Application of Boosting Technique with C4.5 Algorithm to Reduce the Classification Error Rate in Online Shoppers Purchasing Intention

Ading Sunarto[1], Putri Nilam Kencana[2], Baliyah Munadjat[3], Iriana Kusuma Dewi[4], Ali Zaenal Abidin[5] and Robbi Rahim[6*]

[1]Faculty of Economics and Business, Universitas Pamulang, Tangerang Selatan, Indonesia. dosen02153@unpam.ac.id, Orcid: https://orcid.org/0000-0002-6036-7926

[2]Faculty of Economics and Business, Universitas Pamulang, Tangerang Selatan, Indonesia. dosen01877@unpam.ac.id, Orcid: https://orcid.org/0000-0001-6744-7550

[3]Faculty of Economics and Business, Universitas Pamulang, Tangerang Selatan, Indonesia. dosen02162@unpam.ac.id, Orcid: https://orcid.org/0009-0009-1247-7192

[4]Faculty of Economics and Business, Universitas Pamulang, Tangerang Selatan, Indonesia. dosen01729@unpam.ac.id, Orcid: https://orcid.org/0009-0005-4580-2140

[5]Faculty of Economics and Business, Universitas Pamulang, Tangerang Selatan, Indonesia. alizaenalabidin@unpam.ac.id, Orcid: https://orcid.org/0000-0003-2238-6103

[6*]Department of Management, Sekolah Tinggi Ilmu Manajemen Sukma, Medan, Indonesia. usurobbi85@zoho.com, Orcid: https://orcid.org/0000-0001-6119-867X

Received: March 02, 2023; Accepted: May 08, 2023; Published: June 30, 2023

Abstract

The goal of this research is to use the boosting technique with the C4.5 algorithm to lower the number of wrong classifications. One of the techniques for boosting is called Adaboost. It can balance the class by giving more weight to the level of classification error, which can change how the data is spread out. The Online Shoppers Purchasing Intention dataset from the UCI Machine Learning Repository is used to test this method. It has 12330 records and 18 attributes, with a label being one of those attributes. In this study, the results were looked at with the help of software called RapidMiner 9.10. Before the processing stage, the dataset is split into two parts: training and testing, with a 90:10 and 80:20 split, respectively. The results of the experiments with the C4.5 algorithm and the boosting technique at a ratio of 90:10 (accuracy of 95.07 percent and AUC of 0.966) and a ratio of 80:20 (accuracy of 95.07 percent and AUC of 0.966). While the experimental results of the C4.5 algorithm without the boosting technique at a ratio of 90:10 (accuracy of 89.02 percent and AUC of 0.845) and a ratio of 80:20 (accuracy of 88.1 percent and AUC of 0.845). Based on the results of the comparison, boosting on the C4.5 algorithm is much better than the standard C4.5 algorithm, with an average increase of 5.76 percent. We can say that using the boosting technique with the C4.5 algorithm can improve high accuracy and lower the amount of classification error.

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 14, number: 2 (June), pp. 01-11. DOI: *10.58346/JOWUA.2023.I2.001*

*Corresponding author: Department of Management, Sekolah Tinggi Ilmu Manajemen Sukma, Medan, Indonesia.

Keywords: Adaboost, Boosting, Datamining, Classification, Improving Accuracy.

1 Introduction

Data mining is the act of discovering patterns that have the potential to be beneficial for managing massive databases (Syafar et al., 2015; Windarto & Herawan, 2022). Data mining classification approaches, such as the C4.5 algorithm (Aziz & Lawi, 2022; Hasanah et al., 2020), Naive Bayes (Kewsuwun & Kajornkasirat, 2022; Majeed et al., 2021), Neural Network (Hardiyanto et al., 2021; Perdana et al., 2021; Windarto et al., 2018) and K-NN (Atalaya et al., 2022; Varzaneh et al., 2022), are widely utilized by researchers in the solution of problems. Classification approach is also among the most researched algorithms. Neural network (NN) is an algorithm that can produce non-linear predictions, has the ability to accept errors, and is powerful in parallel processing (Deng et al., 2021; Syafar et al., 2014), however in this situation the NN approach has limitations such as over-fitting, convergence, and requires training with huge amounts of data (Buchori et al., 2016; Fasihah et al., 2023; Lin & Wong, 2016; Muhdi et al., 2019). While the C4.5 approach may solve NN problems by addressing overfitting, enhancing computing efficiency, and dealing with continuous characteristics. The C4.5 algorithm In general has a high success rate if the dataset is sufficiently balanced (Katrina et al., 2019; Pratiwi et al., 2020; Sundari et al., 2019; Waluyo et al., 2020). While the success rate is low when the dataset has entropy and gain on class imbalance (Kaur & Gosain, 2022).

Standard classifications typically perform poorly on imbalanced datasets(Ma'arif et al., 2021; Yaman et al., 2021) because they are designed to generalize from training data and are based on the simplest hypotheses that best suit the data. Boosting algorithm (Yaman et al., 2021) can considerably facilitate the reduction of class imbalance in distribution. One of the boosting algorithms is adaboost, which is an ensemble learning method that can reduce variance (Silva et al., 2021), due to the bias impact of the ensemble average reducing variance from a series of classifications (Silhavy et al., 2019). Following are research that employ boosting strategies to improve accuracy, as was done Sonavane & Sonar in the categorization of brain tumor detection. The experimental results for 155 MRI pictures demonstrated 100 percent accuracy. Then, research was undertaken on the classification of Parkinsonian disorders(Tan et al., 2018). The experimental findings demonstrate that the AdaBoost model is superior to KNN and LIBSVM. The classification accuracy of the model is 97.097 percent, or 57.807 percent greater than KNN. Additionally, the computing time is 2.72 times faster than LIBSVM. Considering this, the objective of this study is to apply the Boosting technique to the C4.5 Algorithm in order to reduce its error rate, hence enabling it to perform well on unbalanced datasets.

2 Research Methodology

During the research, a computer with the specs Intel(R) Core (TM) i7-4980HQ 2.80 GHz, 16 GB RAM, and Windows 10 Pro was used to help. For the process of analysis, the Rapid Miner Studio 9.10 software was used. In the Online Shoppers Purchasing Intention(Subagja, 2023; Suherlan, 2023) dataset from the UCI Machine Learning Repository (https://archive.ics.uci.edu/ml/datasets.php, the boosting method was used with the C4.5 algorithm to reduce the level of classification error. This dataset is also used in research on multilayer perceptron and LSTM recurrent neural networks (Sakar et al., 2019) (Sakar et al., 2019). Here are details from the dataset that was used, as shown in Table 1 below.

Table 1: Sets of Data Used for Research

Type	Record	Missing Values
Classification	12330	N/A
Attribute Characteristics	Dataset Characteristics	Class
Integer, Real	Multivariate	18

Source: UCI Machine Learning Repository

Based on the dataset in Table 1, which has 10 numeric attributes and 8 categorical attributes, here are the details of Attribute Information. The name of the class is taken from the "Revenue" attribute. This is shown in Tables 2 and 3, which look like this.

Table 2: Numerical Features Used in the User Behavior Analysis Model

Feature name	Feature description	Min. value	Max. value	Special Day
Administrative	Number of pages about account management that the visitor looked at.	0	27	3,32
Administrative duration	Total amount of time (in seconds) that the visitor spent on pages about account management.	0	3398	176,70
Informational	Number of pages a visitor to a shopping site looked at to find out about the site, how to get in touch with it, and where it is.	0	24	1,26
Informational duration	Total time (in seconds) that a visitor spent on pages with information.	0	2549	140,64
Product related	Number of times a visitor went to a page about a product.	0	705	44,45
Product related duration	Total amount of time (in seconds) that a visitor spent on pages about products.	0	63,973	1912,25
Bounce rate	Average value of the number of pages a visitor stays on before leaving.	0	0,2	0,04
Exit rate	Average exit rate for the pages the visitor looked at.	0	0,2	0,05
Page value	The average value of the pages a visitor looks at.	0	361	18,55
Special day	How close the visit to the site is to a special day.	0	1,0	0,19

Table 3: Categorical Features Used in the User Behavior Analysis Model

Feature name	Feature description	Number of categorical values
Operating Systems	System of operation of the visitor.	8
Browser	The visitor's browser.	13
Region	From what part of the world the visitor started the session.	9
Traffic Type	How the visitor got to the website (traffic source) (e.g., banner, SMS, direct).	20
Visitor Type	There are three types of visitors: "New Visitor," "Returning Visitor," and "Other."	3
Weekend	Boolean value that shows whether the visit date is a weekend or not.	2
Month	Month value of the date of the visit.	12
Revenue	Class label that shows whether the visit ended with a transaction or not.	2

In this study, boosting (Bisri & Wahono, 2015) is suggested as a way to reduce the number of wrong classifications made by the c4.5 algorithm. The adaboost method (Setiawan, 2015) is one of the boosting techniques. It can lower the classification error rate by taking care of class imbalances in the classification. Accuracy will measure the data that comes out of the validation process. Figure 1 shows the framework model for the proposed method.

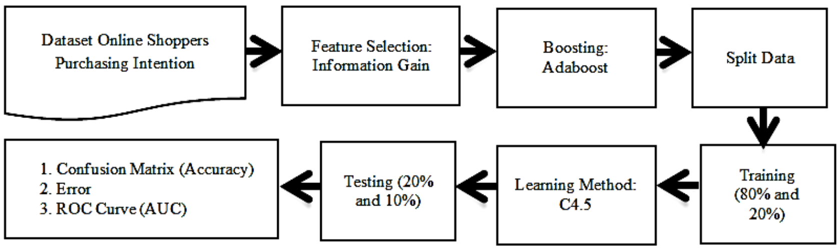

Figure 1: Proposed Model

Figure 1 shows that the UCI Machine Learning Repository (Online Shoppers Purchasing Intention) was used as the input dataset. It has 12330 records and 18 attributes, one of which is a label. With the help of Rapid Miner Studio 9.10 software, the dataset is then turned into an excel (xls) file. By adding the Boosting technique to the C4.5 Algorithm, the classification error rate can be lowered by balancing the dataset and making a model that uses the best parts of both. In a ratio of 80:20 and 90:10, the dataset is also split into training datasets and testing datasets (Pozi, M.S.M., 2020). This is also true for the C4.5 Algorithm without the Boosting technique. Validation of the results of the analysis will produce data that can be measured by accuracy, error, and the ROC Curve (AUC).

According to Wang (Wang & Feng, 2020), the following is a weighting method for the adaboost algorithm. It is shown in Figure 2 as a story:

Algorithm 1 Traditional Adaboost Algorithm

Input: training sample set S, label set Y, weak classifier h, number of iterations T.

Output: the final classifier H.

1. **for** $i = 1{:}1{:}M$
2. initialize the weight of sample s_i: $D_1(i) = 1/M$;
3. **end for**
4. **for** $t = 1{:}\ 1{:}\ T$
5. select a sample subset X from the training set S, train X using h to obtain a weak classifier h_t and calculate the weighted error rate ε_t:

$$\varepsilon_t = \sum_{s_i \in X, h_t(s_i) \neq Y(s_i)} D_t\,(i) \qquad (1)$$

where $h_t(s_i)$ is the predicted label of s_i using classifier h_t, $Y(s_i)$ is the real label of s_i.

6. calculate the classifier weight of h_t:

$$\alpha_t = \frac{1}{2} \ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right) \qquad (2)$$

7. update the weight of each sample in S:
8. **for** $i = 1{:}1{:}M$

$$D_{t+1}\,(i) = \frac{D_t\,(i)}{Z_t} \times \begin{cases} e^{-\alpha_t}, & h_t\,(s_i) = y_i \\ e^{\alpha_t}, & h_t\,(s_i) \neq y_i \end{cases} \qquad (3)$$

where Z_t is the normalization factor, which is defined as follows

$$Z_t = \sum_{s_j \in S} D_t\,(j)\, e^{-\alpha_t y_j h_t(s_j)} \qquad (4)$$

9. **end for**
10. **end for**
11. get the final strong classifier H: for sample s, its class label $H(s)$ is

$$H\,(s) = \mathrm{sign}\left(\sum_{t=1}^{T} \alpha_t h_t\,(s)\right) \qquad (5)$$

Figure 2: Adaboost Algorithm

3 Results and Discussion

Experiments were conducted using Rapidminer 9.10 software to apply the Boosting strategy to the C4.5 algorithm in an effort to reduce the error rate. This study was conducted in multiple phases. The initial step entails the preparation of the Online Shoppers Purchasing Intention dataset. Then, construct an experimental model similar to Figure 2 below.

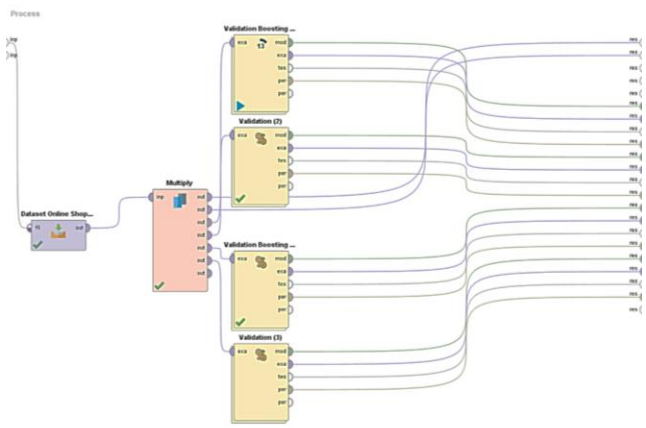

Figure 3: Rapid Miner Model Used

The overall model of the conducted experiments is depicted in Figure 2. In the model, four simultaneous experiments are conducted: the application of the Boosting technique to the C4.5 algorithm (B.C45) with split data (80:20) and (90:10), and the application of the normal C4.5 algorithm with split data (80:20) and (90:10). (90:10). Below is a comprehensive explanation of the typical B.C45 and C4.5 engineering models, as depicted in Figures 3 and 4, respectively.

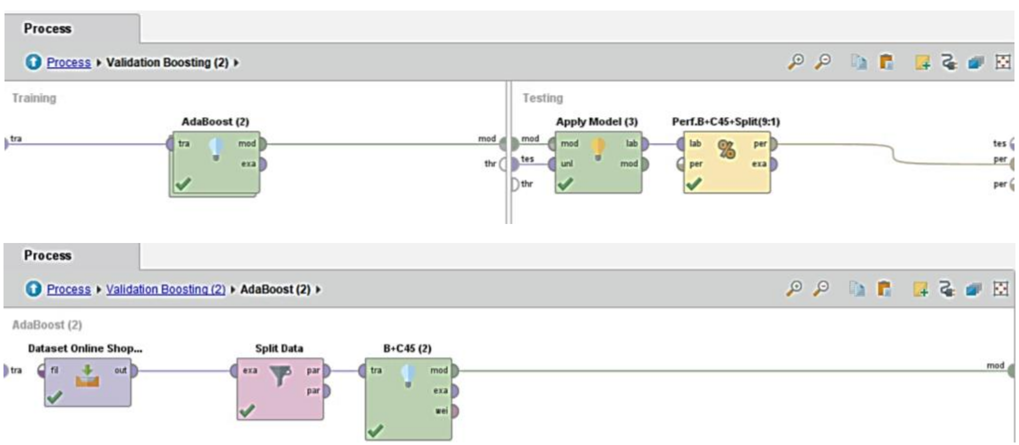

Figure 4: Rapid Miner Model on B.C45

After the data preprocessing stage, which consists of two steps (attribute deletion and label transformation), the feature selection stage comes next. Then use the boosting method, which is called the adaboost algorithm. Before the new dataset is trained and tested, the data is first split in half, with 90% and 80% of the data being used for training and 20% and 10% of the data being used for testing. Through the apply model, the results of the training model will be simulated and put to the test. The results were given as accuracy, error, and the area under the curve (AUC). While the C.45 model without boosting only differs in the use of the adaboost algorithm.

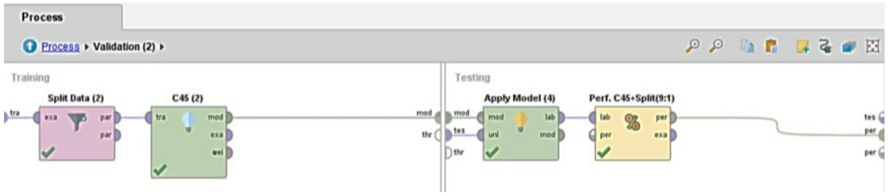

Figure 5: Rapid Miner Model on C4.5 without Boosting

Identify all models in the second step. Table 4 and the next graph show the results of each experiment with the BC45 and C.45 methods without boosting.

Table 4: Model Experimental Results

Parameter	**Accuracy**	**Error**
C4.5 (Split Data 80:20)	88.78%	11.22%
C4.5 (Split Data 90:10)	89.02%	10.98%
Boost+C4.5 (Split Data 80:20)	94.24%	5.76%
Boost+C4.5 (Split Data 90:10)	95.07%	4.93%

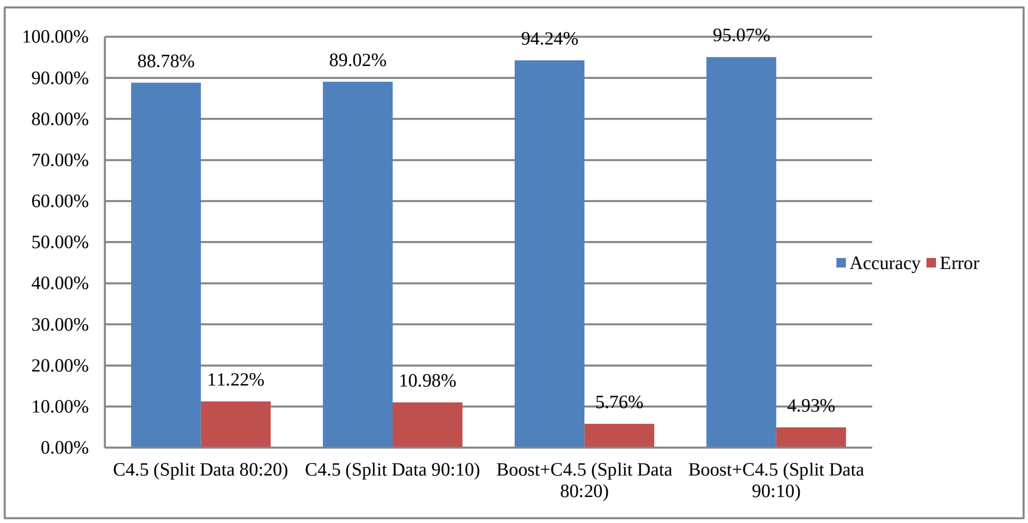

Figure 6: Comparison Chart of all Models

According to the confusion matrix test results in Table 4, the C4.5 algorithm without boosting has an accuracy of 88.78 percent and 89.02 percent, respectively, with a ratio of 80:20 and 90:10. While the C4.5 algorithm with boosting (adaboost technique) achieves a greater level of accuracy, 94.24 percent and 95.07 percent, respectively. In comparison to the regular C4.5 algorithm, the application of boosting on the C4.5 algorithm yields an average improvement of 5.76 percentage points.

Table 5: Comparison of all Models' AUC Values

Parameter	**AUC**
C4.5 (Split Data 80:20)	0.855
C4.5 (Split Data 90:10)	0.845
Boost+C4.5 (Split Data 80:20)	0.97
Boost+C4.5 (Split Data 90:10)	0.966

In addition to the results of the confusion matrix test, the AUC graph shows how well the model is made. Without using a boosting technique, the C4.5 algorithm has gotten AUC values of 0.855 and 0.845. Using the AUC value to analyze the prediction results of the classification, the classification results are of good value (good classification). While the algorithm with the boosting technique gives AUC values of 0.97 and 0.966. This is an increase of 0.118 from the previous and changes the results of the previous classification from good classification to excellent classification. This is shown on the graph in Figure 6 below.

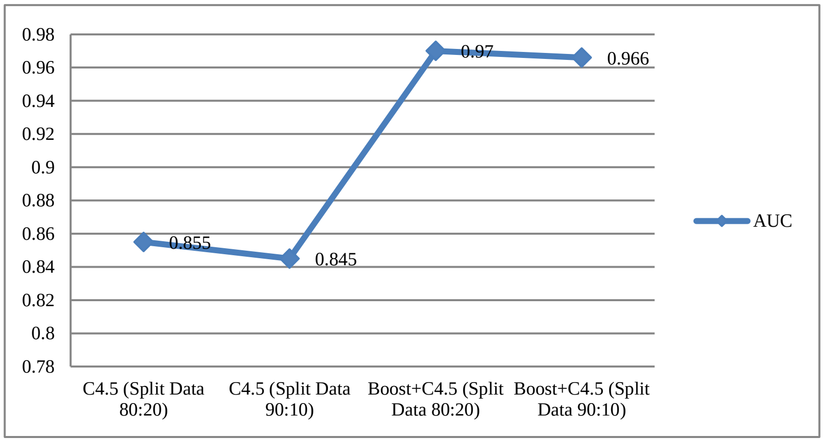

Figure 7: Graph Comparing the AUC Values for Each Model

Following is an illustration of the AUC curve generated by the adaboost method utilizing the boosting technique.

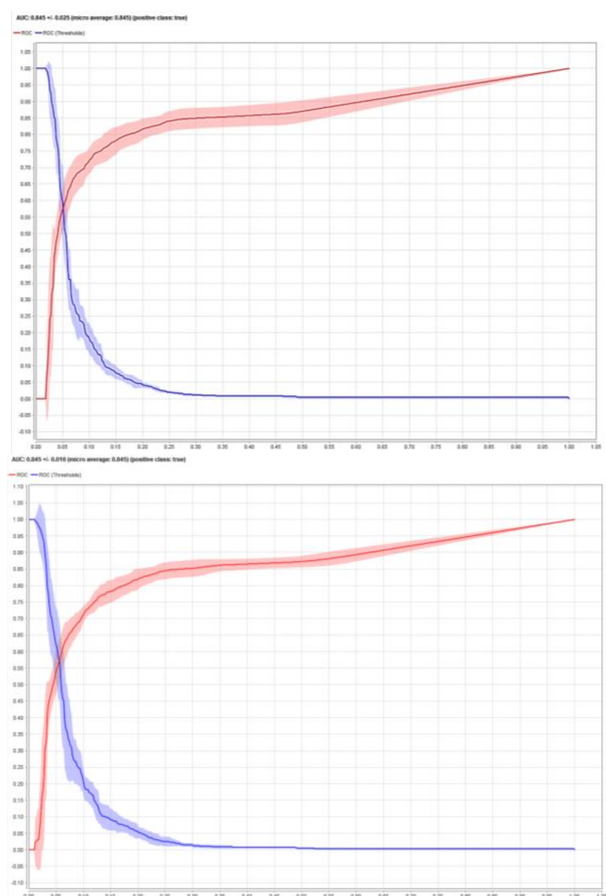

Figure 8: Results of the Boosting Method on the AUC Curve

The boosting strategy, namely the adaboost method, can maximize the classification results among all the tests conducted using the C4.5 classification. The results of studies conducted to assess the accuracy and AUC values outperform the C4.5 model normal.

4 Conclusion

When the boosting technique was used with the C4.5 algorithm on the Online Shoppers Purchasing Intention dataset, it outperformed all other results in terms of both accuracy and area under the curve (AUC). This was done by using a sample of 12330 records that were split into training and testing data with a ratio of 80:20 and 90:10. When compared to using the C4.5 algorithm without boosting, the accuracy value was 5.76 percent higher in all of the experiments. From the AUC, there is also an average increase of 0.118 compared to using the C4.5 algorithm without boosting. We can say that applying the boosting technique to the C4.5 algorithm can increase high accuracy and lower the level of classification error.

References

[1] Atalaya, O.F.C., Morales-Romero, G., Quispe-Andía, A., Caycho-Salas, B., Auqui-Ramos, E. K., Ramos-Salazar, P., & Palacios-Huaraca, C. (2022). K-NN supervised learning algorithm in the predictive analysis of the quality of the university administrative service in the virtual environment. *Indonesian Journal of Electrical Engineering and Computer Science*, *25*(1), 521–528.

[2] Aziz, F., & Lawi, A. (2022). Increasing electrical grid stability classification performance using ensemble bagging of C4.5 and classification and regression trees. *International Journal of Electrical and Computer Engineering*, *12*(3), 2955–2962.

[3] Bisri, A., & Wahono, R.S. (2015). Penerapan Adaboost untuk Penyelesaian Ketidakseimbangan Kelas pada Penentuan Kelulusan Mahasiswa dengan Metode Decision Tree. *Journal of Intelligent Systems*, *1*(1), 27–32.

[4] Buchori, A., Setyosari, P., Dasna, I.W., & Ulfa, S. (2016). Developing character building learning model using mobile augmented reality on elementary school student in Central Java. *Global Journal of Pure and Applied Mathematics*, *12*(4), 3433-3444.

[5] Deng, Y., Zhou, X., Shen, J., Xiao, G., Hong, H., Lin, H., Wu, F., & Liao, B.Q. (2021). New methods based on back propagation (BP) and radial basis function (RBF) artificial neural networks (ANNs) for predicting the occurrence of haloketones in tap water. *Science of the Total Environment*, *772*.

[6] Fasihah, E., Hidayah, N., & Mubarokah, A. (2023). Validation Analysis of Learning Development on Optic Material with Core Learning Model. *SAGA: Journal of Technology and Information System*, *1*(1), 1–4.

[7] Hardiyanto, D., Kristiyana, S., Sartika, D.A., Ahmad, I., & Al_Barazanchi, I. (2021). Training of dual band HT-based emergency communication tools on member of KTB BPBD, yogyakarta. *Jurnal Pengabdian Dan Pemberdayaan Masyarakat Indonesia*, *1*(2), 39–47.

[8] Hasanah, I., Purwanti, E., & Widiyanti, P. (2020). Design and Implementation of an Early Screening Application for Dengue Fever Patients Using Android-Based Decision Tree C4.5 Method. *International Journal on Advanced Science, Engineering and Information Technology*, *10*(6), 2237–2243.

[9] Katrina, W., Damanik, H.J., Parhusip, F., Hartama, D., Windarto, A.P., & Wanto, A. (2019). C.45 Classification Rules Model for Determining Students Level of Understanding of the Subject. *Journal of Physics: Conference Series*, *1255*(1), 1–7.

[10] Kaur, P., & Gosain, A. (2022). Issues and challenges of class imbalance problem in

classification. *International Journal of Information Technology (Singapore)*, *14*(1), 539–545.

[11] Kewsuwun, N., & Kajornkasirat, S. (2022). A sentiment analysis model of agritech startup on Facebook comments using naive Bayes classifier. *International Journal of Electrical and Computer Engineering*, *12*(3), 2829–2838.

[12] Lin, W., & Wong, H. (2016). Polarization Reconfigurable Aperture-Fed Patch Antenna and Array. *IEEE Access*, *4*, 1510–1517.

[13] Ma'arif, A., Mahendra, K.W., Ferbriyanto, S., Habibillah, A., & Nurimam, A. (2021). Training on installing solar water pump for resident of singkar 1 wareng wonosari gunungkidul yogyakarta indonesia. *Jurnal Pengabdian dan Pemberdayaan Masyarakat Indonesia*, *1*(1), 32-39.

[14] Majeed, R., Abdullah, N.A., & Mushtaq, M.F. (2021). IoT-based Cyber-security of Drones using the Naive Bayes Algorithm. *International Journal of Advanced Computer Science and Applications*, *12*(7), 422–427.

[15] Muhdi, Buchori, A., & Wibisono, A. (2019). Whiteboard animation for android design using think talk write model to improve the post graduates students' concepts understanding. *Journal of Advanced Research in Dynamical and Control Systems*, *11*(7), 535–543.

[16] Perdana, A., Defit, S., & Wanto, A. (2021). Optimalisasi Parameter dengan Cross Validation dan Neural Back-propagation Pada Model Prediksi Pertumbuhan Industri Mikro dan Kecil. *Jurnal Sistem Informasi Bisnis*, *01*(11), 34–42.

[17] Pozi, M.S.M., & Mohd. Hasbullah Omar. (2020). A Kernel Density Estimation Method to Generate Synthetic Shifted Datasets in Privacy-Preserving Task. *Journal of Internet Services and Information Security (JISIS)*, *10*(4), 70-89.

[18] Pratiwi, H., Windarto, A.P., Susliansyah, S., Aria, R.R., Susilowati, S., Rahayu, L.K., Fitriani, Y., Merdekawati, A., & Rahadjeng, I.R. (2020). Sigmoid Activation Function in Selecting the Best Model of Artificial Neural Networks. *Journal of Physics: Conference Series*, *1471*(1), 1-7.

[19] Sakar, C.O., Polat, S.O., Katircioglu, M., & Kastro, Y. (2019). Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. *Neural Computing and Applications*, *31*(10), 6893–6908.

[20] Setiawan. (2015). Integrasi Metode Sample Bootstrapping dan Weighted Principal Component Analysis untuk Meningkatkan Performa K Nearest Neighbor pada Dataset Besar. *Journal of Intelligent Systems*, *1*(2), 76–81.

[21] Silhavy, R., Silhavy, P., & Prokopova Zdenk. (2019). *Computational Statistics and Mathematical Modeling Methods in Intelligent Systems*, *2*, 1-423.

[22] Silva, R.R., Caminhas, W.M., & Guimarães, F.G. (2021). A C4. 5 Fuzzy Decision Tree Method for Multivariate Time Series Forecasting. *In IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 1-6.

[23] Subagja, A.D. (2023). Analysis of Factors Leading to E-commerce Adoption. *Apollo: Journal of Tourism and Business*, *1*(1), 1–5.

[24] Suherlan, S. (2023). Buying and Selling Online from an Islamic Legal Poin of View. *Apollo: Journal of Tourism and Business*, *1*(1), 29–35.

[25] Sundari, S., Karmila, Fadli, M.N., Hartama, D., Windarto, A.P., & Wanto, A. (2019). Decision Support System on Selection of Lecturer Research Grant Proposals using Preferences Selection Index. *Journal of Physics: Conference Series*, *1255*(1), 1–8.

[26] Syafar, F., Du, J.T., & Gao, J. (2015). Mobile collaboration technology to support maintenance enterprise systems in large industry. *Americas Conference on Information Systems, AMCIS*.

[27] Syafar, F., Gao, J., & Du, J.T. (2014). Mobile-enabled collaborative maintenance. *Proceedings Pacific Asia Conference on Information Systems, PACIS*.

[28] Tan, Y., Zhao, G., Dai, H., Lin, Z., & Cai, G. (2018). Classification of Parkinsonian Rigidity Using AdaBoost with Decision Stumps. *IEEE International Conference on Robotics and*

Biomimetics, ROBIO, 170–175.
[29] Varzaneh, Z.A., Orooji, A., Erfannia, L., & Shanbehzadeh, M. (2022). A new COVID-19 intubation prediction strategy using an intelligent feature selection and K-NN method. *Informatics in Medicine Unlocked, 28*.
[30] Waluyo, A., Jatnika, H., Permatasari, M.R.S., Tuslaela, T., Purnamasari, I., & Windarto, A.P. (2020). Data Mining Optimization uses C4.5 Classification and Particle Swarm Optimization (PSO) in the location selection of Student Boardinghouses. *IOP Conference Series: Materials Science and Engineering, 874*(1), 1–9.
[31] Wang, Y., & Feng, L. (2020). Improved Adaboost Algorithm for Classification Based on Noise Confidence Degree and Weighted Feature Selection. *IEEE Access, 8,* 153011–153026.
[32] Windarto, A.P., & Herawan, T. (2022). K-Means Algorithm with Rapidminer in Clustering School Participation Rate in Indonesia. In *Lecture Notes in Electrical Engineering*. Springer Nature Singapore Pte Ltd. 2022.
[33] Windarto, A.P., Lubis, M.R., & Solikhun. (2018). Implementation Of Neural Network In Predicting Total Comprehensive Income Of Conventional Commercial Banks Using Backpropagation. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK), 5*(4), 411–418.
[34] Yaman, M.A., Rattay, F., & Subasi, A. (2021). Comparison of Bagging and Boosting Ensemble Machine Learning Methods for Face Recognition. *Procedia Computer Science, 194,* 202–209.

Author Profile

Ading Sunarto

Ading Sunarto received a Bachelor of Economics and a master of management from Universitas Pamulang in 2015 and 2017, and then continue the Doctoral program (S3) at Universitas Pakuan Bogor and in the process of graduation. His research interests include tools for social science analytics, human resource management, economics, management, and entrepreneurship. At the moment the author becomes a permanent lecturer in the Universitas Pamulang Management Program.
E-mail: dosen02153@unpam.ac.id, Orcid: https://orcid.org/0000-0002-6036-7926

Putri Nilam Kencana

Putri Nilam Kencana received a Bachelor of Economics from IPB University in 2011, and a master of management from Pamulang University in 2016, and now still continuing study in Doctoral program at Pakuan University. Her research interests about social science, management, marketing management, human resource management and entrepreneurship.
Email: dosen01877@unpam.ac.id, Orcid: https://orcid.org/0000-0001-6744-7550

Iriana Kusuma Dewi

Iriana Kusuma Dewi received a Bachelor of Development Economics from University of Muhammadiyah Malang in 2009, and a master of Management from Pamulang University in 2016, and now still continuing study in Doctoral program at Syarif Hidayatullah State Islamic University Jakarta. Her research interests about Financial Management, Macroeconomics, Capital Market & Portfolio Management, and Managerial Economics.
Email: dosen01729@unpam.ac.id, https://orcid.org/0009-0005-4580-2140

Baliyah Munadjat

Baliyah Munadjat received a Bachelor of Development Economics from University of Mulawarman in 2002, and a master of Management from Pamulang University in 2017, and now still continuing study in Doctoral program at Pakuan University.Her research interests about Financial Management, Macroeconomics, Capital Market & Portfolio Management, and Managerial Economics.
Email: dosen02162@unpam.ac.id, Orcid: https://orcid.org/0009-0009-1247-7192

Ali Zaenal Abidin

Ali Zaenal Abidin a graduate doctoral program from Universitas Pendidikan Indonesia in 2019. His reseacrh interests include tools for social science analytics, human resources management, management, and entrepreneurship.
Email: alizaenalabidin@unpam.ac.id, Orcid: https://orcid.org/0000-0003-2238-6103

Robbi Rahim

Robbi Rahim is an Indonesian academic with a Doctoral degree in Protocol Cryptography from Universiti Malaysia Perlis. He has expertise in the fields of data mining, big data, and Rapid Miner, all of which are related to the processing and analysis of large datasets. Rahim's doctoral thesis focused on the study of Protocol Cryptography, which involves securing communication protocols using cryptographic techniques. His contributions to research in various fields, including computer science and information technology, have been significant. Since 2017, Rahim has been working as a lecturer at Sekolah Tinggi Ilmu Manajemen Sukma. In his current role, he teaches and mentors students in the areas of data mining, big data, and Rapid Miner. His expertise in these fields has enabled him to bring a unique perspective to his teaching, helping students to develop the skills and knowledge needed to succeed in today's technology-driven world.
Email: usurobbi85@zoho.com, Orcid: https://orcid.org/0000-0001-6119-867X