International Journal of Engineering and Computer Science Applications (IJECSA)
Vol. 3, No. 2, September 2024, pp. 61∼70
ISSN: 2828-5611
DOI: doi.org/10.30812/ijecsa.v3i2.4391

Enhancing Mental Illness Predictions: Analyzing Trends Using Multiple Linear Regression and Neural Network Backpropagation

Riosatria, Hairani Hairani, Anthony Anggrawan, Moch. Syahrir
Universitas Bumigora, Mataram, Indonesia

Article Info

Article history:

Received August 19, 2024
Revised August 28, 2024
Accepted September 5, 2024

Keywords:

Mental Illness Predictions
Analyzing Trends
Multiple Linear Regression
Neural Network Backpropagation

ABSTRACT

The increasing number of mental health cases caused by various factors such as social changes, economic pressures, and technological advancements has made it difficult to accurately predict the number of cases, hindering prevention and early intervention efforts. Therefore, developing more accurate, data-driven predictive models is necessary to improve the effectiveness of prevention and intervention. **This study aims** to develop a predictive model for the number of mental health cases using Multiple Linear Regression and Neural Network Backpropagation methods. The study employs **two predictive methods**, Multiple Linear Regression and Neural Network Backpropagation to forecast future trends in the number of mental health cases. **The findings reveal** that the Neural Network Backpropagation method provides more accurate predictions than Multiple Linear Regression in forecasting mental health case trends. Specifically, the Neural Network Backpropagation method resulted in an Mean Absolute Error (MAE) of 111.39 and a Mean Absolute Percentage Error (MAPE) of 1.77%, while the Multiple Linear Regression method produced an MAE of 115.24 and a MAPE of 1.83%. Thus, **the implication of this study** is that the Neural Network Backpropagation method can be utilized to predict trends in the number of mental health cases due to its ability to provide highly accurate predictions.

Corresponding Author:

Hairani Hairani, 087839793970
Universitas Bumigora, Mataram, Indonesia,
Email: hairani@universitasbumigora.ac.id

How to Cite: R. Riosatria, H. Hairani, A. Anggrawan, and M. Syahrir, Enhancing Mental Illness Predictions: Analyzing Trends Using Multiple Linear Regression and Neural Network Backpropagation, International Journal of Engineering and Computer Science Applications (IJECSA), vol. 3, no. 2, pp. 61 - 70, Sep. 2024. doi: 10.30812/ijecsa.v3i2.4391.

1. INTRODUCTION

Mental health disorders are a growing global health issue, with their prevalence increasing each year [1]. Various factors such as social changes, economic pressures [2], and technological advancements [3] contribute to the rise in mental health cases. This situation calls for efforts to predict trends in mental health cases so that prevention and intervention can be carried out on time [4]. However, the inability to accurately predict the number of mental health cases can hinder early prevention and intervention efforts [5]. Traditional approaches to prediction are often ineffective due to the complexity of mental health disorders, which are influenced by various factors that are difficult to measure and predict. Therefore, the development of more accurate, data-driven predictive models is urgently needed [6, 7].

One approach that can be used in data mining is to analyze complex data and identify hidden patterns [6]. This technique has great potential to improve prediction accuracy and support better decision-making in mental health. One task in data mining that can be used to address this problem is prediction. Some prediction methods that can be used are Linear Regression [8], Neural Network Backpropagation [9, 10], Autoregressive Integrated Moving Average (ARIMA) [11], Seasonal Autoregressive Integrated Moving Average (SARIMA) [12], Gated Recurrent Unit Networks (GRU) [13], and Long Short-Term Memory Network (LSTM) [14].

Previous research on related issues includes a study [15] that used the linear regression method to predict depression levels in final-year students, resulting in a Mean Absolute Error (MAE) of 3.84 and a Root Mean Square Error (RMSE) of 4.7. Another study [16] used linear regression to predict population growth rates, yielding fairly accurate predictions. A study [17] used Multiple Linear Regression to predict trends in Human Immunodeficiency Virus (HIV) cases, achieving an RMSE of 0.816. Additionally, a study [18] predicted the number of diabetes cases using linear regression, resulting in accurate predictions. Lastly, research [19] employed Support Vector Regression (SVR) to predict the number of narcotics users, with an RMSE of 169.53.

Previous studies have shown limitations or gaps, as they only used a single method, and none specifically focused on predicting the number of mental health cases. This creates an opportunity that this research aims to address. This study introduces **a different approach** from previous research by using Multiple Linear Regression and Neural Network Backpropagation methods to predict the number of mental health cases, **which has not been done before**. Therefore, this research **aims** to develop a prediction model for mental health cases using Multiple Linear Regression and Neural Network Backpropagation. The **main contribution** of this study is the development of a more accurate prediction model for estimating mental health cases, which can be utilized by researchers, healthcare practitioners, and policymakers for more effective planning and intervention.

2. RESEARCH METHOD

The research framework is illustrated in Figure 1. The study begins with collecting mental disorder data from a hospital over the past five years, from 2019 to 2023, amounting to 60 data points. The collected data is still raw and requires data processing to be ready for prediction. The data preprocessing technique used is data transformation. This process involves converting raw data into a ready and easy-to-analyze format. In this study, predictions of the number of mental disorders per month are made based on data from the past five years. For example, a prediction is made for data in the 3rd month (March) using data from the 1st month (January) and the 2nd month (February). In other words, the prediction for t+1 is based on data from t-1 and t, where t+1 is Y, t-1 is X2, and t is X1.The research framework is illustrated in Figure 1. The study begins with collecting mental disorder data from a hospital over the past five years, from 2019 to 2023, amounting to 60 data points. The collected data is still raw and requires data processing to be ready for prediction. The data preprocessing technique used is data transformation. This process involves converting raw data into a ready and easy-to-analyze format. In this study, predictions of the number of mental disorders per month are made based on data from the past five years. For example, a prediction is made for data in the 3rd month (March) using data from the 1st month (January) and the 2nd month (February). In other words, the prediction for t+1 is based on data from t-1 and t, where t+1 is Y, t-1 is X2, and t is X1.

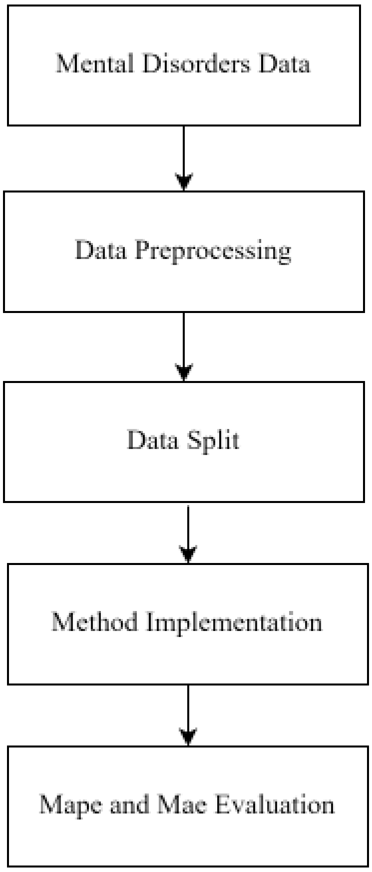

Figure 1. Research Framework

Linear Regression is a method used to model the relationship between a dependent variable and one or more independent variables. Single Linear Regression is used when there is only one predictor variable, while Multiple Regression is used when there are several predictor variables that are considered to influence the response variable. This study uses the Multiple Linear Regression method based on Equation (1), and the steps of this method in making predictions are shown in Figure 2. Here, Y is the dependent variable (response). $X_1 \ldots X_n$ are the independent variables (predictors). β_0 is the intercept. $\beta_1 \ldots \beta_n$ are the regression coefficients corresponding to each independent variable.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots + \beta_n X_n \tag{1}$$

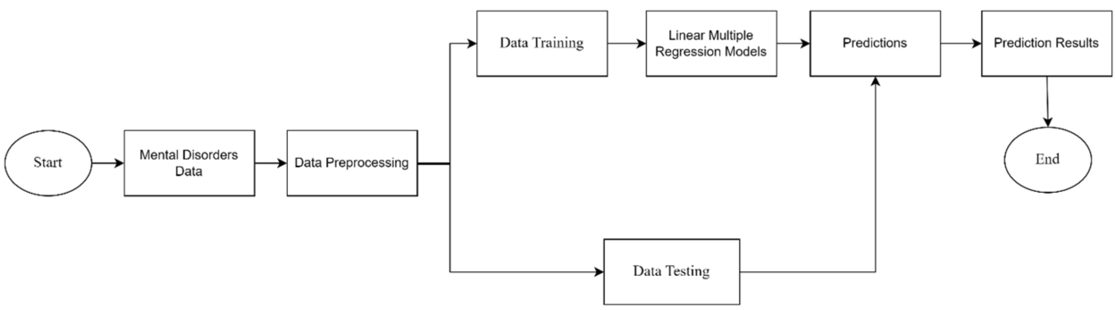

Figure 2. Prediction Flowchart with Multiple Linear Regression

Neural Network Backpropagation is a supervised learning algorithm that trains artificial neural networks such as backpropagation. Backpropagation is crucial for training neural networks, enabling the model to learn from data by minimizing prediction errors. The steps of the backpropagation method are shown in Figure 3. According to Figure 3, the steps are (1) Initialization of Weights and Biases; (2) Forward Propagation, which involves processing calculations between input weights and input attributes using Equation (2), followed by calculating the activation function using Equation (3); (3) Calculate the Loss function to measure the difference between the predicted results and the actual target values. The error calculation function uses Mean Squared Error (MSE) with Equation (4); (4) Backward Propagation aims to adjust weights and biases in the network to minimize the error function. It calculates the error derivative to each weight using the chain rule. The error gradient L to weight wij is calculated using Equation (5); (5) Update weights and biases using Equations (6) and (7); (5) Repeat until convergence.

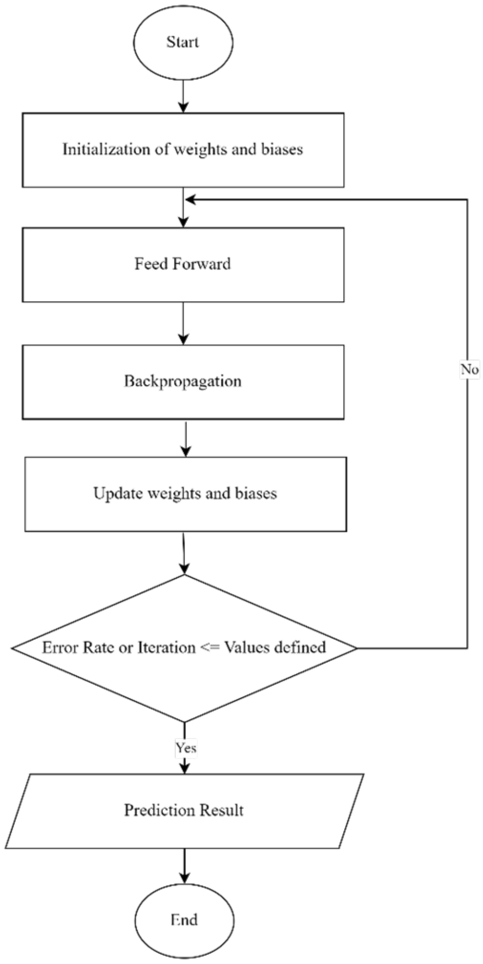

Figure 3. Prediction Flowchart with Neural Network Backpropagation

$$z_j = \sum_{i=1}^{n} w_{ij} .\, x_i + b_j \tag{2}$$

$$f(z) = max(0, z) \tag{3}$$

$$L = \frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}i)^2 \tag{4}$$

$$\frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial \widehat{y_i}} \cdot \frac{\partial \widehat{y_i}}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}} \tag{5}$$

$$w_{ij} \leftarrow \; w_{ij} - \eta \,.\, \frac{\partial L}{\partial w_{ij}} \tag{6}$$

$$b_j \leftarrow \; b_j - \eta \,.\, \frac{\partial L}{\partial b_j} \tag{7}$$

Where z_j is the weighted sum for neuron j. w_{ij} are the weights connecting input x_i to neuron j. b_j is the bias for neuron j. x_i are the input features. Where y_i is the actual target value. $\hat{y}_i$ is the predicted value. m is the number of samples. The gradient concerning the bias b_j is similarly computed. η is the learning rate, a small positive number that controls the step size of the update. The results of predictions using the Multiple Linear Regression and Neural Network Backpropagation methods are evaluated based on the Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). MAE is the average of all absolute errors between the actual values and the values predicted by the model. MAE measures the average error in the same units as the data, using Equation (7). Here, n is the number of observations. y_i is the actual value in the i-th observation. i is the predicted value in the i-th observation. $|y_1 - \hat{y}_i|$ is the absolute error between the actual value and the prediction. MAPE is the average of absolute errors expressed as a percentage of the actual values. MAPE helps measure prediction errors using Equation (8). Here, n is the number of observations. y_i is the actual value in the i-th observation. $\hat{y}_i$ is the predicted value in the i-th observation. $\left|\frac{y_i - \hat{y}_i}{y_i}\right|$ is the absolute error relative to the actual value, expressed as a percentage.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_1 - \hat{y}_i| \tag{8}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{9}$$

3. RESULT AND ANALYSIS

The research begins with collecting mental disorder data from a hospital over the past five years, from 2019 to 2023, with 60 data points (see Table 1). The collected data is raw and requires data processing for prediction. The data preprocessing technique used is data transformation. This process involves converting raw data into a ready and easy-to-analyze format. In this study, predictions of mental disorders per month are made based on five years of data. For example, predictions are made for data in the 3rd month (March) using data from the 1st month (January) and the 2nd month (February). In other words, the prediction for t+1 is based on data from t-1 and t, where t+1 is Y, t-1 is X2, and t is X1. The data transformation results, ready for the prediction process, are shown in Table 2.

Table 1. Dataset of Mental Disorders over 5 Years

Month	Year				
	2019	**2020**	**2021**	**2022**	**2023**
January	3299	2925	3539	5013	6285
February	3353	2949	3488	5062	6290
March	3388	2989	3532	5054	6282
April	3274	2925	3522	5022	6299
May	3349	2935	3477	5043	6292
June	3387	2968	3548	5017	6288
July	3331	2931	3488	5008	6281
August	3380	2957	3505	5058	6304
September	3330	2935	3538	5039	6295
October	3318	2967	3523	5029	6296
November	3376	2929	3494	5024	6293
December	3447	2953	3485	5082	6300

Table 2. Data Transformation Result

No	X1	X2	Y
1	3299	3353	3388
2	3353	3388	3274
3	3388	3274	3349
4	3274	3349	3387
5	3349	3387	3331
	.	...	....
54	6288	6281	6304
55	6281	6304	6295
56	6304	6295	6296
57	6295	6296	6293
58	6296	6293	6300

After transforming the data, it is divided into training and testing sets. In this study, 80% of the data (46 data points) is used for training, and 20% (12 data points) is used for testing. The training data is used to train the model to recognize patterns, allowing the creation of both the multiple linear regression and neural network models. The models are then used to predict the testing data. The multiple linear regression model created from the training data is Y = **112.25 + 0.0181 X1 + 0.96 X2**. The Neural Network Backpropagation method for predicting mental health case trends uses the architecture shown in Figure 4. This architecture includes three hidden layers with different weights in each layer. The parameters used in this Neural Network Backpropagation method are epochs = 500, batch_size = 10, optimizer = Adam, learning_rate = 0.05, loss = 'mean_absolute_error,' and activation functions 'relu' and 'linear.'

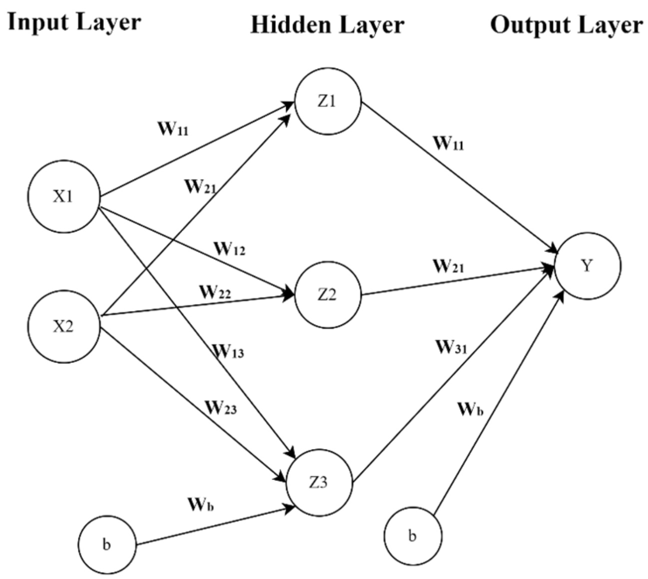

Figure 4. Neural Network Backpropagation Method Architecture

The prediction results for each method are shown in Table 3 for Multiple Linear Regression and Table 4 for Neural Network Backpropagation. Table 3 indicates that the Multiple Linear Regression method has an average Absolute Error of 115.24 and an average MAPE of 1.83%. This shows that the predictions made with this method are quite accurate, although some predictions, especially the first one, have a higher error, with a MAPE of 18.99%. On the other hand, Table 4 shows that the Neural Network Backpropagation method yields a slightly lower average Absolute Error of 111.39 and an average MAPE of 1.77%. This method demonstrates more stable and slightly better performance than Multiple Linear Regression. Although the first prediction also has a high error, it is still lower than Multiple Linear Regression (19.05%). The findings of this study indicate that the Neural Network Backpropagation method provides more accurate predictions compared to Multiple Linear Regression for forecasting mental health case trends (see Figure 5). Although both methods exhibit low error on some data, the Neural Network Backpropagation method demonstrates a greater ability to maintain lower error and more consistent error distribution compared to Multiple Linear Regression. This is supported by studies [20–22], which state that the Neural Network Backpropagation method yields better predictive results than Multiple Linear Regression.

Table 3. Performance Results of Linear Regression Method on Testing Data

Actual	Predicted	Mean Absolute Error	MAPE (%)
6285	5091.34	1193.66	18.99
6290	6249.53	40.47	0.64
6282	6276.09	5.91	0.09
6299	6268.49	30.51	0.48
6292	6284.69	7.31	0.12
6288	6278.27	9.73	0.15
6281	6274.29	6.71	0.11
6304	6267.49	36.51	0.58
6295	6289.49	5.51	0.09
6296	6281.24	14.76	0.23
6293	6282.04	10.96	0.17
6300	6279.18	20.82	0.33
Average		115,24	1,83%

Table 4. Performance Results of Neural Network Backpropagation Method on Testing Data

Actual	Predicted	Mean Absolute Error	MAPE (%)
6285	5087.804	1197.196	19.04846
6290	6267.638	22.36182	0.355514
6282	6297.866	15.86572	0.252558
6299	6290.133	8.866699	0.140764
6292	6306.62	14.62012	0.23236
6288	6300.12	12.12012	0.19275
6281	6296.054	15.05371	0.239671
6304	6289.112	14.88818	0.23617
6295	6311.498	16.49756	0.262074
6296	6303.165	7.164551	0.113795
6293	6303.955	10.95459	0.174076
6300	6301.037	1.036621	0.016454
Average		111,39	1,77%

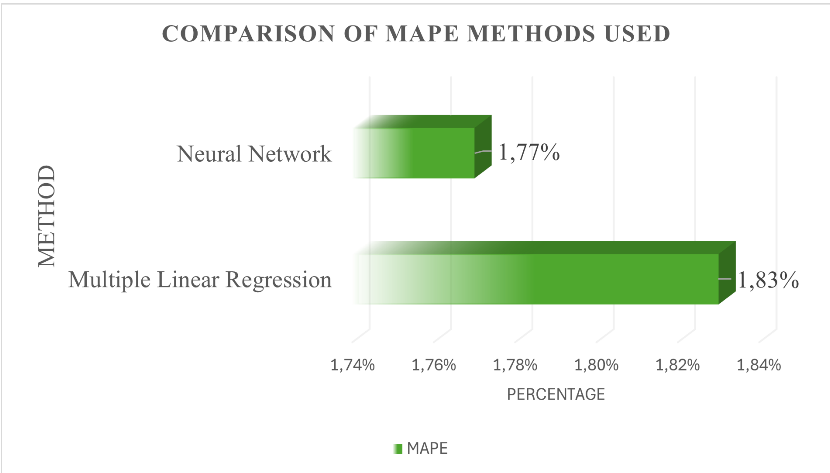

Figure 5. Comparison of MAPE Methods Used

4. CONCLUSION

The results of this study indicate that the Neural Network Backpropagation method produces more accurate predictions than Multiple Linear Regression for forecasting mental health case trends. While both methods show low error rates, Neural Network Backpropagation has a lower average Absolute Error and a more consistent error distribution compared to Multiple Linear Regression. This demonstrates that Neural Network Backpropagation excels in maintaining prediction accuracy. Future research is recommended to focus on tuning the parameters of both methods used in this study to achieve a lower MAPE compared to the current results.

5. DECLARATIONS

AUTHOR CONTIBUTION
All authors contributed to the writing of this article.
FUNDING STATEMENT
-
COMPETING INTEREST
No conflict of interest

REFERENCES

[1] Y. Wu, L. Wang, M. Tao, H. Cao, H. Yuan, M. Ye, X. Chen, K. Wang, and C. Zhu, "Changing trends in the global burden of mental disorders from 1990 to 2019 and predicted levels in 25 years," *Epidemiology and Psychiatric Sciences*, vol. 32, pp. 1–9, 2023. [Online]. Available: https://www.cambridge.org/core/product/identifier/S2045796023000756/type/journal_article

[2] R. Jenkins, F. Baingana, R. Ahmad, D. McDaid, and R. Atun, "Social, economic, human rights and political challenges to global mental health," *Mental Health in Family Medicine*, vol. 8, no. 2, pp. 87–96, Jun. 2011. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3178190/

[3] L. Foulkes and J. L. Andrews, "Are mental health awareness efforts contributing to the rise in reported mental health problems? A call to test the prevalence inflation hypothesis," *New Ideas in Psychology*, vol. 69, p. 101010, Apr. 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0732118X2300003X

[4] J. B. Kirkbride, D. M. Anglin, I. Colman, J. Dykxhoorn, P. B. Jones, P. Patalay, A. Pitman, E. Soneson, T. Steare, T. Wright, and S. L. Griffiths, "The social determinants of mental health and disorder: evidence, prevention and recommendations," *World Psychiatry*, vol. 23, no. 1, pp. 58–90, Feb. 2024. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/wps.21160

[5] M. Colizzi, A. Lasalvia, and M. Ruggeri, "Prevention and early intervention in youth mental health: is it time for a multidisciplinary and trans-diagnostic model for care?" *International Journal of Mental Health Systems*, vol. 14, no. 1, pp. 1–14, Dec. 2020. [Online]. Available: https://ijmhs.biomedcentral.com/articles/10.1186/s13033-020-00356-9

[6] M. M. Islam, S. Hassan, S. Akter, F. A. Jibon, and M. Sahidullah, "A comprehensive review of predictive analytics models for mental illness using machine learning algorithms," *Healthcare Analytics*, vol. 6, p. 100350, Dec. 2024. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2772442524000522

[7] L. J. Vaickus, D. A. Kerr, J. M. Velez Torres, and J. Levy, "Artificial Intelligence Applications in Cytopathology," *Surgical Pathology Clinics*, vol. 17, no. 3, pp. 521–531, Sep. 2024. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1875918124000229

[8] A. Anggrawan, H. Hairani, and N. Azmi, "Prediksi penjualan produk unilever menggunakan metode regresi linear," *Jurnal Bumigora Information Technology (BITe)*, vol. 4, no. 2, pp. 123–132, Dec. 2022. [Online]. Available: https://journal.universitasbumigora.ac.id/index.php/bite/article/view/2416

[9] A. Anggrawan, H. Hairani, and M. A. Candra, "Prediction of Electricity Usage with Back-propagation Neural Network," *International Journal of Engineering and Computer Science Applications (IJECSA)*, vol. 1, no. 1, pp. 9–18, Mar. 2022. [Online]. Available: https://journal.universitasbumigora.ac.id/index.php/IJECSA/article/view/1722

[10] R. Ramadhanti, H. Hairani, and M. Innuddin, “Electric Vehicle Sales-Prediction Application Using Backpropagation Algorithm Based on Web,” *International Journal of Engineering and Computer Science Applications (IJECSA)*, vol. 2, no. 2, pp. 73–80, Sep. 2023. [Online]. Available: https://journal.universitasbumigora.ac.id/index.php/IJECSA/article/view/3388

[11] H. H. M. Hatta, F. M. Daud, and N. Mohamad, “An Application of Time Series ARIMA Forecasting Model for Predicting the Ringgit Malaysia-Dollar Exchange Rate,” *Journal of Data Analysis*, vol. 1, no. 1, pp. 42–48, Sep. 2018. [Online]. Available: http://www.jurnal.unsyiah.ac.id/JDA/article/view/11884

[12] M. Multiningsih, E. Siswanah, and M. Saleh, “Forecasting the Number of Ship Passengers with SARIMA Approach (A Case Study: Semayang Port, Balikpapan City),” *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 6, no. 4, pp. 1060–1080, Oct. 2022. [Online]. Available: http://journal.ummat.ac.id/index.php/jtam/article/view/10211

[13] M. Ridwan, K. Sadik, and F. M. Afendi, “Comparison of ARIMA and GRU Models for High-Frequency Time Series Forecasting.” *Scientific Journal of Informatics*, vol. 10, no. 3, pp. 389–400, Aug. 2023, number: 3. [Online]. Available: https://journal.unnes.ac.id/nju/sji/article/view/45965

[14] A. W. Saputra, A. P. Wibawa, U. Pujianto, A. B. P. Utama, and A. Nafalski, “LSTM-based Multivariate Time-Series Analysis: A Case of Journal Visitors Forecasting,” *ILKOM Jurnal Ilmiah*, vol. 14, no. 1, pp. 57–62, Apr. 2022, number: 1. [Online]. Available: https://jurnal.fikom.umi.ac.id/index.php/ILKOM/article/view/1106

[15] K. F. Khufa and M. Murinto, “Prediksi Kasus Tingkat Depresi Mahasiswa Semester Akhir Menggunakan Regresi Linear Sederhana,” *INTEK : Jurnal Informatika dan Teknologi Informasi*, vol. 7, no. 1, pp. 1–6, May 2024, number: 1. [Online]. Available: https://jurnal.umpwr.ac.id/index.php/intek/article/view/4978

[16] P. Purwadi, P. S. Ramadhan, and N. Safitri, “Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang,” *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika dan Komputer)*, vol. 18, no. 1, pp. 55–61, Feb. 2019. [Online]. Available: https://ojs.trigunadharma.ac.id/index.php/jis/article/view/104

[17] R. N. Azizah, U. K. Nisak, and U. Indahyanti, “Analisis Jumlah Prediksi Penyebaran HIV/AIDS di Kabupaten Sidoarjo menggunakan Metode Multiple Linier Regression,” *Physical Sciences, Life Science andEngineering*, vol. 1, no. 1, pp. 1–11, Jan. 2024. [Online]. Available: https://digital-science.pubmedia.id/index.php/pslse/article/view/163

[18] N. S. Niko, A. Rahman, D. Marini Umi Atmaja, and A. Basri, “Prediksi Penyakit Diabetes Untuk Pencegahan Dini Dengan Metode Regresi Linear,” *Bulletin of Information Technology (BIT)*, vol. 4, no. 3, pp. 313–219, Sep. 2023. [Online]. Available: https://journal.fkpt.org/index.php/BIT/article/view/739

[19] T. A. Setiawan, A. Ilyas, and A. Arochman, “Komparasi Model Prediksi Penanganan Kasus Narkotika,” *IC-Tech*, vol. 17, no. 1, pp. 42–48, Apr. 2022. [Online]. Available: https://ejournal.stmik-wp.ac.id/index.php/ictech/article/view/239

[20] X. Chen, H. Zheng, H. Wang, and T. Yan, “Can machine learning algorithms perform better than multiple linear regression in predicting nitrogen excretion from lactating dairy cows,” *Scientific Reports*, vol. 12, no. 1, pp. 1–13, Jul. 2022. [Online]. Available: https://www.nature.com/articles/s41598-022-16490-y

[21] H. Kablay and V. Gumbo, “Comparison of Multiple Linear Regression and Neural Network Models in Bank Performance Prediction in Botswana,” *Journal of Mathematics and Statistics*, vol. 17, no. 1, pp. 88–95, Jan. 2021. [Online]. Available: https://thescipub.com/abstract/10.3844/jmssp.2021.88.95

[22] M. Li and J. Wang, “An Empirical Comparison of Multiple Linear Regression and Artificial Neural Network for Concrete Dam Deformation Modelling,” *Mathematical Problems in Engineering*, vol. 2019, no. 1, pp. 1–13, Jan. 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1155/2019/7620948

[This page intentionally left blank.]