SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14642

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Implementation of Support Vector Machine Algorithm for Heart Disease Risk Identification Using Signal Electrocardiogram

Bagas Andhika[1], Yennimar[2], Grisela Sangap Damayanti Saragih[3], Fahriza Shiddik[4], Gabriela Br. Surbakti[5]
[1,2,3,4,5]Universitas Prima Indonesia
[1]bagasandhikaa10@gmail.com, [2]yennimar@unprimdn.ac.id, [3]grisela1910@gmail.com, [4]fahrizashiddik@gmail.com, [5]surbaktigabriella@gmail.com

Sumitted : March 13, 2025 | **Accepted** : March 26, 2025 | **Published** : April 14, 2025

Abstract : In the medical world, one of the biggest contributors to death in the world is heart disease. Early detection of the risk of heart disease can increase the chances of recovery and reduce mortality. This research applies the Support Vector Machine (SVM) algorithm to identify the risk of heart disease using Electrocardiogram Signals. The ECG data used was taken from a public database that contained a record of information on the electrical activity of the heart of patients with various heart health conditions. The Support Vector Machine algorithm is applied to classify ECG signals into 2 main classes, namely normal conditions and at-risk conditions. Several methods in data processing, including data normalization and feature selection are used to improve the accuracy and success of the model. The results of the evaluation with this method resulted in accuracy, precision, recall and also F1-score showed that the modeling of this algorithm produced a fairly good classification, with an accuracy of more than 90% in the identification of heart disease risk. This study shows the potential use of this algorithm in automatically detecting the risk of heart disease based on ECG signals, which can be a tool in medical diagnosis. The results show that implementing the SVM strategi with the RBF kernel appears to be a very easy execution when compared to the direct part. An important component that affects the adequacy of an SVM strategy is the parameters of the section and the way the information is handled.

Keywords: Electrocardiogram, Heart Disease, Risk Identification, Support Vector Machine

INTRODUCTION

As one of the organs of the human body, the heart plays an important role in the circulatory system. This part has the obligation to bring blood with oxygen and nutrients to the human body and return blood with carbon dioxide content to the lung organs for purification.(Tsuroyya, Ramadhani, and Ramadhani 2025)

To detect abnormalities in the heart, a device called an Electrocardiogram (ECG) is used. An electrocardiogram is a signal emitted from the movement of the heart muscle, this signal can be translated through the recordings produced by the ECG machine. Early detection of heart disease can minimize mortality rates in people with heart disease(Primadevi and Mardiana 2023). Later, this signal records the rhythm pattern of the heart rate and is used to detect defects in the heart, such as: Arrhythmias, Tachycardia, and also Atrial Fibrillation. Because ECG signals are often compromised, signal processing methods are required to clean, analyze, and extract important features of the signal.(Eka Patria 2020)

Algorithmic Support Vector Machines (SVM) is a classification method for linear and non-linear data. SVM is one of the popular machine learning techniques used for classification and regression. SVM leverages kernel functions that systematically create support vector classifiers in the form of higher dimensions (Chowdhury and Schoen 2020)

SVM is widely used by bioinformatics researchers to predict diseases, for example research predicts diabetes mellitus. (Wafa et al. 2022)

With the development of today's technology, especially in the medical field, it is possible to increase research on heart disease to provide education related to early screening of abnormalities that may identify heart

*Yennimar

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14642
e-ISSN : 2541-2019
p-ISSN : 2541-044X

disease. The purpose of this study is to be able to group indications of abnormalities or abnormalities found in patients through data from ECG scans and can be classified according to existing heart diseases.

LITERATURE REVIEW

The following is an overview of the latest developments in this research through *a literature review* of previous research related to the implementation of the *Support Vector Machine method,* including:

One study showed that SVM is able to classify heart conditions based on the pattern of electrical activity recorded by the ECG. In this study, the ECG data used had gone through a signal cleaning process, including noise removal techniques and data normalization, to improve the quality of input into the model. After the data was ready, the main features representing the characteristics of the heart signal were extracted using methods such as wavelet transform and dimensionality reduction techniques. This study shows that with optimal data processing, SVM can achieve high accuracy in detecting normal and abnormal categories in heart conditions (Primadevi and Mardiana 2023). Another study compared the effectiveness of SVM with other classification algorithms, such as Naïve Bayes and K-Nearest Neighbor (KNN), in analyzing patient data at risk of cardiovascular disease. This study showed that SVM has advantages in handling non-linear data, especially with the use of the Radial Basis Function (RBF) kernel. The data used in this study includes various health parameters, such as blood pressure, cholesterol levels, and heart measurements, which are then processed to determine the relationship between these factors and the risk of heart disease. With parameter optimization techniques such as adjusting the C and gamma values in SVM, the developed model is able to provide predictions with a better level of precision than other methods.(Maulidah et al. 2021)

In addition to being used in the classification of heart disease, the Support Vector Machine (SVM) method is also applied in the detection of financial transaction anomalies, especially in detecting fraud in digital payment systems. In related research, researchers used a financial transaction dataset containing various parameters such as the number of transactions, transaction time, location, and account usage patterns. The data was then analyzed to distinguish between normal and suspicious transactions. To improve the effectiveness of the classification, a feature engineering technique was used that transformed the raw data into a more representative form for the SVM model. The results showed that SVM was able to recognize anomalous patterns with high accuracy, especially with the use of the RBF kernel which is more flexible in handling complex and unstructured transaction patterns.(Muhammadiyah, Aceh, and Nusantara 2024)

In another study, SVM was used to analyze ECG signals with the frequency domain transformation method to improve the quality of features extracted from heart signals. This study compares various signal transformation techniques, such as Short-Time Fourier Transform (STFT) and Stockwell Transform, to identify heart rate patterns associated with cardiovascular disease. The experimental results show that the use of frequency domain transform can improve the ability of SVM to distinguish between normal and abnormal signals more accurately. In addition, this study highlights the importance of proper feature selection, as irrelevant features can reduce the model's performance in heart disease classification (Pratiwi, Rizal, and Magdalena 2020). SVM is also combined with an Artificial Neural Network (ANN) to improve the accuracy of detecting heart disorders based on ECG signals. This study shows that the hybrid approach, which combines the ability of SVM to build optimal hyperplanes with ANN to extract non-linear features, produces better performance than a single approach. With this combination, the model is able to handle large and complex datasets more efficiently. However, this study also notes that the hybrid method has higher computational requirements, so optimization techniques need to be applied to maintain data processing efficiency.(Niendy Alexandra Yosephine and Ratnadewi 2021)

METHOD

Support Vector Machine or abbreviated as SVM is used to partition information into certain levels. SVM is used to find the best hyperplane to partition 2 classes of information with the farthest distance between classes. SVM is used to overcome classification problems, both double (2 classes) and multiclass (more than 2 classes).(Faruk and Nafi'iyah 2020) The purpose of SVM is to find the ideal line or boundary that can divide information, with the aim of providing an edge or separation between the boundary and the information. In addition, SVM can also be used in the relapse problem to anticipate the number of values in a data that has been obtained. One of the advantages of SVM is that it is able to function well in information which has a complicated format (Rusman and Pasae 2023). When the input space information inside cannot be directly differentiated (it is not certain to find a suitable hyperlane), then we can use bitwork to convert that information to a larger room which is often referred to as bitspace. SVM calculation has 4 kinds of bit functions, namely Straight Bit, Basic Spiral Function (RBF), Sigmoid, and Polynomial Bit.(Maneno et al. 2023) There are 3 things to consider when Using the SVM, as follows:

1. Support Vectors are classes that exist when classification is performed.
2. Hyperlanes can be useful for isolating classes from others.

*Yennimar

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14642
e-ISSN : 2541-2019
p-ISSN : 2541-044X

3. The margin is the separation between the main class and the second class.(Manik, Ernawati, and Nurlaili 2021)

This research requires several tools and materials to support the implementation of the desired results. This research material is taken from the Kaggle dataset titled MIT-BIH Arrhythmia Database which will later be processed by the system to get the desired results, here is an example of a dataset in csv format. The dataset below we took from https://www.kaggle.com/datasets/mondejar/mitbih-database

Table 1. Dataset

	'Sample #'	**'MLII'**	**'V5'**
0	0	995	1011
1	1	995	1011
2	2	995	1011
3	3	995	1011
4	4	995	1011
5	5	995	1011
6	6	995	1011

Information:

- Sample # : Data sample number
- MLII, V5 : Amplitude values on MLII and V leads

Research on the identification of heart disease risk with the Support Vector Machine algorithm involves several processes with data testing and data training in obtaining a result. The stages in this algorithm are illustrated in Figure 1.

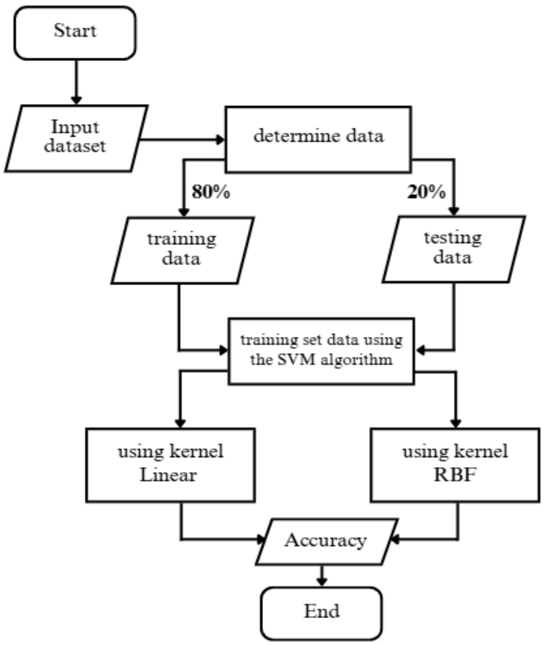

Fig 1. Research Flowchart

In this study, the process we created on the system consisted of the initial stage, namely entering the dataset which is the process of retrieving data from Kaggle. Then after getting the data, it is continued to determine the sample

*Yennimar

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14642
e-ISSN : 2541-2019
p-ISSN : 2541-044X

to be used and then separate the sample into training data and data testing. Next, it enters the processing stage, which is to apply the SVM method with 2 kernels, namely the Linear kernel and the RBF kernel. Some of the metrics that will be looked for in this SVM method are accuracy, precision, recall, and F1-score. After that, we can compare the two kernels to get the best results.

RESULT

In this research, ECG signal information was used which included first amplitude (MLII) and second amplitude data, namely (V5). The data we have consists of 48 .csv files, but we only use 200,000 samples in the first .csv file, This data sample will be divided into 2 elements, namely 80% will be used for data training models and 20% for testing. The process carried out is as follows.
The researcher used the Python programming language and jupyter as applications to perform the process of this SVM method. The first step is to import the necessary libraries. (Muadin and Asnal 2023)

```
import numpy as np
import pandas as pd
from sklearn import svm
from sklearn.preprocessing import StandardScaler
from sklearn.datasets import make_classification
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score
from sklearn.model_selection import cross_val_score
import matplotlib.pyplot as plt
```

Fig 2. Library python

Next, a data load process is carried out which uses the pandas library to read data from a CSV file with the name "ekg.csv" and loads 650,000 of the data into the ekg variable. After that, the program displays the first 5 lines of data using the method head() on the ecg variable. The next process is to clean the data to clean up the noise in the dataset. Then the author displays the data visualization using hyperlane and gets the results as below.

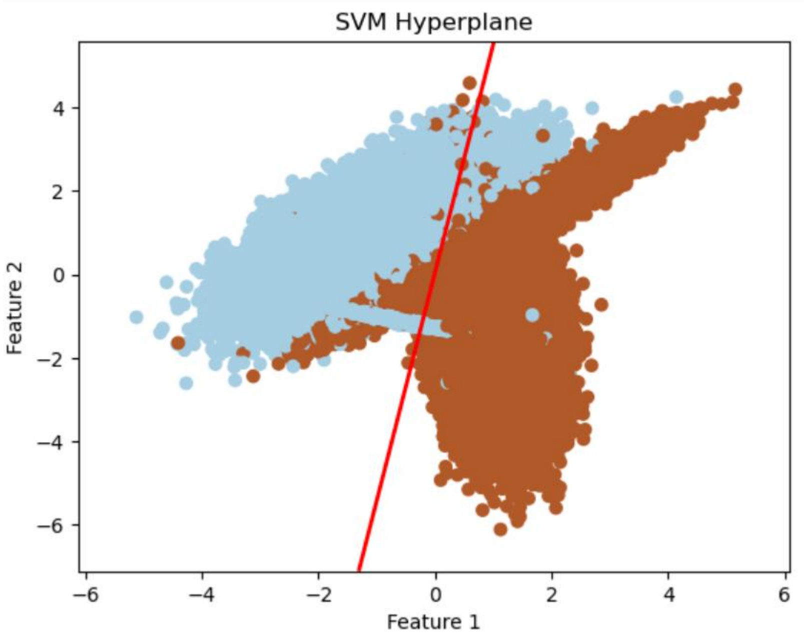

Fig 3. SVM Hypernlane

The use of the SVM algorithm is carried out on a variety of parameters to obtain the best conclusions. After completing a series of training and testing processes, the SVM algorithm is then evaluated into several important metrics such as accuracy, precision, recall and F1-score.

*Yennimar

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 2, April 2025
DOI : https://doi.org/10.33395/sinkron.v9i2.14642
e-ISSN : 2541-2019
p-ISSN : 2541-044X

```
# Model SVM dengan kernel Linear
svm_linear = SVC(kernel='linear', C=1.0, random_state=42)
svm_linear.fit(X_train, y_train)
y_pred_linear = svm_linear.predict(X_test)

# Model SVM dengan kernel RBF
svm_rbf = SVC(kernel='rbf', C=1.0, gamma='scale', random_state=42)
svm_rbf.fit(X_train, y_train)
y_pred_rbf = svm_rbf.predict(X_test)

# Menghitung metrik untuk kernel Linear
precision_linear = precision_score(y_test, y_pred_linear)
recall_linear = recall_score(y_test, y_pred_linear)
f1_linear = f1_score(y_test, y_pred_linear)
accuracy_linear = accuracy_score(y_test, y_pred_linear)

# Menghitung metrik untuk kernel RBF
precision_rbf = precision_score(y_test, y_pred_rbf)
recall_rbf = recall_score(y_test, y_pred_rbf)
f1_rbf = f1_score(y_test, y_pred_rbf)
accuracy_rbf = accuracy_score(y_test, y_pred_rbf)
```

Fig 3. Define metrics on the RBF and Linear kernels

Next, the output of the SVM method uses Base Function (RBF) Radias instead of from kernel linear to determine the best outcome. The table below shows the results of the comparison of the Linear and RBF models.

Table 2. Results of comparison of Linear and RBF models

Kernel SVM	Accuracy	Precision	Recall	F1-score
Linear+	92,66%	93,80%	91,47%	92,62%
RBF	97,18%	97,65%	96,74%	97,19%

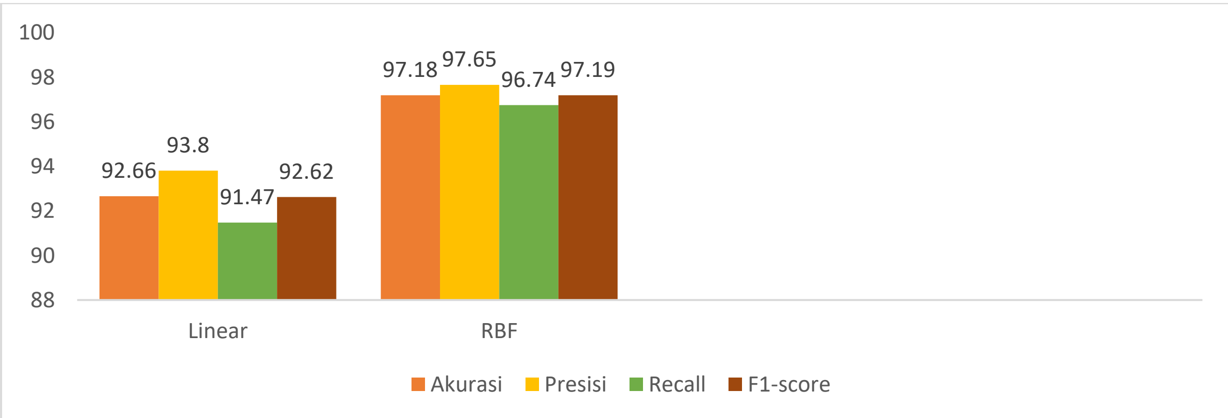

Fig 4. Graph of Accuracy, Precision, Recall and F1-score results

DISCUSSION

The results obtained from this study indicate a significant difference between the two types of Support Vector Machine (SVM) kernels, namely Linear and Radial Basis Function (RBF), in identifying heart disease risk based on Electrocardiogram (ECG) signals. A comparative analysis of these kernels shows that the RBF kernel demonstrated superior classification performance in all evaluation metrics, including accuracy, precision, recall, and F1-score. This indicates its ability to capture more complex patterns in ECG signal variations, making it more effective in identifying potential heart disease cases. The highest accuracy was achieved by the RBF kernel at 97.18%, compared to the Linear kernel at 92.66%. Similarly, the recall and precision values for the RBF kernel

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

outperformed those of the Linear kernel, confirming that this approach is more robust in detecting heart disease risks.

The effectiveness of the SVM model is significantly influenced by data preprocessing techniques, including normalization and feature selection. Proper handling of input data ensures a balanced model, preventing bias toward any particular class and optimizing prediction performance. The results highlight the potential of the SVM algorithm, particularly with the RBF kernel, as a decision-support tool in medical diagnostics. Its high classification accuracy suggests that it can assist medical professionals in the early detection of heart disease risk based on ECG data. However, further studies with larger datasets and additional feature extraction techniques are recommended to enhance its reliability in real-world clinical settings.

The findings emphasize the importance of selecting an appropriate kernel function in SVM-based classification tasks, as well as the necessity of thorough data preprocessing. These results align with prior studies, reinforcing that machine learning techniques, when properly applied, can improve accuracy in medical diagnostics. From the research that has been carried out, there are a few things that will be proposed for future research, including:

1. In the future, for research using the SVM method, it will be easier if the amount of data used is not too large. Datasets are effective on a scale between 1,000-100,000 samples.
2. In the next study, it is suggested that researchers try to use other methods to cover a larger dataset.
3. To make it easier to understand, it's a good idea to add more visualizations in the form of images and graphics.

CONCLUSION

The Support Vector Machine method can in fact be used to identify the risk of heart disease. The use of the SVM method is more inclined to small datasets, because the processing of large datasets can affect the performance of longer SVMs. The reason why researchers use the SVM method is also because they look at previous research. The SVM method has a fairly high level of prediction accuracy compared to other methods but it is possible that the use of the SVM method also gives significant results in prediction results.

Based on the results of this study, there are several conclusions namely first, the adequacy of the SVM calculation, especially using the Radial Basis Function (RBF) component, has had a good impact on identifying the risk of heart disease. Generally presented with a precision value of up to 97.18%, which shows that the algorithm can provide accurate test results. Second, Model Execution where SVM Modeling and RBF components show better results than direct components in accuracy, score, and F1-score. The resulting accuracy (97.65%) shows that the modeling sometimes predicts false positive results, while high recall (96.74%) tells the modeling can identify the risk of heart disease well.

REFERENCES

Alhaq, Z., Mustopa, A., & Santoso, J. D. (2021). *Application of the Support Vector Machine method for Twitter user sentiment analysis*.

Chowdhury, Shovan, and Marco P. Schoen. 2020. “Research Paper Classification Using Supervised Machine Learning Techniques.” *2020 Intermountain Engineering, Technology and Computing, IETC 2020* (October 2020). doi:10.1109/IETC47856.2020.9249211.

Eka Patriya, Eka. 2020. "Implementation of Support Vector Machine in Composite Stock Price Prediction (JCI)." *Scientific Journal of Technology and Engineering* 25(1): 24–38. doi:10.35760/tr.2020.v25i1.2571.

Faruk, Muhammad, and Nur Nafi'iyah. 2020. "Classification of Skin Cancer Based on Texture Features, Image Color Features Using SVM and KNN." *Telematics* 13(2): 100–109.

Made Dwi Dharma Sreya, E. B. S. (2022). *Use of GloVe Method for Feature Expansion in Twitter Sentiment Analysis with Naïve Bayes and Support Vector Machine*.

Maneno, Regolinda, Budiman Baso, Patricia G. Manek, and Kristoforus Fallo. 2023. "Detection of the Maturity Level of Areca Nuts Using the Support Vector Machine Method Based on Color and Texture." *Journal of Information and Technology* 3(2): 60–66. doi:10.32938/jitu.v3i2.5323.

Manik, Geyessella, Iin Ernawati, and Ika Nurlaili. 2021. "SENTIMENT ANALYSIS ON E-COMMERCE USER REVIEWS IN THE FOOD SECTOR USING THE SUPPORT VECTOR MACHINE METHOD (Case Study: Sayurbox and Tanihub Review on Google Play)." *Proceedings of the National Seminar for Students in the Field of Computer Science and Its Applications* 2(2): 64–74. https://conference.upnvj.ac.id/index.php/senamika/article/view/1630.

Maulidah, Nurlaelatul, Riki Supriyadi, Dwi Yuni Utami, Fuad Nur Hasan, Ahmad Fauzi, and Ade Christian. 2021. "Prediction of Diabetes Mellitus Using Support Vector Machine and Naive Bayes Methods." *Indonesian Journal of Software Engineering (IJSE)* 7(1): 63–68. doi:10.31294/ijse.v7i1.10279.

Muadin, Muhammad, and Hadi Asnal. 2023. “Implementasi Metode Support Vector Machine Pada Opinion Mining Masyarakat Terkait Chatgpt.” *JOISIE Journal Of Information System And Informatics Engineering*

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

e-ISSN : 2541-2019
p-ISSN : 2541-044X

7(1): 78–84.

Muhammadiyah, University, Mahakarya Aceh, and Bina Nusantara University. 2024. "The Use of Support Vector Machine (SVM) Algorithms for Fraud Detection in Online Transactions." 13: 1627–32.

Niendy Alexandra Yosephine, and Ratnadewi. 2021. "The Use of Artificial Neural Networks on Electrocardiogram Signals to Detect Supraventricular Arrhythmic Heart Disease." *INFORMATION (Journal of Informatics and Information Systems)* 13(1): 14–23. doi:10.37424/information.v13i1.69.

Pratiwi, Dyah Ayu, Achmad Rizal, and Rita Magdalena. 2020. "Application of Stockwell Transforms and K-Nearest Neighbor for Electrocardiogram Signal Classification." *AITI : Journal of Information Technology* 17(1): 22–32.

PRAMUDITA, D. A. AND MUSDHOLIFAH, A. (2020) 'GSA to Obtain SVM Kernel Parameter for Thyroid Nodule Classification', IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 14(1), p. 11. doi: 10.22146/ijccs.41215.

Primadevi, Ferawidya, and Yahya Mardiana. 2023. "Performance Test of Heart Signal Denoising System or ECG Using Empirical Mode Decomposition (EMD) Algorithm." *Journal of Asymmetrics: Scientific Journal of Engineering & Innovation* 5: 27–34. doi:10.35814/asiimetric.v5i1.3682.

Rifaldi, D., & Fadlil, A. (2023). *DECODE: Journal of Information Technology Education Preprocessing Techniques on Text Mining Using "Mental Health" Tweet Data. 3*(2), 161–171. https://doi.org/10.51454/decode.v3i2.131

Rusman, Juprianus, and Nofrianto Pasae. 2023. "Prototype of Arabica Coffee Fruit Sorting System Based on Maturity Level Using the Support Vector Machine Method." *Technique* 12(1): 65–72. doi:10.34148/teknika.v12i1.602.

Tsuroyya, Feni Atika, Khotimah Nur Ramadhani, and Elsa Oktavia Ramadhani. 2025. "A Review of the Heart Organ as a Center of Life in the Cardiovascular System." 3.

Wafa, Hovi Sohibul, Asep Id Hadiana, Fajri Rakhmat Umbara, Jl Terusan, Jend Sudirman, Cimahi Sel District, and Cimahi City. 2022. "Prediction of Diabetes Using Support Vector Machine (SVM) Algorithm." 1: 40–45.

*Yennimar

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.