Proceeding of International Conference on Advanced Technologies, Innovations and Science
Vol 1, No 1, June 2026, Page 59-66
ISSN XXXX-XXXX (Media Online)
https://journals.adaresearch.or.id/icatis

Classification of Social Assistance Recipients Eligibility Using the K-NN Method

Anas Santriaji[1,*], Ade Fitri Khaerani Rangkuti[1], Dyla Tri Auliah[1], Agus Perdana Windarto[2] , Putrama Alkhairi[3]

[1] Information Systems Study Program, Sekolah Tinggi Ilmu KomputerTunas Bangsa, Pematangsiantar, Indonesia
Jl. Jenderal Sudirman Street, Block A No. 1/2/3, Pematangsiantar 21127, North Sumatra, Indonesia
[2] Master of Information Study Program, Sekolah Tinggi Ilmu Komputer Tunas Bangsa, Pematangsiantar, Indonesia
Jl. Jenderal Sudirman Street, Block A No. 1/2/3, Pematangsiantar 21127, North Sumatra, Indonesia
[3] Information Management Program, Sekolah Tinggi Ilmu Komputer Tunas Bangsa, Pematangsiantar, Indonesia
Jl. Jenderal Sudirman Street, Block A No. 1/2/3, Pematangsiantar 21127, North Sumatra, Indonesia
Email: [1,*]anassantriaji7@gmail.com, [2]adefitrikhaeranirangkuti@gmail.com, [3]dylatriauliah9@gmail.com,
[4]agus.perdana@amiktunasbangsa.ac.id , [5]putrama@amiktunasbangsa.ac.id .
Correspondence Author's Email: anassantriaji7@gmail.com
Submitted: **01/06/2026**; Accepted: **12/06/2026**; Published: **30/06/2026**

Abstract−The classification of the eligibility of social assistance recipients (bansos) was carried out in this study by applying the K-Nearest Neighbor (K-NN) method to support the distribution of assistance that is more targeted, objective, and accurate. The problem behind the research stems from the fact that there is still an element of subjectivity in determining aid recipients, so that the distribution of social assistance has not fully reached the people who need it most. A total of 20 data on prospective recipients of social assistance from Karang Keri Village, Simalungun Regency, were used as research datasets. The variables analyzed included employment, number of dependents, income, and housing conditions, while the status of aid recipients was determined as the target class. Before the classification process is carried out, the data is prepared through a preprocessing stage which includes data selection, data cleaning, encoding, and normalization. Furthermore, the dataset is divided into 70% training data and 30% testing data. The classification process was carried out using the K-NN algorithm with a value of K = 5 and Euclidean Distance calculation. The model's performance was then evaluated using accuracy, precision, recall, F1-score, cross validation, and confusion matrix. The test results showed that the model was able to achieve 83.3% accuracy, 100.0% precision, 75.0% recall, 85.7% F1-score, and 80.0% cross-validation. These findings indicate that the K-NN method has good classification ability and stable performance, so it is suitable for use as a decision-making support in determining social assistance recipients more objectively and on target.

Keywords: Data Mining; K-NN; Social Assistance; Classification; Recipient Eligibility

1. INTRODUCTION

To improve people's welfare and create a more decent life, the government organizes programs in the form of social assistance (bansos) (Fuaddah & Fridayne, 2023). Social assistance is a type of temporary assistance that is selectively determined by the government aimed at benefiting the community, especially the lower middle income group(Pahrudin & Harianto, 2022). The government provides various types of assistance provided in the form of goods and cash(Barat et al., 2023). However, social assistance programs often face a number of problems, especially related to the accuracy of determining aid recipients. Inaccuracy in targeting is the most crucial issue, because social assistance is often not accepted by the right parties and fails to reach the groups most in need(Purbaningrum et al., 2023). In this study, the author took a case study of Karang Keri Village, Simalungun Regency, where the process of determining aid receipts is still carried out based on direct observation and deliberation between village officials. Accurate assessments are often not carried out because the process depends on the personal viewpoints and opinions of the deliberators(Aribowo et al., 2021b). Therefore, a new, more efficient approach is needed so that the determination of social assistance recipients is really on target.

Research related to problems In the context of improving the accuracy of social assistance distribution, various previous studies have been conducted through a data mining-based classification approach. In 2023(Social et al., 2020). Qualitative analysis in villages found that 30-40% of social assistance was mistargeted due to the subjectivity of deliberation, with data mining recommendations. Results: The prediction model increased targeting by 25%. Limitations: Descriptive focus without full algorithm implementation, sample of only 5 villages, and lack of longitudinal data for long-term trends. In 2022(Damuri et al., 2021). The study applied Naïve Bayes to the feasibility of basic food data, obtaining 85% accuracy and 90% recall to detect receivers in low locations. The results prove that this method efficiently processes categorical data such as employment status. Limitations: The attribute of independence assumptions is not always valid on real social data, small sample size (150 data), and minimal integration with real-time government data. In 2022,(Pahrudin & Harianto, 2022). This study used K-NN with Euclidean Distance on data from residents in certain regions, achieving an accuracy of 75.5% to classify social assistance recipients based on income and home ownership variables. Limitations: Sensitivity to optimal k-values (tested only k=3-5), unbalanced data (unbalanced dataset), and lack of external validation outside the local context. In 2021(Alamendah, 2021). The K-NN algorithm was applied to the poor family data, resulting in an accuracy of 9.2% with the confusion matrix showing an F1-score of 0.92 for the eligible class. Optimal results at k=5, minimizing determination errors by up to 12%. Limitations: Slow computation for large datasets, reliance on data normalization, and lack of robust handling of missing value in the field. Then, in 2021(Aribowo et al., 2021a). This study applied the CART algorithm to the recipients of the Family Hope Program (PKH) data in Ngadirejo

Village, resulting in a classification accuracy of 92% in distinguishing eligible and unsuitable families based on income and asset criteria. The results showed a reduction in targeting errors of up to 15% compared to the manual method. Limitations: The dataset is limited to a single village (only 200 samples), is prone to overfitting, and has not taken into account dynamic factors such as post-pandemic economic changes. Therefore, in this study, the application of datamining is very necessary so that errors in determining potential social assistance recipients can be minimized(Nella et al., 2022).

Data mining is a collection of data that is analyzed using techniques, which combine statistics, mathematics, and pattern recognition to produce valuable information(Damuri et al., 2021). Data mining can perform various tasks such as prediction, classification, and information grouping(Issn, 2021). Data mining, which is part of Knowledge Discovery in Database (KDD), consists of a series of steps to explore information in a database that includes data selection, data selection, transformation, data mining and evaluation(Novika et al., 2021). Through the data mining process, various problems can be solved because from the processing of information new insights are created that can be used(Triayudi, 2023). So here the classification process will be carried out first. The classification process is used to form a model or function that is able to describe or differentiate a concept or class of data so that the class of an object whose label is not yet known can be predicted(Trifani, 2022). In classification, patterns are sought to explain or distinguish concepts or categories of data, so that the class of previously unknown objects can be determined(Putri & Wijayanto, 2022). Classification is one of the elements in data mining that aims to understand certain patterns of data so that it can be grouped into several categories(Enyakit et al., 2022). In data mining, data is assessed based on certain requirements so that it can be placed into appropriate categories(Sons & Daughters, 2022). Various methods are used in the classification process in datamining, including decision trees, Naïve Bayesian and Bayesian Trust Networks, Artificial Neural Networks (Backpropagation), techniques based on the discovery of association rules, and other methods (k-Nearest Neighbor, genetic algorithms, and techniques with a rough and fuzzy set approach)(Aziz, 2020). In this study, the classification process uses the K-Nearest Neightbor (K-NN) method. The K-NN method is a method that classifies data by paying attention to the distance between one data and another. The determination of the closest distance is carried out by first separating the data into two categories, namely training and testing data, after the two data are obtained, the distance of each test data to the training data is then calculated using Euclidean Distance (Euclidean Distance). The K-NN method is based on the assumption that data with certain similarities tend to be classified into the same group or have close values(Yolanda & Fahmi, 2021), (Bahtiar, 2023), (Fathah & Juliane, 2025).

In order to improve the accuracy of data-based decision-making, various studies have examined the use of classification in the field of data mining. One of the widely used methods is K-NN, which is known for its conceptual limitations and its ability to group data based on the level of similarity. This study utilizes the K-NN method to solve the problem of receiving social assistance with the aim of obtaining more precise and accurate results. In 2023 by, (Bahtiar, 2023). By applying data mining using the K-Nearest Neighbor method on predictions, 88.89% of the results were obtained using sales data and 80.00% using used material data. With considerable accuracy, it means that the K-Nearest Neighbor method can be used for predictions at the Djaya Twin Frame Shop. Based on an in-depth analysis of the K-Nearest Neighbors (KNN) method for river quality classification, the main conclusion is that KNN is effective in grouping river water quality into 4 classes from 216 samples of 7 IKA parameters. The best accuracy of 78.46% was achieved with 70% training, 30% testing, and k=15. The k value has a big effect on confidence and accuracy(No et al., 2024). The use of the KNN algorithm to determine the majors of new students at SMA N 2 Manokwari resulted in high accuracy of 79% using 60 training data, according to the test with a score of K=9. This algorithm was tested through a matrix confusion formula, which compares the original class and the prediction class, with calculations performed in Microsoft Excel. A total of 74 data samples from SMA N 2 Manokwari were used, and the composition of training, testing, and K value data played a major role in the classification(Nuraeni et al., 2023). The KNN algorithm is a learning supervision for classification/regression based on the closest neighbor of the new data. Performance improvement techniques from tables: Data standardization: Equalize variable scales, improve distribution. Distance weight: Greater weight to close neighbors. Feature selection: Select relevant features, reduce dimensions. Cross-validation: Avoid over/underfitting, evaluate new data. Exact distance metrics: Euclidean, Manhattan, or Cosine(Sakti & Daulay, 2024). The KNN algorithm was applied to the classification of diabetes mellitus using 8 symptoms (age, easy thirst, low BB, high blood pressure, history of diabetes, wounds that are difficult to heal, frequent night baths, high blood sugar). Of the 81 training data & 54 testing data (post-cleaning & normalization), with Euclidean distance K=9, the results: 4 positive, 50 negative (Dwi Fasnuari et al., 2022).

Based on the explanation that has been described above, the purpose of the research is to conduct research on the classification of the eligibility of social assistance recipients. The classification process is carried out to determine whether the community is eligible for social assistance or not. The results to be achieved from the research will no longer be errors in determining families who are eligible for the social assistance.

2. RESEARCH METHODOLOGY

2.1 Research Stages

In this study, the Python programming language was used to implement the K-Nearest Neighbor (K-NN) method with the support of the Scikit-Learn library as a machine learning framework. Through Python, the entire research process is

carried out, including data preprocessing, division of datasets into training and test data, classification using K-NN, and evaluation of model performance based on various performance measurement metrics. Before the research is carried out, the framework needs to be understood first. The research framework is designed to show each stage in the research process, so that the flow of activities becomes more directed and every step from problem definition, method application, to drawing conclusions can be clearly seen. The framework used in this study is seen in the following Figure 1:

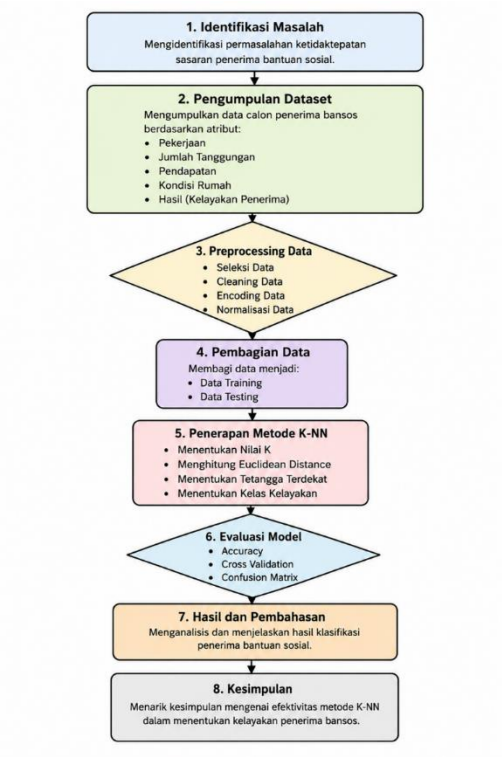

Figure 1. Research Framework

Figure 1 describes the structured process in determining the eligibility of social assistance recipients through the K-Nearest Neighbor (K-NN) method. The research stage begins with the identification of problems related to the inaccuracy of social assistance distribution, then continues with the collection of data on prospective recipients based on several attributes, such as employment, number of dependents, income, housing conditions, and eligibility status. The data that has been obtained is then processed through a preprocessing stage which includes selection, cleaning, encoding, and normalization of data to make it suitable for use in the classification process. Furthermore, the dataset is divided into training data and testing data to train and test the model's performance. In the application of the K-NN method, the K value was determined, Euclidean Distance calculation, the search for the nearest neighbors, and the determination of the eligibility class of aid recipients. Once the model is formed, its performance is tested using accuracy, cross validation, and confusion matrix to measure the level of classification accuracy. The results obtained were then analyzed and discussed as a basis for drawing conclusions about the effectiveness of the K-NN method in determining social assistance recipients more accurately and on target.

2.2 Research Dataset

The research data used is displayed in detail. The data includes 20 alternatives or objects that are used as the basis for new data processing. The decision-making process is supported by four main attributes, namely employment, number of dependents, income, and home conditions. In addition, results that act as class targets are also included for grouping purposes, with two objective categories, namely Accept and No. After the initial data is obtained, the data processing stage is carried out to adjust the characteristics of the data to the method used. The following is a sample of the dataset used.

Table 1. Social Assistance Fund Recipient Data

Yes	Objects	Jobs	Number of dependents	Revenue	Home Condition	Results
1	A1	Factory Workers	1	IDR 1,00,000	Own Property	Receive
2	A2	Freelance Workers	1	IDR 500,000	Own Property	Receive
3	A3	Farmer	2	IDR 1,300,000	Own Property	Receive
4	A4	Private Employees	4	IDR 2,000,000	Own Property	No
5	A5	Freelance Workers	4	IDR 500,000	Own Property	Receive

Yes	Objects	Jobs	Number of dependents	Revenue	Home Condition	Results
6	A6	Farmer	4	IDR 1,100,000	Own Property	Receive
7	A7	Entrepreneurship	1	IDR 2,500,000	Rent	No
8	A8	Freelance Workers	2	IDR 600,000	Own Property	Receive
9	A9	Self-employed	1	IDR 1,500,000	Own Property	No
10	A10	Farmer	6	IDR 1,200,000	Rent	Receive
11	A11	Farmer	3	IDR 1,300,000	Own Property	Receive
12	A12	Self-employed	4	IDR 1,500,000	Own Property	No
13	A13	Entrepreneurship	2	IDR 2,000,000	Own Property	No
14	A14	Factory Workers	3	IDR 1,300,000	Own Property	Receive
15	A15	Entrepreneurship	1	IDR 3,000,000	Own Property	No
16	A16	Factory Workers	4	IDR 1,200,000	Own Property	Receive
17	A17	Private Employees	1	IDR 1,500,000	Own Property	No
18	A18	Private Employees	1	IDR 1,700,000	Own Property	No
19	A19	Factory Workers	1	IDR 1,300,000	Own Property	Receive
20	A20	Self-employed	4	IDR 1,250,000	Own Property	No

2.3 K-NN Method

The K-Nearest Neighbor (K-NN) method is a method that utilizes a number of closest neighbors from the training data to determine the results of calcification, so that the data grouping process is carried out based on the proximity of a data to its neighbors(Said et al., 2022). The determination of the closest distance is carried out by first separating the data into two categories, namely training and test data, after the two data are obtained, the distance of each test data to the training data is then calculated using Euclidean Distance(Bahtiar, 2023). Euclidean Distance is the distance between two points or coordinates calculated with the Pythagorean formula(Akbar, 2024). The Euclidean formula is used to calculate the distance between two points, namely the point in the training data (x) and the point in the test data (y), as shown in the following equation.Q1

$$D_q = \sqrt{(a_1} - b_1)^2 + (a_2 - b_2)^2 + (a_n - b_n)^2 \quad (1)$$

Description :
ai = i-i attribute value on new data
bi = the value of the i-i attribute on the old data
n = number of attributes
Dq = distance between new data and old data

The formula is Euclidean Distance, which is a way of calculating the distance between two data, namely new data (query) and old data (dataset). In K-NN, the data that is the smallest distance means the most similar, so it is used as the closest neighbor to determine the class.

3. RESULTS AND DISCUSSION

3.1 Preprocessing Data

Based on the *results of preprocessing* in table 2, attributes relevant to the research objective, namely determining the eligibility of aid recipients, were selected through the data selection stage. The independent variables used include Employment, Number of Dependents, Income, and House Conditions, while Results are determined as the target variable or decision class. The identity of the data (A1–A20) is still listed as a marker, but is not involved in the modeling process. After that, data *cleaning* is carried out to ensure that there are no *missing values*, duplications, or inconsistencies of the data. The results show that all data are complete, consistent, and in a uniform format so that they are ready to be used in the next stage of analysis.

The next stage is data *encoding*, which is converting categorical data into numerical forms so that it can be processed by *data mining* or *machine learning* algorithms. In the table, the categories in the Occupation, Income, and Home Conditions attributes have been changed to specific numeric codes, while the Results target variables are differentiated into two classes, namely 1 Accepted and 0 No. Furthermore, data normalization is applied to align the range of values between attributes so that no variables dominate the analysis process due to differences in scale. This process results in a more uniform value and is ready to be used as input data at the modeling stage. Thus, the data in the table has gone through the stages of selection, cleaning, coding, and normalization so that it has better quality, consistency, and is suitable for use in the development of research classification models.

Table 2. Data Preprocessing Results

No	Objects	Jobs	Number of dependents	Revenue	Home Condition	Results
1	A1	5	1	1	2	1

Proceeding of International Conference on Advanced Technologies, Innovations and Science
Vol 1, No 1, June 2026, Page 59-66
ISSN XXXX-XXXX (Media Online)
https://journals.adaresearch.or.id/icatis

No	Objects	Jobs	Number of dependents	Revenue	Home Condition	Results
2	A2	6	1	1	2	1
3	A3	4	2	1	2	1
4	A4	1	4	2	2	0
5	A5	6	4	1	2	1
6	A6	4	4	1	2	1
7	A7	2	1	3	1	0
8	A8	6	2	1	2	1
9	A9	3	1	1	2	0
10	A10	4	6	1	1	1
11	A11	4	3	1	2	1
12	A12	3	4	1	2	0
13	A13	2	2	2	2	0
14	A14	5	3	1	2	1
15	A15	2	1	3	2	0
16	A16	5	4	1	2	1
17	A17	1	1	1	2	0
18	A18	1	1	1	2	0
19	A19	5	1	1	2	1
20	A20	3	4	1	2	0

3.2 Data Sharing

Figure 3. Training and Testing

Based on the data distribution diagram, the research dataset is allocated to 70% Data Training and 30% Data Testing. The model is trained using Data Training to study patterns, relationships between attributes, and characteristics that affect the target variable (*Outcome*). The 70% proportion was chosen so that the model obtained adequate data to form a representative classification pattern. In contrast, the model's performance was evaluated using 30% Data Testing which was not included in the training process. Thus, the model's ability to predict new data can be measured more objectively. The 70:30 ratio is used because it has become a common standard in *data mining* and *machine learning research*, so that the balance between the training and testing processes can be maintained and results in more accurate evaluations.

3.3 Application of the K-NN Method

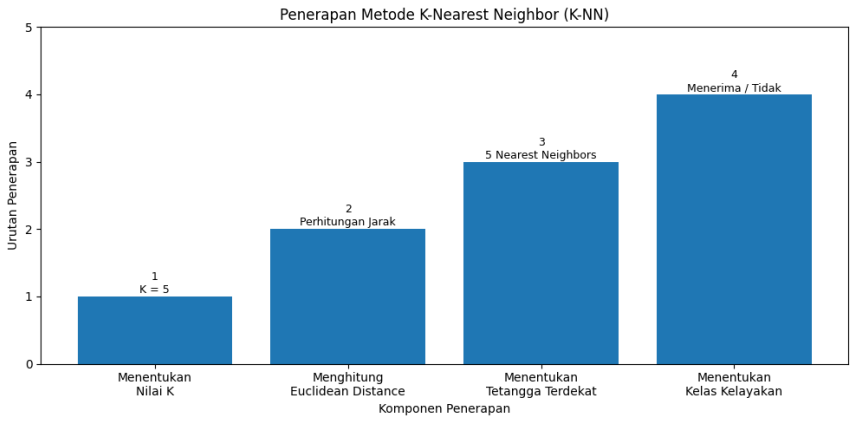

Figure 4. Application of the K-NN Method

Proceeding of International Conference on Advanced Technologies, Innovations and Science
Vol 1, No 1, June 2026, Page 59-66
ISSN XXXX-XXXX (Media Online)
https://journals.adaresearch.or.id/icatis

Based on the diagram of the Implementation of the K-Nearest Neighbor (K-NN) Method, the classification of the eligibility of aid recipients is carried out through four successive stages. The initial stage is in the form of determining the value of K as the number of closest neighbors used in the classification process, with a value of K = 5 in this study. Furthermore, the distance between the test data and all training data is calculated using Euclidean Distance to measure the degree of similarity between the data. After that, five training data with the closest distance to the test data (5 Nearest Neighbors) were selected as the basis for determining the decision. The final stage is the determination of the eligibility class based on the majority of the classes from the closest neighbors that have been obtained, resulting in the category of "Accept" or "No". The diagram illustrates that the K-NN method is applied in a structured manner, starting from determining parameter K to producing a classification of the eligibility of aid recipients based on the proximity of the data.

3.4 Model Evaluation

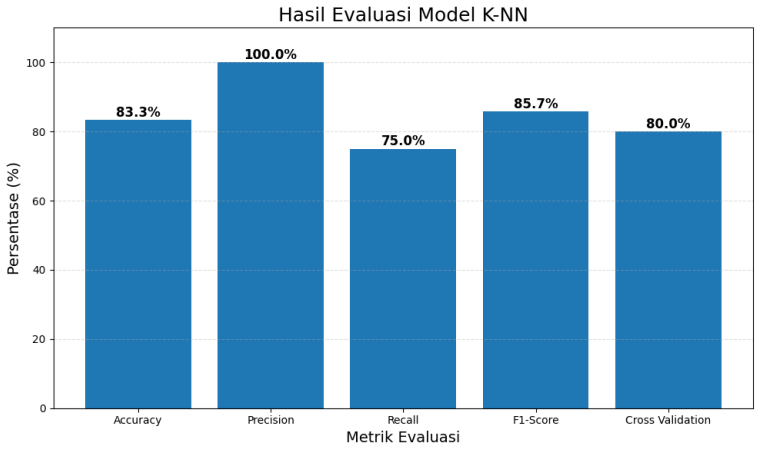

Figure 5. Model Evaluation

The results of the evaluation showed that the K-Nearest Neighbor (K-NN) method was able to classify the eligibility data of social assistance recipients effectively. This ability is reflected in the accuracy value of 83.3%, which indicates that most of the test data has been correctly classified. The model's performance is also supported by a precision value of 100.0%, recall of 75.0%, and an F1-Score of 85.7%, so that the model's ability to distinguish between feasible and unsuitable categories of social assistance recipients can be assessed well. In addition, a cross validation value of 80.0% shows that the model's performance remains stable in various data sharing scenarios. To obtain a more detailed picture of the classification results, including the number of predictions that are appropriate and that are still error-prone, the evaluation is then analyzed through the Confusion Matrix.

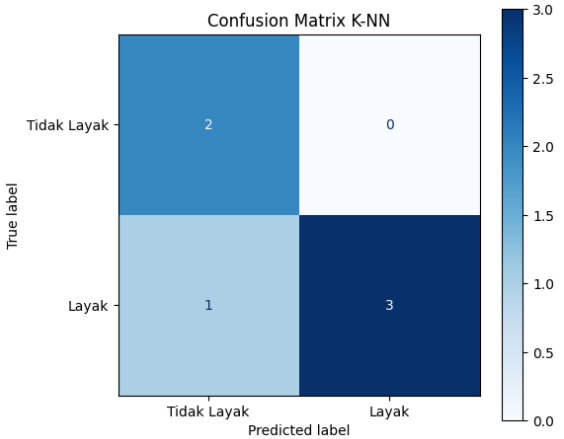

Figure 6. Cofusion Matrix

The results of the evaluation using the K-NN Confusion Matrix on 6 test data showed that the model was able to perform classification well. From the Unfeasible class, all 2 data were successfully recognized according to the actual conditions without any prediction errors to the Eligible class. In the Eligible class, 3 data were correctly classified, while 1 data was incorrectly categorized as Inappropriate. Overall, 5 correct and 1 false prediction were obtained, so that the accuracy level reached 83.3%. A precision value of 100% was obtained because no data was found that was incorrectly predicted as Feasible, while the recall was 75% due to the existence of one Decent data that failed to be identified. Based on these results, it can be concluded that the K-Nearest Neighbor (K-NN) method is effective in classifying the eligibility

of social assistance recipients because the majority of the test data was successfully predicted according to their actual conditions.

4. CONCLUSION

The application of the K-Nearest Neighbor (K-NN) method has been proven to be able to classify the eligibility of social assistance recipients based on job attributes, number of dependents, income, and home conditions. To improve the quality of the data, the preprocessing process is first carried out through the stages of selection, cleaning, encoding, and normalization. The model was built using a value of K = 5, a data distribution of 70% for training and 30% for testing, as well as the calculation of distance with Euclidean Distance. Based on the results of the evaluation, the model produced an accuracy of 83.3%, precision of 100.0%, recall of 75.0%, F1-score of 85.7%, and cross validation of 80.0%. Testing through the confusion matrix showed that 5 out of 6 test data were correctly classified, indicating a good level of model accuracy and consistency. Thus, the K-NN method can be used as an alternative decision support in the selection process of social assistance recipients because it is able to increase the objectivity, accuracy, and effectiveness of aid distribution.

ACKNOWLEDGMENT

The implementation of this research is inseparable from the support of various parties. Therefore, the author expresses his gratitude to Allah SWT for His mercy and guidance so that the research can be completed properly. Awards and thanks were also given to Dr. Agus Perdana Windarto, S.T., M.Kom. for the guidance, direction, and input provided during the research process. Support, prayers, and motivation from parents and families also played a role in the timely completion of this research. In addition, appreciation was expressed to the colleagues of the Information Systems Study Program and Master of Computer Science STIKOM Tunas Bangsa Pematangsiantar, especially Anas Santriaji, Ade Fitri Khaerani Rangkuti, and Dyla Tri Auliah, for their contribution in data collection and discussion during the research.

REFERENCES

Akbar, F. M. N. (2024). KNN (K-Nearest Neighbor) method to determine water quality. *Compact Techno Journal*, *18*(1), 28. Https://Doi.Org/10.33365/Jtk.V18i1.3241

Alamendah, D. (2021). *Optimizing Social Assistance Funds During the Covid-19 Pandemic*. *88* (December).

Aribowo, A., Kuswandhie, R., & Primadasa, Y. (2021a). *Implementation of the CART Algorithm in Determining the Eligibility of PKH Assistance Recipients in Ngadirejo Village Implementation of the CART Algorithm in Determining the*. *7*(1), 40–51.

Aribowo, A., Kuswandhie, R., & Primadasa, Y. (2021b). Application and implementation of the CART algorithm in determining the eligibility of PKH assistance recipients in Ngadirejo Village. *Cogito Smart Journal*, *7*(1), 40–51. Https://Doi.Org/10.31154/Cogito.V7i1.293.40-51

Aziz, R. Z. A. (2020). *Implementation of Rough Set and Naive Bayes Algorithm to Get Rules in Selecting Applicants for Housing Facility Assistance (Case Study: Pringsewu Regency Government). 03*(02).

Bahtiar, R. (2023). Implementation of data mining for the prediction of best-selling frame sales using the K-Nearest Neighbor method. *Journal of MULTI Informatics*, *1*(3), 203–214.

West, J., Communication, J. I., Science, F., & Soedirman, U. J. (2023). *Community Digital Literacy as a Response to the Problem of Social Assistance That Is Not On Target (Case Study in Urban Villages Provides Assistance to People Who Are Experiencing Social Risks). Assistance that groups and communities with financial categories k*. *4*(2), 74–82.

Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementation of Data Mining with Naïve Bayes Algorithm for Classification of Eligibility of Recipients of Basic Food Assistance. *JURIKOM (Journal of Computer Research)*, *8*(6), 219. https://doi.org/10.30865/jurikom.v8i6.3655

Dwi Fasnuari, H. A., Yuana, H., & Chulkamdi, M. T. (2022). Application of K-Nearest Neighbor (K-NN) Algorithm for Classification of Diabetes Mellitus Case Study: Residents of Jatitengah Village. *Antivirus: Scientific Journal of Informatics Engineering*, *16*(2), 133–142.

Enyakit, K. U. K. L. P., Elitus, D. I. M., Asus, S. T. K., & Esa, W. A. D. (2022). *P a k-n n (k-nn) k p d m s k : w d j*. *16*(2), 133–142.

Fathah, A., & Juliane, C. (2025). Data Mining Study for Gender Classification Using Facial Data with Naive Bayes and K Nearest Neighbor (KNN) Algorithms. *Journal of Information Technology and Computer Science*, *12*(1), 99–110. https://doi.org/10.25126/jtiik.2025128724

Fuaddah, A., & Fridayne, R. (2023). Community Digital Literacy as a Response to the Problem of Social Assistance That Is Not On Target (Case Study in Nanggewer Village, Cibinong, Bogor, West Java). *Journal of Paradigm: Multidisciplinary Journal of Indonesian Postgraduate Students*, *4*(2), 74–82.

Issn, P. (2021). Implementation Of The Decision Tree Algorithm For. *7*(2), 45–51.

Nella, N. I., Setiawan, N. Y., & Ratnawati, D. E. (2022). *Classification of Recipients of Family Hope Program Assistance using Decision Tree C4 Algorithm . 5 (Case Study: Mlirip Village, Mojokerto Regency)*. *6*(3), 1332–1339.

No, V., Hal, O., Sandra, M., & Erny, D. (2024). *Water quality classification uses K-Nearest Neighbors, Naïve Bayes, and Logistic Regression*. *6*(4), 851–858.

Novika, T., Okprana, H., Windarto, A. P., & Siahaan, H. (2021). *Application of Data Mining Classification of Students' Level of Understanding in Mathematics Lessons*. *5*:9–17. https://doi.org/10.30865/mib.v5i1.2498

Nuraeni, S., Putri, S., Syam, A., Wajdi, M. F., & Malkan, M. (2023). *G-Tech: Journal of Applied Technology Implementation of the K-NN Method to Determine Students' Majors at SMAN 02 Manokwari*. *7*(1), 89–95.

Pahrudin, P., & Harianto, K. (2022). Application of the K-Nearest Neighbor algorithm for the classification of social assistance

recipients. *Building of Informatics, Technology and Science (BITS), 4*(3), 1241–1245. https://doi.org/10.47065/bits.v4i3.2276

Purbaningrum, D., Adinugraha, & Adinugraha, H. H. (2023). Inaccuracy of social assistance targets in villages. *Journal of Development and Public Policy*, *15*(2), 31–44.

Putra, M. Y., & Putri, D. I. (2022). Utilization of Naïve Bayes and K-Nearest Neighbor Algorithms for Classification of Grade XI Students' Majors. *Journal of Compact Technology*, *16*(2), 176. https://doi.org/10.33365/jtk.v16i2.2002

Putri, N. B., & Wijayanto, A. W. (2022). Comparative Analysis of Data Mining Classification Algorithms in Phishing Website Classification. *Computing : Journal of Computer Systems*, *11*(1), 59–66. https://doi.org/10.34010/komputika.v11i1.4350

Said, H., Matondang, N. H., & Irmanda, H. N. (2022). Application of the K-Nearest Neighbor algorithm to predict the quality of consumable water. *Techno.Com*, *21*(2), 256–267. https://doi.org/10.33633/tc.v21i2.5901

Sakti, R., & Daulay, A. (2024). *Critical Analysis and Development of the K-Nearest Neighbor (KNN) Algorithm : A Literature Review*. *4*(2), 131–141.

Social, T., Distribution, S., Through, I., & Social, I. (2020). *Innovation in the distribution of social security is on target through the Social Welfare Integrated Data Budget Management Policy (DTKS) and the use of the "Social Assistance Check" application*. 204–215.

Triayudi, A. (2023). The application of data mining for the classification of recipients of social assistance funds using the K-Nearest Neighbor algorithm. *Building of Informatics, Technology and Science (BITS), 5*(2), 532–542. https://doi.org/10.47065/bits.v5i2.3972

Trifani, A. (2022). *Application of Data Mining C4 Classification .5 in Determining the Stress Level of Final Students*. *1*(2).

Yolanda, I., & Fahmi, H. (2021). *The Application of Data Mining for the Prediction of Sales of the Best-Selling Bread Products at PT. Nippon Indosari Corpindo Tbk uses the K-Nearest Neighbor method*. *3*(3), 9–15.