IJISIT

http://ejournal.gevivapublisher.org/index.php/ijisit

Classification of Teacher Certification Eligibility Using the C4.5 Algorithm

[1]Agnes Irene Silitonga, Universitas Negeri Medan, Indonesia
[2]Mismauli Nainggolan, Universitas Negeri Medan, Indonesia
[3]Tasya Arcinta, Universitas Negeri Medan, Indonesia
[4]Yoakim Simamora, Universitas Negeri Medan, Indonesia
[5]Ferry Indra Sakti H. Sinaga, Universitas Negeri Medan, Indonesia
Correspondence: E-mail: agnesirenesilitonga@unimed.ac.id

Article Info

Article history:

Received January 27, 2026
Revised March 15, 2026
Accepted June 10, 2026

Keywords:

Machine Learning
C4.5 Algorithm
Decision Tree
Classification
Teacher Certification Eligibility

ABSTRACT

Determining teacher certification eligibility is a crucial process in improving the quality of education. The C4.5 algorithm is a decision tree-based machine learning algorithm. This algorithm offers a systematic approach to data analysis and provides accurate results for decision-making. This study aims to develop a predictive model using the C4.5 algorithm to assess teacher certification eligibility based on relevant data such as teaching experience, education, and competency exam results. This study reveals that the C4.5 algorithm is capable of producing transparent decision rules and enabling clear interpretation of the results. This research is expected to make a significant contribution to supporting a more objective and efficient teacher certification policy.

1. INTRODUCTION

Improving the quality of education in Indonesia depends heavily on teacher competence and professionalism. One government effort to ensure this quality is through the teacher certification program, which aims to ensure that educators meet established competency standards. However, the certification eligibility assessment process often faces challenges, such as subjectivity in evaluations and limited human resources for conducting comprehensive assessments. This can lead to inaccuracies in determining which teachers are eligible for certification, necessitating a system that can assist in objective and efficient decision-making.

In recent years, the application of information technology and data mining methods has shown great potential in supporting decision-making processes in various fields, including education. One widely used algorithm in data classification is the C4.5 algorithm, a development of the ID3 algorithm and known to be effective in

constructing decision trees based on relevant attributes [1]. This algorithm is capable of handling data with both numeric and categorical attributes and produces easily interpretable models, making it suitable for use in the context of assessing teacher certification eligibility. Several studies have applied the C4.5 algorithm to various cases, such as predicting student achievement [2], measuring student satisfaction with teacher performance [3], and classifying community personality [4]. However, the specific application of this algorithm in the context of determining teacher certification eligibility remains limited, requiring further research to explore its potential use in this area.

This research aims to develop a predictive model using the C4.5 algorithm to assess teacher certification eligibility based on relevant data, such as teaching experience, education level, and competency test results. By building a decision tree based on this data, the system is expected to provide objective and accurate recommendations regarding a teacher's certification eligibility. This model is expected to be an effective tool for relevant parties in the decision-making process and increase transparency and accountability in the teacher certification process.

The main contribution of this research is the development of a data-driven teacher certification eligibility evaluation model based on machine learning algorithms, specifically C4.5. Unlike conventional approaches, this model offers transparency in the decision-making process through visualization of the resulting decision tree. Furthermore, this research has the potential to serve as a basis for developing decision support systems for human resource management in the education sector and contribute to the literature on the application of the C4.5 algorithm in the context of education evaluation in Indonesia.

2. LITERATURE REVIEW

2.1. Data Mining

Data mining is the process of exploring and analyzing large amounts of data to discover meaningful patterns, relationships, and hidden knowledge that can be used to support decision-making. This process is a core component of Knowledge Discovery in Databases (KDD), a systematic process for acquiring knowledge from raw data, encompassing data cleaning, feature selection, transformation, data mining, and interpretation of results [5].

More broadly, data mining is used in various fields, such as finance, healthcare, education, and social protection. The main techniques used in data mining include classification, clustering, association, and anomaly detection. Classification, a popular technique, allows systems to map data into predetermined groups based on specific characteristics [6].

2.2 Decision Tree

A decision tree is a tree-based machine learning method used for classification and regression purposes. This tree structure consists of a root node, internal nodes, and leaf nodes. Each internal node represents a test against a specific attribute, each branch shows the result of that test, and each leaf node represents the final decision or class label [7].

The decision tree method is considered highly intuitive and easy to understand because it is able to model decisions in a clear and systematic visual format. The decision tree algorithm divides the dataset into subsets based on specific attributes using criteria such as information gain, the Gini index, or the gain ratio, depending on the algorithm used [8].

One of the main advantages of this method is its ability to handle mixed data types (numerical and categorical), handle missing values, and generate if-then rules that can be easily applied or extracted. Therefore, decision trees are widely used in social studies, including the analysis of risk factors for child abuse, which has many causal variables and complexity between attributes.

2.3. C4.5 Algorithm

The C4.5 algorithm is a decision tree-based machine learning algorithm developed by Ross Quinlan as an improvement on the ID3 algorithm. This algorithm is widely used for classification tasks due to its ability to produce easily interpretable decision trees [9]. C4.5 has several advantages over its predecessors, such as being able to handle numeric and categorical attributes, handling missing values, performing pruning to prevent overfitting, and using a more stable gain ratio approach in attribute selection [10].

In building a decision tree, C4.5 selects the best attributes to be used as nodes based on the information gain value, which is a measure of the reduction in uncertainty (entropy) after the dataset is divided based on a specific attribute. The entropy value is calculated using equation 1.

$$Entropy\ (S) = \sum_{i=1}^{n} -p_i \log_2 p_i \qquad (1)$$

Where:
S = dataset
n = number of partitions
p_i = The proportion of class *i* to the total data S.

Next, the gain is calculated to determine how much entropy is reduced when the dataset is divided based on a particular attribute A:

$$Gain\ (S, A) = Entropy\ (S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} x Entropy\ (S_i) \qquad (2)$$

Where:
S = dataset
A = attribute
n = number of partitions of attribute A
$|S_i|$ = number of cases in the i-th partition
|S| = number of cases in S

3. METHODS

This study uses a quantitative approach with a data mining method based on the C4.5 algorithm to classify teacher certification eligibility. The data used comes from the Basic Education Data (DAPODIK) of elementary and junior high schools in Pangaribuan and Garoga Districts, North Tapanuli Regency. This data includes information such as age, length of service, education level, class, and teacher competency test results. The research stages include:

1. Data Acquisition
   The initial stage of this study was data collection from the Basic Education Data (Dapodik) managed by the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia. The data collected consisted of 50 databases covering information from elementary and junior high schools in Pangaribuan and Garoga Districts, North Tapanuli Regency. Information obtained included teacher profiles such as age, length of service, education level, employment status, and certification. Data collection was carried out to ensure that the data collected was up-to-date and relevant for analyzing teacher certification eligibility.
2. Data Preprocessing
   This step involves cleaning the data to remove duplication, handle missing values, and ensure consistency of data format. For example, if there are duplicate entries for a teacher or incomplete information, the data is checked and adjusted to prevent it from affecting the analysis results. This process is crucial for improving data quality and ensuring that models built based on it can produce accurate results.
3. Data Transformation
   Data transformation is performed to convert raw data into a format suitable for further analysis. In this context, attributes such as age and length of service, which are numeric, can be categorized into specific intervals. This transformation helps build a more effective classification model and facilitates the interpretation of the analysis results.
4. Application of the C4.5 Algorithm
   Once the data is prepared, the C4.5 algorithm is applied to build a decision tree that will be used in classifying teacher certification eligibility. This algorithm was chosen because of its ability to handle data with both numeric and categorical attributes and produce an easily interpretable model. This process involves

calculating the information gain for each attribute and selecting the attribute with the highest gain as the decision node. The resulting decision tree is then used to classify teacher data based on the predetermined attributes.

5. Model Evaluation
   The final step is evaluating the developed model. This evaluation is conducted using metrics such as accuracy, precision, and recall to assess the model's performance in classifying data. Accuracy measures how often the model makes correct predictions, precision measures the proportion of correct positive predictions, and recall measures the proportion of positive data successfully identified by the model. This evaluation is crucial to ensure that the developed model can be used effectively in determining teacher certification eligibility.

4. RESULTS AND DISCUSSION

The data used is national data on violence against children from 2024 from 38 provinces in Indonesia, as listed in Table 1. This data contains the number of teachers graduating from ASN, which includes seven data points such as teacher code, school of origin, employment status, age, education, years of service, and certification eligibility. Each row in the dataset represents the number of teachers graduating from state civil servants.

Table 1. Teacher Data for Civil Servants Graduation Analysis

No	Teacher Code	School of Origin	Employment Status	Age	Education	Years of Service	Eligibility
1	G-SMP2-01	SMP N 2 Garoga	Civil Servants	59	Bachelor	29	Yes
2	G-SMP2-02	SMP N 2 Garoga	Civil Servants	56	Bachelor	29	Yes
3	G-SMP2-03	SMP N 2 Garoga	Civil Servants	51	Bachelor	29	Yes
4	G-SMP2-04	SMP N 2 Garoga	Civil Servants	43	Bachelor	20	Yes
5	G-SMP2-05	SMP N 2 Garoga	Civil Servants	59	Bachelor	6	Yes
6	G-SMP2-06	SMP N 2 Garoga	Civil Servants	45	Bachelor	20	Yes
7	G-SMP2-07	SMP N 2 Garoga	Civil Servants	42	Bachelor	14	Yes
8	G-SMP2-08	SMP N 2 Garoga	Civil Servants	56	Bachelor	14	Yes
9	G-SMP2-09	SMP N 2 Garoga	Civil Servants	42	Bachelor	27	No
10	G-SMP2-10	SMP N 2 Garoga	Civil Servants	54	Bachelor	15	No
11	G-SMP2-11	SMP N 2 Garoga	Civil Servants	45	Bachelor	19	Yes
12	G-SMP2-12	SMP N 2 Garoga	Civil Servants	45	Bachelor	13	Yes
13	G-SMP2-13	SMP N 2 Garoga	Civil Servants	42	Bachelor	8	Yes
14	G-SMP2-14	SMP N 2 Garoga	Civil Servants	41	Bachelor	8	No
15	G-SMP2-15	SMP N 2 Garoga	Honorary	27	Bachelor	5	No
16	G-SMP2-16	SMP N 2 Garoga	Honorary	24	Bachelor	4	No
17	G-SMP2-17	SMP N 2 Garoga	Honorary	46	Bachelor	2	No
18	G-SMP2-18	SMP N 2 Garoga	Honorary	45	Bachelor	14	No
19	G-SMP2-19	SMP N 2 Garoga	Honorary	27	Bachelor	3	No
20	G-SMP5-01	SMP N 5 Garoga	Civil Servants	59	Bachelor	2	Yes
21	G-SMP5-02	SMP N 5 Garoga	Civil Servants	57	Bachelor	30	Yes
22	G-SMP5-03	SMP N 5 Garoga	Civil Servants	56	Bachelor	28	Yes
23	G-SMP5-04	SMP N 5 Garoga	Civil Servants	54	Bachelor	28	Yes
24	G-SMP5-05	SMP N 5 Garoga	Civil Servants	42	Bachelor	25	Yes
25	G-SMP5-06	SMP N 5 Garoga	Civil Servants	44	Bachelor	15	Yes
26	G-SMP5-07	SMP N 5 Garoga	Civil Servants	42	Bachelor	15	Yes
27	G-SMP5-08	SMP N 5 Garoga	Civil Servants	35	Bachelor	6	Yes

28	G-SMP5-09	SMP N 5 Garoga	Civil Servants	34	Bachelor	6	Ya
29	G-SMP5-10	SMP N 5 Garoga	Honorary	25	Bachelor	6	No
30	G-SMP5-11	SMP N 5 Garoga	Honorary	25	Bachelor	3	No
31	G-SMP5-12	SMP N 5 Garoga	Honorary	26	Bachelor	2	No
32	G-SMP5-13	SMP N 5 Garoga	Honorary	27	Bachelor	2	No
33	G-SMP7-01	SMP N 7 Pangaribuan	Civil Servants	57	Bachelor	3	Yes
34	G-SMP7-02	SMP N 7 Pangaribuan	Civil Servants	44	Bachelor	30	Yes
35	G-SMP7-03	SMP N 7 Pangaribuan	Civil Servants	59	Bachelor	19	Yes
36	G-SMP7-04	SMP N 7 Pangaribuan	Civil Servants	38	Bachelor	18	Yes
37	G-SMP7-05	SMP N 7 Pangaribuan	Civil Servants	33	Bachelor	10	Yes
38	G-SMP7-06	SMP N 7 Pangaribuan	Civil Servants	39	Bachelor	10	No
39	G-SMP7-07	SMP N 7 Pangaribuan	Civil Servants	56	Bachelor	4	No
40	G-SMP7-08	SMP N 7 Pangaribuan	Civil Servants	42	Bachelor	3	No
41	G-SMP7-09	SMP N 7 Pangaribuan	Honorary	25	Bachelor	15	No
42	G-SMP7-10	SMP N 7 Pangaribuan	Civil Servants	59	Bachelor	32	Yes
43	G-SMP7-11	SMP N 7 Pangaribuan	Civil Servants	56	Bachelor	28	Yes
44	G-SMP7-12	SMP N 7 Pangaribuan	Civil Servants	50	Bachelor	22	Yes
45	G-SD-01	SD Inpres 175765 Padangsiandomang	Civil Servants	59	Bachelor	28	Yes
46	G-SD-02	SD Inpres 175765 Padangsiandomang	Employment Agreements	42	Bachelor	16	No
47	G-SD-03	SD Inpres 175765 Padangsiandomang	Employment Agreements	30	Bachelor	6	No
48	G-SD-04	SD Inpres 175765 Padangsiandomang	Civil Servants	30	Bachelor	22	No
49	G-SD-05	SD Inpres 175765 Padangsiandomang	Civil Servants	54	Bachelor	24	Yes
50	G-SD-06	SD Inpres 175765 Padangsiandomang	Honorary	48	Bachelor	3	No

Table 2. Employment Status Attributes

Employment Status	Pre-processing
Civil Servants	G1
Honorary	G2
Employment Agreements	G3

Table 3. Age Attributes

Age	Pre-processing
0-20	U1
25-30	U2
35-40	U3
45-50	U4
55-60	U5

Table 4. Years of Service Attributes

Years of Service	Pre-processing
0-10	M1
15-20	M2
25-30	M3

35-40	M4

Table 5. Eligibility Certification Attributes

Eligibility Certification	Pre-processing
Yes	S1
No	S2

After preprocessing, the data is then transformed based on the respective attributes in table 8.

Table 8. Data After Pre-processing

No	Teacher Code	School of Origin	Employment Status	Age	Education	Years of Service	Eligibility
1	G-SMP2-01	SMP N 2 Garoga	G1	U5	Bachelor	M3	Yes
2	G-SMP2-02	SMP N 2 Garoga	G1	U5	Bachelor	M3	Yes
3	G-SMP2-03	SMP N 2 Garoga	G1	U2	Bachelor	M3	Yes
4	G-SMP2-04	SMP N 2 Garoga	G1	U4	Bachelor	M2	Yes
5	G-SMP2-05	SMP N 2 Garoga	G1	U5	Bachelor	M1	Yes
6	G-SMP2-06	SMP N 2 Garoga	G1	U4	Bachelor	M2	Yes
7	G-SMP2-07	SMP N 2 Garoga	G1	U4	Bachelor	M2	Yes
8	G-SMP2-08	SMP N 2 Garoga	G1	U5	Bachelor	M2	Yes
9	G-SMP2-09	SMP N 2 Garoga	G1	U4	Bachelor	M3	No
10	G-SMP2-10	SMP N 2 Garoga	G1	U5	Bachelor	M2	No
11	G-SMP2-11	SMP N 2 Garoga	G1	U4	Bachelor	M2	Yes
12	G-SMP2-12	SMP N 2 Garoga	G1	U4	Bachelor	M2	Yes
13	G-SMP2-13	SMP N 2 Garoga	G1	U4	Bachelor	M1	Yes
14	G-SMP2-14	SMP N 2 Garoga	G1	U4	Bachelor	M1	No
15	G-SMP2-15	SMP N 2 Garoga	G2	U2	Bachelor	M1	No
16	G-SMP2-16	SMP N 2 Garoga	G2	U2	Bachelor	M1	No
17	G-SMP2-17	SMP N 2 Garoga	G2	U4	Bachelor	M1	No
18	G-SMP2-18	SMP N 2 Garoga	G2	U4	Bachelor	M2	No
19	G-SMP2-19	SMP N 2 Garoga	G2	U2	Bachelor	M1	No
20	G-SMP5-01	SMP N 5 Garoga	G1	U5	Bachelor	M1	Yes
21	G-SMP5-02	SMP N 5 Garoga	G1	U5	Bachelor	M3	Yes
22	G-SMP5-03	SMP N 5 Garoga	G1	U5	Bachelor	M3	Yes
23	G-SMP5-04	SMP N 5 Garoga	G1	U5	Bachelor	M3	Yes
24	G-SMP5-05	SMP N 5 Garoga	G1	U4	Bachelor	M3	Yes
25	G-SMP5-06	SMP N 5 Garoga	G1	U4	Bachelor	M2	Yes
26	G-SMP5-07	SMP N 5 Garoga	G1	U4	Bachelor	M2	Yes
27	G-SMP5-08	SMP N 5 Garoga	G1	U3	Bachelor	M1	Yes
29	G-SMP5-10	SMP N 5 Garoga	G2	U2	Bachelor	M1	Ya
30	G-SMP5-11	SMP N 5 Garoga	G2	U2	Bachelor	M1	No
31	G-SMP5-12	SMP N 5 Garoga	G2	U2	Bachelor	M1	No
32	G-SMP5-13	SMP N 5 Garoga	G2	U2	Bachelor	M1	No
33	G-SMP7-01	SMP N 7 Pangaribuan	G1	U5	Bachelor	M1	No

34	G-SMP7-02	SMP N 7 Pangaribuan	G1	U4	Bachelor	M3	Yes
35	G-SMP7-03	SMP N 7 Pangaribuan	G1	U5	Bachelor	M2	Yes
36	G-SMP7-04	SMP N 7 Pangaribuan	G1	U3	Bachelor	M2	Yes
37	G-SMP7-05	SMP N 7 Pangaribuan	G1	U3	Bachelor	M1	Yes
38	G-SMP7-06	SMP N 7 Pangaribuan	G1	U3	Bachelor	M1	Yes
39	G-SMP7-07	SMP N 7 Pangaribuan	G1	U5	Bachelor	M1	No
40	G-SMP7-08	SMP N 7 Pangaribuan	G1	U4	Bachelor	M1	No
41	G-SMP7-09	SMP N 7 Pangaribuan	G2	U2	Bachelor	M2	No
42	G-SMP7-10	SMP N 7 Pangaribuan	G1	U5	Bachelor	M4	No
43	G-SMP7-11	SMP N 7 Pangaribuan	G1	U5	Bachelor	M3	Yes
44	G-SMP7-12	SMP N 7 Pangaribuan	G1	U4	Bachelor	M3	Yes
45	G-SD-01	SD Inpres 175765 Padangsiandomang	G1	U5	Bachelor	M3	Yes
46	G-SD-02	SD Inpres 175765 Padangsiandomang	G3	U4	Bachelor	M2	Yes
47	G-SD-03	SD Inpres 175765 Padangsiandomang	G3	U2	Bachelor	M1	No
48	G-SD-04	SD Inpres 175765 Padangsiandomang	G1	U2	Bachelor	M3	No
49	G-SD-05	SD Inpres 175765 Padangsiandomang	G1	U5	Bachelor	M3	No
50	G-SD-06	SD Inpres 175765 Padangsiandomang	G2	U4	Bachelor	M1	Yes

Table 8 displays the results of the data transformation, where out of a total of 50 data, 30 certifications are classified as Yes, while 20 certifications are included in the No category. The next step is to calculate the entropy and gain values for each attribute, as presented in Table 9.

Table 9. Entropy and Gain Results

Attribute	Information	Total Case	Yes	No	Entropy	Gain
Total		50	30	20	0.970950594	
Employment Status						0.45311915
	G1	37	30	7	0.699772222	
	G2	11	0	11	0	
	G3	2	0	2	0	
Age						0.28105947
	U1	0	0	0	0	
	U2	11	1	10	0.439496987	
	U3	5	4	1	0.721928095	
	U4	18	11	7	0.964078765	
	U5	16	14	2	0.543564443	
Years of Service						-6.4111586
	M1	21	7	14	0.918295834	
	M2	14	10	4	0.863120569	
	M3	14	12	2	0.591672779	
	M4	1	1	0	0	

Based on Table 9, the attributes with the highest gain value, with a value of 0.4531192, are employment status and age. Thus, all three attributes can be the root node. Therefore, a decision tree can be formed as shown in Figure 1.

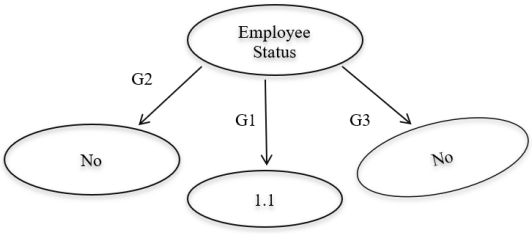

Figure 1. Decision Tree Root Node Employment Status

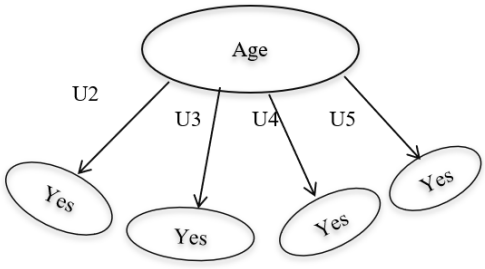

Figure 2. Decision Tree Root Node Age

The decision tree formation process stops because there are no more root attributes to be formed based on the entropy and gain calculations. Figure 1 yields three rules derived from the gain and entropy calculations:

1. Employment Status with G1 is considered Yes
2. Employment Status with G2 is considered No
3. Employment Status with G3 is considered No

Based on Figure 2, four rules were obtained from the gain and entropy calculations:

1. Age with U2 is considered Yes
2. Age with U3 is considered Yes
3. Age with U4 is considered Yes
4. Age with U5 is considered Yes

5. CONCLUSIONS

This study demonstrates that the application of data mining techniques, particularly the C4.5 decision tree algorithm, can effectively support the process of determining teacher certification eligibility in an objective and systematic manner. Based on the analysis of teacher data obtained from the Dapodik system in Pangaribuan and Garoga Districts, North Tapanuli Regency, the C4.5 algorithm is capable of classifying certification eligibility using several key attributes, including employment status, age, education level, and years of service. The results indicate that machine learning approaches are highly relevant for addressing challenges related to subjectivity and inefficiency in conventional certification assessment processes.

The entropy and information gain calculations reveal that employment status and age are the most influential attributes in determining certification eligibility. Teachers with civil servant status (G1) tend to have a higher likelihood of being classified as eligible, while honorary and employment agreement teachers (G2 and G3) are predominantly classified as not eligible. Similarly, age categories U2 to U5 show a strong tendency toward eligibility, indicating that maturity and experience play an important role in certification outcomes. These findings confirm that the decision tree produced by the C4.5 algorithm can generate clear and interpretable decision rules, which is a major advantage for stakeholders who require transparency in policy implementation.

Furthermore, the resulting decision tree model provides an intuitive representation of the classification process, enabling education administrators and policymakers to better understand the underlying factors that influence certification decisions. This transparency not only enhances trust in the system but also supports accountability in the teacher certification program. The model can be used as a decision support tool to assist education authorities in making data-driven and consistent decisions.

Despite its promising results, this study is limited by the relatively small dataset and the scope of the study area. Future research is recommended to involve larger and more diverse datasets, include additional relevant attributes such as competency test scores and teaching performance evaluations, and compare the C4.5 algorithm with other classification methods. Overall, this research confirms that the C4.5 algorithm is a viable and effective approach for supporting teacher certification eligibility assessment and has the potential to be developed further into a comprehensive decision support system in the education sector.

6. REFERENCES

[1] J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, 1993

[2] A. P. Windarto, E. Irawan, W. Saputra, and H. Okprana, "Penerapan Data Mining Menggunakan Algoritma C4.5 Dalam Mengukur Tingkat Kepuasan Pasien BPJS," *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 2, pp. 376–385, 2020.

[3] M. B. Ardima, S. C. Bilqis, and M. Idris, "Penerapan Algoritma C4.5 Untuk Pengukuran Tingkat Kepuasan Siswa Terhadap Kinerja Guru," *Jurnal Informatika Medis*, vol. 2, no. 1, pp. 33–37, 2024.

[4] K. F. L. Simanjuntak, A. C. B. Barus, and A. Anita, "Implementasi Metode Decision Tree dan Algoritma C4.5 Untuk Klasifikasi Kepribadian Masyarakat," *JOISIE (Journal of Information Systems and Informatics Engineering)*, vol. 5, no. 1, pp. 51–59, 2021.

[5] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, 2012.

[6] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From Data Mining to Knowledge Discovery in Databases," *AI Magazine*, vol. 17, no. 3, pp. 37–54, 1996.

[7] S. R. Safavian and D. A. Landgrebe, "A Survey of Decision Tree Classifier Methodology," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 3, pp. 660–674, May 1991.

[8] L. Rokach and O. Maimon, *Data Mining with Decision Trees: Theory and Applications*, Singapore: World Scientific, 2008.

[9] A. I. Silitonga, F. H. Sihotang, V. S. Silaban, T. F. Fahros, and Haryadi, "Penentuan Kelayakan Penerimaan Bantuan Program Indonesia Pintar (PIP) menggunakan Algoritma C4.5", *CESS (Journal of Computer Engineering, System and Science)*, *10*(1), 25–38, 2025.

[10] A. I. Silitonga, L. M. Ginting, Y. Simamora, and F. I. S. H. Sinaga, "Analisis Preferensi Pemilihan Sistem Operasi Smartphone Menggunakan Algoritma C4.5", JATI (Jurnal Mahasiswa Teknik Informatika), *9*(4), 5874–5879, 2025.