Classification of Students' Academic Performance Using Neural Network and C4.5 Model

Sulika[1], Ririen Kusumawati[2], Yunifa Miftachul Arif[3]
[1,2,3] Faculty of Science and Technology, Master's Program Specifications in Computer Science (Master Informatics), University Islam Maulana Malik Ibrahim, Indonesia

Article Info

Article history:

Received Nov 22, 2023
Revised Mar 11, 2024
Accepted Mar 25, 2024

Keywords:

Classification
Students' academic
Neural networks
C4.5 algorithm
Data mining

ABSTRACT

Education involves deliberately creating an environment and learning process to empower students to fully utilize their academic and non-academic potential. It encompasses fostering spiritual qualities, religious understanding, self-discipline, cognitive abilities, and skills necessary for personal, societal, national, and state development. Madrasah Aliyah, in particular, emphasizes preparing participants for higher studies in areas of their interest, thereby showcasing their academic prowess. The evaluation of educational models like Neural Networks is crucial for ensuring their effectiveness in problem-solving. This involves testing and assessing the performance of the Neural Network model to ensure its accuracy and reliability. Similarly, the C4.5 method, based on condition data mining, is utilized to measure classification performance by assessing accuracy, precision, and recall. Research findings indicate that the neural network algorithm is more adept at accurately classifying students' academic abilities compared to the C4.5 algorithm. With an accuracy of 92.6% for the neural network algorithm and 80.6% for the C4.5 algorithm, it is evident that the former is more precise in determining the classification of students' academic abilities. This highlights the suitability of the neural network approach for classifying academic abilities in Madrasah Aliyah. Furthermore, the insights gained from this classification process can be extrapolated to benefit other madrasas.

Corresponding Author:

Sulika
Faculty from Science and Technology,
Specifications of the Master of Computer Science Study Program (Master of Informatics), University Islam Maulana Malik Ibrahim, Indonesia
Jl. Gajayana No. 50, Dinoyo, Lowokwaru, Malang City, East Java 65144
E-mail: sulika.idris@gmail.com

1. INTRODUCTION

The process of selecting students based on their academic aptitude is a crucial aspect of educational institutions, particularly evident in the diverse subject groups offered at Madrasah Aliyah Nur Divine, situated in Malang Regency, for Class X students. This selection process involves assessing students' interests and talents, often through interest and ability tests or questionnaires. Academic performance, including factors such as grade point average, plays a significant role in determining eligibility for specific fields of study. Moreover, discussions between students and teachers or supervisors are commonplace, where informed decisions are made based on an understanding of individual strengths and weaknesses [1].

Additionally, considerations extend beyond immediate academic interests to encompass future career plans [2]. Students often explore potential job prospects and career opportunities within their preferred fields, necessitating adequate facilities and teaching staff support within the madrasah aliyah education system [3], [4]. However, resource limitations may sometimes require the institution to restrict student enrollment in certain subjects.

Education, as perceived in Madrasah Aliyah, is a deliberate effort to create an environment conducive to the active development of students' academic abilities and spiritual qualities. Emphasis is placed on preparing students for higher education through specialized academic skills development. Academic achievements, such as Science Olympiad victories or excellence in other academic domains, serve as benchmarks for assessing students' abilities in specific fields.

In an era marked by rapid technological advancements and evolving societal needs, there arises an imperative to continually enhance human resources' quality to remain competitive in the global labor market. It is within this context that the present research seeks to contribute. By employing the Neural Network algorithm and the C4.5 method, the study aims to classify the academic abilities of Madrasah Aliyah students meticulously. Through this endeavor, the research endeavors to provide a comprehensive analysis, considering the strengths and limitations of each method, ultimately striving for the highest precision in academic ability classification.

The choice of utilizing the Neural Network algorithm and the C4. 5 method is rooted in their proven efficacy in handling complex datasets and discerning intricate patterns [5], [6], [7], [8], [9]. Neural networks, drawing inspiration from the human brain's functioning, exhibit a remarkable ability to capture nonlinear relationships within data, providing a nuanced understanding of students' academic abilities. Conversely, the C4.5 method, based on decision tree algorithms [10], [11], offers transparent classification rules, enhancing interpretability and facilitating informed decision-making [12].

By leveraging these advanced computational techniques, this study aims to transcend the limitations of traditional assessment methods, providing educators and policymakers with actionable insights to optimize student placement and support strategies. Furthermore, the research underscores the importance of embracing technological innovation in educational assessment, paving the way for a more data-driven and student-centered approach to academic classification [13], [14], [15].

In essence, the significance of this study lies in its potential to revolutionize the educational assessment landscape, empowering educational institutions like Madrasah Aliyah Nur Divine to tailor interventions and support mechanisms with unprecedented precision. As we navigate through the intricacies of these computational methodologies, we embark on a journey towards a more equitable, inclusive, and effective educational system, where every student's academic potential is nurtured and realized to the fullest extent possible.

2. RESEARCH METHOD

Madrasah Aliyah Nur Divine, located in Malang Regency, serves as a dynamic hub for various learning activities aimed at fostering knowledge and skill acquisition among learners. Central to this process is the interactive engagement between teachers and students, where teaching and learning transpire within an interactive context, leading to changes in learners' understanding, skills, and attitudes.

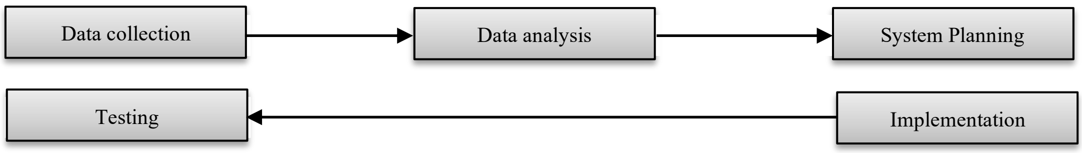


Figure 1. Research Design

The research methodology employed in this study involves several sequential steps, as illustrated in Figure 1. Data collection, the initial phase, entails gathering information necessary to address the research question or problem statement. Data may originate from diverse sources such as databases and text files, as depicted in Figure 2. Furthermore, data collection includes insights obtained from interviews conducted with Madrasah Aliyah representatives to acquire essential information for system development. Specifically, data pertaining to Class X students, categorized into Mathematics and Science (MIPA) and Social Sciences (IPS) groups, serves as the primary research material.

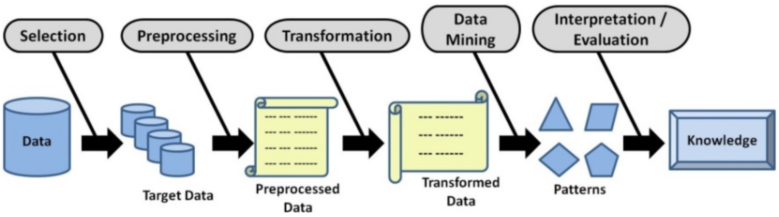


Figure 2. Process Knowledge Discovery in Databases

Attribute selection is a critical step aimed at ensuring that the data utilized aligns with the core research process, facilitating subsequent testing and training data preparation. Attributes considered during this stage include student names, Family Identification Numbers (NIK), student questionnaires, parent recommendations, and report card scores, with the objective of generating the desired average score.

Subsequently, system design involves delineating the utilization and management of data within the system through case usage descriptions, as delineated in Figure 3. This stage employs training data to process the algorithm, followed by the utilization of test data to evaluate accuracy and perform comparisons. The system then classifies students' academic abilities based on the achieved accuracy.

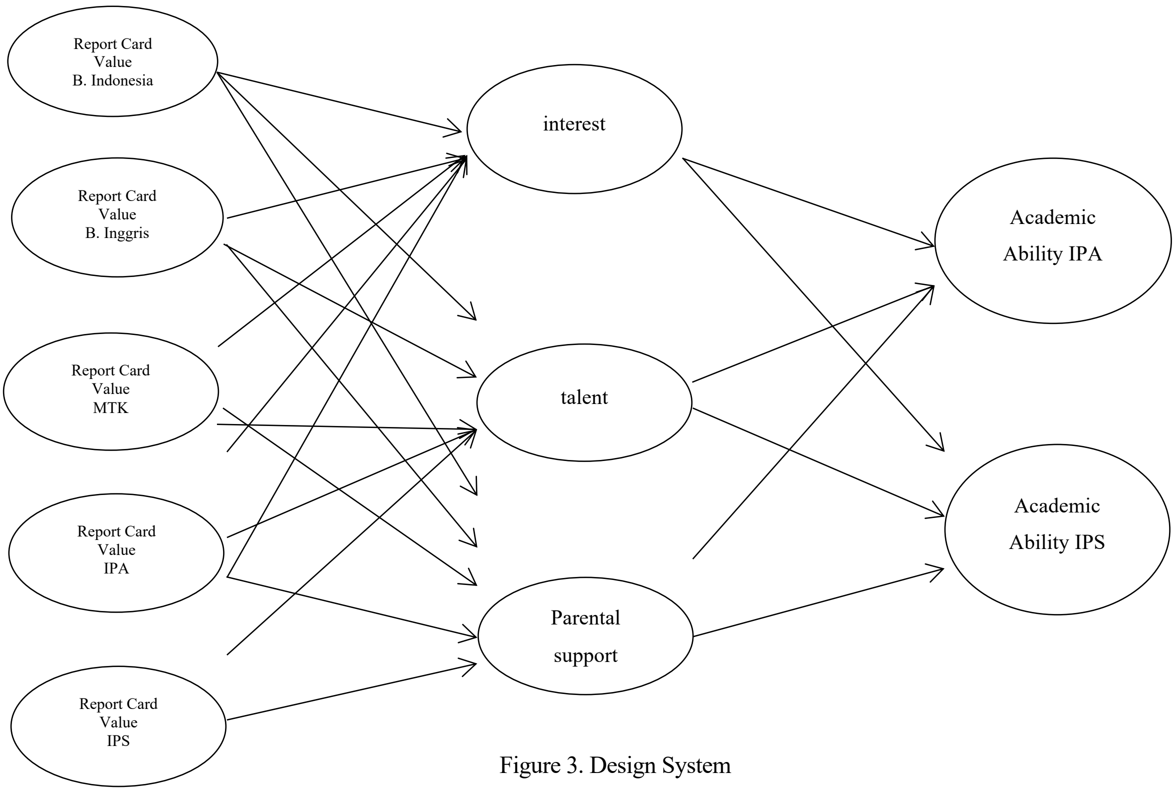


Figure 3. Design System

The system design involves the utilization of training data to process the algorithm, followed by the utilization of test data for accuracy assessment and comparison. Based on the accuracy achieved, the system proceeds to classify students' academic abilities. The testing stage is executed using machine learning software, specifically Orange, which facilitates data analysis through the arrangement of widgets in a workflow. Each widget is assigned tasks related to data retrieval, initial data processing, visualization, modeling, or evaluation.

The testing and evaluation phase aims to ensure the correctness, accuracy, and reliability of the developed Neural Network model in addressing the research problem. Performance measures are assessed using a confusion matrix, aiding the algorithm in categorizing results into two primary categories commonly referred to as positive or negative outcomes. This meticulous evaluation process is depicted in the flow of the research process outlined in Figure 4.

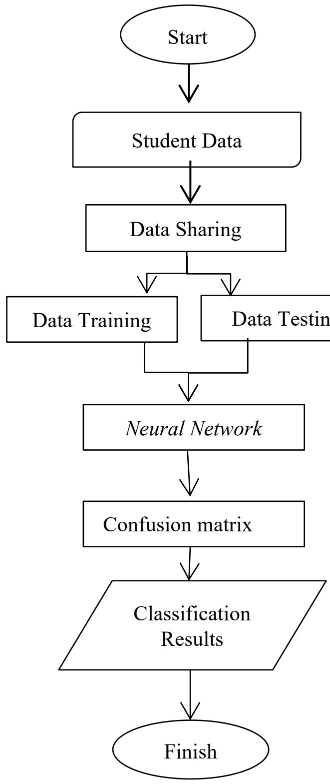


Figure 4. Neural Network flowchart

The neural network model employed in this study represents a common type of network, characterized by a multilayer structure comprising input, hidden, and output layers. The feed-forward topology ensures the sequential transmission of information from input to output layers, as outlined in the third edition of Han et al. [16]' book.

$$\mathrm{Ij} = \sum_i \mathrm{wij\ i\ Oi} + \theta\mathrm{j} \quad Ij$$

j = hidden or output layer
Ij = Net Input
wij = connection weight from unit i in the previous layer to unit j
Oi = Output from unit i
θj = bias

Additionally, the C4.5 method serves as a rigorous testing protocol for data mining, facilitating the measurement of classification effectiveness through accuracy, recall, and precision calculations, which are formulated using the best confusion matrix. The design process for the C4.5 model is illustrated in Figure 5.

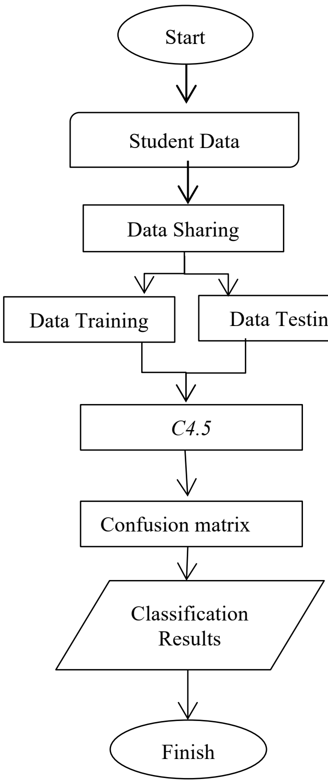


Figure 5. C4.5 Flowchart

By general Algorithm C4.5 build decision tree decision as following (17):

a. Calculation *Entropy* and *Gains*
b. Election *Gains* highest as root (*Nodes*)
c. Repeat the *Entropy* and *Gain calculation process.* For look for branch until all case on branches have the same class, namely when all variables have become part from tree decision or eachvariables have own decision.
   Create *a Rule* based on a decision tree. Before get mark *gain,* moreover used to search *Entropy.* Calculation *Entropy* formulatedon equality following

$$\text{Entropy (s)} = \sum_{i=1}^{n} - \text{ pi} \log_2 \text{pi}$$

Information:
S = Set of cases
n = Amount partition S
P_i = proportion S_i to S

Election an attribute as a root is based on the highest gain value of the existing attributes and is formulated as follows :

$$Gain = Entropy\ (S) - \sum_{i=1}^{n} \frac{|si|}{S} * Entropy\ (si)$$

description :

S = Set cases
A = Features
n = amount partition attributes A

|Si| = proportion Si to S
|S| = amount cases in S

3. RESULTS AND DISCUSSION

The system was implemented on a computer with Core i3 specifications, 4 GB RAM, and utilizing the Orange application. A dataset comprising 134 student records was prepared in CSV format and meticulously examined, with detailed scrutiny of variables outlined in Table 1.

Table 1. Dataset variables

No	Attributes	Type	Variables
1	Name	Text	Independent
2	VIN	Numerical	Independent
3	Date of birth	Date	Independent
4	Place of birth	Text	Independent
5	Mother's name	Text	Independent
6	Gender	Text	Independent
7	Interest	Text	Independent
8	Talent	Text	Independent
9	Parent Support	Text	Independent
10	Science Value	Numerical	Independent
11	IPS score	Numerical	Independent
12	Mathematics Value	Numerical	Independent
13	Grade B. Indonesia	Numerical	Independent
14	Grade B English	Numerical	Independent
15	Average	Numerical	Independent
16	Major	Categorical	Independent

During the application of the Orange application, the gathering of training data facilitated the identification of features that exerted the most significant impact on the classification process. This analysis involved the utilization of two primary methods: neural networks and the C4.5 method, both of which demonstrated the highest accuracy levels.

Subsequently, in the data analysis stage, following data collection, an analysis process was undertaken to ascertain the requisite criteria for classifying students' academic abilities. These criteria were harnessed as attributes, incorporating variables such as first and second-semester report card grades, student interests, and academic test scores in subjects like Mathematics and Social Studies. Each attribute was characterized by specific types and values, categorized into low (0-50), medium (51-75), and high (76-100) ranges, as elucidated in Table 3.

Table 3. Explanation variables which needed

Attributes	**Type**	***Value* (Mark)**
AverageMark	Polynomials	a. Low (0-50) b. Currently (51-75) c. Tall (76-100)
Interest	Polynomials	a. MIPA b. Social Sciences
Mark MIPA	Polynomials	a. Low (0-50) b. Currently (51-75) c. Tall (76-100)
IPS scores	Polynomials	a. Low (0-50) b. Currently (51-75) c. Tall (76-100)
Major	Label	a. MIPA b. Social Sciences

The selection of attributes was meticulously carried out to ensure consistency with the primary classification process, enabling the execution of testing and training data procedures seamlessly. Notably, attributes such as student names, Family Identification Numbers (NIK), student questionnaires, parent recommendations, and report card results were instrumental in indexing the desired academic performance.

The performance of the model was assessed using the confusion matrix index, a technique commonly employed in data mining and machine learning to evaluate model predictive capabilities. With 134 datasets utilized for training, the MIPA section comprised 45 datasets, while the Social Sciences section comprised 89 datasets.

As shwon in Figure 6, the confusion matrix serves as a vital tool for assessing model performance, particularly in classification tasks where outcomes are divided into distinct categories. It aids in evaluating True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) predictions, thereby offering insights into the model's predictive accuracy and potential errors [17]. True Positives (TP) represent the instances where the model correctly predicts the positive class. True Negatives (TN) indicate the instances where the model correctly predicts the negative class. False Positives (FP) occur when the model predicts a positive class incorrectly, falsely labeling a negative instance as positive. False Negatives (FN) occur when the model predicts a negative class incorrectly, falsely labeling a positive instance as negative. These metrics are crucial in assessing the performance of classification models, providing insights into their ability to accurately classify instances into their respective classes. TP and TN reflect the model's accuracy, while FP and FN highlight its errors, helping in fine-tuning and evaluating model performance in various applications such as healthcare, finance, and marketing.

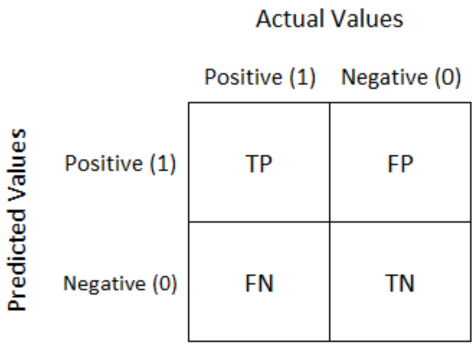


Figure 6. Confusion matrix

The following section elucidates the utilization of the confusion matrix in evaluating model performance and delineates the calculation methodology for evaluating classification accuracy, recall, precision, and other performance metrics based on the confusion matrix outcomes.

Accuracy

Accuracy is a testing method that based on how much close the predicted value to the actual value. Know how much a lot of properly classified data can help You determine accuracy of prediction results You.

$$Accuracy = \frac{TP + TN}{\mathrm{TP} + FP + FN + TN}$$

Precision

Precision is a testing method that compares the amount of relevant information received by the system with the total amount of information received system, regardless of whether it's relevant or not or not information the.

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

Recall

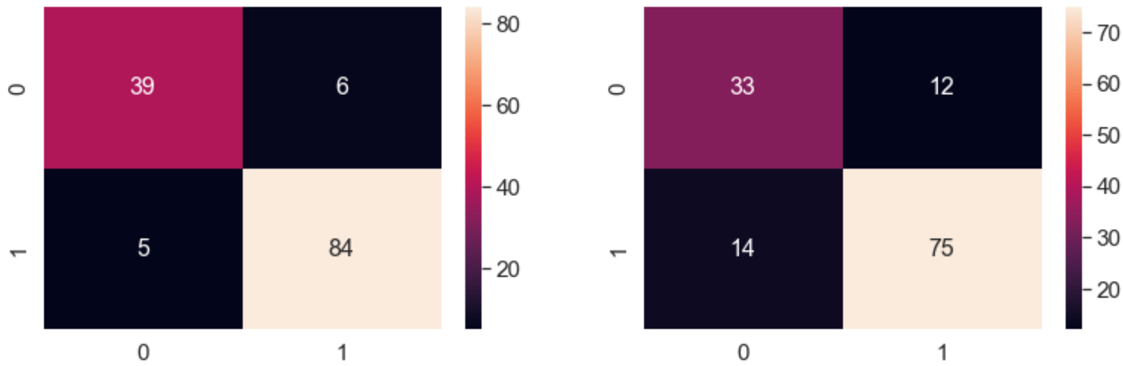
Recall is a testing method that compares the amount of relevant information retrieved by the system with the sum of all relevant information in the information set (either retrieved by system or not).

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

F-Measure or F1-Score

F-Measure or also known as F1- Score describes a weighted average comparison between precision and recall. We use precise precision as a measure performance algorithm when our data set contains very similar numbers of false negatives and false positives (symmetrical). However, if the numbers are not close together, it is better to use them score F1.

$$F1 = \frac{2 * \mathrm{Recall} * \mathrm{precision}}{\mathrm{Recall} + Precision}$$



a. Neural network b. C4.5

Figure 7. Confusion matrix results

The confusion matrix resulting from testing the neural network and C4.5 model is as follows:

Table 4. Experiment results

Model	AUC	C.A	F1-score	Precision	Recall	MCC
Neural Networks	0.977	0.918	0.918	0.918	0.918	0.815
C4.5	0.811	0.806	0.807	0.808	0.806	0.570

The analysis of the model's performance reveals intriguing insights into its predictive capabilities. As demonstrated in Table 4, the model exhibits an overall accuracy of 80.6%, with a precision of 70.2% for predicting Spam labels and a recall of 73.3% for Spam labels. Conversely, for the Not Spam label, the model achieves a precision of 86.2% and a recall of 84.3%.

Upon closer examination of the test results, it becomes apparent that the Neural Network algorithm outperforms the C4.5 algorithm in accurately classifying students' academic abilities based on recommended directions. Specifically, the Neural Network algorithm achieves an impressive accuracy rate of 91.8%, surpassing the C4.5 algorithm's accuracy of 80.6%. This significant disparity underscores the Neural Network algorithm's efficacy in this particular application.

The superior performance of the Neural Network algorithm suggests its suitability for guiding students and identifying appropriate fields of study. By leveraging the algorithm's predictive capabilities, educators and policymakers can tailor educational interventions to optimize students' learning experiences. Moreover, the Neural Network algorithm presents an avenue for enhancing the learning process, providing students with ample opportunities for success.

As the research progresses, there are avenues for further exploration and refinement. For instance, the addition of new variables and columns to the dataset could enrich the model's predictive capabilities, offering a more comprehensive understanding of students' academic aptitudes. Furthermore, comparative analyses incorporating alternative algorithmic techniques hold promise for enhancing the accuracy and robustness of the classification process.

In conclusion, while the Neural Network algorithm demonstrates remarkable efficacy in classifying students' academic abilities, there remains scope for refinement and expansion. By embracing ongoing research and development initiatives, educators can harness the power of advanced computational techniques to foster a more personalized and effective learning environment, ultimately empowering students to realize their full potential.

4. CONCLUSION

In the realm of education, the classification of students' academic abilities holds paramount importance as it serves to enhance academic prowess, thereby indirectly bolstering the pool of human resources. By discerning students' academic strengths and weaknesses, educators can tailor instructional strategies to optimize learning outcomes, paving the way for students to acquire knowledge applicable in both professional settings and further educational pursuits.

The findings of this study underscore the efficacy of the Neural Network method in classifying students' academic abilities, boasting an impressive accuracy rate of 91.8% compared to the C4.5 method's accuracy of 80.6%. This highlights the Neural Network method as a suitable choice for academic classification within Madrasah Aliyah institutions. Furthermore, the application of this classification model holds promise for adaptation and implementation in other madrasas, contributing to the advancement of educational practices across the board.

Looking ahead, there exists ample potential for further development and refinement of this research. Future endeavors could involve the incorporation of more comprehensive datasets, including variables such as economic indicators of the local community and environmental factors, which undoubtedly influence students' specialization abilities in their chosen subject areas. By enriching the dataset, researchers can glean deeper insights into the multifaceted factors shaping students' academic trajectories, thus facilitating more nuanced and targeted interventions.

In essence, this research serves as a stepping stone towards a more data-driven and personalized approach to educational classification, with implications extending beyond Madrasah Aliyah Nur Divine to encompass educational institutions at large. By harnessing the power of advanced computational techniques and leveraging comprehensive datasets, educators can unlock new avenues for optimizing student learning experiences and fostering academic excellence.

ACKNOWLEDGEMENT

The author would like to thank Madrasah Aliyah Malang Regency for helping or providing support related to the research carried out such as facilities accompaniment, data and information supplies.

REFERENCES

[1] S. K. Slick, “Assessing versus assisting: The supervisor’s roles in the complex dynamics of the student teaching triad,” *Teach Teach Educ*, vol. 13, no. 7, pp. 713–726, Oct. 1997, doi: 10.1016/S0742-051X(97)00016-4.

[2] L. P. Khobragade and P. Mahadik, “Students academic failure prediction using data mining,” *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, 2015.

[3] O. Taylan and B. Karagözo\uglu, “An adaptive neuro-fuzzy model for prediction of student’s academic performance,” *Comput Ind Eng*, vol. 57, no. 3, pp. 732–741, 2009.

[4] R. M. Tawafak, A. M. Romli, and M. J. Alsinani, “Student assessment feedback effectiveness model for enhancing teaching method and developing academic performance,” *International Journal of Information and Communication Technology Education*, vol. 15, no. 3, pp. 75–88, 2019, doi: 10.4018/IJICTE.2019070106.

[5] G. Zhang, B. Eddy Patuwo, and M. Y. Hu, “Forecasting with artificial neural networks: The state of the art,” *Int J Forecast*, vol. 14, no. 1, pp. 35–62, 1998, doi: 10.1016/S0169-2070(97)00044-7.

[6] Z. Zhang, “Neural networks: further insights into error function, generalized weights and others,” *Ann Transl Med*, vol. 4, no. 16, pp. 300–300, 2016, doi: 10.21037/atm.2016.05.37.

[7] G. P. Zhang, “Neural networks for classification: a survey,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 30, no. 4, pp. 451–462, 2000.

[8] E. Budiman, Haviluddin, N. Dengan, A. H. Kridalaksana, M. Wati, and Purnawansyah, “Performance of Decision Tree C4.5 Algorithm in Student Academic Evaluation,” *Lecture Notes in Electrical Engineering*, vol. 488, pp. 380–389, 2018, doi: 10.1007/978-981-10-8276-4_36/COVER.

[9] A. Cherfi, K. Nouira, and A. Ferchichi, “Very Fast C4.5 Decision Tree Algorithm,” *Applied Artificial Intelligence*, vol. 32, no. 2, pp. 119–137, Apr. 2018, doi: 10.1080/08839514.2018.1447479.

[10] S. B. Kotsiantis, “Decision trees: a recent overview,” *Artif Intell Rev*, vol. 39, no. 4, pp. 261–283, 2013.

[11] D. Saputra, W. Irmayani, D. Purwaningtias, J. Sidauruk, and B. Gurbuz, “A Comparative Analysis of C4.5 Classification Algorithm, Naïve Bayes and Support Vector Machine Based on Particle Swarm

Optimization (PSO) for Heart Disease Prediction," *International Journal of Advances in Data and Information Systems*, vol. 2, no. 2, pp. 84–95, Oct. 2021, doi: 10.25008/IJADIS.V2I2.1221.

[12] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining explanations: An overview of interpretability of machine learning," *Proceedings - 2018 IEEE 5th International Conference on Data Science and Advanced Analytics, DSAA 2018*, pp. 80–89, Jul. 2018, doi: 10.1109/DSAA.2018.00018.

[13] E. Sutoyo and A. Almaarif, "Educational Data Mining for Predicting Student Graduation Using the Naïve Bayes Classifier Algorithm," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 1, pp. 95–101, 2020, doi: 10.29207/RESTI.V4I1.1502.

[14] D. Y. T. Liu, K. Bartimote-Aufflick, A. Pardo, and A. J. Bridgeman, "Data-driven personalization of student learning support in higher education," *Studies in Systems, Decision and Control*, vol. 94, pp. 143–169, 2017, doi: 10.1007/978-3-319-52977-6_5/COVER.

[15] R. Van Campenhout, B. Jerome, and B. G. Johnson, "Engaging in Student-Centered Educational Data Science Through Learning Engineering," pp. 3–40, 2023, doi: 10.1007/978-981-99-0026-8_1.

[16] J. Han, M. Kamber, and J. Pei, *Data Mining, Concepts and Techniques*. 2012.

[17] J. Han and M. Kamber, *Data Mining, Southeast Asia Edition: Concepts and Techniques*. Morgan kaufmann, 2006.