Institute of Research and Publication Indonesia (IRPI)
Public Research Journal of Engineering, Data Technology and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/predatecs
Vol. 2 Iss. 1 July 2024, pp: 44-51
ISSN(P): 3024-921X | ISSN(E): 3024-8043

Classification of Diabetes Mellitus Sufferers Eating Patterns Using K-Nearest Neighbors, Naïve Bayes and Decission Tree

Ayuni Fachrunisa Lubis[1*], Hilmi Zalnel Haq[2], Indah Lestari[3], Muhammad Iltizam[4], Nitasnim Samae[5], Muhammad Aufi Rofiqi[6], Sakhi Hasan Abdurrahman[7], Balqis Hamasatiy Tambusai[8], Puja Khalwa Salsilah[9], Bilal Shahid[10]

[1,2,3]Department of information System, Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[4]Sarjana Pendidikan, Fakultas Pendidikan, Universiti Pendidikan Sultan Idris, Malaysia
[5]Department of Teaching Islamic Studies, Faculty of Islamic Science,
Prince of Songkla University - Pattani Campus, Thailand
[6]Department of Hadith Science, Faculty of Ushuluddin, Universitas Al-Azhar, Egypt
[7]Department of Tafsir and Science of the Koran, Faculty of Ushuluddin, Universitas Al-Azhar, Egypt
[8]Department of Human Resources Management, Faculty of Economics and Administration,
Süleyman Demirel University, Turkey
[9]Department of Sociology, Faculty of literature, Süleyman Demirel University, Turkey
[10]Department of Sales and Marketing, Faculty of Management Sciences,
Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Pakistan

E-Mail: [1]12150322141@students.uin-suska.ac.id, [2]12150311054@students.uin-suska.ac.id,
[3]12150324262@students.uin-suska.ac.id, [4]ijampetra@gmail.com, [5]nitasnimsamae12@gmail.com,
[6]Aufirofiqi@gmail.com, [7]udelgateltv@gmail.com, [8]balqishamasatiy@gmail.com,
[9]khalwaspuja@gmail.com, [10]bilalshahid1226@gmail.com

Received Dec 30th 2023; Revised Mar 20th 2024; Accepted Apr 20th 2024
Corresponding Author: Ayuni Fachrunisa Lubis

Abstract

The study investigates three classification algorithms, namely K-Nearest Neighbor (K-NN), Naïve Bayes, and Decision Tree, for the classification of Diabetes Mellitus using a dataset from Kaggle. K-NN relies on distance calculations between test and training data, using the Euclidean distance formula. The choice of k, representing the nearest neighbor, significantly influences K-NN's effectiveness. Naïve Bayes, a probabilistic method, predicts class probabilities based on past events, and it employs the Gaussian distribution method for continuous data. Decision Trees, form prediction models with easily implementable rules. Data collection involves obtaining a Diabetes Mellitus dataset with eight attributes. Data preprocessing includes cleaning and normalization to minimize inconsistencies and incomplete data. The classification algorithms are applied using the Rapidminer tool, and the results are compared for accuracy. Naïve Bayes yields 77.34% accuracy, K-NN performance depends on the chosen k value, and Decision Trees generate rules for classification. The study provides insights into the strengths and weaknesses of each algorithm for diabetes classification.

Keyword: Classification Algorithms, Decision Tree, Diabetes Mellitus, K-Nearest Neighbor, Naïve Bayes

1. INTRODUCTION

The progression of technology and the ongoing evolution of time have had a significant impact on human lifestyles, which are rapidly shifting from traditional to modern. These alterations also bring about alterations in the emergence pattern of diseases, particularly those associated with an individual's lifestyle [1]. An unhealthy lifestyle contributes to the development of conditions like obesity, hypertension, coronary heart disease, and diabetes mellitus. Diabetes mellitus, more commonly referred to as diabetes, is a long-term metabolic illness that is typified by blood sugar levels that are higher than normal [2]. The high sugar levels are caused by the body not being able to process food into energy [1].

The high prevalence of Diabetes Mellitus (DM) in Indonesia highlights the need to evaluate how much patients' quality of life might be impacted by their understanding of DM and how to manage it. Patients should educate themselves about DM as this will impact how they manage their disease to control blood sugar and avoid complications[3]. Lack of awareness regarding DM can frequently be a barrier to managing the illness

since it affects choices made about how to help someone with DM overcome obstacles to engaging in self-care activities like exercise, good eating, and lifestyle[4].

Diabetes is a long-term problem that involves many risk factors, complexity, and a high mortality rate. There are four types of diabetes: type 1 diabetes, type 2 diabetes, prediabetes, and gestational diabetes [3]. Diabetes disease affects approximately 6.6% of the entire adult population. According to global health reports, it is estimated that the number of people with diabetes will increase from 376 million to 490 million by 2030 [4] [5].

In DM disease there are 4 attributes as a reference for calculating sugar levels in the patient's blood. these attributes Later it will form a pattern with existing food menu classes. From data In this way, doctors can detect patterns with using several methods, one of which is using data mining. Data mining is a method for analyzing patterns or knowledge from a collection of data automatic. Data mining techniques that can used for classification include: clustering techniques and classification techniques [6].

One of the methods used in this research is decision trees. A decision tree is a prediction system that can be used to describe an issue and identify a solution because it resembles a branching tree structure or a hierarchical structure [7]. Applied disciplines including finance, marketing, engineering, and medical regularly use decision trees. Decision tree classifiers are widely used for disease diagnosis, including prediction of diabetes, breast cancer, ovarian cancer, diagnosis of heart sounds, and other conditions [8].

In determining whether someone has diabetes, it can also be done using the K-Nearest Neighbor (K-NN) classification method, in its application the best K value (number of nearest neighbors) will be looked for[9]. Technique is based on training data with objects that have the closest distance [10]. Next, use the Naïve Bayes algorithm to find out the classification results between the 2 categories and find out the level of accuracy of the existing samples[11]. This technique utilizes the principles of probability, which is a branch of mathematics, to find the highest probability of a classification by considering how often each classification appears in the training data. The Naive Bayes algorithm is known for its efficiency, which allows it to reduce the time required in classification [12].

The research conducted by [13] in predicting Diabetes using Random Forest, K-NN, Naïve Bayes, and J48 algorithms resulted in NB and J48 prediction algorithms are better for large datasets large dataset analysis. K-NN technique is not good for large dataset analysis. The same research was conducted by [14] Diabetes prediction using the K-NN and Naïve Bayes algorithms produced an accuracy of 76.07% using Naive Bayes, while K-NN produced an accuracy of 73.33%. Further research conducted [5] using the Support Vector Machine (SVM), Naive Bayes (NB), K-NN and C4.5 Decision Tree resulted in C4.5 Decision Tree getting the highest accuracy.

Meanwhile, research conducted by [3] resulted in the prediction of Diabetes shows that the accuracy produced by each algorithm is K-NN with 78.5%, Naive Bayes with 75.9% accuracy, SVM with 78.5%, Decision Tree with 72% accuracy, and Random Forest with 79.8% accuracy. Random Forest produces the highest accuracy value. Another research conducted by [15] conducted a comparison of the Decision Tree, Random Forest, Naïve Bayes, SVM and K-NN algorithm methods to detect leaf diseases in rice plants resulting in the best method among the five, namely the K-NN method with an accuracy value of 87%.

This research aims to develop a classification method diabetes using an algorithm Decision Tree, K-NN, and Naive Bayes. Clinical data from patients who diagnosed with diabetes and those who are not diagnosed with diabetes will be analyzed for identify related patterns with this disease. Attributes such as age, body mass index (BMI), blood pressure, and the results of laboratory tests related to blood sugar will be used as inner variables this analysis [16]. By using an algorithm Decision Tree, K-NN and Naive Bayes are expected that method classification developed in research This can provide accurate results in diagnose diabetes [17].

2. MATERIAL AND METHOD

This research is experimental in nature and follows the specific methodology outlined in Figure 1. This research stage begins with reviewing literature studies from scientific articles, books and other sources of information related to the research topic, then the diabetes mellitus data collection stage, after the data is collected it goes to the initial data processing stage and data classification using the K-Nearest classification algorithm. Neighbors (K-NN), Naïve Bayes Classifier, and Decision Tree algorithm, then enter the Accuracy testing stage, then compare the accuracy results of the three algorithms used, and the final stage is a conclusion from the results of the research that has been carried out.[18]

2.1. Diabetes

Diabetes is a long-term medical condition marked by elevated blood glucose levels. [19]. Diabetes occurs when the body is no longer able to take sugar (glucose) into cells and use it as energy[20]. Some symptoms of diabetes include frequently feeling thirsty, frequently feeling hungry, frequent urination, especially at night, weight loss for no apparent reason, and reduced muscle mass. Diabetes is a lifelong condition that cannot be cured. Nonetheless, diabetes can be kept under control and managed with regular

blood sugar checks, a nutritious diet, and regular exercise. Diabetes prevention can be done with a healthy lifestyle, such as maintaining an ideal body weight, eating healthy food, and exercising regularly.

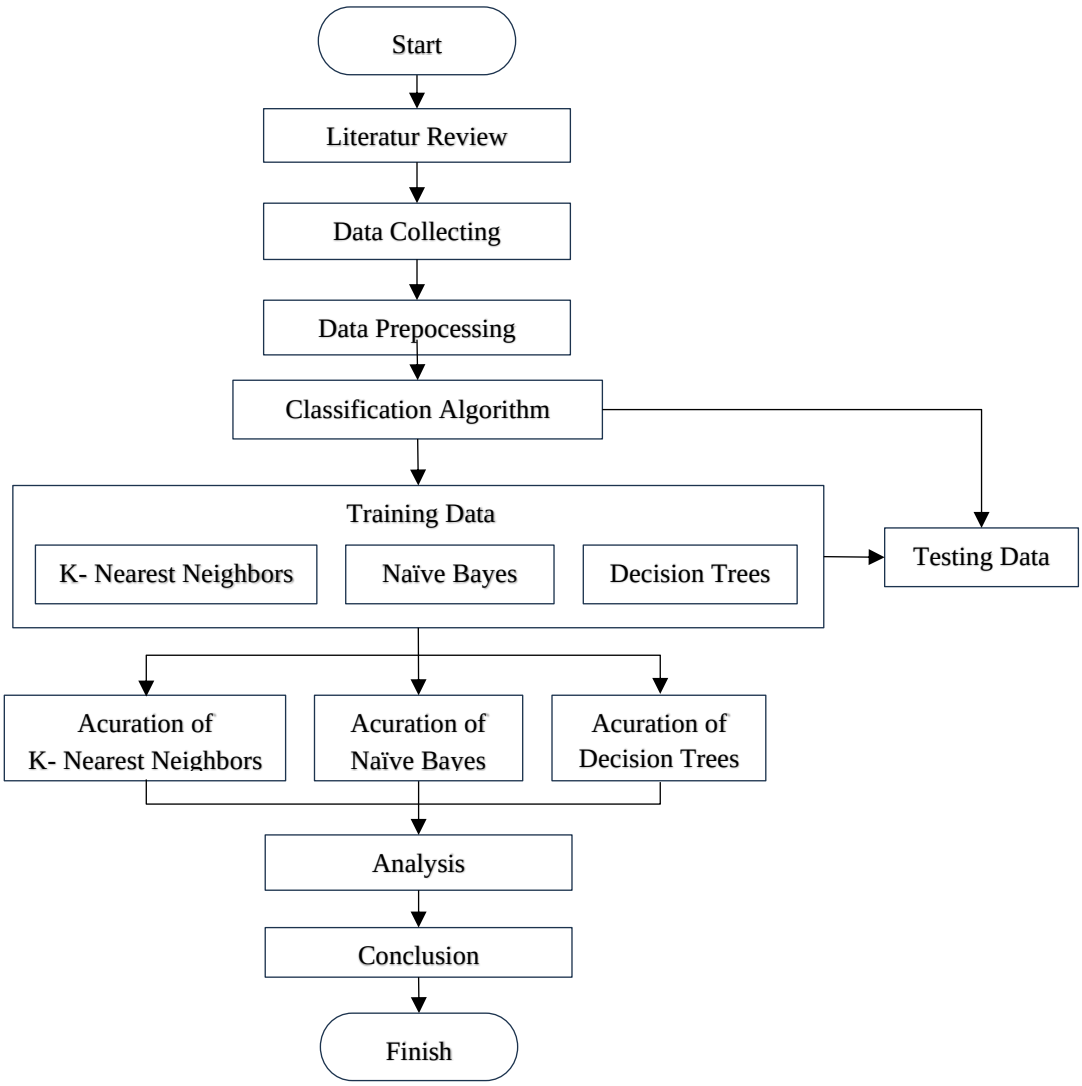


Figure 1. Research Methodology

2.2. K-Nearest Neighbors (K-NN)

The K-NN algorithm is a nearest neighbor algorithm. algorithm, which is calculated from the distance value on testing data with training data, from the value of the smallest nearest neighbor [21]. The purpose of this algorithm to classify new objects based on attributes and training samples. and training samples. In this method, the value of K in K-NN is K-data closest to the test data. K-Nearest Neighbor can produce good classification if using a large amount of data [22]. The stages of the K-NN algorithm is explained as follows:

1. Preparation of training and testing data
2. Determining the k-value
3. Calculating the distance of testing data to each training data [23].

Training data is calculated using the formula Euclidean distance calculator as equation 1.

$$\text{Euclidean Distance} = \sqrt{(x - x_{terbaru})^2 + (y - y_{terbaru})^2} \quad (1)$$

X : Value of the previous attribute
Xnew : The value of the attribute to be classified
Y : The value of the second previous attribute
Ynew : The new second attribute's valueThe title ought to be succinct and obvious [24]

2.3. Naïve Bayes Classifier

Naïve Bayes is a method that method that uses pre-existing data or method that predicts from past events past events. The drawback of this method is that it does not supports continuous-valued attributes (numeric).[25] In brief, the naïve bayes algorithm classification algorithm is a data set classifier which predicts all the probabilities of each member of a class. probability of each member of a class[26]. Because the data in this research is continuous, the Gaussian distribution method will be used. because the Gaussian Distribution method will be used which is commonly used for numeric type data. The following is the formula used in Naïve Bayes method as equation 2, 3 and 4.

$$g(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi.\sigma}} \exp^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (2)$$

$$\mu = \frac{\sum_i^n x_i}{n} \quad (3)$$

$$\sigma = \sqrt{\frac{\sum_i^n (x_i - \mu)^2}{n-1}} \quad (4)$$

g : Gausian
μ : means
σ : Deviation standard

2.4. Decision Trees

A supervised machine learning approach called a decision tree is used to address classification issues. The main purpose of using the decision tree algorithm, because the C4.5 algorithm is capable of produce a specific prediction model in the form of rules that are easy to implemented. In a decision tree has a root node and internode to perform prediction and classification. There are several stages in making a decision tree c4.5 [6], namely:

1. Determine the attribute that will be used as root by determining the lowest entropy value and the highest gain value. Determine the entropy value with the equation 5.

$$\text{Entropy}(y) = \sum_{i=1}^{n} - p_i \log_2 p_i \quad (5)$$

2. Determine the gain value, with the equation 6.

$$\text{Gain}\ (y, A) = \text{Entropy}\ (y) - \sum_{c \in \text{nilai}\ (A)} \frac{y_c}{y} \text{entropy}\ (y_c) \quad (6)$$

3. Create branches for each value.
4. Divide each case into branches.
5. Repeating the process for each branch, so that all cases have the same the same class

3. RESULTS AND ANALYSIS

3.1. Data Collecting

The first stage in this research is data collection. The dataset used in this study is a 769-record dataset on diabetes mellitus that was obtained from Kaggle (https://www.kaggle.com/datasets/salihacur/diabetes). This dataset consists of 8 attributes, including Pregnancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI, Diabetes Pedigree, Function, Age, and Outcome. Furthermore, before entering the classification process, data analysis will be carried out to understand the variables before processing the data thoroughly, namely by preprocessing the data.

3.2. Data Prepocessing

The data preprocessing stage is the initial stage in the data mining process which is useful for minimizing misunderstandings when inputting data. This data preprocessing stage includes data cleaning and data normalization, which aims to eliminate data inconsistencies, incomplete data and minimize data or attributes that will be used in the classification process. The datasets that have been carried out in the data preprocessing stage can be seen in table 1.

Table 1. Preprocessing Data Result

No	Pregnancies	Glucose	Blood Pressure	Skin Thickness	Insulin	BMI	Diabetes Pedigree Function	Age	Outcome
1	6	148	72	35	0	33.6	0.627	50	1

No	Pregnancies	Glucose	Blood Pressure	Skin Thickness	Insulin	BMI	Diabetes Pedigree Function	Age	Outcome
2	1	85	66	29	0	26.6	0.351	31	0
3	8	183	64	0	0	23.3	0.672	32	1
4	1	89	66	23	94	28.1	0.167	21	0
5	0	137	40	35	168	43.1	2.288	33	1
6	5	116	74	0	0	25.6	0.201	30	0
7	3	78	50	32	88	31	0.248	26	1
8	10	115	0	0	0	35.3	0.134	29	0
9	2	197	70	45	543	30.5	0.158	53	1
10	0	123	72	0	0	36.3	0.258	52	1
..	…	…	…	…	…	…	…	…	…
762	9	170	74	31	0	44	0.403	43	1
763	9	89	62	0	0	22.5	0.142	33	0
764	10	101	76	48	180	32.9	0.171	63	0
765	2	122	70	27	0	36.8	0.34	27	0
767	5	121	72	23	112	26.2	0.245	30	0
768	1	126	60	0	0	30.1	0.349	47	1
769	1	93	70	31	0	30.4	0.315	23	0

3.3. Classification Algorithm

Rapidminer tools are used in the data processing procedure at this classification level. Three classification algorithms—Naive Bayes, K-NN, and Decision Tree—were employed in the study to process the information. The objective was to compare accuracy findings and generate conclusions based on the classification of individuals with diabetes mellitus. The classification results for each algorithm can be seen in Figures 2, 3, and 4.

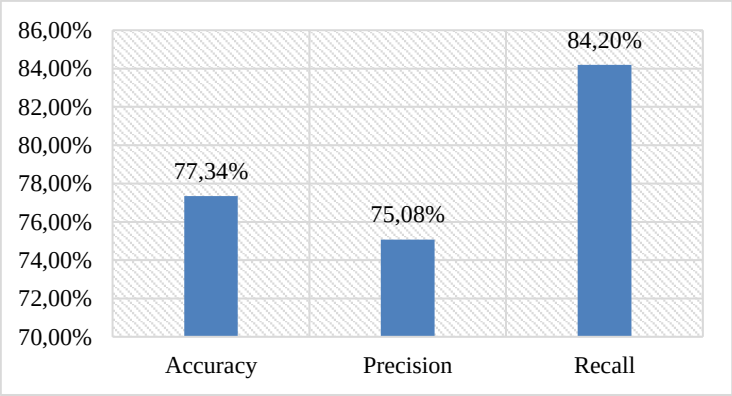


Figure 2. Naïve Bayes Classification Result

Based on Figure 2, it can be seen that the results of the naïve bayes algorithm classification model produce a classification accuracy of diabetes mellitus patients of 76.17%, with a precision of 80.19% and a recall of 84.20%. The classification results show the performance of the modeling used is good, but it needs improvement to improve the results of the modeling capabilities of the naïve bayes algorithm so as to produce a consistent and high accuracy value on classification.

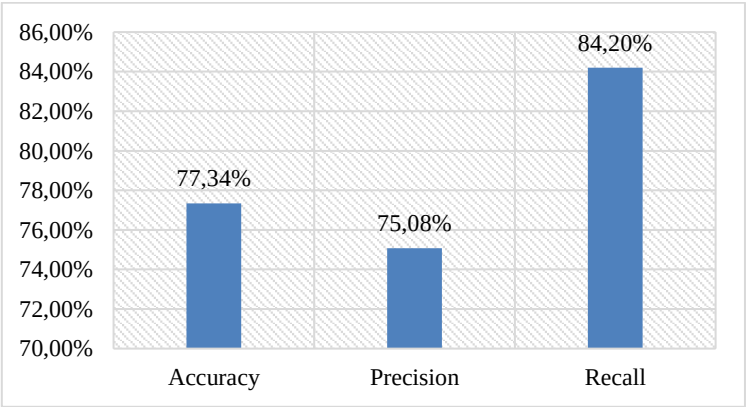


Figure 3. K-NN Classification Result

Based on Figure 3, it can be seen that the use of the K-NN algorithm in the classification of diabetes mellitus sufferers got an accuracy value of 80.34%, this result shows a good accuracy value, while the precision result was 82.62% which was higher than the accuracy result, and the recall was 88.40 % which shows good results in terms of accuracy. The K-NN classification results show that this model is good at classifying, because the resulting precision and recall values are high compared to the accuracy values.

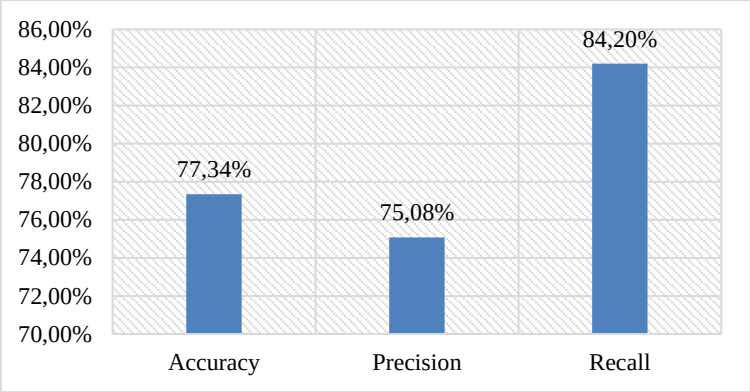
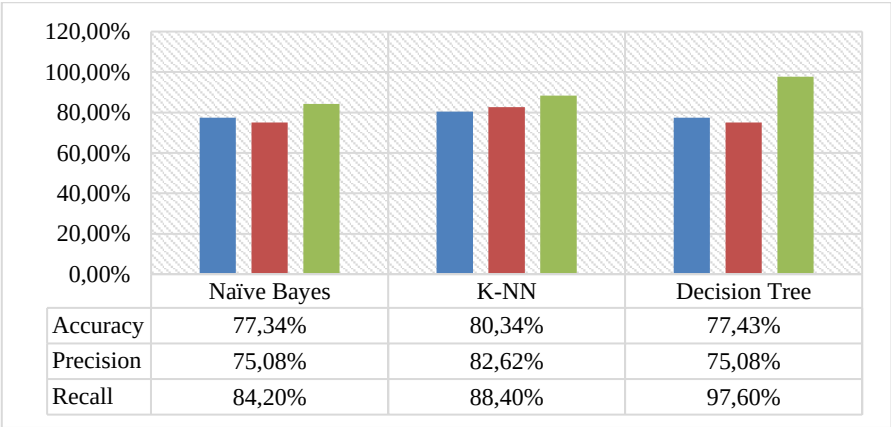


Figure 4. Decision Tree Classification Result

Figure 4 shows the results of classification of diabetes mellitus sufferers using the Decision Tree algorithm with an accuracy result of 77.43% which is quite a good result, followed by a precision of 75.08% which is a lower result compared to the accuracy value, and for a recall value of 97.60%, the value The results are very high compared to the accuracy and precision values. The modeling performance using Decision Tree can be said to be not good, this is because the difference in results between accuracy, precision and recall is quite large. So improvements need to be made to produce consistent and good modeling.

3.4. Analysis

In this analysis section will discuss the comparison of classification results using the Naïve Bayes, K-NN and Decision tree algorithms on the classification of people with Diabetes Mellitus. From a series of experiments conducted using these 4 algorithms, researchers have found the best algorithm that gets the highest accuracy results for classifying people with Diabetes Mellitus. The comparison of the results of the Naïve Bayes, K-NN, and Decision Tree algorithms in classifying data for people with Diabetes Mellitus can be seen in Figure 5.



	Naïve Bayes	K-NN	Decision Tree
Accuracy	77,34%	80,34%	77,43%
Precision	75,08%	82,62%	75,08%
Recall	84,20%	88,40%	97,60%

Figure 5. Comparative Evaluation of Naïve Bayes, K-NN and Decision Tree Algorithms

Figure 5 shows the Classification Algorithm experiment in the classification of diabetes mellitus patients resulted in K-NN as the best algorithm with accuracy and recall values of 80.34% and 82.62%, respectively. Then followed by the Decision tree algorithm with an accuracy of 77.43%, and produced the highest recall of 97.60%. It is also seen that the Naïve Bayes algorithm produces the lowest accuracy, precision, and recall compared to the other two algorithms. It can be said that the Naïve Bayes algorithm is not suitable for the classification of people with Diabetes Mellitus.

4. CONCLUSION

Based on the findings of a study that classified a dataset of individuals with diabetes mellitus using three different algorithms—Naive Bayes, K-NN, and Decision Tree—using data mining techniques; the accuracy values of each algorithm were 77.34% for Naive Bayes, 80.34% for K-NN, and 77.43% for Decision Tree. These findings demonstrate that the K-NN method yields the greatest accuracy value, followed by Decision Tree and Naïve Bayes. Thus, it can be said that the K-NN algorithm is the most effective method for categorizing individuals with diabetes mellitus.

REFERENCES

[1] K. M. Varma and D. B. s Panda, "Comparative analysis of Predicting Diabetes Using Machine Learning Techniques," JETIR, vol. 6, no. 6, pp. 522–530, 2019.

[2] A. Nila Yuliawati, P. Made Desy Ratnasari, and P. Riski Rosalina, "Knowledge and Quality of Life in Type 2 Diabetes Mellitus Patients also its Related Factors," JMPF, vol. 12, no. 1, pp. 14–27.

[3] H.Muthukrishnan, T.Abarna, K.L.Gherija, and K.L.Gherija, "Comparison of Machine Learning Algorithms in Predicting Diabetes Mellitus," vol. 25, no. 5, pp. 621–634, 2021.

[4] I. H. Sarker, F. Faruque, H. Alqahtani, and A. Kalim, "K-Nearest Neighbor Learning based Diabetes Mellitus Prediction and Analysis for eHealth Services," EAI Endorsed Transactions on Scalable Information Systems, vol. 7, no. 26, pp. 1–9, 2020, doi: 10.4108/eai.13-7-2018.162737.

[5] Md. F. Faruque, Asaduzzaman, and I. H. Sarker, "Performance Analysis of Machine Learning Techniques to Predict Diabetes Mellitus," IEEE.

[6] M. Tombokan et al., "HUBUNGAN DUKUNGAN KELUARGA DENGAN MOTIVASI DALAM MENGONTROL KADAR GULA DARAH PADA PENDERITA DIABETES MELITUS DI WILAYAH KERJA PUSKESMAS PAMPANG KECAMATAN PANAKKUKANG KOTA MAKASSAR Family Support Relationship With Motivation In Control Drinking Blood Sugar To Patients Diabetes Melitus At Pampang Public Health Center Panakkukang of Makassar City," 2017.

[7] N. Sakagraha Kuspinta, A. W. Widodo, and M. T. Furqon, "Penentuan Menu Makanan Untuk Penderita Diabetes Menggunakan Metode Iterative Dichotomizer Tree (ID3)," 2018. [Online]. Available: http://j-ptiik.ub.ac.id

[8] R. Patel, "Diabetes Prediction using Data Mining Techniques - A Comparative Study," JETIR, vol. 10, no. 6, pp. 818–824, 2023.

[9] N. Azizah, M. R. Firdaus, R. Suyaningsih, F. Indrayatna, and U. Padjadjaran, "Penerapan Algoritma Klasifikasi K-Nearest Neighbor pada Penyakit Diabetes," 2023, [Online]. Available: http://prosiding.snsa.statistics.unpad.ac.id

[10] F. Sholekhah, A. D. Putri, and L. Efrizoni, "Comparison of Naive Bayes and K-Nearest Neighbors Algorithms for Metabolic Syndrome Classification Perbandingan Algoritma Naïve Bayes dan K-Nearest Neighbors untuk Klasifikasi Metabolik Sindrom," vol. 4, no. April, pp. 507–514, 2024.

[11] L. U. Khasanah, Y. N. Nasution, F. Deny, and T. Amijaya, "Klasifikasi Penyakit Diabetes Melitus Menggunakan Algoritma Naïve Bayes Classifier," vol. 1, no. 1, pp. 41–50, 2022, [Online]. Available: http://jurnal.fmipa.unmul.ac.id/index.php/basis

[12] B. A. Maulana and M. J. Fahmi, "Sentiment Analysis of Pluang Applications With Naive Bayes and Support Vector Machine (SVM) Algorithm Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM)," vol. 4, no. April, pp. 375–384, 2024.

[13] M. Alehegn, R. R. Joshi, and P. Mulay, "Diabetes Analysis And Prediction Using Random Forest , K-NN , Naïve Bayes , And J48 : An Ensemble Approach," vol. 8, no. 09, 2019.

[14] M. E. Febrian, K. M. Suryanigrum, and R. Yunanda, "ScienceDirect ScienceDirect 7th International Conference on Computer Science and Computational Intelligence 2022 Diabetes prediction using supervised machine learning Diabetes prediction using supervised machine learning," Procedia Computer Science, vol. 216, no. 2022, pp. 21–30, 2023, doi: 10.1016/j.procs.2022.12.107.

[15] A. Purnamawati, W. Nugroho, D. Putri, and W. F. Hidayat, "Penyakit Daun pada Tanaman Padi Menggunakan Algoritma Decision Tree , Random Forest , Naïve Bayes , SVM dan K-NN," InfoTekJar : Jurnal Nasional Informatika dan Deteksi, vol. 1, no. 1, pp. 212–215, 2020.

[16] K. J. Rani, "Diabetes Prediction Using Machine Learning," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 294–305, Jul. 2020, doi: 10.32628/CSEIT206463.

[17] E. Novianto, A. Hermawan, and D. Avianto, "KLASIFIKASI ALGORITMA K-NEAREST NEIGHBOR, NAIVE BAYES, DECISION TREE UNTUK PREDIKSI STATUS KELULUSAN MAHASISWA S1," Rabit : Jurnal Teknologi dan Sistem Informasi Univrab, vol. 8, no. 2, pp. 146–154, Jul. 2023, doi: 10.36341/rabit.v8i2.3434.

[18] D. Sisodia and D. S. Sisodia, "Prediction of Diabetes using Classification Algorithms," in Procedia Computer Science, Elsevier B.V., 2018, pp. 1578–1585. doi: 10.1016/j.procs.2018.05.122.

[19] J. P. Kandhasamy and S. Balamurali, "Performance analysis of classifier models to predict diabetes mellitus," in Procedia Computer Science, Elsevier B.V., 2015, pp. 45–51. doi: 10.1016/j.procs.2015.03.182.
[20] A. Zaharim and WSEAS (Organization), Applied computational science : proceedings of the 13th international conference on applied computer and applied computational science (ACACOS '14) : Kuala Lumpur, Malaysia, April 23-25 2014.
[21] A. M. Argina, "Indonesian Journal of Data and Science Penerapan Metode Klasifikasi K-Nearest Neigbor pada Dataset Penderita Penyakit Diabetes," vol. 1, no. 2, pp. 29–33, 2020.
[22] A. H. Khaleel, G. A. Al-Suhail, and B. M. Hussan, "International Journal of Computer Science and Mobile Computing A Weighted Voting of K-Nearest Neighbor Algorithm for Diabetes Mellitus," 2017. [Online]. Available: www.ijcsmc.com
[23] "Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM Sumarlin STIKOM Uyelindo Kupang."
[24] M. Soni and S. Varma, "Diabetes Prediction using Machine Learning Techniques." [Online]. Available: www.ijert.org
[25] D. Susilawati and D. Riana, "Optimization the Naive Bayes Classifier Method to diagnose diabetes Mellitus," 2019.
[26] A. Sachdeva, P. Kumar, O. P. Yadav, R. K. Garg, and A. Gupta Editors, "Operations Management and Systems Engineering." [Online]. Available: http://www.springer.com/series/15734