Analysis Indonesia's Export Value Forecasting to G20 Countries Using Long Short-Term Memory Neural Network Method

Veronica[1]*, Herlan Silaban[2], Syafrani Putri Nasution[3] Evta Indra[4]
[1,2,3,4] Universitas Prima Indonesia
[1]veronicasihombing@outlook.com, [2]herlansilaban14@gmail.com,
[3]syafraniputri2001@gmail.com , [4]evtaindra@unprimdn.ac.id

Abstract: Export is one of the most important ways for the country to generate income, which can have an impact on the country's economic stability. This research aims to forecast the value of Indonesian exports to G20 member countries. The Long Short-Term Memory method is used in this research to examine historical data on Indonesian exports from the previous 16 years. Experimental results show that the LSTM Neural Network method has the ability to predict the value of Indonesian exports to G20 member countries with a sufficient level of accuracy. The predictions generated by the model provide insight into trends and fluctuations in the value of exports in the future. The results of this study provide insight into the potential application of artificial intelligence techniques in economic and trade analysis. The results demonstrate that the LSTM model is capable of producing relatively accurate predictions, with an average score of Root Mean Square Error (RMSE) on training data is 0.10 and on testing data is 0.13, as well as graphs of prediction results demonstrating that the LSTM model can capture patterns and trends from Indonesia's export data to G20 countries. According to the prediction results, the highest export value to China is expected to be $6,100,000 in the 200th month (or in the year 2039), while the lowest export value to Mexico is expected to be $27,000 in the 135th month (or in the year 2034).

Keywords: Forecasting; Export; G20; LongShort Term Memory; Neural Networks

INTRODUCTION

In Predictions in estimating systematically about something that is most likely to happen in the future based on past and present data or information, with the aim of preventing errors; the difference between something that happens and the estimated result can be minimized. Predictions do not have to give 100% correct results for future events; instead, aim to find or get a result that is as close as possible to what is going to happen. The term prediction has the same meaning as forecast (Husni, et. al, 2022). International trade plays a crucial role in the economic development of a country. In this context, exports have a significant impact on the country's economic growth and financial stability. As one of the countries with the largest economy in Southeast Asia, Indonesia has an important role in the dynamics of global trade. Therefore, understanding trends and patterns in the value of Indonesia's exports to its main trading partners, especially G20 member countries, is essential in economic planning and trade policy efforts.

In recent decades, developments in technology and computing have provided new opportunities in the analysis and prediction of economic data. Methods involving artificial intelligence, such as Neural Networks, have been proven capable of processing complex data and providing accurate prediction results.

*name of corresponding author

One type of Neural Network that has received attention is the Long Short-Term Memory Neural Network. LSTM has the ability to overcome long-term dependencies on sequential data, such as time series data which is generally related to economic trends. However, although developments in predictive analysis using artificial intelligence technology are increasingly rapid, its application in the context of predicting the value of Indonesian exports to G20 member countries is still limited. Therefore, this research aims to analyze and predict the value of Indonesia's exports to G20 member countries using the Long Short-Term Memory (LSTM) Neural Network method. With this research, it is hoped that more in-depth insight will be gained into trends and fluctuations in the value of Indonesia's exports to G20 member countries. Export itself is an activity or activity of sending goods or commodities from one country to another (Hidayah et al., 2019).

This activity is generally carried out by small, medium to large scale companies as a form of strategy to compete in the international market. Export activities also generate foreign exchange for the country of origin, where foreign exchange is the value of wealth owned by a country in the form of foreign currency (foreign currency), which then functions as a means of foreign payment that can be exchanged for foreign money to finance international trade transactions (Khairu Nissa et al., 2020.) On the other hand, according to exports are shipments of merchandise abroad: goods sent abroad, both financially and individually. Research on predicting a price or something has been carried out a lot, some of which are research (Ismail et al., 2009) which uses the Long Short-Term Memory Recurrent Neural Network method in predicting Stock Market Price, research (Fadli & Hidayatullah, 2019) which uses a Regression algorithm to predict gold price movements , and also research (Rizki et al., 2020) which uses Linear Regression, XGB, and LSTM to predict coffee prices in the market. In addition, one more example is the research entitled "Implementation of the Long Short Term Memory Method to Predict Gold Price Movements".

The research was conducted using data taken from Yahoo Finance to perform an analysis of gold price predictions (Owen et al., 2022). During the pandemic and post-pandemic times, all countries in the world, including Indonesia, experienced various kinds of crises, especially the economic crisis. There were even several countries that went bankrupt because their governments were unable to handle or take quick and wise steps in dealing with the economic crisis. With the economic crisis, it is necessary to have several steps or strategies made by the Indonesian government to increase the country's economic growth, and one of them is export. In exports themselves there are problems, where there is uncertainty in the value of Indonesia's exports to foreign countries.

LITERATURE REVIEW

Several Research can be defined as the activity of collecting, processing and summarizing data using certain techniques to find solutions to questions that arise in various aspects of life. Collection, analysis, and drawing conclusions from empirical data used in research activities can also be considered as an attempt to answer questions systematically using certain methods (Hwase & Fofanah, 2021) In this study, researchers made predictions using the Long Short-Term Memory method, which is a variation of the Recurrent Neural Network that makes changes involving the addition of memory cells which can later store data for a very long time (Orpa et al., 2019). Literature study is a method used by researchers to find and collect data or sources related to the topic raised in a research by collecting a number of books and journals related to research problems and objectives (Meliyana & Latifah, 2022.) To find new methods or research existing ones related to predictions or forecasting using the Long Short-Term Memory method, researchers conduct literature studies by looking for references to national and international journals, theses, articles and books. Data acquisition or data acquisition system is a system that acts to retrieve, collect and prepare data before processing it to produce requirements (Liliana et al., 2021)

In this research, researchers obtained a dataset from the Trademap.org site, where this site is a portal that provides information regarding import & export values, volumes, growth rates, market shares, and so on. The aim of taking this dataset is to collect information and find out developments in the movement of Indonesia's export value to G20 countries over several years. ok studies have been conducted in the field of political behavior and characteristics of social media users. Research (NM, 2016) explores the influence of political behavior and the factors associated with social media use. Meanwhile, research (Habsy et al., 2017) proposed a method to characterize the behavior of corporate social media users. In the study, it was found that users in different groups have different topic interests but still have similarities

*name of corresponding author

within the group, known as homophilic. Research (FRHAN, 2017) also contributed with a new approach called Event WebClickviz, which utilizes agglomerative hierarchical clustering and TF-IDF feature extraction for visualization and analysis of detected to produce the target data set, the researcher carried out data selection. Meanwhile, to change the data into an appropriate form, data transformation is carried out (Rizki et al., 2020).

At this stage, researchers carry out selection by selecting and grouping exported products into 9 export categories, namely agricultural products, animal products, mining products, manufactured goods, consumer goods, chemical products, food and beverage products, medicinal products and tools. medical, and also textile products. Data visualization is needed to process data whose form has been changed to make it easier to understand (Rizki et al., 2020). Data visualization itself is an important aspect of EDA Exploratory Data Analysis because it helps convey data analysis results quickly and effectively. The application of EDA equals the conditions and description of the analysis. Data is explored from various points of view, so that EDA does not focus on just one basic method (Habsy, 2017). The forms of data visualization with EDA are bar charts, lines, circles, and other visual forms. Data visualization aims to show where important information is provided clearly, as well as explaining the data in diagrammatic forms (Bhardwaj & Kwatra, 2022). Long Short-Term Memory is made as a modified form of Recurrent Neural Network to be able to predict things accurately. Predictions are obtained based on the estimate error rate, because the smaller the error rate, the more precise the prediction will be (Rizki et al., 2020).

One of the advantages of using LSTM is the ability to remember long terms (data size) which is difficult to achieve with previous existing feature techniques. And LSTM can also use large data sizes with lots of information or data entered (Tamba et al., 2019).

METHOD

This research was conducted by following the stages shown in the flowchart in Figure 1. The following is a flowchart of system design for clustering content types and
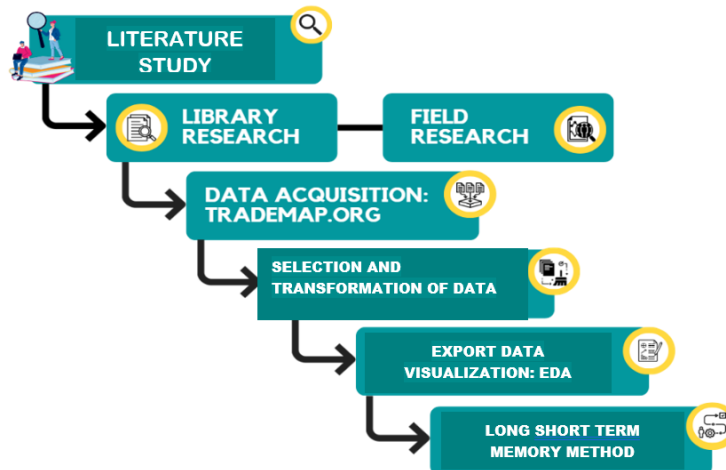

Figure 1. Flowchart System Design

In the research listed in Figure 1, several stages will be carried out, Data acquisition is data collection carried out by researchers to obtain data to be processed in research. In this study, researchers collected data by downloading a dataset on the value of Indonesia's exports to G20 countries that have been provided on the Trademap.org website, where the data is provided in Microsoft Excel format (.xls) which will later be converted into csv format. The data obtained is data for 2002 - 2021 for Export categories/commodities, while for data on the value of Indonesia's exports to G20 countries starting from January 2006 to November 2022.
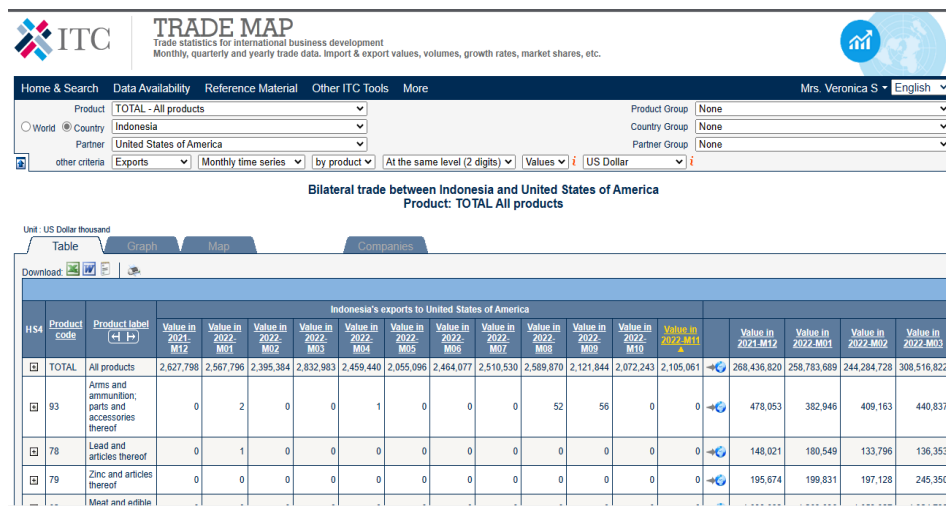
*name of corresponding author

Figure 2. Data
Acquisition

Pre-Processing

Pre-processing is data pre-processing that is done before the data is analyzed or modeled. The purpose of pre-processing is to clean, transform, and prepare data so that it can be used more effectively and efficiently. At this stage the researcher discusses the research stages, starting from Data Cleaning, Data Transformation (Scaling, Reshaping), Data Splitting, and Data Normalization. In processing this data, researchers used the Python programming language through Google Collab with several libraries. okesscollect the data, this study used a Python module called snscrape. The data collected came from the Twitter platform using various hashtags related to the presidential candidate, Anies Baswedan. Each collected data consists of the Twitter user's username and the content of the related tweets.

Data Cleaning

Data cleaning is one of the stages of data pre-processing which is carried out to identify, correct and delete incomplete, inaccurate or irrelevant data in a dataset. At this stage, data cleaning is carried out by deleting unnecessary data and adding 9 commodity codes at the beginning.

Data Transformation

Data transformation is the process of changing or converting data from one form or format to another to facilitate analysis or processing. In this study, researchers changed the form of the dataset into csv format.

A.   Scaling

This Scaling process normalizes data within a certain range, such as changing the data scale to a range of 0 to 1 [0, 1].

B.   Reshaping

Reshaping in the context of data analysis is the process of changing the layout or shape of the data.

C.   Data Cleaning

The purpose of this data cleansing process is to ensure the cleanliness and consistency of the data. During this stage, various actions are performed, such as eliminating punctuation, marks, removing number, link, emoticon. Any excessive spaces in the text are also addressed and removal as part of the cleansing process.

D.   Data Spliting

At this stage the researcher broke the dataset into two parts, namely the training set and the testing set. The training subset is used to train the model, while the testing subset is used to test the model's performance on data it has never seen before.

E.   Data normalization

Data normalization is the process of changing data values in a dataset into a more standard or normal

*name of corresponding author

form.

F.   Data Visualization

In this research, researchers used time series graphs in the form of line plots to display the total value of Indonesia's exports to G20 countries in US Dollars.
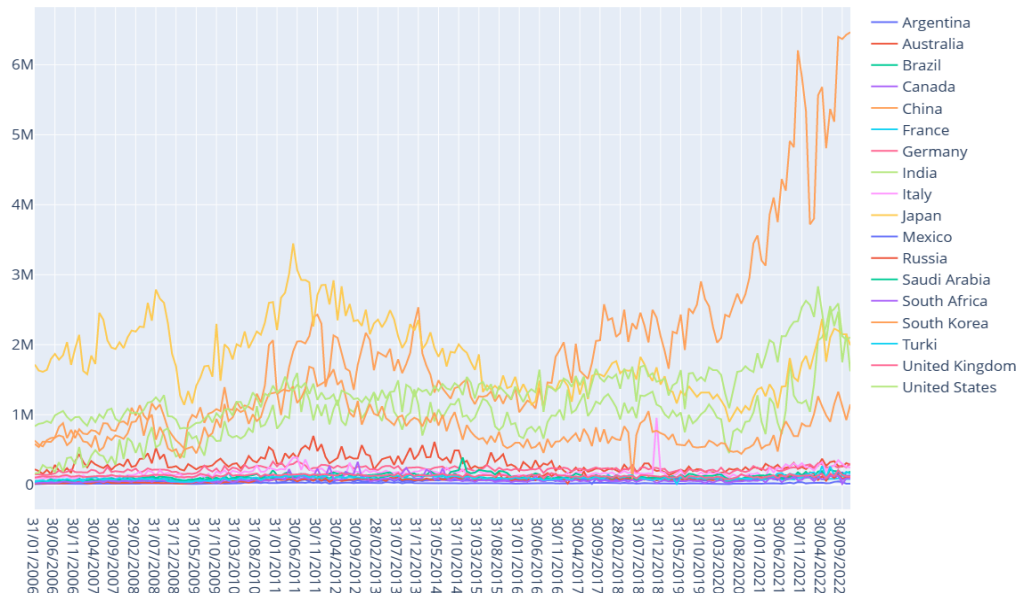

Figure 3. Data Visualization

It can be seen in the picture above that China ranks first as the country with the highest export value, then Japan ranks second, and America ranks third. Meanwhile, the country with the lowest export value is Argentina. The highest export value for commodities is Mining Products, with a Tool value of almost $60,000,000. And for export value the lowest was Drugs and Medical, at $448,000.

Long Short-Term Memory Method
Model Building
At this stage, the LSTM model is built using the Sequential architecture from Keras. Build, train, and evaluate models using Keras using LSTM, Dropout, and Dense layers. This model is compiled using the mean squared error as a loss function, the adam optimizer, and the mean squared error metric. After that, a model summary is printed, including the number of parameters that can be changed and the total parameters in the model.

Training and Testing
Relates to the evaluation of prediction results after inverse scale transformation. Using the 'scaler' object, the 'trainPredict' (the previously rescaled prediction result) is converted back to its original scale. 'trainY' is converted into a 2-dimensional array for the 'inverse_transform ()' method. Next, the RMSE (Root Mean Squared Error) value is calculated using 'trainY' and 'trainPredict [: ,0]' for the training data. After making predictions using the LSTM model, the following are the Train Score and Test Score results for each G20 country.

The RMSE (Root Mean Square Error) value produced by the LSTM model that has been created tends to have a value below 0.30, which means that the LSTM model created has succeeded in producing good accuracy values because it shows that the LSTM model has a prediction error that is smaller or closer to true value. In the context of LSTM models, RMSE is used to measure how well the model can accurately predict target data.

*name of corresponding author

Train Score and Test Score Results

Table 1. Train Score and Test Score Results

Country	Train Score	Test Score
Argentina	0.11 RMSE	0.16 RMSE
Australia	0.15 RMSE	0.10 RMSE
Brazil	0.14 RMSE	0.13 RMSE
Canada (Kanada)	0.09 RMSE	0.20 RMSE
China (Tiongkok)	0.04 RMSE	0.10 RMSE
France (Prancis)	0.11 RMSE	0.14 RMSE
Germany (Jerman)	0.16 RMSE	0.15 RMSE
India	0.08 RMSE	0.11 RMSE
Italy (Italia)	0.05 RMSE	0.16 RMSE
Japan (Jepang)	0.10 RMSE	0.08 RMSE
Korea	0.10 RMSE	0.08 RMSE
Mexico	0.10 RMSE	0.17 RMSE
Russia	0.15 RMSE	0.18 RMSE
Saudi Arabia (Arab Saudi)	0.08 RMSE	0.09 RMSE
South Africa (Afrika Selatan)	0.14 RMSE	0.08 RMSE
Turkiye (Turki)	0.07 RMSE	0.20 RMSE
United Kingdom (Inggris)	0.12 RMSE	0.19 RMSE
United States (Amerika)	0.07 RMSE	0.16 RMSE
Mean	**0.10 RMSE**	**0.13 RMSE**

This visualization uses appropriate scales for the export value and time (month) axes to make the graph easy to understand. Actual data is displayed as a blue line, while train data is displayed as a red line and test data is displayed as a green line. In the picture below, the lowest predicted value of exports from China to Indonesia is in 2025 with a value of $600,000 and the highest value reaches $6,100,000, namely in the 200th month or 2039.

*name of corresponding author

RESULT

In this section a visual representation by one of the 18 G20 countries of the extent to which LSTM models can follow actual data trends and patterns. This visualization uses appropriate scales for the export value and time (month) axes to make the graph easy to understand. Actual data is shown as a blue line, while training data is shown as a red line and test data is shown as a green line. In the image below the predicted value of exports from China to Indonesia is the lowest in 2025 with a value of $600,000 and the highest value reaching $6,100,000, namely in the 200th month or 2039. The results of the data analysis show trends and patterns in the value of Indonesia's exports to G20 member countries over a certain period of time. This data is used as input for training and testing the LSTM model. The performance of the LSTM model is evaluated based on the resulting predictions and comparison with actual export data. Evaluation metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error are used to measure how accurate the model's predictions are. The results of this evaluation indicate the extent to which the model can generalize patterns from historical data to make accurate predictions. The findings from this research can have significant implications in the context of economics and trade. A more accurate prediction of export value can help the government and business people in planning and making decisions. For example, information on demand forecasts from G20 member countries can help manage production and supplies. The following is a table of prediction results from all G20 countries, namely the highest and lowest values and their times.

Table 2. Predicted results for all G20 countries

Argentina	The highest score	35th month / year 2025 (up to $12,000)
	Lowest Value	200th month / year 2039 (up to $38,000)
Australia	The highest score	73rd month / year 2029 (reach $520,000)
	Lowest Value	155th month / year 2035 (up to $210,000)
Brazil	The highest score	60th month / year 2027 (reach $165,000)
	Lowest Value	33rd month / year 2025 (up to $60,000)
Canada	The highest score	197th month / year 2039 (up to $100,000)
	Lowest Value	200th month / year 2039 (up to $35,000)
France	The highest score	49th month / year 2027 (up to $130,000)
	Lowest Value	5th month / year 2023 (up to $63,000)
Germany	The highest score	60th month / year 2027 (up to $280,000)
	Lowest Value	5th month / year 2023 (reach $185,000)
India	The highest score	190th month / year 2038 (up to $3000,000)
	Lowest Value	2nd month / year 2023 (up to $350,000)
Italy	The highest score	55th month / year 2027 (reach $700,000)
	Lowest Value	23rd month / year 2024 (reach $160,000)

*name of corresponding author

Japan	The highest score	65th month / year 2028 (reach $3,400,000)
	Lowest Value	172nd month / year 2037 (reached $1,400000)
Mexico	The highest score	190th month / year 2038 (reaches $135,000)
	Lowest Value	Month-to 135 / year 2034 (to $27,000)
Russia	The highest score	Month-135 / year 2034 (to $135,000)
	Lowest Value	2nd month / year 2023 (reach $30,000)
Saudi Arabia	The highest score	110th month / year 2032 (reaches $320,000)
	Lowest Value	3rd month / year 2023 (reach $70,000)
South Africa	The highest score	2nd month / year 2023 (reach $50,000)
	Lowest Value	80th month / year 2029 (reaches $180,000)
South Korea	The highest score	65th month / year 2028 (reach $1,500,000)
	Lowest Value	140th month / year 2034 (reaches $350,000)
Turki	The highest score	185th month / year 2038 (reaches $220,000)
	Lowest Value	160th month / year 2036 (reaches $40,000)
United Kingdom	The highest score	200th month / 2039 (reaches $145,000)
	Lowest Value	170th month / year 2037 (reaches $120,000)
United States	The highest score	190th month / year 2038 (reached $2,300,000)
	Lowest Value	35th month / year 2025 (reach $900,000)

DISCUSSIONS

Actual data is displayed as a blue line, while train data is displayed as a red line and test data is displayed as a green line. In the picture below, the lowest predicted value of exports from China to Indonesia is in 2025 with a value of $600,000 and the highest value reaches $6,100,000, namely in the 200th month or 2039.
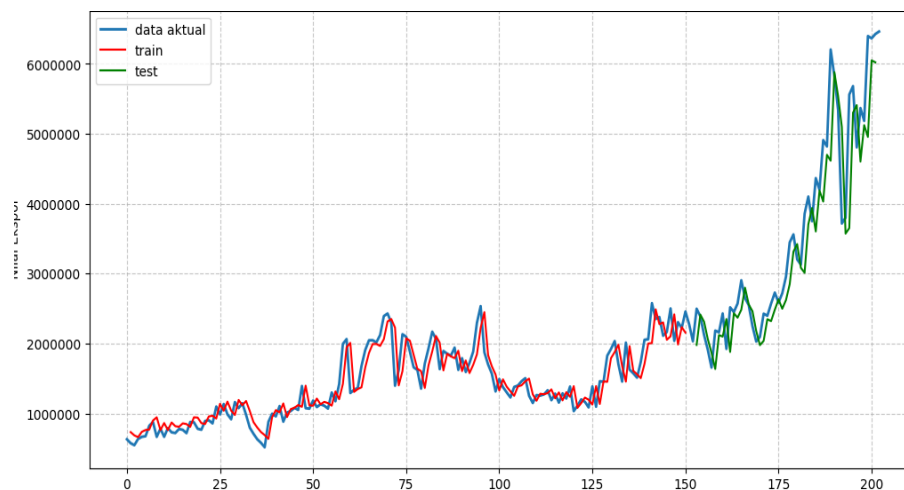
*name of corresponding author

Figure 4. Prediction Result Graph

CONCLUSION

After conducting research research results regarding Predicting the Value of Indonesian Exports to G20 Member Countries Using the Long Short-Term Memory Method, the following are several conclusions that can be drawn on training and testing data the LSTM model is able to produce relatively accurate predictions with an average RMSE value below 0.20. Predictions produced by the LSTM model tend to follow the upward and downward trends as in the actual data. Based on the prediction results, there is the highest export value to China of $6,100,000 which is predicted in the 200th month or in 2039. The lowest value of Indonesian exports is to Mexico, namely $27,000 which is predicted in the 135th month or in 2034.

REFERENCES

Astuti, I., & Ayuningtyas, F. (2018). PENGARUH EKSPOR DAN IMPOR TERHADAP PERTUMBUHAN EKONOMI DI INDONESIA. *Jurnal Ekonomi & Studi Pembangunan, 19*(1), 1-10. doi:https://doi.org/10.18196/jesp.19.1.3836

Bhardwaj, P., & Kwatra, J. (2022). Stock Market Price Forecasting Using Recurrent Neural Network . *Indonesia Journal on Computing (Indo-JC)*, *7*(1), 51-60. https://doi.org/10.34818/INDOJC.2022.7.1.612

Fadli, H. F., & Hidayatullah, A. F. (2021). Identifikasi cyberbullying pada media sosial twitter menggunakan metode lstm dan bilstm. *AUTOMATA*, *2*(1).

Farhan, A., Obaid, S. N., & Azlan, H. (2017). Corporate governance effect on firms' performance – evidence from the UAE. Journal of Economic and Administrative Sciences, 33(1), 66–80. https://doi.org/10.1108/jeas-01-2016-0002

Fredilio, F., Rahmad, J., Sinurat, S. H., Sitompul, D. R. H., Ziegel, D. J., & Indra, E. (2023). Perbandingan Algoritma K-Nearest Neighbors (K-NN) dan Random forest terhadap Penyakit Gagal Jantung. *Jurnal Teknologi Informatika dan Komputer*, *9*(1), 471-486.

Fofanah, A. J. (2021). Machine learning model approaches for price prediction in coffee market using linear regression, XGB, and LSTM techniques. *International Journal of Scientific Research in Science and Technology*, (6).

Hasby, B. A. (2017). Seni Memahami Penelitian Kualitatif Dalam Bimbingan dan Konseling: Studi Literatur. JURKAM: Jurnal Konseling Andi Matappa, 1 (2), 90.

Hendra, S., & Kusumadewi, S. (2015). Fungsi Similaritas Pada Sistem Berbasis Kasus Penyelesaian Masalah Akademik Mahasiswa. In *Seminar Nasional Informatika Medis (SNIMed)* (pp. 115-123).

Hidayah, N., Sulfahmi, S., Zairani, I., Yusuf, M., & Sufiati, S. (2020). Combine Assurance Dalam Konteks Pengendalian. *Equilibrium: Jurnal Ilmiah Ekonomi, Manajemen dan Akuntansi*, *8*(2).

Holst, J., Brock, A., Singer-Brodowski, M., & de Haan, G. (2020). Monitoring progress of change: Implementation of Education for Sustainable Development (ESD) within documents of the German

*name of corresponding author

education system. *Sustainability*, *12*(10), 4306.

Husni, D. T., Sitompul, D. R. H., Sinurat, S. H., Ruben, R., Situmorang, A., Ziegel, D. J., ... & Indra, E. (2022). ANALISIS BIG DATA PENJUALAN VIDEO GAMES MENGUNAKAN EDA. *Jurnal Tekinkom (Teknik Informasi dan Komputer)*, *5*(1), 43-48.

Kafil, M. (2019). Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *3*(2), 59-66.

Liliana, D. Y., Hikmah, N. N., & Harjono, M. (2021). PENGEMBANGAN SISTEM PEMANTAUAN SENTIMEN BERITA BERBAHASA INDONESIA BERDASARKAN KONTEN DENGAN LONG SHORT-TERM MEMORY. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, *8*(5).

Meliyana, D., & Latifah, K. (2022). Pengelolaan Dan Visualisasi Data Pada Dinas Komunikasi Dan Informatika Kabupaten Pati. In *Proceeding Science and Engineering National Seminar* (Vol. 7, No. 1).

Nissa, N. K., Nugraha, Y., Finola, C. F., Ernesto, A., Kanggrawan, J. I., & Suherman, A. L. (2020). Evaluasi Berbasis Data: Kebijakan Pembatasan Mobilitas Publik dalam Mitigasi Persebaran COVID-19 di Jakarta. *Jurnal Sistem Cerdas*, *3*(2), 84-94.

N. Hidayah, S. Sulfahmi, I. Zairani, M. Yusuf, and Sufiati, "Combine Assurance in the Context of Control," J. Ilm. Econ. Management, and Accounting, vol. 08, no. 02, pp. 32–37, 2019.

Understanding, CONTENTS., Qualitative., Guidance, AND., Counseling, D., & Habsy, B. A. (2017). Counseling View project Social Psychology View project. https://doi.org/10.235678/25271987

Orpa, E. P. K., R. E. F., T. (n.d.). (2019) Prediction Model for the Beginning of Student Study Period. Journal of Information Systems and Technology (Vol.7, No.4)

Owen, M., Vincent, V., Br Ambarita, R., & Indra, E. (2022). IMPLEMENTASI METODE LONG SHORT TERM MEMORY UNTUK MEMPREDIKSI PERGERAKAN NILAI HARGA EMAS. *Jurnal Tekinkom (Teknik Informasi Dan Komputer), 5*(1), 96-104. doi:10.37600/tekinkom.v5i1.507

Tamba, S. P., Batubara, M. D., Purba, W., Sihombing, M., Siregar, V. M. M., & Banjarnahor, J. (2019, July). Book data grouping in libraries using the k-means clustering method. In *Journal of Physics: Conference Series* (Vol. 1230, No. 1, p. 012074). IOP Publishing.

Tyasnurita, R., Luthfiansyah, R. R., & Brameswara, M. R. (2023). GOLD PRICE FORECASTING USING MULTIPLE LINEAR REGRESSION METHOD. *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, *9*(3), 339-344.

Rizki, M., Basuki, S., & Azhar, Y. (2020). Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory Untuk Prediksi Curah Hujan Kota Malang. *Repositor*, *2*(3), 331-338.

*name of corresponding author