Predicting Sugar Import Quantity Using Multi-Layer Perceptron Regressor Method

Mas Syahdan Filsafan
Department of InformaticsFaculty of
Intelligent and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
syahdanfilsafan58@gmail.com

Riyanarto Sarno
Department of Informatics, Faculty of
Intelligent and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
riyanarto@if.its.ac.id

Shintami Chusnul Hidayati
Department of Informatics, Faculty of
Intelligent and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
shintami@its.ac.id

Bernadetta Raras
Department of Informatics, Faculty of
Intelligent and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
rarasbernadetta@gmail.com

Agus Tri Haryono
Department of Informatics, Faculty of
Intelligent and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
7025231020@student.its.ac.id

2024 7th International Conference on Informatics and Computational Sciences (ICICoS) | 979-8-3503-7588-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICICoS62600.2024.10636884

***Abstract*—This paper presents a novel approach for predicting the quantity of sugar imports using the MLP Regressor method. The study compares various Machine learning and Deep Learning techniques, including Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Ensemble Methods, with the MLP Regressor model demonstrating superior performance in terms of key evaluation metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R-squared (R2) scores. The dataset's missing values challenge is addressed by implementing the K-Nearest Neighbors (KNN) algorithm for data imputation, ensuring data integrity. The findings contribute to developing an accurate predictive model for sugar imports and emphasise the significance of handling missing values in the dataset. Further research efforts should focus on refining and enhancing the proposed model's accuracy and reliability**

Keywords—Sugar imports, MLP (Multi Layer Perceptron) Regressor, Predictive modelling.

I. INTRODUCTION

Granulated sugar is an essential commodity in people's daily lives and is a crucial ingredient for food and beverage manufacturers and sellers. It serves as the primary sweetener in the food and beverage industry, meeting direct consumption and production requirements. According to sucden.com, global sugar consumption has significantly increased over the years, driven by population growth and changing consumption patterns. In the early 20th century, with a population of 1.6 billion, sugar consumption reached eight million tons. However, with the global population exceeding seven billion, sugar consumption has surpassed 170 million tons. This remarkable increase signifies a per capita consumption rise of 4-5 times during this period, with 70% of the world's sugar production derived from sugar cane and the remaining from beets. In 2008, Indonesia imported 2.3 million tones of sugar, including white, granulated, and raw sugar [1]. The Badan Assosiasi Gula Indonesia (AGI) data shows that the most significant sugar imports will occur in 2023, estimated at 3,181,710 tons of sugar production, and domestic demand exhibits annual fluctuations. Notably, there have been years without sugar imports, as depicted in Fig. 1.

Artikanur's focus research revealed that sugar production, sugar consumption, domestic sugar prices, international sugar prices, the exchange rate between the Indonesian Rupiah and the US Dollar, gross domestic product, and import tariffs significantly influence sugar imports in Indonesia. Higher sugar production and lower domestic sugar prices reduce the need for imports, while increased sugar consumption and higher international sugar prices increase imports. Similarly, the exchange rate, gross domestic product, and import tariffs impact sugar imports [2]. However, their study did not include a sugar forecasting method for the future, which is vital for policymakers to anticipate and address the sugar needs of Indonesian society. One suitable method for this purpose is the LSTM (Long Short-Term Memory) model, which can accommodate up to 17 variables.

Machine learning is a powerful approach for extracting knowledge from existing data. While traditional Machine learning methods may face challenges in tasks such as image classification due to the need for large datasets, Deep Learning emerges as a branch of Machine learning capable of processing vast amounts of data. Deep learning, which employs Artificial Neural Networks (ANN) inspired by the human brain, has proven effective in various domains. In the context of this research, ensuring data completeness is crucial for the Machine learning and Deep Learning processes. The dataset obtained from the Badan Pusat Statistik (BPS-Statistics Indonesia) often contains missing values, and addressing this issue requires specialised algorithms. One practical approach is utilising the K-Nearest Neighbors (KNN) Imputer with a K-value of 5, which leverages nearest neighbour values to estimate and fill in the missing data. The dataset can be effectively completed by applying the KNN Imputer with a K-value of 5, providing a more comprehensive and reliable foundation for subsequent analysis [2], [3].

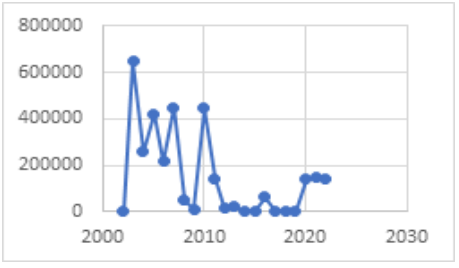

Fig. 1. Indonesia Sugar Import

Previous Studies used several algorithms in this study, including LSTM, CNN, GRU, ExtraTrees Regressor, AdaBoost Regressor, GradientBoosting Regressor, RandomForest Regressor, Bagging Regressor, and mixed learning with the ensemble method. Four scenarios were used in this study. The first scenario involved normalisation without denormalisation. The second scenario involved normalisation and denormalisation. The third scenario involved normalisation and denormalisation only on the target (import). The fourth scenario did not involve normalisation or denormalisation. Of these four scenarios, the MLP method showed its superiority because it combined the results of LSTM, CNN, GRU, ExtraTrees Regressor, AdaBoost Regressor, GradientBoosting Regressor, RandomForest Regressor, and Bagging Regressor. This study is also the first to use a method like Mixed Learning, which no previous researchers have done.

The proposed method in the previous study has some common limitations associated with using deep learning for crop yield prediction. These limitations include a reliance on large amounts of data for practical model training, the complexity of the model architecture, difficulties in interpreting the model's predictions, dependence on high-quality data, and the risk of overfitting [4]. It is important to note that these limitations are general challenges in utilising deep learning for crop yield prediction.

This research aims to analyse the factors influencing sugar imports in Indonesia over the past two decades using Panel Data from 2002 to 2023. Additionally, the study will employ Machine learning and Deep Learning methods to forecast future sugar imports based on previous research findings by comparing the forecasting results obtained from Panel Data with those from Machine learning. The research aims to gain deeper insights into the factors influencing sugar imports in Indonesia and provide a comprehensive overview of future sugar import forecasts.

II. Background Study

The literature review is a crucial reference to provide a concise overview of previous research closely related to the topic. It helps identify existing gaps and informs the design of this research. In various fields, including healthcare, dealing with missing values has been a pervasive challenge that significantly impacts decision-making[5]. The debate surrounding imputation methods has been ongoing since 1998, but recent advancements in machine learning have brought substantial progress, particularly in the healthcare industry [6]. Imputation techniques can be approached through statistical methods and machine learning, each with strengths and limitations. However, statistical imputation techniques still exhibit bias and information loss when estimating missing values [7].

This study aims to integrate deep learning methods into data panel analysis to address this gap in future sugar forecasting. While previous research has employed data panel analysis, there is a need for more studies incorporating Deep Learning methods in this context [8]. By adopting this approach, the study strives to enhance the accuracy and reliability of sugar forecasting by leveraging the power and potential of Deep Learning.

The forecasting techniques for non-linear and non-stationary time series are already comprehensively analysed, especially in manufacturing and health informatics [9]. The findings highlighted a growing interest in nonparametric models for forecasting non-linear and non-stationary time series. This research suggested that integrating principles from non-linear dynamic systems into nonparametric modelling approaches holds great promise for further enhancing forecasting accuracy.

A. Machine learning

As shown in Fig. 2, machine learning is a branch of artificial intelligence that allows computers to learn from data and make predictions or decisions without explicit programming. It involves the development of algorithms that enable systems to recognize patterns, extract meaningful insights, and improve their performance through iterative learning processes[10]. With supervised learning, models are trained on labelled data to predict future outcomes, while unsupervised learning uncovers hidden patterns in unlabeled data. Machine learning applications are widespread, spanning fields such as image and speech recognition, medical diagnosis, financial forecasting, and more [11].

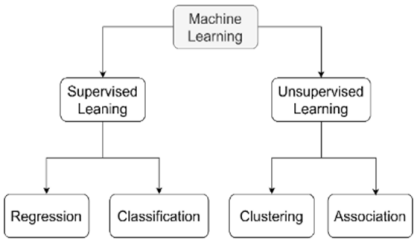

Fig. 2. Depths of Machine Learning Algorithms

B. Data Collection Techniques

This research employs documentation techniques and gathers data from multiple sources to ensure a comprehensive and robust analysis, including sources such as BPS and data from indonesia.id, which provides access to reliable and official statistical information [12]. Additionally, utilizing AGI and literature studies allows for a broader understanding of the subject matter. Furthermore, incorporating statistical data specifically focused on Java enables a more localized analysis. Including online sources and other pertinent references ensures a diverse range of perspectives and up-to-date information, contributing to a well-rounded and informed discussion in the research [13].

C. Contribution

The paper makes three key contributions:

1. This research incorporates machine learning and deep learning analyses that build upon prior studies [14].
2. The paper you mentioned appears to be the first one that examines the combination of results from several predictive models, including Extra Trees, AdaBoost, Gradient Boosting, Random Forest, and Bagging, using the Multi-Layer Perceptron (MLP) method [15], [16].
3. Modified regression modeling approach to forecast import volume [17].

III. Proposed Method

The literature review is used as a reference to summarise previous research and is close to the topic of this research. Reference of earlier studies and is used as a reference for the design of this research for existing gaps. Train and test represent the accuracy of the training and test data, respectively, in percentage.

The data modelling process workflow starts with data preparation and collection, followed by data cleansing to clean the data. Next, the data is split into training, validation, and testing sets. Then, machine learning modelling LSTM and deep learning modelling: Extra Tree Regressor, Adaboost Regressor, Bagging Regressor, Random Forest Regressor, and Gradient Boosting Regressor. are generated. The data from these models is then evaluated using MAE and MSE metrics. Once the evaluation is done, the data modelling process is completed.

A. Dataset

Sources of datasets in this study can be obtained from the Directorate General of Plantations, as well as agencies outside the Ministry of Agriculture such as the Central Statistics Agency (BPS), the Food and Agriculture Organization (FAO) and the World Bank. Datasets are compiled by referring to secondary data and information.

B. Splitting Data

Split the dataset into training and testing sets using random sampling without replacement. About 60% of the rows are randomly selected and placed into the training set. In comparison, the remaining 20% is allocated to the test set, and the remainder is used to validate the data model. It is important to note that the colours in the "Features", "Target", and "Validation" sections indicate the data to be assigned as "X_train", "X_test", "X_val", "y_train", "y_test", "y_val," for division specific training-testing.

C. Handling Missing Value

Missing data imputation techniques using statistical algorithms and machine learning have been employed to address missing values in datasets [18]. Performance evaluation of these methods can be assessed using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Table 1 shows that the KNN method outperforms others in terms of MAE and RMSE. Euclidean distance, a simple and parameter-free measurement, offers minimal increase in time complexity for time series analysis. Additionally, Complexity-Invariant Distance (CID) maintains efficiency in various algorithms used for classification, clustering, motif discovery, and outlier detection [19]. This study confirms the suitability of these distance measurements for time series data with the following formula:

$$ED(Q,C) = \sqrt{\sum_{i=1}^{n}(q_i - c_i)^2} \qquad (1)$$

This method provides deeper insight into the extent to which predictions from various approaches to dealing with missing data are close to the actual values. With the combination of MAE, RMSE, and accuracy rate measurements, this research provided a more holistic understanding of the performance of various missing data handling techniques in the context of the Machine learning problems discussed in this research.

D. Handling Missing Value CNN

CNN is a neural network architecture designed explicitly for processing visual data, such as images and videos. This architecture is inspired by how the human vision system works and uses convolutional layers to extract essential features from the data. With their ability to automatically learn and recognise complex patterns, CNN has become the basis of many image processing applications, including image classification, object detection, facial recognition, and significant research topics in scientific journals.

The CNN MobileNetV2 architecture has the highest accuracy in predicting lung X-ray images related to COVID-19, with an accuracy value of 0.96. The MobileNetV2 CNN effectively makes predictions on the dataset used in the study [18]. However, it is essential to note that the effectiveness of CNNs in prediction can vary depending on the dataset used, the complexity of the problem, and other settings. It is always important to comprehensively evaluate and validate the developed model before implementing it in broader use [19].

E. LSTM

LSTM is a modified Recurrent Neural Network (RNN) model that enhances the ability to capture long-term dependencies. Unlike traditional RNN, LSTM incorporates gating mechanisms that enable it to selectively store and discard information over time. LSTM can overcome the vanishing gradient problem encountered in RNN by effectively combining and processing information through these gates [20]. LSTM is particularly well-suited for tasks involving large-scale data analysis, such as text classification, natural language processing, and sequential data processing, including sensor signals and biomedical data [21].

The LSTM method has been tested for classification on asthma datasets. Experimental results show that the LSTM method is of good quality, with an accuracy of 96.6%, precision of 96.1%, recall of 95.5%, and F1-Score of 95.6%. The LSTM method is effective in classifying asthma data with a high level of accuracy.

F. Ensemble Methode

Ensemble decision-making is relevant both in our daily lives and in machine learning. In real-life situations, we seek diverse opinions to make informed decisions. Similarly, in machine learning, ensemble learning combines weak classifiers to create a stronger one, using techniques like voting or averaging [22]. In upcoming experiments, various ensemble methods will be explored to demonstrate this concept further [23].

G. MLP-Regressor

Previous research proposes using MLP-Regressor to predict stock prices by integrating sentiment scores from financial news. The model achieves a high accuracy of 0.90 for predicting trends and future stock prices over ten days, based on experiments conducted on ten years of historical stock price data and financial news from four different companies. The research highlights the impact of financial news on stock prices and explores the effects of each sentiment algorithm. The MLP Regressor can be combined with other models, such as LSTM, CNN, and Ensemble, to create a more accurate and stable prediction model. However,

thorough experimentation and research are necessary to ensure improved results compared to previous models. Furthermore, integrating natural language processing techniques can enhance the model's performance by extracting more nuanced sentiment features from financial news [24]. Future studies can also explore the application of this model in real-world trading scenarios to evaluate its practical effectiveness [23].

This paper uses Multi-layer Perceptron (MLP) as a machine learning model for data processing in the context of 5G cellular networks. MLP has advantages in handling complex and non-linear data and the ability to learn patterns embedded in the data. In the context of Service Function Chain (SFC) placement and scalability in 5G networks, MLP performs traffic measurements and predicts the required number of virtual network function instances based on traffic demands. By utilising MLP, accurate estimates can be obtained to optimise network resource allocation and ensure the end-to-end latency requirements of services are consistently met [25].

The number of layers and neurons are defined as hyperparameters in MLP models. However, studies such as auto-adaptive MLP have attempted to determine the number of neurons in the hidden layers automatically based on the nature of the training time series data [26]. That allows the network to adapt to the training data's characteristics and optimize performance.

In MLP Regressor, the initial inputs are obtained from the predictions of various sklearn libraries (Extra Trees, Gradient Boosting, Bagging, Random Forest, Ada Boost), which are then combined with deep learning methods such as LSTM and CNN as explained in the scenario stages. The results of each model can be displayed before using MLP, allowing for a comparison to determine whether MLP has the best results. The proposed writing style for the paper is to present this information from a scenario point of view.

H. Build Scenario

The approach used in modelling, namely modelling with Neural Networks and ensemble methods, is shown in Fig. 3. In modelling with Neural Networks, various activation functions and optimiser combinations were tried to produce 49 different models. Meanwhile, the ensemble method uses several models and evaluates the results of each model. This research gained insight into the model's prediction performance through this experiment. Apart from that, there is also an explanation of the scenarios and variations used in modelling for the MLP Regressor.

1) Modelling with Neural Network: In the LSTM and CNN experiments, various activation functions and optimisers were combined in this study to produce 49 models. The activation functions used include 'relu', 'tanh', 'sigmoid', 'leaky_relu', 'elu', 'selu', and 'softmax', while the optimisers used include 'SGD', 'RMSprop', ' Adagrad', 'Adadelta', 'Adam', 'Adamax', and 'Nadam'. By using this combination, valuable insights into the model's prediction performance can be gained.
2) In ensemble methods experiments, models such as AdaBoost Regressor, Bagging Regressor, ExtraTrees Regressor, Gradient Boosting Regressor, and Random Forest Regressor were employed in this research. By utilising this combination of models in previous research, different results were evaluated, providing valuable insights into the performance of each model in making predictions. The ensemble model was trained and evaluated using test data, and evaluation metrics, including MSE, MAE, MAPE, and R-squared, were calculated for each model. By leveraging the strengths of each model through the ensemble method, valuable insights into the overall prediction performance were obtained in this international journal paper on computer science.
3) Modelling for Machine Learning: Extra Trees Regressor, Gradient Boosting Regressor, Bagging Regressor, Random Forest Regressor, Ada Boost Regressor, and Deep Learning: LSTM, CNN. In each of these scenarios, there are four different approaches. The first scenario involves data normalisation without denormalization. The second scenario involves normalisation and denormalization. The third scenario involves only data import without normalisation, but denormalization is still performed. The fourth scenario involves neither normalisation nor denormalization.
4) Modelling for MLP Regressor, where the four scenarios above are still used. However, there are several new scenarios adapted to the needs of experiments, which are, MLP Regressor with all scenarios combined equally without optimisation (MLP 1). MLP Regressor with all scenarios combined and optimised (MLP 2). MLP Regressor with three best scenarios and no optimisation (MLP 3). MLP Regressor with three best and optimised scenarios (MLP 4). MLP Regressor with one best scenario and no optimisation (MLP 5). MLP Regressor with one best and optimised scenario (MLP 6). MLP Regressor with different Activation, Optimizer, Learning Rate in each row of the three best scenarios. MLP Regressor with different Activation, Optimizer, Learning Rate in each row, one best scenario. MLP Regressor with different Activation, Optimizer, Learning Rate in each row for all scenarios.

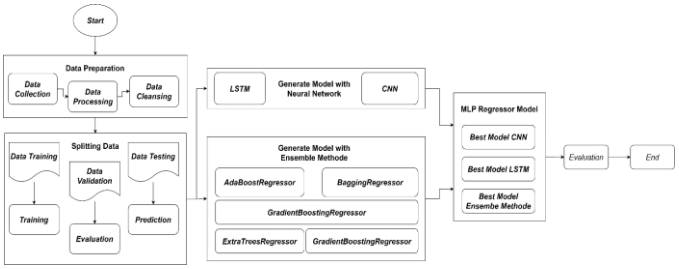

Fig. 3. Flowchart Model Building (For optimal clarity and detail, it is recommended to view this figure in its digital format)

I. Evaluation Method

MSE, MAE, and R2 are commonly used evaluation methods for forecasting accuracy [27]. MSE focuses on the average squared difference between forecasted and actual values, giving more weight to significant errors. It can be minimised using gradient descent, Newton's method, or Bayesian optimisation. On the other hand, MAE treats all errors equally, calculating the average absolute difference. Minimising MAE helps determine optimal model parameters

for predicting new data. R2 measures the proportion of variance explained by independent variables in a regression model, ranging from 0 to 1. Higher R2 scores indicate a better fit, while negative values suggest worse performance than a simple horizontal line.

J. Presentation of Data

Data separation is crucial in data pre-processing, particularly when developing data-based models. This technique plays a significant role in enhancing the accuracy of machine learning models by ensuring that the data used for modelling is appropriately organised and distinct. The used dataset obtains a dataset that includes information about the import of white sugar circulating on the market. Fill in empty columns using the KNN imputation method to handle missing values [28]. The target column for prediction is the import column.

TABLE I. MEAN PERCENTAGE VALUES

Metrics	Result
Median	13168294.790
Most Frequent	27895562.979
Constant	30876712.270

TABLE II. MEAN PERCENTAGE VALUES

Algorithm	Mean Percentage
KNN	94.734
Most Frequent	94.715
Median	94.701
Constant	93.990

IV. RESULT AND DISCUSSION

Before proceeding with the machine learning and deep learning models, the experiments on handling missing values using different algorithms. Table I summarises the Mean Percentage values obtained from each algorithm. The results show the mean percentage values, which indicate the degree of similarity between the imputed values and the original data. Table II shows that the KNN algorithm achieved the highest mean percentage value of 94.734, suggesting it closely resembles the original data based on the Euclidean distance measurement. The values indicate that the KNN method is particularly effective in handling missing values, ensuring the integrity of the dataset for subsequent modelling steps. Therefore, based on these results and the Euclidean distance metric, it would be recommended that the KNN algorithm be selected for imputation.

After handling missing values, comparing four modelling scenarios with Machine Learning regressors unveils distinct performance trends. The results are shown in Table III. Scenario 1 emerges as notably superior, showcasing models like Extra Trees with a low MSE of 0.029 and Bagging Regressor maintaining reasonable metrics despite a negative R-squared (-0.109). Conversely, Scenario 2 encounters substantial hurdles, notably seen in Extra Trees and Ada Boost exhibiting notably high MSE and MAPE values, signalling diminished prediction accuracy. In Scenario 3, while improvements surface in models like Random Forest compared to Scenario 2, others, such as Extra Trees, falter with exceedingly high MSE and impractical MAPE values. Scenario 4 presents a blend of outcomes, with Extra Trees showcasing progress but models like Ada Boost and Gradient Boosting witnessing deteriorated performance. Overall, Scenario 1 is the most promising, showcasing enhanced predictive capabilities within this experimental framework.

TABLE III. PERFORMANCE RESULTS MACHINE LEARNING REGRESSOR

S	Model	MSE	MAE	MAPE	R^2
1	**Extra Trees**	**2.9E-02**	**9.4E-02**	**3.1E+05**	**-3.3E-01**
	Gradient Boosting	2.3E-02	9.7E-02	7.6E+06	-8.3E-02
	Bagging	2.4E-02	8.7E-02	2.9E+07	-1.1E-01
	Random Forest	2.2E-02	8.7E-02	3.0E+07	-3.8E-02
	AdaBoost	3.2E-02	1.2E-01	1.1E+08	-4.7E-01
2	**Extra Trees**	**1.9E+17**	**2.4E+08**	**5.5E+04**	**-3.2E-01**
	AdaBoost	1.7E+17	2.7E+08	2.7E+05	-2.0E-01
	Gradient Boosting	1.6E+17	2.5E+08	7.8E+05	-8.7E-02
	Random Forest	1.4E+17	2.3E+08	8.2E+06	3.0E-03
	Bagging	1.4E+17	2.4E+08	1.1E+07	5.5E-02
3	**Extra Trees**	**2.1E+17**	**3.2E+08**	**8.8E+01**	**-9.0E-01**
	Random Forest	1.8E+17	3.1E+08	1.5E+02	-6.4E-01
	AdaBoost	1.8E+17	3.0E+08	1.7E+02	-6.5E-01
	Bagging	1.6E+17	3.2E+08	2.6E+02	-4.3E-01
	Gradient Boosting	1.8E+17	3.4E+08	2.8E+02	-6.8E-01
4	**Extra Trees**	**1.9E+17**	**2.4E+08**	**7.3E+04**	**-2.9E-01**
	AdaBoost	1.9E+17	3.1E+08	2.4E+05	-3.3E-01
	Gradient Boosting	1.6E+17	2.5E+08	1.3E+06	-8.8E-02
	Random Forest	1.6E+17	2.2E+08	6.3E+06	-8.5E-02
	Bagging	1.5E+17	2.3E+08	6.7E+06	-2.6E-02

The study tested KNN, LSTM, and ensembles for sugar import prediction. After initial success, the best from each were compared using MLP Regressor (Table IV). MLP Regressor outperformed all other models in terms of MSE, MAE, MAPE, and R-squared, showing promise for accurate sugar import prediction. Even the negative R-squared, suggesting weak correlation, was better than other models. This indicates MLP Regressor's potential as a reliable method for predicting sugar import.

TABLE IV. PERFORMANCE RESULTS

Model	MSE	MAE	MAPE	R^2
LSTM Norm	555686702 870921000	7403 52900	100	-72.46
CNN Norm	182734503 209336000	3780 38272	554 125	-0.75
Ensemble Learning Norm	167224711 882720000	3689 89271	690 956	-0.12
MLP Norm	146908017 460261000	2285 01429	673 3396	-0.04
LSTM	135241644 886948000	2386 94830	1066 3777	0.05
CNN	22443781 892722300	1215 47471	15	-1.88
Ensemble Learning	12433927 027081700	887 32785	11	-0.52
MLP Regressor	**62846063 56592130**	**717 23032**	**10**	**-0.18**

V. CONCLUSION

The MLP Regressor, chosen from comparisons with LSTM, CNN, and ensembles, shows inconsistent performance. While some cases achieve good fit and low errors (MAPE, MSE), others have concerning issues (negative R-squared, high MSE). Further work is needed to improve

accuracy and capture data patterns. This could involve hyperparameter tuning, feature selection refinement, data collection, exploring different algorithms/architectures, and data quality checks for outliers or missing information.

AUTHOR CONTRIBUTIONS

This research can run well and successfully because of the following research contributions: Conceptualization by Prof. RS and MSF; methodology by Prof. RS, ATH, and SCH, MSF; validation and formal analysis by Prof. RS, ATH, MSF; data, visualization, and editing by MSF, and ATH; Funding by BR; DAtaset from BR; supervision by Prof. RS.

ACKNOWLEDGMENT

This research was funded by Institut Teknologi Sepuluh Nopember (ITS) under Penelitian Dana Departemen Program, Penelitian Keilmuan, Penelitian Flagship, Penelitian Kolaborasi Pusat and under the Project Scheme of the Publication Writing and Intellectual Property Rights (IPR) Incentive (Penulisan Publikasi dan Hak Kekayaan Intelektual/PPHKI) Program 2024.

REFERENCES

[1] S. D. Artikanur, Widiatmaka, Y. Setiawan, and Marimin, "AHP-GeoTOPSIS method to analyze priority areas for sugarcane plantation development in Lamongan Regency," IOP Conf. Ser. Earth Environ. Sci., vol. 950, no. 1, p. 12078, Jan. 2022.

[2] W. P. Inayaturrohmah, I. Istiqomah, and R. S. Gunawan, "Determinants of Indonesian Sugar Import," SIGNIFIKAN, vol. 12, no. 2, pp. 275–286, Oct. 2023.

[3] T. Handhayani, J. Hendryli, and L. Hiryanto, "Comparison of shallow and deep learning models for classification of Lasem batik patterns," in 2017 1st International Conference on Informatics and Computational Sciences (ICICoS), IEEE, Nov. 2017.

[4] N. V Dharwadkar, V. H. Kalmani, and V. Thapa, "Crop yield prediction using deep learning algorithm based on CNN-LSTM with Attention Layer and Skip Connection," Oct. 2023.

[5] O. F. Ayilara, L. Zhang, T. T. Sajobi, R. Sawatzky, E. Bohm, and L. M. Lix, "Impact of missing data on bias and precision when estimating change in patient-reported outcomes from a clinical registry," Health Qual. Life Outcomes, vol. 17, no. 1, p. 106, Jun. 2019.

[6] M. Mohri, A. Rostamizadeh, and A. Talwalkar, Foundations of machine learning, 2nd ed. in Adaptive Computation and Machine Learning series. London, England: MIT Press, 2018.

[7] N. Solomon, Y. Lokhnygina, and S. Halabi, "Comparison of regression imputation methods of baseline covariates that predict survival outcomes," J. Clin. Transl. Sci., vol. 5, no. 1, 2021.

[8] B. Jan, H. Farman, M. Khan, M. Imran, I. U. Islam, A. Ahmad, S. Ali, et al., "Deep learning in big data Analytics: A comparative study," Comput. Electr. Eng., vol. 75, pp. 275–287, 2017, doi: 10.1016/J.COMPELECENG.2017.12.009.

[9] C. Cheng, A. Sa-Ngasoongsong, O. Beyca, T. Le, H. Yang, Z. Kong, and S. Bukkapatnam, "Time Series Forecasting for Nonlinear and Nonstationary Processes: A Review and Comparative Study," IIE Transactions, Sep. 2015, doi: 10.1080/0740817X.2014.999180.

[10] D. Nemirovsky, T. Arkose, N. Marković, M. Nemirovsky, O. Unsal, A. Cristal, and M. Valero, "A General Guide to Applying Machine Learning to Computer Architecture," Supercomput. Front. Innov., vol. 5, pp. 95–115, 2018, doi: 10.14529/JSFI180106.

[11] I. A. Sabilla, Z. A. Cahyaningtyas, R. Sarno, A. Al Fauzi, D. R. Wijaya, and R. Gunawan, "Classification of human gender from sweat odor using electronic nose with machine learning methods," in 2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob), IEEE, Apr. 2021.

[12] G. A. Bowen, "Document Analysis as a Qualitative Research Method," Qualitative Research Journal, vol. 9, pp. 27–40, 2009, doi: 10.3316/QRJ0902027.

[13] J. Hofhuis, P. Schafraad, D. Trilling, N. Luca, and B. van Manen, "Automated content analysis of cultural Diversity Perspectives in Annual Reports (DivPAR): Development, validation, and future research agenda.," Cultur Divers Ethnic Minor Psychol, 2021, doi: 10.1037/cdp0000413.

[14] G. I. Arastika, "Analisis faktor-faktor yang mempengaruhi Impor gula di indonesia", Accessed: Jul. 02, 2023. [Online]. Available: https://repository.uinjkt.ac.id/dspace/handle/123456789/52507

[15] D. Dhande, C. Choudhari, D. Gaikwad, N. Sinaga, and K. Dahe, "Prediction of spark ignition engine performance with bioethanol-gasoline mixes using a multilayer perception model," Pet Sci Technol, vol. 40, pp. 1–25, Jun. 2022, doi: 10.1080/10916466.2022.2025832.

[16] S. Subiyanto, A. Sukmono, N. Bashit, and F. Amarrohman, "The use of a MLP neural network for analysis and aodeling of land use changes with variations variable of physical and economic social," IOP Conf Ser Earth Environ Sci, vol. 389, p. 12029, Jun. 2019, doi: 10.1088/1755-1315/389/1/012029.

[17] C.-C. Chou, C.-W. Chu, and G.-S. Liang, "A modified regression model for forecasting the volumes of Taiwan's import containers," Math Comput Model, vol. 47, no. 9, pp. 797–807, 2008, doi: https://doi.org/10.1016/j.mcm.2007.05.005.

[18] P. J. G. Laencina, J.-L. S. Gomez, A. R. F. Vidal, and M. Verleysen, "K Nearest Neighbours with Mutual Information for Simultaneous Classification and Missing Data Imputation," Neurocomputing, vol. 72, pp. 1483–1493, 2009.

[19] D. Cao, Y. Tian, and D. Bai, "Time Series Clustering Method Based on Principal Component Analysis," 2015, doi: 10.2991/ICIMM-15.2015.163.

[20] D. M. Ibrahim, N. M. Elshennawy, and A. Sarhan, "Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases," Comput Biol Med, vol. 132, pp. 104348–104348, 2021, doi: 10.1016/j.compbiomed.2021.104348.

[21] H. M. Ramadhani, A. P. Pratiwi, S. C. Hidayati, and D. Herumurti, "CNN architecture comparison for covid-19 image classification process," in 2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE), IEEE, Jan. 2022.

[22] R. A. Priyantina and R. Sarno, "Sentiment Analysis of Hotel Reviews Using Latent Dirichlet Allocation, Semantic Similarity and LSTM," International Journal of Intelligent Engineering and Systems, 2019, [Online]. Available: https://api.semanticscholar.org/CorpusID:198313965

[23] K. Yamamoto and T. Ohtsuki, "Non-Contact Heartbeat Detection by Heartbeat Signal Reconstruction Based on Spectrogram Analysis With Convolutional LSTM," IEEE Access, vol. 8, pp. 123603–123613, Dec. 2020, doi: 10.1109/ACCESS.2020.3006107.

[24] C. Latha and S. Jeeva, "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques," Inform Med Unlocked, 2019, doi: 10.1016/J.IMU.2019.100203.

[25] J. Maqbool, P. Aggarwal, R. Kaur, A. Mittal, and I. A. Ganaie, "Stock Prediction by Integrating Sentiment Scores of Financial News and MLP-Regressor: A Machine Learning Approach," Procedia Comput Sci, vol. 218, pp. 1067–1078, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.086.

[26] J. Song, L. Zhang, G. Xue, Y. Ma, S. Gao, and Q. Jiang, "Predicting hourly heating load in a district heating system based on a hybrid CNN-LSTM model," Energy Build, 2021, doi: 10.1016/J.ENBUILD.2021.110998.

[27] J. Ren, G. Yu, Y. Cai, and Y. He, "Latency Optimization for Resource Allocation in Mobile-Edge Computation Offloading," IEEE Trans Wirel Commun, vol. 17, pp. 5506–5519, 2017, doi: 10.1109/TWC.2018.2845360.

[28] F. A. Campo, M. C. G. Neri, O. O. V Villegas, V. G. C. Sánchez, H. D. J. O. Dom\'\inguez, and V. G. Jiménez, "Auto-adaptive multi-layer perceptron for univariate time series classification," Expert Syst Appl, vol. 181, 2021.

[29] E. Pekel, M. Gul, E. Celik, and S. Yousefi, "Metaheuristic Approaches Integrated with ANN in Forecasting Daily Emergency Department Visits," Math Probl Eng, 2021, doi: 10.1155/2021/9990906.

[30] X. Liu, X. Lai, and L. Zhang, "A Hierarchical Missing Value Imputation Method by Correlation-Based K-Nearest Neighbors," pp. 486–496, 2019, doi: 10.1007/978-3-030-29516-5_38.