Journal of Theoretical and Applied Information Technology

15th November 2023. Vol.101. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# PERFORMANCE COMPARISON OF K-NEAREST NEIGHBOR AND NAIVE BAYES ALGORITHMS FOR HEART DISEASE PREDICTION

**ROHMAT TAUFIQ[1*], ANGGA ADITYA PERMANA[2], ARSANAH[1], ANALEKTA TIARA PERDANA[3]**

[1*]Department of Informatic Engineering, Universitas Muhammadiyah Tangerang, Indonesia

[2]Department of Informatic, Universitas Multimedia Nusantara, Indonesia

[3]Department of Biology, UIN Sultan Maulana Hasanuddin, Banten, Indonesia

E-mail: [1*]rohmat.taufiq@umt.ac.id, [2]angga.permana@umn.ac.id, [1]arsanah@ft-umt.ac.id, [3]analekta.tiara@uinbanten.ac.id

**ABSTRACT**

Heart disease is leading causes of death in several countries, including Indonesia. Accurate and precise prediction of heart disease is so important. There are lots of machine learning algorithms that can be used to make predictions. In this study, 2 classification algorithms: K-Nearest Neighbors and Naive Bayes were used to compare the performance of algorithms for better prediction. Heart disease dataset UCI Machine Learning Repository Center for Machine Learning and Intelligent Systems dataset was used for generating confusion matrix values, analyzing accuracy values in predicting heart disease based on 14 attributes. The results showed that the KNN algorithm has an accuracy value of 91.25% while Naive Bayes is 88.7%.

**Keywords:** *KNN, Naïve Bayes, Comparison, Heart Disease*

## 1. INTRODUCTION

In the current digital era, accurate information is needed in any case, even in the world of health. It is very useful to find out medical or preventive things that must be done against a disease especially metabolic and degenerative disease including heart disease. Data from the World Health Organization (WHO) mention, more than 17 million people in the world die from heart disease, in Indonesia heart disease reaches 651,481 people per year.

The number of deaths caused by heart disease is due to the lack of undertanding about the symptoms of heart disease [1]. Early diagnosis of heart disease becomes difficult, treartments of heart disease are complex and prediction of the heart disease risk goes a long way [2]. So that, it is important to do early prevention or prediction of the disease to safe many life.

Data for disease prediction were very large. However, the effective tool to discover sheltered patterns from those data was deficient. Data mining is a promising tool to invent relation from vast amount of data. In prediction of heart disease, data mining appears to be necessary providing knowledge for the physicians, moreover intensify the accuracy of prediction [3].

Machine learning (ML) algorithms can be used to identify and predict heart disease as the early detection of the disease. Machine learning is a technology that has an accuracy of up to 100% to analyze data and draw conclusions based on this data analysis [4]. ML have been applied to various medical datasets to extract hidden pattern of the data as a decision support for the medical profession [2].

Research on heart disease has been done before. Researchers have tried to use many classification algorithms that can predict with higher accuracy values. Naive Bayes, neural networks and others have been used for the classification of heart disease with datasets obtained from Kaggle [5]. K-nearest neighbor (KNN) [6] and naïve Bayes classifier [7] are excellent classification algorithms for the data in this study using data mining. Data mining will process the dataset to create a data model and make predictions on the dataset.

Asfaw [8] used 2 classification techniques of data mining in prediction of heart disease such that KNN and Gaussian Naïve Bayes. Among both, Naïve Bayes Classification performs better

accuracy of 77.05%. Mentari and Agustina [1] compared five classification algorithms: Naïve Bayes Algorithm, KNN, Neural Network, Decision Tree, and Random Forest using RapidMiner 5 software in predicting heart disease. The result showed that Naïve Bayes as the highest accuracy rate of 82.23% and KNN as the lowesr accuracy rate of 62.69%. Damayunita *et al*. [2] tested 3 classification algorithms: Naive Bayes, KNN, and Support Vector Machine (SVM). From the results of trials conducted KNN has higher accuracy value of 91% than Naive Bayes with an accuracy value 88%.

Correct prediction of certain disease implies to quality of services. It can lead to effective treatments for patients, so the prediction system with the most effective classifier appears to be vital. The proposed system to predict whether the patient is undergo heart disease or not using two classifiers. Lately, researchers were discovering diverse classifiers that can be effective to make valid predictions. Besides that, the systems will be prospected to user friendly, scalable and reliable [3].

This study conducts performance comparison of classification algorithms: KNN and Naive Bayes for heart disease prediction from UCI Machine Learning Repository Center for Machine Learning and Intelligent Systems dataset with 14 attributes. The results of this study are expected to present a novel approach for predicting heart disease. The diagnostic results hopefully can be used as the advice for the diagnosis made by experts [2].

## *2. LITERATURE REVIEW*

### 2.1 Artificial Intelligence

Artificial Intelligence is a simulation of human intelligence that is modeled on machines and programmed to think like humans. Machine Learning is the most popular technique because it is widely used to replace or imitate human behavior or develop new things to simplify and speed up work or solve problems. [9]

Broadly speaking, there are 3 types of machine learning learning methods used, namely:

1. Supervised learning is a program supervised learning method given several examples of data whose type/classification is known as learning or training materials or classifying an object after being trained.
2. Unsupervised learning is an unsupervised learning method using procedures that attempt to find partitions of a pattern.
3. Reinforcement learning is a learning method that studies control rules by interacting with an unfamiliar environment.

### 2.2 Data Mining

Data mining Data mining is the process of collecting and processing data that aims to extract important information from data. The process of collecting and extracting this information can be done using software with the help of statistical calculations, mathematics, or Artificial Intelligence (AI) technology. Data mining is often called Knowledge Discovery in Database (KDD). [10]

Data Mining is an analysis of reviewing data sets to find unexpected relationships and summarize data in a different way than before, which is understandable and useful to the data owner. analyze a large and large set of data, then process it to get the results of the analysis in the form of predictive data which are a source of information for making decisions.

The datamining process has its stages as follows:

Data Selection

a. Creating a target data set, selecting a data set, or focusing on a variable subset or sample of data, where the discovery will be made.
b. Selection of data from a set of operational data needs to be done before the information gathering stage begins. The selected data to be used for the data mining process is stored in a file, separate from the operational database.

Based on their functionality, data mining tasks can be grouped into the following six groups. Classification: generalizes known structures to be applied to new data. For example, the classification of disease into a number of types.

1. Clustering : grouping data whose class label is unknown into a certain number of groups according to the size of their similarity.
2. Regression: find a function that models the data with the minimum possible prediction error.
3. Anomaly detection: identifying unusual data
4. Association rule mining or dependency modeling
5. Summarization: provides a simpler data representation, including visualization and report generation.

### 2.3 Heart Disease

Heart disease is a non-communicable disease but can also be inherited from parents with heart disease. If you often have problems with irregular

heartbeats, weak heart muscles, congenital heart defects, these can be symptoms of heart disease. [11]

At present the death rate of heart disease is also one of the highest in the world. In the field of health sciences, it is very dependent on disease prediction tools, for precise and accurate and timely predictions. The function of the heart is to pump oxygen-rich blood to all organs of the body, then blood with oxygen from the organs returns to the heart, right atrium and then forwards to the right ventricle.

Sources for predicting the probability of future disease, therefore there is a system for analysis to find hidden information and patterns in heart disease medical datasets by finding patterns of approach originating from large data sets to extract useful information for making decisions by understanding the level of accuracy of each each data that has been tested detects heart disease. With the criteria according to medical record data including the following : Age, Sex, Cp, Trestbps, Chol, Fbs, Restecg, Thalach, Exang, Olpeak, Slope, Ca and Thal.

### 2.4 Confusion Martix

The confusion matrix performs tests to predict whether the object being analyzed is true or false. In the confusion matrix test where the predicted class is displayed at the top of the matrix and the class observed on the left, contains a number indicating how many actual cases of the observed class are to be predicted. [12] Confusion Matrix is a method that is usually used to perform accuracy calculations on data mining concepts. There are 4 terms in measurement using a confusion matrix, as a representation of the results of the classification process (Table 1)

*Table 1. Confusion Matrix*

| Prediction | Accuration | |
|---|---|---|
| | TP | FN |
| | FP | FT |

Information
TP = classified as positive.
TN = classified as negative.
FP = classified as negative.
FN = classified as positive.

To calculate the level of accuracy on the matrix used:
The results of the values for accuracy, precision, and recall are usually displayed as a percentage.

1. Accuracy (Accuracy)
Accuracy can illustrate how accurate a model is to classify correctly. Predictive accuracy is the comparison of the number of true positive and true negative data with the overall data. Accuracy is the degree of closeness of the predicted value to the actual value.

2. Precision (Precision)
Precision can be described as the level of accuracy of the requested data with the predicted results provided by the model. Then it gets the ratio of the correctly positive predictions divided by the overall positive predicted outcome. In other words, of all the positive classes that have been correctly predicted, how many of the data are actually positive.

3. Recall
Recall or sensitivity can be described as the success of the model in obtaining information. So it can be said that the recall is the ratio of the correct positive predictions divided by all the correct positive data.

### 2.5 K- Nearest Neighbor (KNN)

KNN is a supervised learning algorithm that classifies objects whose class is unknown and looks for the closest object that has the same class. The number of closest objects is determined based on the value of K. The K-Nearest Neighbor Algorithm is an algorithm used to classify an object based on the k pieces of training data that are closest to the object (test data). The condition for the value of k is that it cannot be greater than the amount of training data. [6][13]

KNN is a supervised algorithm that classifies objects whose class is unknown and looks for the closest object that has the same class. The number of closest objects is determined based on the value of K. The following is the formula for determining the Euclidean distance:

$$D_i = \sqrt{\sum_{i=1} n \left( p_i - q_i \right)^2}$$

Information:
Di = Distance of the i-variable
i = data variable (i=1,2,3,...n)
n = data dimension
pi = testing data

qi = training data

First determine the value of K first. Then, the distance between the testing data and the training data uses the formula above. Sort the distance from closest to farthest. Take objects that have the closest distance as many as K values. Determine the class that has the most number based on the closest and then assign that class to the object that was just tested.

### 2.6 Naïve Bayes Classifier

Naïve Bayes classifier is one of the statistical classifier methods, where this classifier can predict the probability of class membership of a data that will enter a certain class, according to probability calculations. [7], [14]–[16]. Another definition says naive Bayes is a classification using the probability and statistical method put forward by a British scientist named Thomas Bayes, namely predicting future opportunities based on previous experience. The Naive Bayes Classifier method is used because too many datasets must be used. requires a method that has fast performance in classification as well as good and high accuracy.

The advantage of the Naive Bayes Classifier method is that it only requires a small amount of data, such as training data, to determine the estimation of the parameters needed in the process of classifying a data. The Naive Bayes Classifier has a method in classification, there are two stages, namely the training and probability stages, the term in the data being classified. expressed in an equation with the Naïve Bayes Theorem formula:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

Information:

| | |
|---|---|
| X | : Data with unknown class |
| H | : Hypothesis data is a specific class |
| P(H\|X) | : Probability of hypothesis H based on condition X (posterior probability) |
| P(H) | : Probability of the hypothesis H (probability prior) |
| P(X\|H) | : Probability of X Based on the conditions in hypothesis H |
| P(X) | : Probability of X |

### 2.7 Python

Python is a dynamic, very flexible and simple programming language that is often used in analytical development or application development, Python has several features that can be used for software. Python is a multipurpose interpretive programming language with a design philosophy that focuses on code readability. ( MA ' ARIF, 2020)

### 2.8 Website

A website is a collection of pages that are displayed according to the information you want to find, in the appearance of the website there are text, images, videos, documents digitally, you have to use the internet network, can be accessed anytime and anywhere, computer users or cellphone users, most importantly there is a browser that will access a website. [18]

The website is also one of the computer user services connected to the internet which has the advantage of connecting with links with other documents that will later be accessed using a browser, a protocol that conveys information from the website server to be displayed to users via a web browser.

So a website is a collection of pages that are displayed according to the information you want to find, the display contains text, images, videos, digital documents, must use the internet network, can be accessed whenever and wherever computer users or mobile users, most importantly there is a browser that will access a website.

### 2.9 Unified Modeling Language (UML)

UML is a visual language for modeling and communicating about a system using diagrams and supporting texts and consistent communication tools to support application system developers. [19]

## 3. RESEARCH METHODS

In this study, several stages were carried out (Figure 1) including data collection, data processing, data modelling and classification, evaluation, mockup design, application development and application testing. Data collection was started with secondary data. The secondary data was obtained from the Kaggle website with the name UCI Machine Learning Repository Center for Machine Learning and Intelligent Systems dataset. In this study, we took 399 datasets with 14 attributes. After that, data processing was carried out on the Jupyter Lab Notebook to input and normalize the datasets including data training and testing before data modelling and classification steps. Data modelling and classification were utilized

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

KNN and Naive Bayes classification algorithms in Python. We evaluated data modelling and classification results using confusion matrix. Mockup design was done to design the user interface of the application. After that we designed the application using PHP programming language. Then, we carried out application testing black box testing approach .

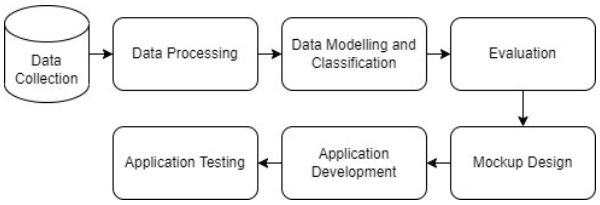


*Figure 1. Research Stages*

## 4. RESULTS AND DISCUSSION

### 4.1 Data Collection and Processing

The results of the analysis in this study applied the prediction system development method with the php framework and the dataset analysis process using a jupyter lab notebook to be processed with the two algorithms used, namely the k-nearest nighbor algorithm and the naive bayes classifier used to find a comparison of the prediction results of accuracy, precision, recall on heart disease datasets.

From the results of the dataset showing different images according to the attributes in heart disease, there are 14 attributes of heart disease, including (1) age, (2) sex, (3) cp, (4) trestbps, (5) chol, (6) fbs, (7) restecg, (8) thalach, (9) exang, (10) olpeak, (11) slope, (12) ca, (13) thal and (14) target (Figure 2), and for the age frequency of people with heart disease according to the data analyzed there are 399 (Figure 3) datasets, that this barplot shows at the age of 58 years the most lots of heart disease (Figure 4). In the next process of dividing training and testing data or x and y data, before making predictions or testing with algorithmic methods, the python system will quickly divide the amount of training and testing data (Table 2).

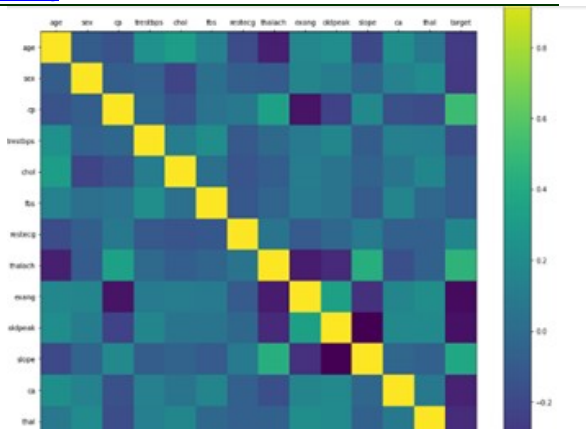

*Figure 2. Attributes Of Heart Disease*

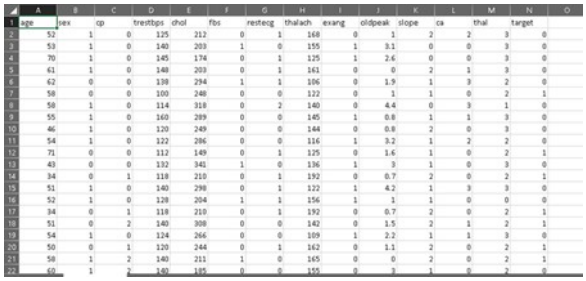

*Figure 3. Dataset*

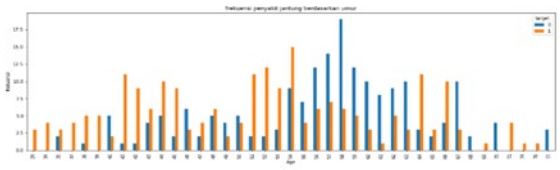

*Figure 4. Age Of Heart Disease*

*Tabel 2. Training Data And Testing Data*

| | |
|---|---|
| Data training | 319 |
| Data testing | 80 |

### 4.2 Data Modelling and Classification

In the barplot Figure 5, it explains the results of the number of datasets in blue which means that women are worth 0 with a total of 175 and orange with 200 are worth 1 man. The percentage of patients who do not have heart disease is 47.1% and those who have heart disease are 52.88%. in Figure 6 the red color image means positive heart disease and blue negative heart disease.

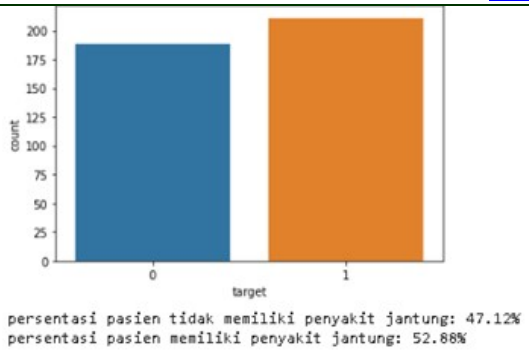


*Figure 5. Gender Of The Patient With Heart Disease*

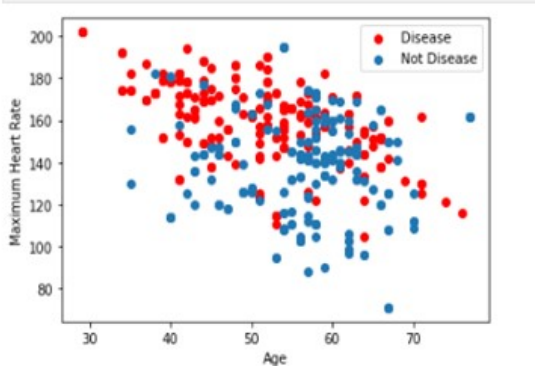


*Figure 6. Distribution Of Datasets*

### 4.3 Evaluation

On the results of the accuracy of the tests carried out with the k-nearest nightbors and naive bayes classifier algorithm, the average score is 89-94, close to 100% (Figure 7). For k-nearest nights, the accuracy is obtained with a score of 91.25%. and with the calculation of the value of k 7 also has a high graph (Figure 8). Results Naïve Bayes accuracy produces a score of 88.75 % of the 399 dataset, with an average value of 0.89 (Figure 9)

| Jenis kelamin | Precision | Recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.94 | 0.86 | 0.95 | 37 |
| 1 | 0.89 | 0.95 | 0.92 | 43 |
| Accuracy | | | 0.91 | 80 |
| maxro avg | 0.92 | 0.91 | 0.91 | 80 |
| weighted avg | 0.91 | 0.91 | 0.91 | 80 |
| score 91.25% | | | | |

*Figure 7. Confusion Matrix Knn*

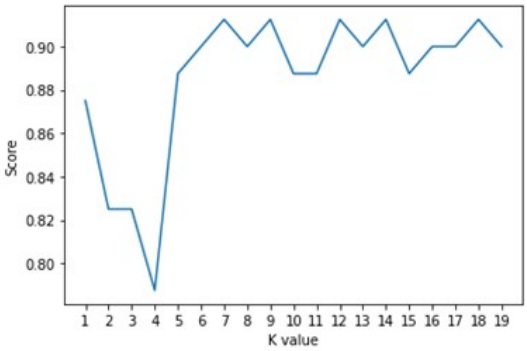


*Figure 8. Knn Score*

| Jenis kelamin | Precision | Recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.95 | 0.84 | 0.89 | 43 |
| 1 | 0.83 | 0.95 | 0.89 | 37 |
| Accuracy | | | 0.89 | 80 |
| maxro avg | 0.89 | 0.89 | 0.89 | 80 |
| weighted avg | 0.89 | 0.89 | 0.89 | 80 |
| score 88.75% | | | | |

*Figure 9. Confusion Matrix Naïve Bayes*

### 4.4 Mockup Design

A web-based application created to quickly classify data, the excel dataset will be tested first according to the class value which will become the testing data and get the classification results. This is a use case diagram which will give an overview of what admin interactions can be done with the website system that has been created (Figure 10).

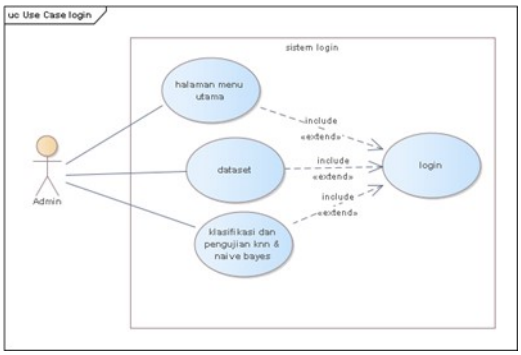


*Figure 10. Usecase Admin*

*For the proposed system design there are four pages, namely: login page, dashboard, dataset page, and test page (Figure 11-14).*

Journal of Theoretical and Applied Information Technology
15th November 2023. Vol.101. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

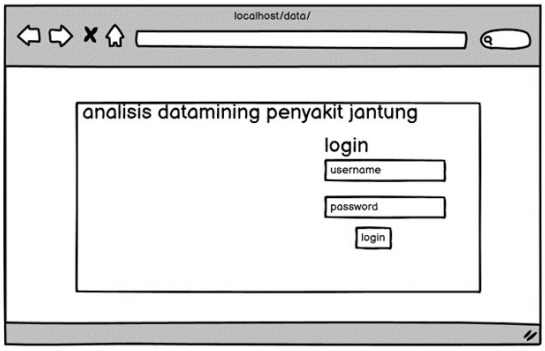


*Figure 11. Mockup Login Page*

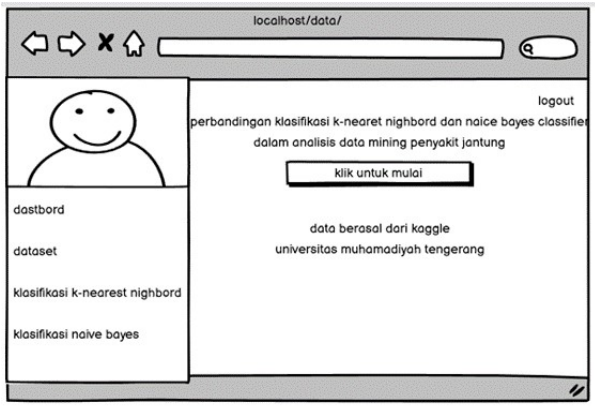


*Figure 12. Mockup Dashboard*

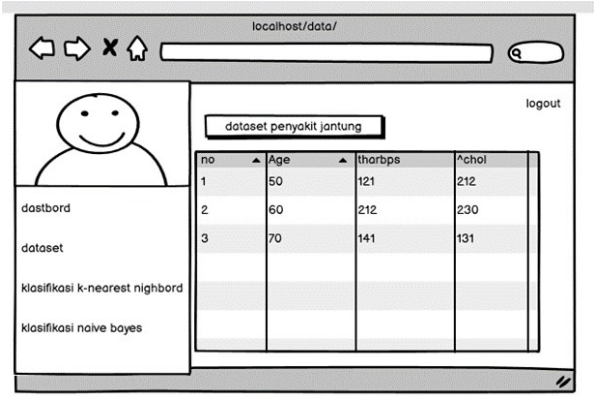


*Figure 13. Mockup Dataset*

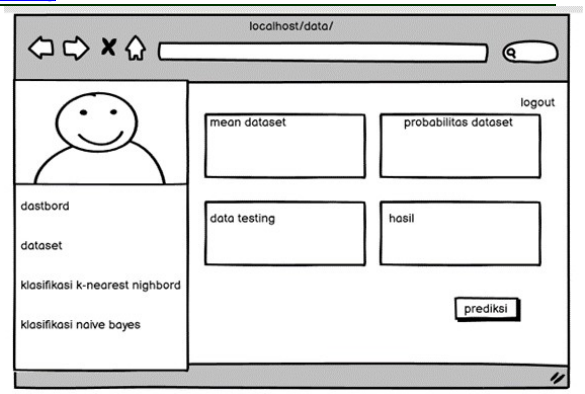


*Figure 14. Mockup Test Page*

### 4.5 Application Design

*We designed user interface system based on the previously designed mockup, the system user interface that will be created also has four pages: login page, dashboard, dataset page, and test page (Figure 15-18).*

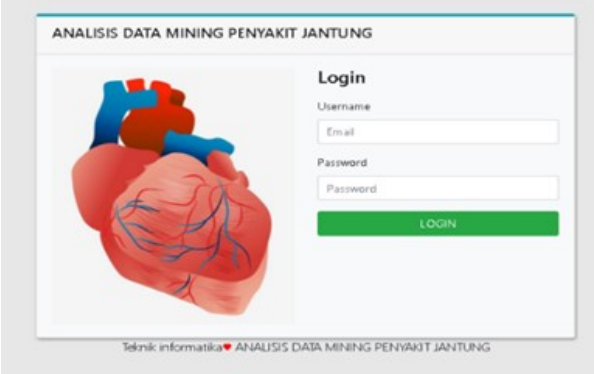


*Figure 15. User Interface Login*

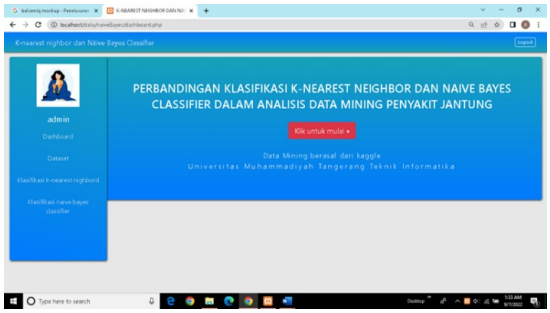


*Figure 16. User Interface Dashboard*

Journal of Theoretical and Applied Information Technology

15th November 2023. Vol.101. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

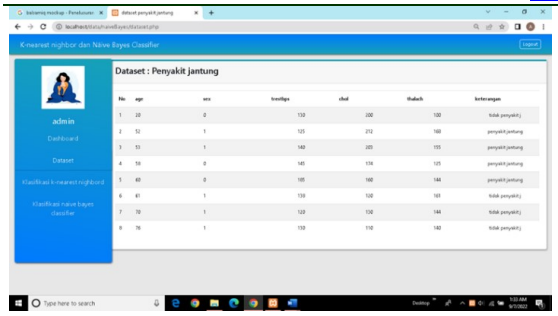


*Figure 17. User Interface Dataset*

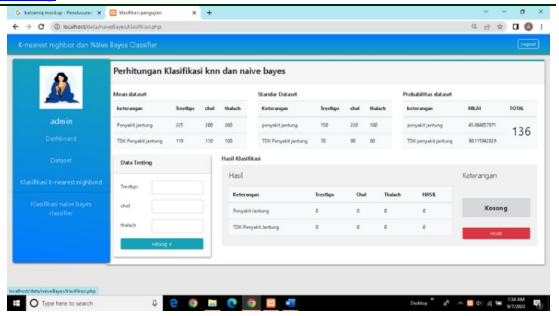


*Figure 18. User Interface Test Page*

### 4.6 Application Testing

The results of system testing using black box testing approach were shown in Table 3.

*Table 3. Black Box Testing*

| No | Requirement | Test Scenario | Test result | Conclusion |
|---|---|---|---|---|
| 1 | Has a login feature | Login using the username and password that matches the database | Users can log in | Valid |
| 2 | Has a login feature | Login using a username or password that does not match the database | The user cannot login and an error message appears | Valid |
| 3 | Can display heart disease dataset information | The user clicks on the dataset menu | Displays the dataset information page | Valid |
| 4 | Can perform testing and classification with the KNN and Naive Bayes algorithms | The user clicks on the classification menu | Input values for classification testing and get results | Valid |

This paper compares the performances moreover finds out the best of the classification algorithms in the prediction of heart disease. KNN performs better than naïve bayes because of its accuracy. Htay and Min [3] also reported effective heart disease prediction system by using KNN can quickly produce the result and precisely fece the problem about probability calculation of each attributes to confirm decision during heart disease prediction. Furthermore, this system will be customized to predict not only heart disease but also the risk factor in order to avoid xtra care for *heart failure.*

## 5. CONCLUSION

From the results of research on the comparison of the k-nearest nighbor and naive bayes classifier classifications in the analysis of heart disease, it can be concluded as follows that the highest accuracy results are KNN with an accuracy value of 91.25%, the results of the accuracy of the naive bayes classifier with an accuracy value of 88.7 %.

## ACKNOWLEDGMENTS

Acknowledgments This research can be carried out properly thanks to the help of various parties, for this reason the researchers would like to thank Universitas Multimedia Nusantara and Universitas Muhammadiyah Tangerang for the support and assistance that has been given during the process of writing this article.

## REFERENCES:

[1] L. N. Mentari and F. Agustina, “Comparative Analysis of Naïve Bayes Algorithm , K-Nearest Neighbours , Neural Network , Decision Tree , and Random Forest to Classify Data Mining in Predicting Heart Disease,” vol. 6, no. 3, pp. 248–255, 2021.

[2] A. Damayunita, R. S. Fuadi, and C. Juliane, “Comparative Analysis of Naive Bayes, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) Algorithms for Classification of Heart Disease Patients,” *J. Online Inform.*, vol. 7, no. 2, pp. 219–225, 2022, doi: 10.15575/join.v7i2.919.

[3] Y. Y. Htay and Y. Min, “Performance

comparison between K-Nearest Neighbor and Naive Bayesian Classifiers by using heart disease dataset," *Int. J. Adv. Res. Dev.*, vol. 5, no. 5, pp. 28–30, 2020.

[4] P. Waldmann, "Approximate Bayesian neural networks in genomic prediction," *Genet. Sel. Evol.*, vol. 50, no. 1, 2018, doi: 10.1186/s12711-018-0439-1.

[5] D. S. Permana and A. Silvanie, "Prediksi Penyakit Jantung Menggunakan Support Vector Machinedan Pythonpada Basis Data Pasiendi Cleveland," *J. Nas. Inform.*, vol. 2, no. 1, pp. 30–34, 2021.

[6] L. andiani, J. Ilmu Komputer, and D. Palupi Rini, "Analisis Penyakit Jantung Menggunakan Metode KNN Dan Random Forest," *Pros. Annu. Res. Semin.*, vol. 5, no. 1, pp. 978–979, 2019.

[7] A. N. Sari and S. Alfionita, "Klasifikasi Penyakit Jantung Menggunakan Metode Naïve Bayes," *AMRI (Analisa Metod. Rekayasa Inform.*, vol. 1, no. 1, pp. 22–26, 2022, doi: 10.12487/AMRI.v1i1.xxxxx.

[8] T. Asfaw, "Performance Comparison of K-Nearest Neighbors and Gaussian Naïve Bayes algorithms for Heart Disease Prediction," *Int. J. Eng. Sci. Invent.*, vol. 8, no. 8, pp. 45–48, 2019, [Online]. Available: www.ijesi.org%7C%7CVolume8www.ijesi.org

[9] V. Amrizal and Q. Aini, *Naskah Kecerdasan Buatan*. 2013.

[10] Yuli Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database ( KDD ) . Jurnal Edik Informatika," *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2019.

[11] Kemenkes RI, "Situasi kesehatan jantung," *Pus. data dan Inf. Kementeri. Kesehat. RI*, p. 3, 2014, doi: 10.1017/CBO9781107415324.004.

[12] S. Visa, B. Ramsay, A. Ralescu, and E. Van Der Knaap, "Confusion matrix-based feature selection," *CEUR Workshop Proc.*, vol. 710, pp. 120–127, 2011.

[13] C. Steven and W. Wella, "The Right Sentiment Analysis Method of Indonesian Tourism in Social Media Twitter," *IJNMT (International J. New Media Technol.*, vol. 7, no. 2, pp. 102–110, 2020, doi: 10.31937/ijnmt.v7i2.1732.

[14] X. Jiang, B. Cai, D. Xue, X. Lu, G. F. Cooper, and R. E. Neapolitan, "A comparative analysis of methods for predicting clinical outcomes using high-dimensional genomic datasets.," *J. Am. Med. Inform. Assoc.*, vol. 21, no. e2, 2014, doi: 10.1136/amiajnl-2013-002358.

[15] A. A. PERMANA and W. A. NOVIYANTO, "Comparison of the Accuracy of the Lexicon-Based and Naive Bayes Classifier Methods To Public Opinions About Removing Masks on Social Media Twitter," *J. Theor. Appl. Inf. Technol.*, vol. 101, no. 3, pp. 1174–1183, 2023.

[16] A. A. Permana, A. T. Perdana, N. Handayani, and R. Destriana, "A Stunting Prevention Application 'nutrimo' (Nutrition Monitoring)," *J. Phys. Conf. Ser.*, vol. 1844, no. 1, 2021, doi: 10.1088/1742-6596/1844/1/012023.

[17] "BUKU AJAR PEMROGRAMAN LANJUT BAHASA PEMROGRAMAN PYTHON Oleh : ALFIAN MA ' ARIF," 2020.

[18] A. A. Permana, R. Taufiq, and S. Ramadhina, "Prototype design of mobile application 'hydrolite' for hydroponics marketplace," *Int. Conf. Electr. Eng. Comput. Sci. Informatics*, vol. 2020-Octob, no. October, pp. 45–48, 2020, doi: 10.23919/EECSI50503.2020.9251303.

[19] Haviluddin, "Memahami Penggunaan UML ( Unified Modelling Language )," *Memahami Pengguna. UML (Unified Model. Lang.*, vol. 6, no. 1, pp. 1–15, 2011, [Online]. Available: https://informatikamulawarman.files.wordpress.com/2011/10/01-jurnal-informatika-mulawarman-feb-2011.pdf