.2	1.75	3.27	4.73	3.08	2.56	2	1.64	1.59	1.75	1.47	1.67	1.48	1.87	1.78	
Aug	1.05	1.23	1.6	2.18	2.11	2.93	2.06	3.32	4.68	2.79	2.57	1.86	1.45	1.55	1.84	1.47	1.5	1.7	1.9	1.61	
Sep	1.11	1.24	1.7	1.81	2.16	2.83	2.17	3.53	4.55	3.04	2.64	1.67	1.35	1.57	1.86	1.44	1.5	1.86	1.79	1.48	
Oct	1.3	1.27	1.7	1.82	2.33	1.92	2.35	3.92	4.06	3.2	2.53	1.63	1.33	1.66	1.64	1.43	1.43	2.19	1.87	1.5	
Nov	1.28	1.23	1.6	1.62	2.48	1.65	2.54	4.31	3.37	2.97	2.49	1.64	1.24	1.87	1.57	1.35	1.54	2.3	1.93	1.43	
Dec	1.26	1.18	1.68	1.73	2.48	1.2	2.8	4.75	3.38	3.11	2.56	1.61	1.27	2.23	1.65	1.44	1.66	2.33	1.92	1.54	



METHODOLOGY

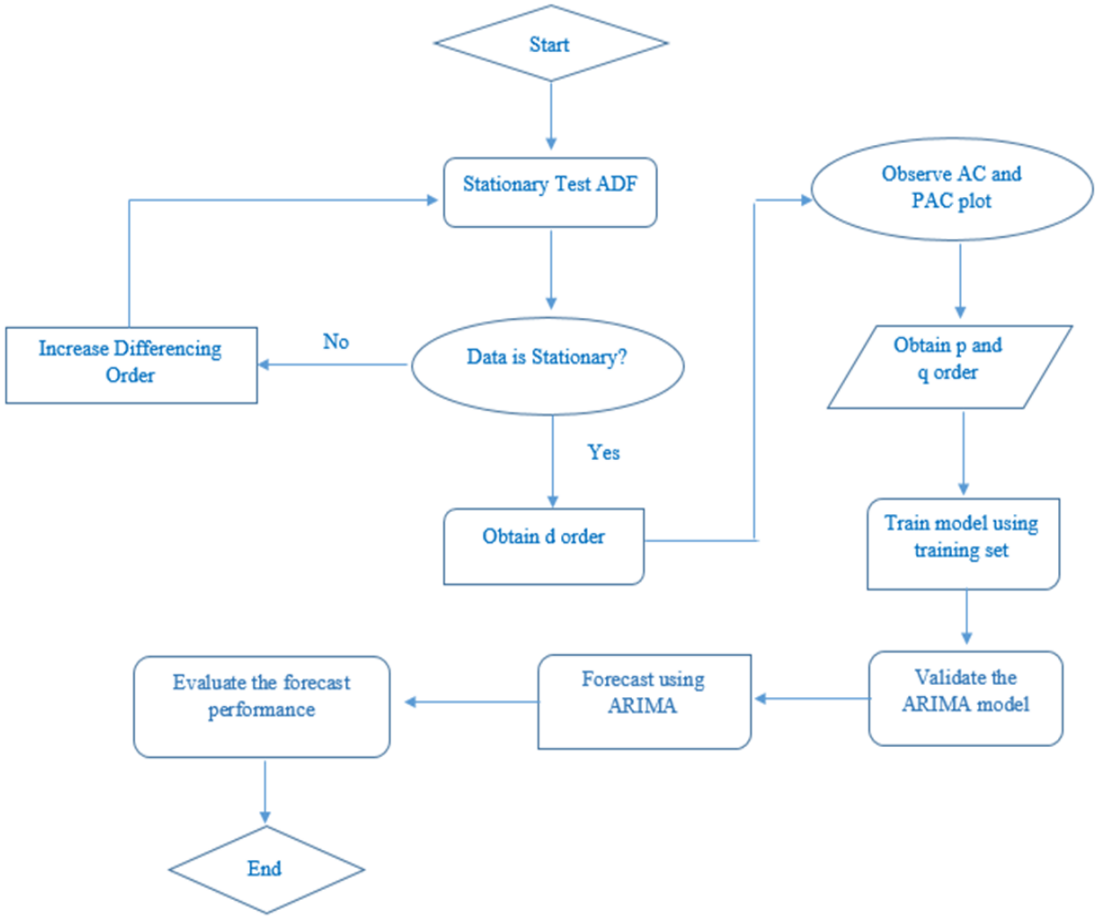
Autoregressive Integrated Moving Average Model (ARIMA)

- Autoregressive Integrated Moving Average Model (ARIMA) is a generalized model of Autoregressive Moving Average (ARMA) that combines Autoregressive (AR) process and Moving Average (MA) processes and builds a composite model of the time series.

$$x_t = c + \sum_{i=1}^{p} \phi_i x_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

As acronym indicates, ARIMA(p, d, q) captures the key elements of the model:

- p==> autoregressive lags
- q== moving average lags
- d==> difference in the order



METHODOLOGY

Augmented Dickey-Fuller (ADF) Test or Unit Root Test

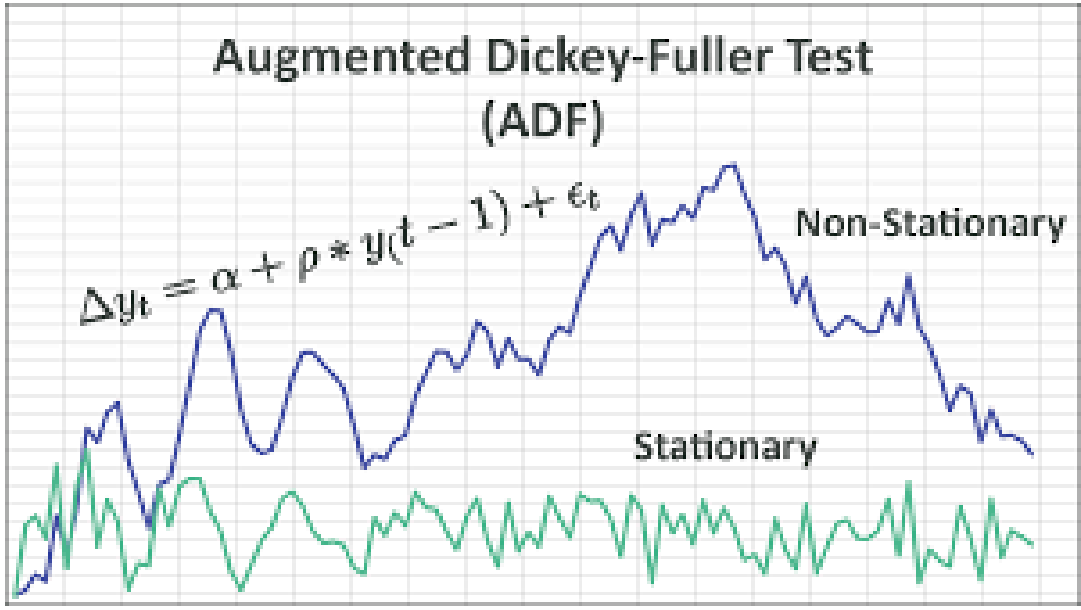
The ADF test is the most popular statistical test for stationary. It is done with the following assumptions:

Null Hypothesis (H0): Series is non-stationary

Alternate Hypothesis (H1): Series is stationary

p-value >0.05 Fail to reject (H0)

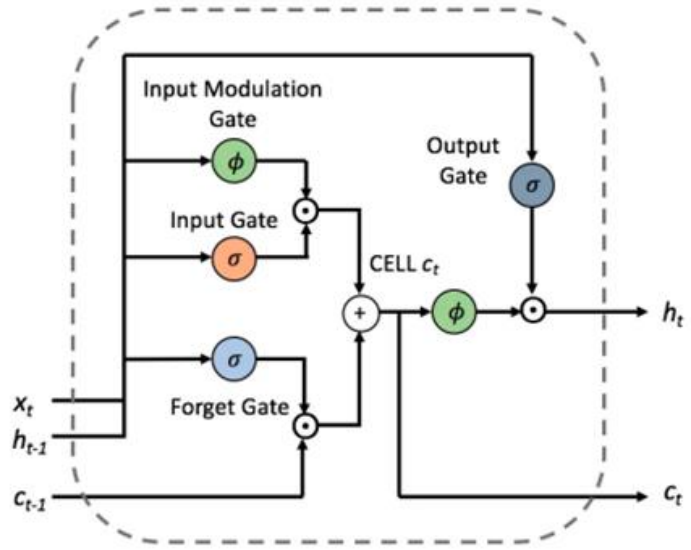
p-value <= 0.05 Accept (H1)

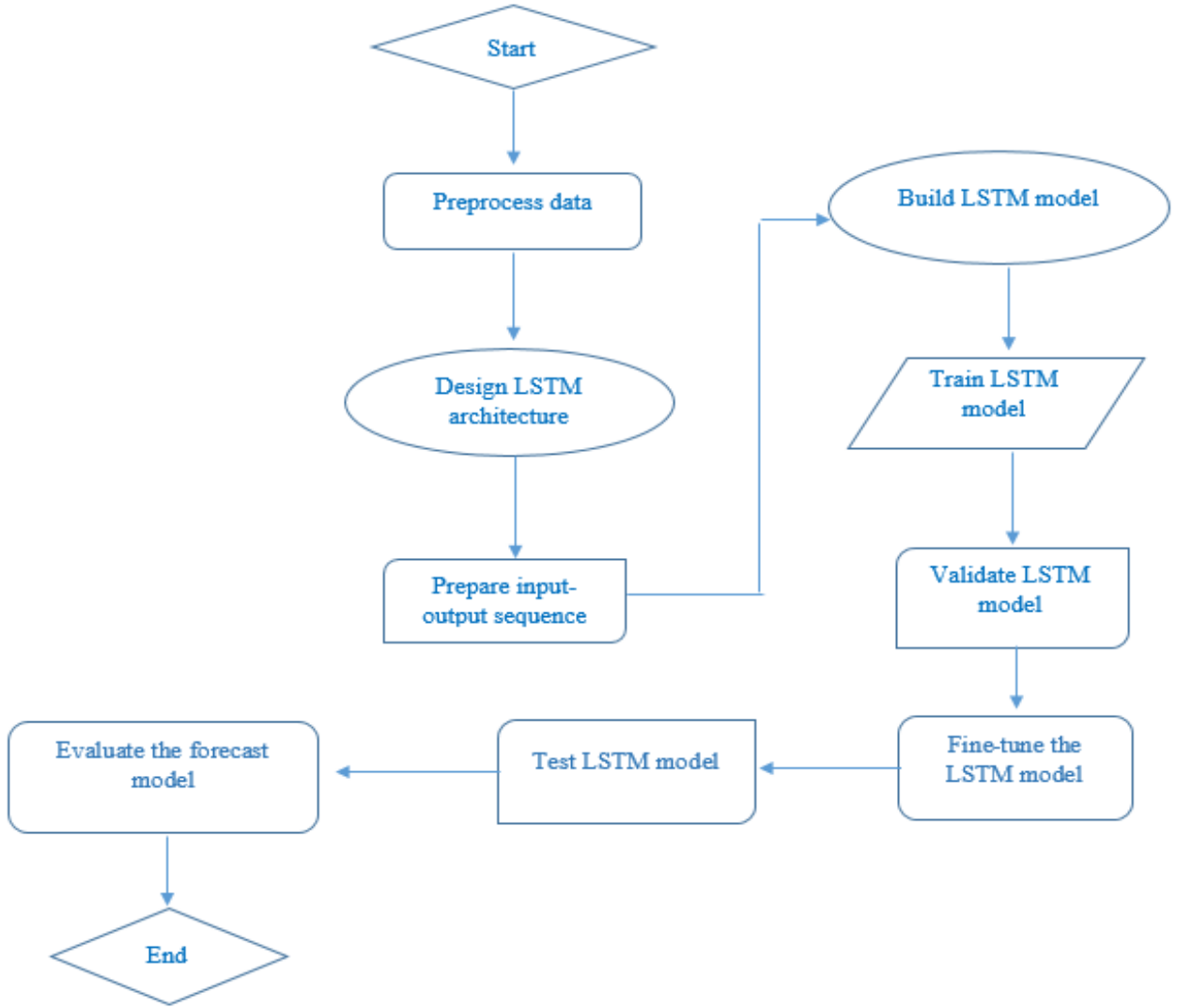


METHODOLOGY (Cont.)

Long-Short Term Memory (LSTM)

Long-Short Term Memory (LSTM) is a type of recurrent neural network (RNN) that can remember information from previous step for the purpose of future uses.





METHODOLOGY (Cont.)

Performance Metric

- **Root-mean-square error (RMSE)** is a measure of how accurate a model's predictions are. It is calculated by taking the square root of the average of the squared errors between the predicted values and the actual values.
- RMSE is a scale-dependent measure, so it is only meaningful to compare RMSE values for models that have been trained on the same dataset.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}$$

- **The Mean Absolute Error (MAE)** is defined as the average of the absolute difference between forecasted and true values.
- The MAE shows us how much inaccuracy we should expect from the forecast on average. The lower the MAE value, the better the model.

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

- **Mean Absolute Percentage Error (MAPE)** is the proportion of the average absolute difference between projected and true values divided by the true value.
- The model is better if the MAPE is low.

$$\mathrm{MAPE} = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

Expected Result

The expected result of this study is to develop a time series model that can accurately forecast the price of RSS3.

The model will be trained on historical monthly data from the Singapore Commodity Exchange (SICOM) for the past 20 years.

The model will be evaluated using the root mean square error (RMSE), mean absolute error (MAE) and the mean absolute percentage error (MAPE).

The model is expected to have a low RMSE, MAE and MAPE, which will indicate that it is accurate. The model is also expected to be able to forecast the price of RSS3 for a reasonable period into the future.

CONCLUSION

The model developed in this study can be used to forecast the price of RSS3 for a reasonable period into the future.

The model developed in this study is not without its limitations.

1. The model is only as good as the data it is trained on. If the data is not representative of the future, the model will not be accurate.
2. The research will primarily focus on the application of ARIMA and LSTM models and may not cover other advanced time series techniques.
3. The findings and recommendations of this research should be interpreted within the context of Ribbed Smoke Sheet No. 3 natural rubber and may not be directly applicable to other rubber grades or commodities.
4. The model is only a tool and should not be used as the sole basis for making decisions. Businesses and consumers should always consider other factors, such as market conditions and their own financial situation, when making decisions.

References

[1] Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2018). A Comparison of ARIMA and LSTM in Forecasting Time Series. *In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 555-559). IEEE. doi: 10.1109/ICMLA.2018.00227.

[2] Sukiyono, K., Nusril, ., Sumartono, E., Cahyadinata, I., Yuliarso, M., Mulyasari, G., Nabiu, M., & Annisa, F. (2020). Possible Method for Monthly Natural Rubber Price Forecasting. In Proceedings of the 2nd International Research Conference on Economics and Business (IRCEB 2018) (pp. 172-179). *SCITEPRESS – Science and Technology Publications, Lda*. DOI: 10.5220/0008785301720179. ISBN: 978-989-758-428-2.

[3] Vijaysinh, L.(2021, November 1). A Guide to Different Evaluation Metrics for Time Series Forecasting Models. *Mystery Vault*. https://analyticsindiamag.com/a-guide-to-different-evaluation-metrics-for-time-series-forecasting-models/

[4] Phoksawat, K., Phoksawat, E., & Chanakot, B. (2023). Forecasting smoked rubber sheets price based on a deep learning model with long short-term memory. *International Journal of Electrical and Computer Engineering*, 13(1), 688-696. doi:10.11591/ijece.v13i1.pp688-696

Thank You for Your Attention!!