Infortech, Vol. 8, No. 1, June 2026, E-ISSN: 2715-8160, Page. 1-7
Published: LPPM Universitas Bina Sarana Informatika
JURNAL INFORTECH
website: http://jurnal.bsi.ac.id/index.php/infortech/

Breast Cancer Prediction Optimization Using Support Vector Machine and Naive Bayes Algorithms

Devi Wulandari [1], Qudsiah Nur Azizah[2], Diah Puspitasari[3], Kresna Ramanda[4], Erma Delima Sikumbang[5], Sulaeman Hadi Sukmana[6]

[1,2,3,5]Sistem Informasi, Universitas Bina Sarana Informatika, Jakarta, Indonesia
[4,6]Teknologi Informasi, Universitas Bina Sarana Informatika, Jakarta, Indonesia

ARTICLE INFORMATION

Artikel History:
Received: 22-04-2026
Revised: 08-05-2026
Accepted: 15-06-2026
Available Online: 22-06 -2026

Keyword:

Data mining
Breast Cancer
Naïve Bayes
Support Vector Machine
Machine Learning

ABSTRACT

Breast cancer is ranked as the second most common cause of death for women worldwide. Breast cancer is often found when it has entered the final stage. In general, this is due to slow handling and treatment, so it is very necessary to detect the disease early. The purpose of this study is to determine the performance of the two algorithms, namely Naïve Bayes and Support Vector Machine (SVM) in classifying breast cancer types which will then be analysed and compared the accuracy of the two algorithms. The dataset used in this study, Breast Cancer Wisconsin, is public data originating from UCI Machine Learning, has a total of 683 data with 10 attributes and has two classes, namely benign class with 458 data and malignant class with 241 data. The dataset was split 80:20, with 80% used as training data and 20% as testing data, and then evaluated using cross-validation. The results of the study show that Support Vector Machine (SVM) has the best performance with an accuracy of 96.89% while Naïve Bayes 96.15%. With this accuracy, These results indicate that the SVM model provides better classification performance than Naïve Bayes for the Breast Cancer Wisconsin dataset.

Corresponding Author:

Qudsiah Nur Azizah,
Sistem Informasi,
Universitas Bina Sarana Informatika,
Jl. Kramat Raya No.98, Jakarta, Indonesia, 10450
Email: qudsiah.qna@bsi.ac.id

INTRODUCTION

Information technology is currently developing rapidly and is beneficial to society as a whole. Information technology has emerged as a crucial necessity for humans in everyday life because it can make work easier. Therefore, information technology can be used in many fields, such as medicine and health. This includes medical services for detecting diseases. One such service is detecting breast cancer, a disease that is common in women (Ketut, 2022).

According to the World Health Organization, malnutrition encompasses undernutrition, micronutrient deficiencies, overweight, and obesity, which have substantial impacts on health and increase mortality risk, especially among children under five (World Health Organization, 2024). Breast cancer ranks second as the most common cause of death for women worldwide (Karamti et al., 2024). Breast cancer is a global health challenge that results in high mortality rates, where in 2020 in Indonesia alone there were more than 22 thousand cases of death. (Rahmatia et al., 2024). According to the Ministry of Health in Indonesia itself, when compared with other types of cancer, the number of breast cancer cases ranks highest and is one of the main causes of death. (Abdul Jabbar et al., 2022). Lifestyle factors and lack of knowledge indicate that there is a possibility of an increase in cancer cases (Qurrotaayun et al., 2025). As many as 1 million women are diagnosed with breast cancer each year, and many of them ultimately die. This is usually due to delayed treatment, which means the cancer is only detected when it has reached an advanced stage. Therefore, early detection of this disease is crucial.

DOI: https://doi.org/10.31294/infortech.v8i1.

Faster and more accurate early diagnosis of breast cancer can be achieved using datsa mining and machine learning.

Data mining is a method for extracting new information by analyzing specific patterns from large amounts of data. This process consists of several stages aimed at discovering hidden insights or knowledge that cannot be identified directly or manually (Azizah et al., 2025). Machine learning is also gaining attention due to its improved results. Machine learning is reliable, accurate, and fast. Due to these aforementioned benefits, this study proposes a data mining-based approach to breast cancer detection to achieve high accuracy (Karamti et al., 2024). To determine whether a patient who has undergone an examination has a benign or malignant cancer diagnosis, researchers will apply classification using the Naive Bayes algorithm and Support Vector Machine (SVM).

The Naïve Bayes algorithm is widely used to solve prediction problems such as classification. Furthermore, Naïve Bayes offers the advantages of simplicity, speed, and high accuracy (Sinaga et al., 2022). The Support Vector Machine (SVM) was chosen because it is an excellent algorithm and produces high levels of accuracy compared to other algorithms. Support Vector Machine is a method for data analysis and pattern recognition that can be used for classification. This method requires labeling. (Laela & Baihaqi, 2024).

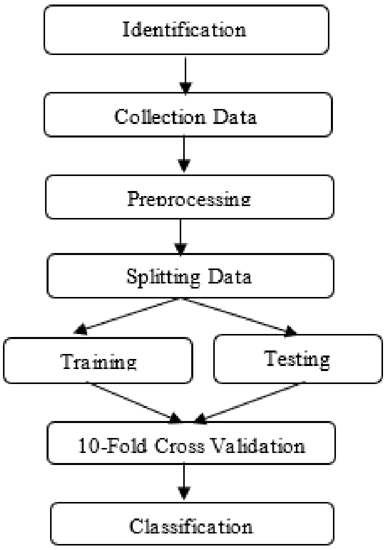
In previous research conducted by (Soer et al., 2023) with the title "Diagnosis Prediction of Typhoid Fever Disease Using Data Mining with the Naive Bayes Classifier Algorithm Method" using early 2018 data at Budi Asih Hospital using the Naive Bayes algorithm using 250 patient data of typhoid fever cases. The results of the analysis also showed that the accuracy of classifying typhoid fever cases using the naive Bayes method in this study was 92%. Previous research conducted by (Arifin et al., 2021) with the title "Classification of Cervical Cancer Disease Using the Support Vector Machine (SVM) Algorithm" This research aims to simplify the classification of cervical cancer data. The data taken from the Archive.com website consisted of 72 data sets and 19 attributes, using 59 training data sets and 4 attributes. The data was divided into two, namely testing and training data, with an 80:20 split. The test results, using an 80:20 data split, had an accuracy value of 92.9%. In research conducted by (Muntiari & Hanif, 2022) Comparing several algorithms regarding breast cancer to find an algorithm that has high accuracy, the results obtained from several algorithms, namely the neural network algorithm obtained a system accuracy value of 94.25%, the decision tree obtained a system accuracy value of 95.00%, the k-nearest neighbor obtained a system accuracy value of 95.00%, the random forest algorithm obtained a system accuracy value of 94.75%, the naïve bayes algorithm obtained a system accuracy value of 95.00%, the support vector machines algorithm obtained a system accuracy value of 94.75% and several other algorithms.

The Naïve Bayes algorithm and Support Vector Machine (SVM) method for breast cancer classification. This study compares the performance of Naïve Bayes and SVM on the Breast Cancer Wisconsin dataset to identify which algorithm provides better classification performance under the selected evaluation scheme. Furthermore, this study also compares the performance of Naïve Bayes and SVM on the Breast Cancer Wisconsin dataset to identify which algorithm provides better classification performance based on the applied evaluation scheme.

RESEARCH METHODS

This study was conducted to classify breast cancer types using data processing techniques by implementing the Naïve Bayes and Support Vector Machine (SVM) algorithms. The data processing and testing procedures were carried out using the Rapidminer application. The dataset was divided using an 80:20 ratio, in which 80% of the data were used as training data and the remaining 20% were used as testing data. Furthermore, the model evaluation process was performed using the Cross Validation method, specifically 10-Fold Cross Validation, in order to obtain more accurate and stable evaluation results.

In the classification stage, the Naïve Bayes algorithm implemented in this study was Gaussian Naïve Bayes, as it is suitable for datasets containing numerical attributes. In addition, this study also applied the Support Vector Machine (SVM) algorithm using a linear kernel function. During the SVM model training stage, the initial parameters were initialized with $\alpha = 0$, $C = 1$, epsilon = 0.001, gamma = 0.5, and lambda = 0.5 to regulate the optimization process and improve the performance of the SVM model.



Source: Research (2026)

Figure 1. Research Framework

1. Identification

In this study, identifying the problem is the initial step taken by reading literature such as journals to find the problem issues that can be raised from this problem, the problem taken is classifying breast cancer, then looking for an algorithm that has high accuracy, this study applies the Naïve Bayes algorithm and Support Vector Machine (SVM) in predicting the type of cancer. By conducting analysis using the Naïve Bayes and Support Vector Machine (SVM) methods, it is possible to determine the type of benign and malignant breast cancer early on, with the aim of allowing patients to quickly receive appropriate treatment for the type of cancer they are suffering from, with precise accuracy and a high accuracy value.

2. Collection Data

Data collection was conducted by searching websites that provide public datasets, for example, UCI Machine Learning. The dataset used in this study is available at the following link https://archive.ics.uci.edu/dataset/15/breast+cancer+wisconsin+original . This study used a dataset called Breast Cancer Wisconsin (Original). There are 699 data points with 11 attributes: ID, Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli, Mitoses, Class.

Table 1. Attribute Description

Attribute	Description
ID	Sample identifier
Mitoses	Cell division activity
Clasess	Tumor diagnosis class
Clump Thickness	Thickness of the tumor cell clusters
Uniformity of Cell Size	Degree of uniformity in cancer cell size
Uniformity of Cell Shape	Degree of uniformity in cancer cell shape
Marginal Adhesion	Adhesion between cancer cells and surrounding tissue
Single Ephithelial Cell Size	Size of individual epithelial cells
Bare Nuclei	Nuclei not surrounded by cytoplasm (cell remnants)
Bland Chromatin	Uniformity and regularity of chromatin distribution within the nucleus
Normal Nucleoli	Presence, uniformity, and regularity of nucleoli within the nucleus

Source: Research (2026)

3. Preprocessing

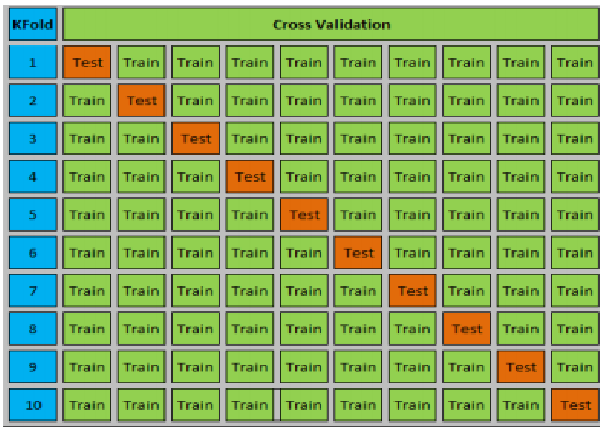
At this stage, attribute selection, data cleaning, and data transformation are performed. Data preprocessing is necessary because the system cannot use the raw data directly. To improve the accuracy of the data used, some processing must be performed to slightly modify the data (Hibatullah & Apriandari, 2024). Data cleaning involves checking for missing values in each attribute in the Wisconsin Breast Cancer dataset. The transformation stage involves transforming numerical variables into polynomials for the Naïve Bayes test, and nominal-to-numerical transformation for the SVM test.

4. Splitting Data

Data splitting is the process of separating training data and testing data. According to (Musu et al., 2021) A crucial component of supervised learning requires two different types of data: training data and testing data. Training data is a set of data with class labels or attributes used to train a classification model. Testing data is the data used to assess the performance of the model. The data is divided into 80% training data and 20% testing data. At this stage, the data is split using Rapidminer.

5. 10-Fold Cross Validation

One of the advantages of using K-Fold Cross Validation is its ability to improve the stability and accuracy of performance estimation by calculating the average results from “k” iterations, thereby achieving lower empirical risk compared to merely dividing the dataset into training data and testing data (Pradana & Atina, 2024). The Cross Validation method applied in this study was 10-Fold Cross Validation, whereby the dataset was iteratively partitioned and evaluated through 10 testing processes to obtain more reliable and robust model evaluation results (Septiani et al., 2022).



KFold	Cross Validation									
1	Test	Train	Train	Train	Train	Train	Train	Train	Train	Train
2	Train	Test	Train	Train	Train	Train	Train	Train	Train	Train
3	Train	Train	Test	Train	Train	Train	Train	Train	Train	Train
4	Train	Train	Train	Test	Train	Train	Train	Train	Train	Train
5	Train	Train	Train	Train	Test	Train	Train	Train	Train	Train
6	Train	Train	Train	Train	Train	Test	Train	Train	Train	Train
7	Train	Train	Train	Train	Train	Train	Test	Train	Train	Train
8	Train	Train	Train	Train	Train	Train	Train	Test	Train	Train
9	Train	Train	Train	Train	Train	Train	Train	Train	Test	Train
10	Train	Train	Train	Train	Train	Train	Train	Train	Train	Test

Source: (Pradana & Atina, 2024)

Figure 3. 10-Fold Cross Validation Scheme

6. Classification

At the classification stage, the Naïve Bayes and Support Vector Machine (SVM) algorithm methods are applied.

a. Naïve Bayes

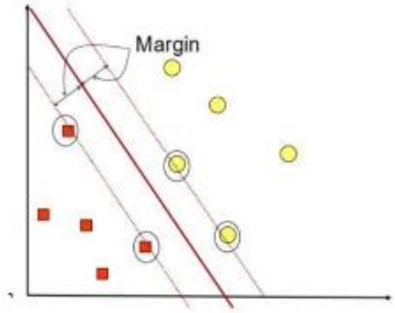
Naive Bayes is a simple classification method that calculates all probabilities based on Bayes' theorem combined with a combination of database frequency values. In classification tasks, the goal is to predict the class label of a given sample based on a set of features or characteristics. (Syahril Dwi Prasetyo et al., 2023). The Naïve Bayes algorithm can be formulated using the equation (Putra et al., 2024) :

$$P(H|X) = \frac{P(X|H)}{P(X)} P(H) \qquad (1)$$

Information:
X : Data with unknown class
H : Data hypothesis
P(H|X) : Probability of hypothesis H based on conditions X (posteriori probability)
P(H) : Hypothesis Probability
P(X|H) : Probability of X based on the conditions in Hypothesis H
P(X) : Probability X

b. Support Vector Machine

Support Vector Machine is a vector-based supervised machine learning algorithm. SVM classifies data by dividing it into two classes using a hyperplane. The hyperplane is located between the two classes with a distance d between the closest points of each class. This distance d is called the margin, and the points on this margin are called support vectors. SVM aims to determine the best hyperplane with the maximum margin relative to the data. (Ipmawati et al., 2024).



Source : (Ipmawati et al., 2024)

Figure 2. Support Vector Machine Representation

7. Model Evaluation

Model performance evaluation was conducted using Cross Validation and confusion matrix analysis based on accuracy, recall, precision, F1-score, and ROC analysis to determine the most optimal classification method between Naïve Bayes and Support Vector Machine (SVM) in predicting cancer types. The dataset was divided using an 80:20 ratio, where 80% of the data were used as training data and 20% as testing data. Furthermore, 10-Fold Cross Validation was applied to the training data to obtain more stable and accurate model evaluation results before conducting the final testing process.

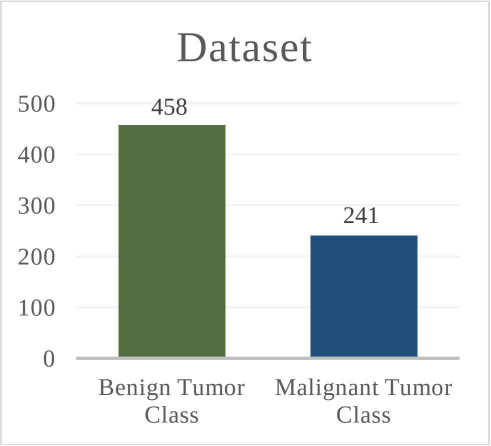
RESULTS AND DISCUSSION

1. Identification

The initial stage involved identifying the problem and identifying issues that could be used for this research. One method involved conducting a literature review of articles, journals, and books to identify the issues. After conducting the literature review, the problem was identified as a disease frequently experienced by women with a high mortality rate, namely breast cancer. Therefore, early diagnosis of breast cancer, including predicting the type of breast cancer, is necessary using data mining methods with machine learning. The algorithm used is Naïve Bayes and Support Vector Machine (SVM). By using this method, it is hoped that patients can be treated and given treatment quickly and according to the type of cancer.

2. Collection Data

In the data collection stage, this study utilized the *Breast Cancer Wisconsin* dataset obtained from the *UCI Machine Learning Repository*. The dataset consists of 699 samples and 11 attributes, with two target classes: *benign* (458 samples) and *malignant* (241 samples). Data cleaning was subsequently performed by removing records containing missing values to ensure data quality. In addition, the ID attribute was excluded because it serves only as a sample identifier and does not contribute to the classification process. After data cleaning and attribute selection, the final dataset comprised 683 records and 10 attributes with no remaining missing values.



Source: Research (2026)

Figure 3. Visualization of the Number of Datasets

3. Preprocessing

The data pre-processing stage is used to improve the accuracy of the classification technique comparison process on the dataset. The first step in this stage is attribute selection. Only 10 attributes are used out of a total of 11 attributes. Then, the data cleaning stage is carried out by checking for missing values. In this data, several records that do not have values or missing values are removed because they are meaningless information. The table below shows data that has gone through attribute selection and data cleaning, where there are no missing values. Therefore, the data used in this study includes 683 records and 10 attributes.

Table 3. Marginal Adhesion, Normal Nucleoli, Single Epithelial Cell Size.

Name	Type	Missing
Clump thickness	*int*	0
uniformity of cell size	*int*	0
uniformity of cell shape	*int*	0
marginal adeshion	*int*	0
single epithelial cell size	*int*	0
bare nuclei	*int*	0
bland chromatin	*int*	0
normal nuchleoli	*int*	0
mitoses	*int*	0
class	*int*	0

Source: Research (2026)

4. Splitting Data

The next step involves partitioning the dataset using the Split Data operator in the RapidMiner application with a ratio of 80%:20%. A total of 80% of the data is allocated as training data to develop the model, while the remaining 20% is used as testing data to evaluate the model's performance. This data partitioning process results in 546 training instances and 137 testing instances. A more detailed distribution of the data is presented in the table below:

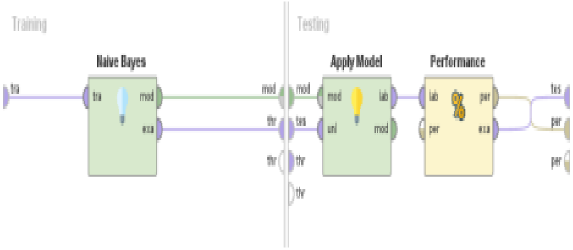
Table 4. Data Distribution

Persentase (Training : Testing)	Training	Testing
80% : 20%	546	137

Source: Research (2026)

5. 10 K-Fold Cross Validation

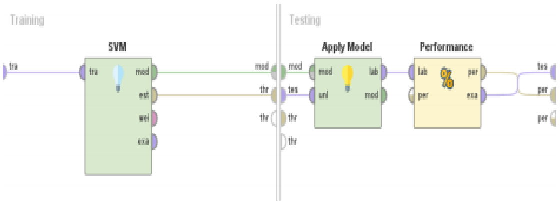
At the model evaluation stage, the Cross Validation method was applied using K = 10 folds, commonly referred to as 10-Fold Cross Validation. This method was employed to obtain more accurate and stable evaluation results through the alternating division of training data and testing data. In this study, the evaluation process was conducted on the Gaussian Naïve Bayes algorithm to measure the model performance in classifying breast cancer types based on the numerical data used. The implementation process of Cross Validation in the RapidMiner application is shown in the figure below:



Source: Research (2026)

Figure 4. The Cross Validation Process Applied to The Naïve Bayes Algorithm

At the model evaluation stage, the Cross Validation method was applied using K = 10 folds, commonly referred to as 10-Fold Cross Validation. This method was employed to obtain more accurate and stable evaluation results through the alternating division of training data and testing data. In this study, the evaluation process was conducted on the Support Vector Machine (SVM) algorithm to measure the model performance in classifying breast cancer types based on the numerical data used. The implementation process of Cross Validation in the RapidMiner application is shown in the figure below:

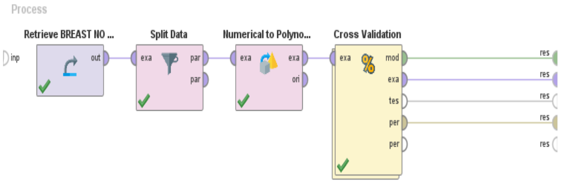


Source: Research (2026)

Figure 4. The Cross Validation Process Applied to The Support Vector Machine Algorithm

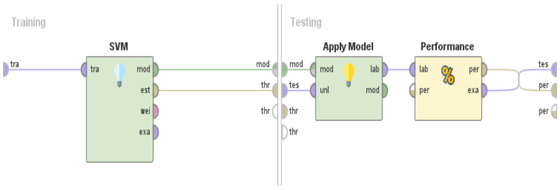
6. Classification

At the algorithm classification stage, computations are performed by applying the Naïve Bayes and Support Vector Machine (SVM) algorithms to the model using the RapidMiner application. This process aims to develop a classification model capable of identifying patterns within the data and generating accurate predictions, as illustrated in the figure below:



Source: Research (2026)

Figure 4. Initial SVM Process



Source: Research (2026)

Figure 5. Cross Validation Process

7. Model Evaluation

Evaluation is a stage that aims to assess the performance of the classification model. At this stage, the results obtained from the previous processes are analyzed using several evaluation metrics, including accuracy, precision, recall, and Area Under the Curve (AUC), in order to measure the effectiveness and reliability of the model in performing classification.

Table 5. Naïve Bayes Confusion Matrix

Accuracy : 96,15%

	True Benign	**True Malignant**	***Precision***
Pred. Benign	388	4	98.83%
Pred. Malignant	17	187	91.67%
Recall	95.21%	97.91%	

Source: Research (2026)

Table 5 explains that the results of the confusion matrix evaluation of the Naïve Bayes algorithm, based on training data, are known from 546 data, 388 data were classified as benign cancer according to the prediction of the Naïve Bayes algorithm, 4 data were predicted as malignant but turned out to be benign, 17 data were predicted as benign but turned out to be malignant and 187 malignant classes were predicted accordingly. The Naïve Bayes algorithm has an accuracy of 96.15%.

Table 6. SVM Confusion Matrix

Accuracy : 96.89%

	True Benign	**True Malignant**	***Precision***
Pred. Benign	352	6	98.32%
Pred. Malignant	11	177	94.15%
Recall	96.97%	96.72%	

Source: Research (2026)

Table 6 explains that out of 546 data, 352 data were classified as benign cancer according to the Support Vector Machine (SVM) prediction, 6 data were predicted to be malignant but turned out to be benign, 11 data were predicted to be benign but turned out to be malignant and 177 data turned out to be malignant according to the prediction. The Naïve Bayes algorithm has an accuracy of 96.89%.

CONCLUSION

Based on experiments conducted on the breast cancer dataset, the Naive Bayes and Support Vector Machine (SVM) algorithms were employed to classify breast cancer diagnoses. The Naive Bayes algorithm achieved an accuracy of 96.15%, while the Support Vector Machine (SVM) achieved a higher accuracy of 96.89%. In a comparative analysis of both methods, the Support Vector Machine (SVM) was demonstrated to be the superior approach for breast cancer classification, as it yielded a higher accuracy in classifying the dataset. Therefore, the Support Vector Machine (SVM) can be concluded as the most effective method for breast cancer classification in this study.

REFERENCES

Abdul Jabbar, M., Hasmin, E., Susanto, C., Musu, W., & Artikel, I. (2022). Komparasi Algoritma Decision Tree, Naive Bayes, dan K-Nearest Neighbors dalam Klasifikasi Kanker Payudara Comparison of Decision Tree Algorithms, Naive Bayes, and K-Nearest Neighbors in Breast Cancer Classification. *Oktober*, *14*(3), 258–270. https://www.doi.org/10.22303/csrid.14.3.2022.258-270

Arifin, S. S., Siregar, A. M., Ratna, A., & Mudzakir, T. Al. (2021). Klasifikasi Penyakit Kanker Serviks Menggunakan Algoritma Support Vector Machine (SVM). *The 4th Conference on Innovation and Application of Science and Technology (CIASTECH 2021)*, (4), 521–528.

Azizah, Q. N., Puspitasari, D., Sukmana, S. H., Sikumbang, E. D., & Ramanda, K. (2025). Pendekatan Algoritma Naive Bayes Dalam Memprediksi Penyakit Diabetes. *Jurnal Infortech*, *7*(2). http://ejournal.bsi.ac.id/ejurnal/index.php/infortech89

Hibatullah, A. A., & Apriandari, W. (2024). *KLASIFIKASI KUALITAS JENIS KOPI HALUS ROBUSTA MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) DAN MOBILENET-V2*. *8*(5), 8650–8657.

Ipmawati, J., Saifulloh, S., & Kusnawi, K. (2024). Analisis Sentimen Tempat Wisata Berdasarkan Ulasan pada Google Maps Menggunakan Algoritma Support Vector Machine: Sentiment Analysis of Tourist Attractions Based on Reviews on Google Maps Using the Support Vector Machine Algorithm. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(1), 247–256.

Karamti, H., Alharthi, R., Umer, M., Shaiba, H., Ishaq, A., Abuzinadah, N., Alsubai, S., & Ashraf, I. (2024). Breast cancer detection employing stacked ensemble model with convolutional features. *Cancer Biomarkers*, *40*(2), 155–170. https://doi.org/10.3233/CBM-230294

Ketut, S. (2022). Stadium ca mamae. *Ganesha Medicine Journal*, *2*(1), 2–7.

Laela, I. N., & Baihaqi, W. M. (2024). Analisis Klasifikasi Penyakit Multiple Sclerosis Menggunakan Algoritma Logistic Regression dan SVM. *Generation Journal*, *8*(1), 27–33. https://doi.org/10.29407/gj.v8i1.20646

Muntiari, N. R., & Hanif, K. H. (2022). Klasifikasi Penyakit Kanker Payudara Menggunakan Perbandingan Algoritma Machine Learning. *Jurnal Ilmu Komputer Dan Teknologi*, *3*(1), 1–6. https://doi.org/10.35960/ikomti.v3i1.766

Pradana, A. I., & Atina, V. (2024). *Teknik K-Fold Cross Validation untuk Mengevaluasi Kinerja Mahasiswa*. 239–248. https://doi.org/10.33364/algoritma/v.21-1.1618

Putra, I. M. A. A. D., Sunarya, I. M. G., & Gunadi, I. G. A. (2024). Perbandingan Algoritma Naive Bayes Berbasis Feature Selection Gain Ratio dengan Naive Bayes Kovensional dalam Prediksi Komplikasi Hipertensi. *JTIM : Jurnal Teknologi Informasi Dan Multimedia*, *6*(1), 37–49. https://doi.org/10.35746/jtim.v6i1.488

Qurrotaayun, S. F., Lestari, D. M., Fitrianto, R., Simanjuntak, S. D., & Ariyanto, J. (2025). Peningkatan Kasus Kanker : Kontribusi Human Papilloma Virus (HPV) dan Gaya Hidup. *GALENICAL : Jurnal Kedokteran Dan Kesehatan Mahasiswa Malikussaleh*, *4*(3), 117–129. https://doi.org/10.29103/jkkmm.v4i3.2202

Rahmatia, A. S., Rijal, S., Mutmainnah, M. I., Purnamasari, R., & Irsandy, F. (2024). *FAKUMI MEDICAL JOURNAL Analisis Faktor Risiko Kanker Payudara*. *04*(06), 473–484.

Septiani, R. K., Anggraeni, S., & Saraswati, S. D. (2022). Klasifikasi Sentimen Terhadap Ibu Kota Nusantara (IKN) pada Media Sosial Menggunakan Naive Bayes. *Jurnal Teknika*, *16*(2), 245–254.

Sinaga, S., Sembiring, R. W., & Sumarno, S. (2022). Penerapan Algoritma Naive Bayes untuk Klasifikasi Prediksi Penerimaan Siswa Baru. *Journal of Machine ...*, *1*(1), 55–64. https://journal.fkpt.org/index.php/malda/article/view/162%0Ahttps://journal.fkpt.org/index.php/malda/article/download/162/115

Soer, U. D., Fauziah, S., & Aggita, M. (2023). Diagnosa Prediksi Penyakit Thypoid Fever Menggunakan Data Mining Dengan Metode Algoritma Naive Bayes Classifier. *Jurnal Teknologi Informatika Dan Komputer*, *9*(1), 523–536. https://doi.org/10.37012/jtik.v9i1.1611

Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, & Fitri Nurapriani. (2023). Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN. *Jurnal KomtekInfo*, *10*, 1–7. https://doi.org/10.35134/komtekinfo.v10i1.330

World Health Organization. (2024). *Malnutrition*. https://www.who.int/news-room/fact-sheets/detail/malnutrition