Inform : Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi
Vol.10 No.1 January 2025, P-ISSN : 2502-3470, E-ISSN : 2581-0367

Comparative Analysis of Decision Tree and Artificial Neural Network Methods for Predicting Potential Heart Disease

Farrel Muhammad Raihan Akhdan[1], Ade Ismail[2], Irsyad Arif Mashudi[3], Anastasia Lidya Maukar[4]

[1]*Informatics Department, Universitas Islam Indonesia, Indonesia*

[2,3]*Informatics Department, Politeknik Negeri Malang, Indonesia*

[4]*Industrial Engineering Department, President University, Indonesia*

[1]farrel.akhdan76@gmail.com(*)

[2,3][aismail, irsyad.arif]@polinema.ac.id, [4]almaukar@president.ac.id

Received: 2024-07-17; Accepted: 2025-01-09; Published: 2025-01-18

Abstract— Prediction models have been used in various fields such as health, education, and industry. This system can connect various data collected to be used as learning for the system in solving a problem similar to the data used as learning. The prediction model involves various elements such as mathematics, machine learning, and statistics. Heart disease remains a leading cause of mortality globally, and accurate prediction models are crucial for early detection and treatment. However, existing models often struggle with dataset imbalance, leading to suboptimal performance. This study aims to compare the performance of Decision Tree and Artificial Neural Network (ANN) models, including the Elman and Jordan variants, to identify the most suitable prediction model for heart disease with a quantitative study. The type of ANN used is multi-layer with the Elman and Jordan models. However, a comparative analysis of heart disease objects was carried out using the K-Nearest Neighbor (KNN) and Naïve Bayes methods, which resulted in Naïve Bayes being better than KNN. From all the processes that have been carried out, the researchers obtained results from precision, recall, and F1-score, which were classified as poor, with an average of 55%. The Decision Tree model achieved an average accuracy of 79%, while the Elman and Jordan networks achieved 87% and 86%, respectively. However, precision, recall, and F1-scores were relatively low, averaging 55%, likely due to dataset imbalance. The accuracy results obtained are also not always directly proportional to the amount of data used. There is a significant decline at the beginning of the process, but the accuracy obtained continues to increase until all the data is used. Apart from that, there was a spike in precision, up to 80%, in several implementation processes with prediction models. Based on the results obtained in the implementation process, it can be said that the Elman Network is superior to other methods when using accuracy benchmarks. However, the relatively low precision, recall, and F1-score results indicate the model's performance is lacking.

Keywords— Decision Tree; Artificial Neural Network; Confusion Matrix; Heart Disease; Prediction; Data Mining.

I. INTRODUCTION

Heart disease is one of the deadliest diseases to date. This disease is caused by blood vessel blockage that leads to the heart due to fat and cholesterol deposits [1]. There is a need for a method that can warn individuals that their current health status has the potential to cause coronary heart disease during the next ten years. This is necessary to prevent individuals from developing this disease [2]. One approach that can be used in this problem is a prediction model.

The prediction model is often used in scientific research, namely data mining [3][4]. Data mining is a process for connecting various data [5][6], which involves various technical elements such as mathematics, machine learning, and statistics [7], [8], [9]. Data mining is divided into 3 types, namely classification or what can be called prediction, clustering, and association [10]. Data mining is extracting data to obtain recommendations in the form of predictions through new knowledge [11].

The prediction has been widely applied in various fields such as [12], which discusses the comparative analysis of the Naive Bayes and K-Nearest Neighbor algorithms for heart disease prediction, [13] which discusses the comparison of 3 prediction models with the object of heart disease, and [1] which uses Naive Bayes for predicting heart disease. By studying previous research, there is a similar problem, namely that the data used is still relatively small, and it is possible that using other methods can increase the accuracy obtained [1] [9][10][11][16]. In research [12], there is a suggestion to compare the Decision Tree and Artificial Neural Network (ANN) methods using a larger dataset to get better results than previous research. Research [17] conducted a comparative analysis of the Decision Tree and ANN methods with different objects. In this context, ANN is relevant to datasets in the form of images, while Decision Trees are relevant to datasets in numerical form. Even though ANN has advantages in terms of its learning model compared to Decision Trees, can ANN outperform Decision Trees, which are already relevant to non-image or numerical datasets? Thus, comparing the two algorithms using different data quantities and varying the training and testing data ratios, compared to previous studies with smaller data sets, can impact the analysis results.

II. RESEARCH METHODOLOGY

This research compares the two algorithms represented by Fig. 1, illustrating the research process starting with raw data to the comparative analysis process of the two algorithms. Raw data must go through pre-processing to filter out

DOI : https://doi.org/10.25139/inform.v10i1.8631

unnecessary or invalid datasets. After this process, the original 3,300 data entries were reduced to 2,927 due to the presence of entries with empty or invalid values. Next, the dataset will be divided into training data and testing data with a ratio of 70:30, 80:20, and 90:10. In The next step, the training data will be implemented into algorithm models, namely Decision Tree, Elman Network and Jordan Network. After the model has been successfully trained, data testing will test the model's performance using the confusion matrix method to obtain accuracy, precision, recall, and F1-score. This process will be repeated with parameters for the number of datasets used, starting from 500 data until all data is used. The training and testing data positions are exchanged to obtain gaps in the training data results and testing data at each processing path.

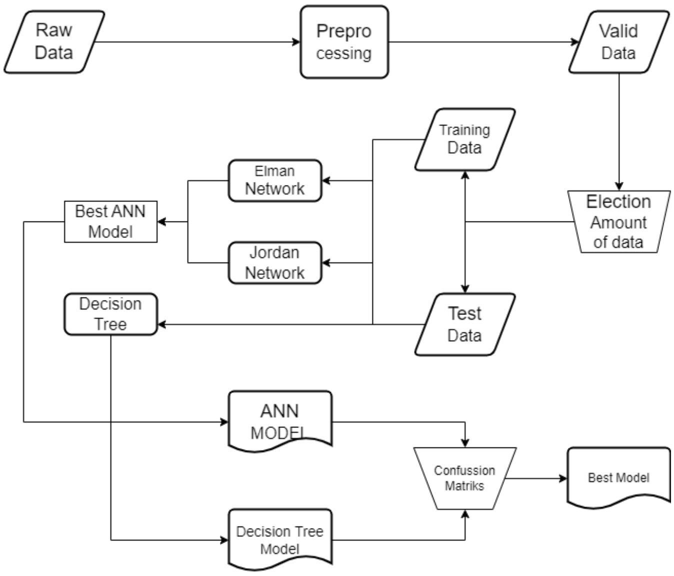

Fig.1. Research Stage

A. *Decision Tree*

The Decision Tree algorithm is a commonly used Data Mining algorithm [18]. Like prediction algorithms in Data Mining, Decision Trees are also used to build classification systems based on several covariates. Decision Trees can also be used to develop prediction models for targeted variables. One of the advantages of the Decision Tree algorithm is the efficiency level in handling complex and large data sets. This algorithm classifies a pattern into segments resembling branches on an inverted tree [18].

Mathematically, the entropy D for a dataset with n classes is defined as where pi is the proportion of class i in dataset D. The entropy value reaches a maximum when all classes are evenly distributed, indicating high uncertainty and reaches a minimum (i.e., zero) when the dataset contains only one class, indicating low uncertainty. This formula can be seen in Equation (1).

$$Info(D) = \sum_{i=1}^{n} -p_i \, log_2(p_i) \quad (1)$$

Decision Tree is a type of algorithmic approach using divide and conquer. This allows Decision Trees to learn problems from independent data sets. The following are the stages of implementing the Decision Tree model on the dataset.

The flow of the Decision Tree method is illustrated in Fig.2. Starting from raw data, the data normalization process is carried out with pre-processing until the data becomes valid. Next, divide the data into training data and testing data with the previous ratio, namely 70:30, 80:20, and 90:10. Meanwhile, researchers will also build a Decision Tree model using a library in Python. Next, the model that has been built will be ready for model training and testing with the Confusion Matrix.

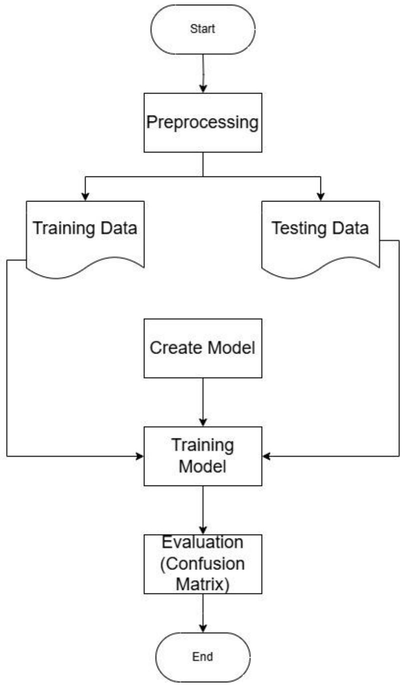

Fig.2. Flow Decision Tree

B. *Artificial Neural Network*

Artificial Neural Network (ANN) is a computational model inspired by the structure and function of human biological neural networks. ANN models complex relationships between input and output, learns patterns, and performs tasks such as pattern recognition, classification, regression, etc. ANN has several types: single-layer, multi-layer, and recurrent-layer. The three types of ANN have differences in their modelling architecture. Single-layer uses two input layers and an output layer. The input-output layer in the single layer acts as a recipient of input data, while the output layer acts as a medium for providing output results. Meanwhile, multi-layer has an additional layer, namely a hidden layer whose function is to carry out calculations from the input layer, which is continued to the output layer. In a multi-layer architecture, more than one hidden layer can be based on the data used.

The last type of ANN is Recurrent. This architecture differs from the previous architectures, fed forward or moving forward. This architecture is unique because there must be at

DOI : https://doi.org/10.25139/inform.v10i1.8631

least one feedback loop to improve the model's ability to learn the data set. This architecture itself has been developed by several researchers, two of whom are Elman and Jordan. These two researchers have different ways of developing this architecture, namely in original learning.

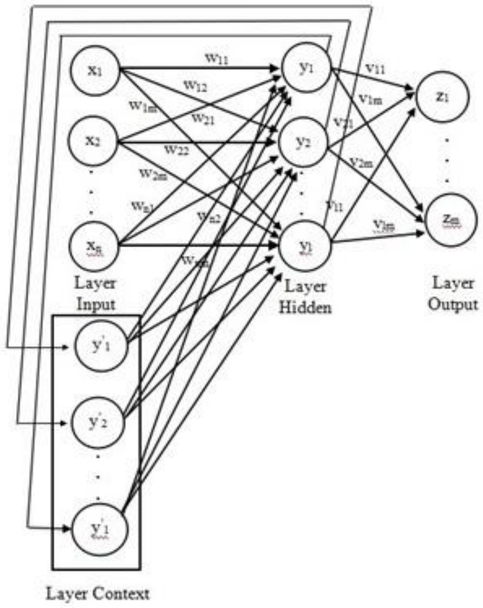

Fig.3 Elman Network

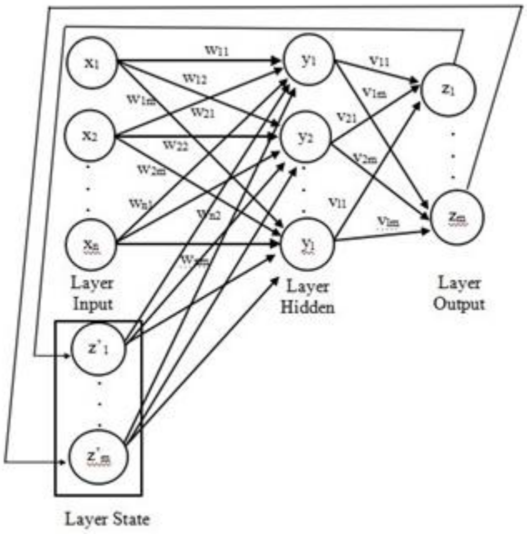

Fig.4. Jordan Network

In Fig.3, Elman carries out the learning process by making a copy of the input layer, which functions as an extension of the input layer called the context layer. The context layer functions as a place to store the status of the hidden layer, which will later be returned to the hidden layer. While the development carried out by Jordan in Fig.4 was different from that of Elman, who copied the input layer, Jordan chose to copy the output layer. The copy of the output layer is called the state layer. In the architecture developed by Jordan, the output results in the previous iteration will become part of the next iteration.

C. *Confusion Matrix*

Confusion matrix is an important performance evaluation tool in classification modeling in machine learning [19]. This matrix summarizes the prediction results from a model by comparing the actual classes from the test data with the classes predicted by the model. In the confusion matrix, the four main categories include True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). This information can calculate various evaluation metrics such as accuracy, precision, recall, and F1-score [20].

III. Result and Discussion

The data used in this research comes from the Kaggle repository. This data set, "Cardiovascular Risk Factor Data," contains patient information. The total dataset researchers use consists of 3,390 observation data points, each with 16 criteria. The criteria include age, education, gender, is_smoking, cigsPerDay, BPMeds, prevalentStroke, prevalentHyp, diabetes, totChol, sysBP, diaBP, BMI, heart rate, glucose, and TenYearCHD. Example data can be seen in Table I.

TABLE I
RAW DATA

Index	Age	Sex	…	Heartrate	Glucose	TenyearCHD
1	64	F	…	90	80	1
2	36	M	…	72	75	0
3	46	F	…	88	94	0
4	50	M	…	68	94	1
…	…	…	…	…	…	…
3386	46	F	…	70	103	0
3387	44	M	…	80	84	0
3388	60	M	…	73	72	1
3389	54	F	…	80	85	0

A. *Pre-processing*

At this stage, the dataset obtained by researchers will undergo processing to get a proposed model for comparison.

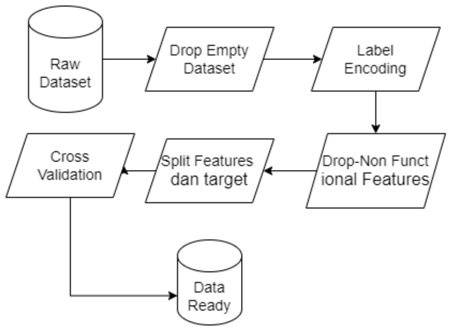

Fig.5 Pre-processing Stage

DOI : https://doi.org/10.25139/inform.v10i1.8631

In Fig.5, the raw dataset will undergo pre-processing techniques to obtain a more valid dataset. During this process, checks will be carried out for missing or zero data, labeling, and selection. This process is very important in designing predictive models because if the dataset is invalid, the level of accuracy cannot be accurately assessed. An example of validated data can be seen in Table II.

TABLE II
VALID DATA

Index	Age	Sex	…	Heartrate	Glucose	TenyearCHD
1	64	0	…	90	80	1
2	36	1	…	72	75	0
3	46	0	…	88	94	0
4	50	1	…	68	94	1
…	…	…	…	…	…	…
3386	46	0	…	70	103	0
3387	44	1	…	80	84	0
3388	60	1	…	73	72	1
3389	54	0	…	80	85	0

B. Classification Using Decision Tree and ANN Algorithms

The classification results from the decision tree method using the Confusion Matrix obtained an average accuracy of 79%. Meanwhile, the ANN method with the Elman model obtained an average accuracy of 86% and the Jordan model of 85%. This average was obtained through training and testing data percentages of 70:30, 80:20, and 90:10. In Fig.6, there is an example of the implementation results using testing data. In contrast, Fig.7 uses training data. These results are based on a percentage of 80% training data and 20% testing data.

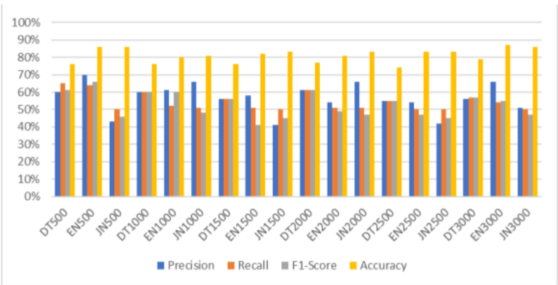

Fig.6 Graph of Hypothesis Testing Data Results

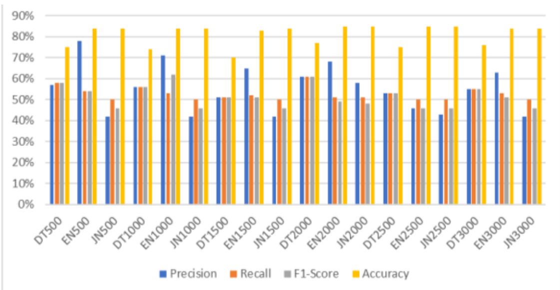

Fig.7 Graph of Hypothesis Training Data Results

Furthermore, to add analytical benchmarks, researchers obtained the average precision, recall, and F1-score for both methods using the Confusion Matrix. However, the results of obtaining precision, recall, and F1-score for both algorithms cannot meet the minimum limit set by the researchers, which is 70%. This is caused by an imbalance in the dataset used. There is a ratio of 3 to 5 to 1 between non-potential and potential results.

C. Comparative Analysis of Decision Tree and ANN Algorithms

Based on the results of the hypothesis, researchers compared the gap between the accuracy obtained from training data and testing data. It can be seen in the diagram above that the Elman Network model is superior, indicated by the small average gap between testing and training data that has been illustrated in Fig.8. Meanwhile, the Decision Tree algorithm tends to have a weak gap between training and testing results, this indicates that there is no balance between the two results obtained. Jordan Network, a sibling of Elman Network, almost outperforms the performance regarding the gap between training and testing data.

Fig.8 Graph Comparison GAP Testing and Training

In this study, researchers obtained average precision, recall, and F1-score results, which tended to be very low. This can be caused by various factors, such as comparing the final results of unbalanced data shown in each hypothesis in the implementation and using non-potential values 3 to 5 times more than the potential value. Even though the precision obtained was more than 80% at a certain point, it could not cover the other two benchmarks. Obtaining precision, recall, and F1-score results is very important for the performance of an algorithm apart from accuracy.

If determining the best algorithm for the dataset researchers use is based only on the accuracy results obtained, the Elman Network is at the forefront. In this case, it can answer the first problem formulation. Meanwhile, if the determination also involves precision, recall, and F1-score results, then no model can meet the minimum threshold because it is too low. However, in its application on a mobile platform, Decision Tree has performance that can outperform the others with 80% accuracy when ANN gets 50% accuracy. This could be a question for further research.

Based on the results of this research, several contributions were obtained that can be used for further research. Even though researchers have used more datasets and variables than previous studies, they cannot perform better. Therefore, it is

DOI : https://doi.org/10.25139/inform.v10i1.8631

necessary to conduct a review to test the two algorithms using a larger dataset and to have a balanced comparison of results.

IV. CONCLUSION

Heart disease is a leading cause of death worldwide, necessitating the development of predictive models that can assist in early diagnosis and intervention. This study aims to evaluate various machine learning models for predicting the potential of heart disease in humans using a dataset obtained from Kaggle. When applied to varying amounts of data, the focus is on determining the accuracy and reliability of different models, including Decision Tree, Elman Network, and Jordan Network.

This study's dataset is related to the potential for human heart disease. This dataset was obtained via Kaggle and included 3000 CSV data. As a form of the process carried out before designing the model, the data was normalized through a pre-processing process and then divided into training data and testing data with a ratio of 70:30, 80:20, and 90:10. At the model testing stage. The researcher carried out a hypothesis using 500 data points first and then added them to the entire data set. This is done to test the model's performance with different amounts of data. Based on the results, the number of datasets used is not always directly proportional to the accuracy obtained by the model through the confusion matrix process. The average accuracy results obtained by the Decision Tree model were 79%, the Elman Network was 86%, and the Jordan Network was 85%.

Looking at the accuracy results obtained by the two models, the Elman Network shows superiority over the other models. This is supported by the average gain in precision, recall, and F1-score for each output, which tends to be higher than that of other models. However, there is a too high gap in the results between output 0 and 1. Therefore, even though the average of the Elman Network model is higher, the results obtained through precision, recall, and F1-score still need to be improved because they are less than the minimum required limit.

In conclusion, while the study provides valuable insights into heart disease prediction using machine learning, further efforts are required to address the identified gaps and enhance model robustness, ensuring its applicability in real-world medical diagnostics.

REFERENCES

[1] D. Larassati, A. Zaidiah, and S. Afrizal, "Sistem Prediksi Penyakit Jantung Koroner Menggunakan Metode Naive Bayes," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 7, no. 2, pp. 533–546, 2022, doi: 10.29100/jipi.v7i2.2842.

[2] D. Saraswati, "Inovasi Pelayanan Kesehatan : Deteksi Dini Penyakit Jantung Koroner melalui Posbindu PTM," vol. 2, pp. 10–16, 2024.

[3] E. Haryatmi and S. Pramita Hervianti, "Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 386–392, 2021, doi: 10.29207/resti.v5i2.3007.

[4] M. Rizq Daffa Jodi, "Fakultas Komputer Algoritma dan Struktur data," *Fak. Kompiter*, vol. 1, 2020.

[5] R. Ghorbani and R. Ghousi, "Predictive data mining approaches in medical diagnosis: A review of some diseases prediction," *Int. J. Data Netw. Sci.*, vol. 3, no. 2, pp. 47–70, 2019, doi: 10.5267/j.ijdns.2019.1.003.

[6] Sylvia Rheny, "Algoritma adalah: Pengertian, Fungsi, 5 Ciri, dan Contohnya," Ekrut Media.

[7] R. S. Wahono, *Data Mining Data mining*, vol. 2, no. January 2013. 2005.

[8] S. Aulia, "Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja)," *Djtechno J. Teknol. Inf.*, vol. 1, no. 1, pp. 1–5, 2021, doi: 10.46576/djtechno.v1i1.964.

[9] R. Ordila, R. Wahyuni, Y. Irawan, and M. Yulia Sari, "Penerapan Data Mining Untuk Pengelompokan Data Rekam Medis Pasien Berdasarkan Jenis Penyakit Dengan Algoritma Clustering (Studi Kasus : Poli Klinik PT.Inecda)," *J. Ilmu Komput.*, vol. 9, no. 2, pp. 148–153, 2020, doi: 10.33060/jik/2020/vol9.iss2.181.

[10] F. Fahrianto, "Data Warehouse dan Data Mining," *GAES-PACE B. Publ.*, p. 193, 2016.

[11] A. Journal *et al.*, "Penggunaan Data Mining Sebagai Pengambilan Keputusan Penerimaan Bantuan Terhadap Rumah Ibadah (Studi Kasus : Rumah Ibadah Di Kabupaten Tanggamus)," *Aisyah J. Informatics Electr. Eng.*, vol. 4, no. 1, pp. 55–65, 2022.

[12] D. A. Langga and Dkk, "Perbandingan Algoritma Naive Bayes Dengan Algoritma K-Nearest Neighbor Untuk Prediksi Penyakit Jantung," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2019.

[13] I. Alhabib, "Komparasi Metode Deep Learning, Naïve Bayes Dan Random Forest Untuk Prediksi Penyakit Jantung," *INFORMATICS Educ. Prof. J. Informatics*, vol. 6, no. 2, p. 176, 2022, doi: 10.51211/itbi.v6i2.1881.

[14] D. Derisma, "Perbandingan Kinerja Algoritma untuk Prediksi Penyakit Jantung dengan Teknik Data Mining," *J. Appl. Informatics Comput.*, vol. 4, no. 1, pp. 84–88, 2020, doi: 10.30871/jaic.v4i1.2152.

[15] S. Sulastri, K. Hadiono, and M. T. Anwar, "Analisis Perbandingan Klasifikasi Prediksi Penyakit Hepatitis Dengan Menggunakan Algoritma K-Nearest Neighbor, Naïve Bayes Dan Neural Network," *Dinamik*, vol. 24, no. 2, pp. 82–91, 2020, doi: 10.35315/dinamik.v24i2.7867.

[16] R. Annisa, "Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Penderita Penyakit Jantung," *J. Tek. Inform. Kaputama*, vol. 3, no. 1, pp. 22–28, 2019.

[17] K. Algoritma, D. Tree, and S. V. M. Ann, "Komprasi Algoritma," vol. 16, pp. 21–27, 2023.

[18] Y. Y. Song and Y. Lu, "Decision tree methods: applications for classification and prediction," *Shanghai Arch. Psychiatry*, vol. 27, no. 2, 2015, doi: 10.11919/j.issn.1002-0829.215044.

[19] N. V. Pusean, N. Charibaldi, and B. Santosa, "Comparison of Scenario Pre-processing Performance on Support Vector Machine and Naïve Bayes Algorithms for Sentiment Analysis," *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 8, no. 1, pp. 57–63, 2023, doi: 10.25139/inform.v8i1.5667.

[20] K. Izzah, K. Khalid, and D. Rolliawati, "Decision Support System for Determining the Feasibility of a Program Keluarga Harapan Receiver Using the Analytic Network Process Algorithm," *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 5, no. 1, pp. 45–53, 2020, doi: 10.25139/inform.v5i1.2300

DOI : https://doi.org/10.25139/inform.v10i1.8631