International Journal of Informatics and Computation (IJICOM)
Vol. 7, No. 2, 2025
ISSN: 2685-8711, E-ISSN: 2714-5263 DOI: 10.35842/ijicom

Classification of Heart Disease Using the Ensemble SVM Method

Ammara Desma Marzooqa[1], Dewi Pramudi Ismi[2]

Abstract

Cardiovascular disease (CVD), particularly coronary heart disease, remains the leading cause of global mortality. It makes early detection essential for effective prevention of the diseases. Machine learning offers a promising alternative for rapid and accurate prediction. This study investigates the performance of Support Vector Machine (SVM) classifiers enhanced through an ensemble stacking approach. In this study, we employed three SVM kernels, including linear, RBF, and polynomial, using GridSearchCV to obtain accuracies of 97.1%, 97.2%, and 96.3%, respectively. Experimental results show that the optimized stacking ensemble achieved the highest accuracy of 97.5%, with TP=91, FN=4, FP=1, and TN=104. This model outperformed individual SVM kernels and surpassed several existing methods, including ANN and hybrid SVM–NN approaches. The findings confirm that integrating multiple optimized SVM kernels enhances classification accuracy, stability, and robustness for heart disease prediction. The proposed ensemble-based SVM model provides a valuable contribution to medical diagnostics by improving early detection reliability and supporting preventive strategies for cardiovascular diseases.

Keywords:
Classification, Heart Disease, Ensemble, SVM, Stacking

1. Introduction

Cardiovascular disease (CVD) is a global threat that endangers many lives. The World Health Organization (WHO) notes that more than 17.9 million people die each year from cardiovascular disease [2]. Based on 2020 data, the number of CHD patients in Indonesia was recorded at 352,618 men and 442,674 women, with the highest rate of 3.6% in the 65-74 age group, which means that around 3-4 out of every 100 people in that age range suffer from CHD [3]. Coronary heart disease is one of the leading causes of death in Indonesia. Coronary heart disease occurs when the heart is unable to pump blood or when there is a blockage in the blood vessels, resulting in heart failure. Risk factors for coronary heart disease are divided into modifiable and non-modifiable risk factors. Non-modifiable risk factors include age, gender, family history, and others. Meanwhile, modifiable risk factors can certainly prevent heart disease. Examples include unhealthy lifestyles such as tobacco use or smoking, an unbalanced diet such as one that is high in fat, alcohol, and a lack of physical activity [4].

Given the seriousness of the situation, early detection of heart disease is essential. The use of machine learning is an alternative solution that provides faster and more accurate predictions. Machine learning plays an important role in everyday life and is widely used in various sectors, including medicine. Machine learning is a learning technology developed to learn independently without instruction, thereby simplifying work. Machine learning offers

---

Corresponding Author: Dewi Pramudi Ismi (dewi.ismi@tif.uad.ac.id)
1 Ammara Desma Marzooqa, Universitas Ahmad Dahlan, 2200018269@webmail.uad.ac.id
2 Dewi Pramudi Ismi, Universitas Ahmad Dahlan, dewi.ismi@tif.uad.ac.id

many benefits, such as time efficiency and medical workforce efficiency [5]. Other benefits include reducing costs and speeding up diagnosis time so that action can be taken more quickly. The machine learning process begins with data collection and ends with model evaluation [6]. Machine learning is divided into three types: supervised learning, unsupervised learning, and reinforcement learning [7]. Early detection of disease falls under supervised learning because it uses classification. Classification is the grouping of objects or concepts that have distinctive features and characteristics based on similarities and differences into several classes [8]

There are various classification algorithms, such as K-Nearest Neighbor (KNN), Random Forest, C45, Multilayer Perceptron (MLP), Logistic Regression, Naïve Bayes, SVM, and DT [9] [10] [11]. Previous studies predominantly used single classifiers, while those using ensemble learning had different case studies and lacked information about meta models. Single classifier models are prone to overfitting and underfitting, sensitive to noise, and have limited performance, so the choice of algorithm greatly affects the model results. Single classifiers are simple, fast, and clear, but fragile when using complex data, lots of noise, and requiring high accuracy. Because single classifiers have problems with disease detection and do not perform optimally, this study uses an ensemble learning approach. Ensemble is an approach that uses several algorithms combined into one to produce better and more accurate performance than a single classifier [12].

Combining two or more models allows the system to learn from multiple data sources and make more accurate, precise, and robust predictions because one model corrects the other [13]. This study uses a stacking ensemble in classifying models because it excels in flexibility and the ability to capture complex patterns from a combination of various models. Ensemble SVM with the stacking method chosen because the stacking method is suitable for combining the outputs of several simple models. The SVM algorithm was chosen because of its strong generalization ability, robustness to small and high-dimensional medical data, and flexibility in handling nonlinear patterns through various kernels. Moreover, SVM tends to be less affected by outliers and overfitting, which are common challenges in medical datasets. By applying an Ensemble SVM with a stacking approach, this study leverages the strengths of multiple SVM kernels (linear, polynomial, and RBF) to enhance model stability and reduce false negatives, making it more reliable for early detection of heart disease.

Therefore, this study applies an Ensemble SVM with a stacking approach to enhance model stability and minimize false negatives in heart disease detection. By leveraging the unique strengths of multiple SVM kernels, the model aims to improve predictive reliability and support early, accurate diagnosis of cardiovascular diseases, which can ultimately assist healthcare professionals in timely intervention and treatment planning.

2. Related Works

Research on heart disease prediction is widely found using various methods and in various case studies. Study [14] applied only one algorithm, namely the *Artificial Neural Network* (ANN) algorithm. This study produced an accuracy rate of 73.77%. High recall indicates that the model is quite good at detecting cases of heart disease (positive). The study shows the important role of pre-processing in model training. The study also shows that ANN is a fairly good method for classifying heart disease.

Another study, namely study [15] on heart disease prediction, uses the SVM method. The study applies several types of SVM kernels to be tested. These types of kernels include *linear, Radial Basis Function* (RBF), and *polynomial*. The study found that the *linear* kernel was the most effective among the other kernels, with an accuracy of 86.92%. This study shows that the selection of the right kernel can determine the accuracy and success of a model. The dataset used was Heart Failure Prediction with the link www.kaggle.com/datasets/fedesoriano/heart-failure-prediction.

Another study [16] compared *Neural Network* algorithms with SVM in classifying heart disease using the *Heart Disease Dataset* (a combination of two sub-datasets: Hungary and Switzerland) from the UCI *Machine Learning Repository*. The study found that the SVM algorithm outperformed the Neural Network with an accuracy of 83%. The SVM algorithm proved to be accurate with a relatively small amount of data. The selection of appropriate features is also very important for successful classification. A recent study applied the *Ensemble Learning* approach with the *stacking* method using various algorithms. The base learners used the *Gaussian Naïve Bayes* (GNB), *K-Nearest Neighbour* (KNN), and SVM algorithms with a *meta-learner* that was not described in detail. The highest accuracy achieved was 97.19% from results with a dataset split percentage of 80:20. The study shows that the *Ensemble stacking* method can improve the accuracy of anemia prediction models [17].

From these studies, there have not been many studies on ensemble learning for heart disease prediction models. Most studies only compare algorithms or single classifiers. Therefore, this study uses *Ensemble* SVM with the *stacking* method because the *stacking* method is suitable for combining the outputs of several simple models. SVM was chosen because the algorithm is individually strong in classification. This advantage makes SVM more stable in classification and capable of recognizing complex patterns in medical data with limited quantities but many attributes. SVM with the right kernel combination has been proven to reduce FN values, thereby increasing *recall*, making it a more reliable algorithm in supporting early diagnosis, and has been proven to improve the performance of heart classification models.

3. Proposed Method

3.1 SVM Algorithm

SVM is one of the *supervised learning* algorithms commonly used for classification [18]. SVM is divided into two types, namely *linear* SVM and *non-linear* SVM [34]. For *non-linear* data, SVM uses a *kernel trick* that transforms the data so that it becomes separate [19]. In SVM, the classification process is carried out by finding the best separating line *(hyperplane)* between two data groups. This line is determined by important data located at the edge of each group, called support vectors.

In this study, the main objective of SVM is to find this separating line by maximizing the distance between the two groups so that the classification results are more accurate [20]. SVM also has several types of kernels, including RBF, *Sigmoid, Linear, Non-linear*, *Polynomial*, and others. The simplest kernel used is the linear kernel for linear data classification, while the RBF kernel is used for non-linear data [21]. SVM excels because it can work well for linear and non-linear classification with the help of the *kernel trick*. In data with clear margins between classes, SVM can also improve accuracy, making it superior in predicting heart disease. For further clarification, please refer to the following mathematical formula.

$$f(x) = \operatorname{sign}\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right) \tag{1}$$

Table 1 – Mathematical formula SVM method

Notation	Description
$f(x)$	Decision function that determines the class of new data x
α_i	Training result: Lagrange multiple (determines the weight of each support)
y_i	Class label of training i $(+1$ or $-1)$
$K(x_i, x)$	Kernel function that transforms data into a higher-dimensional space
b	Bias
sign(.)	Sign function that determines whether the result is positive (+1) or negative (-1)

3.2 Ensemble learning

Ensemble learning is a method in machine learning that combines two or more models to produce predictions that are more accurate, stable, and reliable than a single model (single classifier). The main concept of ensemble is that the collective decision of several models is usually better than relying on a single model, because each model has its own strengths and weaknesses. By combining these models, ensemble learning can reduce the risk of overfitting, improve generalization, and enhance performance in terms of accuracy and resilience to varying data. Ensemble techniques are generally divided into several approaches, such as bagging, boosting, and stacking, each of which has a different strategy for combining models to strengthen prediction results.

3.3 *Stacking*

Stacking is a technique in ensemble learning that combines several different models (base models). The prediction results from these models are then combined using a meta-learning algorithm to produce the final output. In this way, stacking utilizes the strengths of various types of models to improve prediction accuracy. For further clarification, please refer to the following mathematical formula.

$$\hat{y} = \mathrm{sign}\left(\sum_{m=1}^{M} w_m f_m(x)\right) \tag{2}$$

This formula shows that the final prediction $(\hat{y})$ is generated from a weighted sum of the prediction results of several basic models $f_m(x)$, where each model has a weight w_m That indicates how much influence it has on the final decision. The sum is then fed into the sign (·) function to determine the final class (+1 or −1). In other words, this formula describes how the meta model in stacking combines and balances the results of all base models (SVM with linear, RBF, and polynomial kernels) to produce a more accurate and stable final prediction.

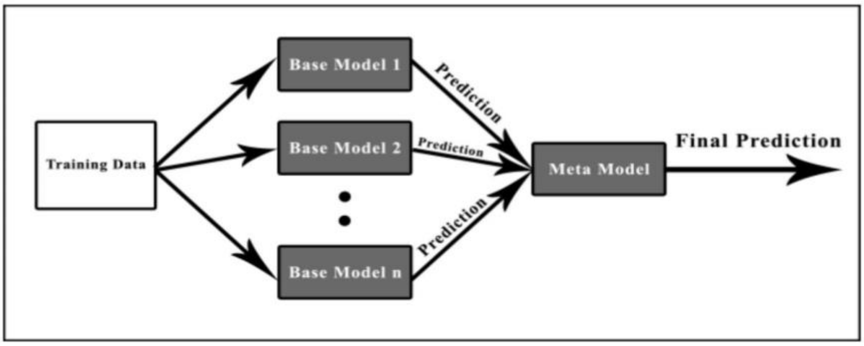

Figure 1 - *Stacking*

4. Experimental Setup

4.1 Dataset

In this study, we gathered a dataset from the India-Cardiovascular-Dataset(https://data.mendeley.com/datasets/dzz48mvjht/1). It consists of 1,000 patient medical records with 14 relevant health attributes. In this study, the data collection process was carried out using this secondary dataset. From the total available attributes, this study used 12 attributes as predictor variables, namely age, gender, chest pain, resting BP, serum cholesterol, fasting blood sugar, resting electrocardiogram, max heart rate, exercise angina, old peak, slope, and no of major vessels. Meanwhile, the 'target' attribute was used as a label or class variable in the classification process for heart disease prediction. This study discarded the 'patientid' attribute and divided the dataset into training data and testing data at a ratio of 80:20.

Table 2 - Details of the Dataset

No.	Attribute	Type of Data	Range (Minimum – Maximum)
1.	*Age*	Numeric	In Years
2.	*Gender*	Binary	0,1
3.	*chestpain*	Nominal	0,1,2,3
4.	*restingBP*	Numeric	94-200 (mm HG)
5.	*serumcholestrol*	Numeric	126-546 (mg/dl)
6.	*fastingbloodsugar*	Binary	0,1 > 120 mg/dl
7.	restingrelectro	Nominal	0,1,2
8.	*maxheartrate*	Numeric	71-202
9.	*exerciseangia*	Binary	0,1
10.	*oldpeak*	Numeric	0-6.2
11.	*slope*	Nominal	1,2,3
12.	*noofmajorvessels*	Numeric	0,1,2,3
13.	*target*	Binary	0,1

This study conducted several stages to build a machine learning model for detecting heart disease. These stages begin with data loading, pre-processing, creating a single classifier model, combining it with the stacking method, and obtaining the final prediction results. This can be seen in the following image.

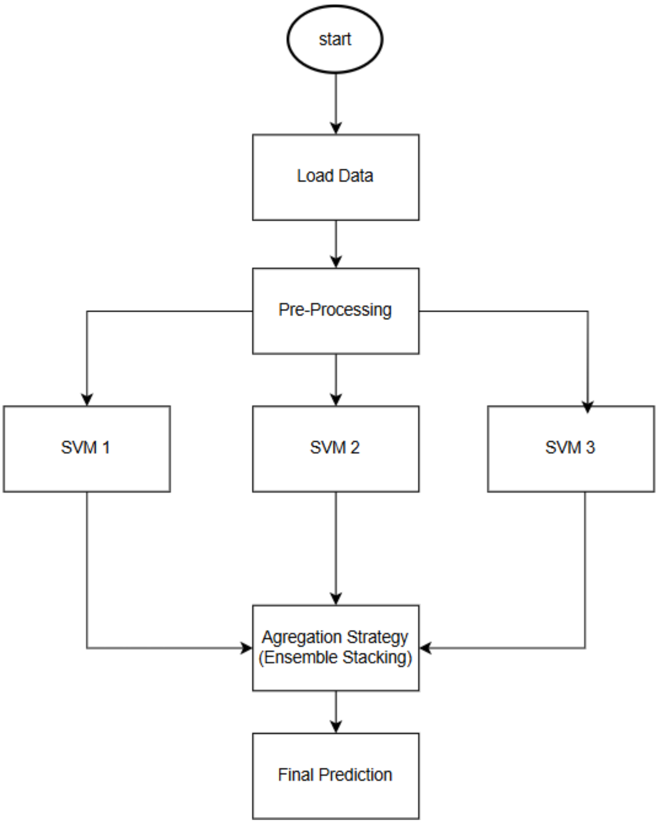

Figure 2 – Research Stage

4.2 Pre-Processing

Data pre-processing is a stage carried out to process raw data into data that is ready to be used for model creation. The output of this stage will be training data input that is ready to be used during model training. Pre-processing at this stage is carried out to check whether there are missing values or attributes that are not used in this study. This study has no attributes with missing values. The 'patientid' attribute is dropped because it has no specific meaning in this study or is only used as an identifier. The target attribute was discarded because it is the attribute to be predicted in this study. Then, the dataset was divided into two parts, namely training data and testing data. The purpose of dividing the data is to train the machine learning model to recognize patterns in the training set and measure the performance of the machine learning model with the testing set. This study will divide the dataset into an 80:20 ratio for training data and testing data.

4.3 Implementation Model

At this stage, the first SVM model will be built using the previously processed training data. The model architecture designed at this stage uses a linear kernel. To maximize the linear kernel, this study will use GridSearchCV. The second model will be trained using an RBF kernel. To maximize the RBF kernel, two parameters will be used, namely the C and gamma parameters. To determine these parameters, this study will use GridSearchCV to facilitate the research. The final model training will use a polynomial kernel. It is almost the same as the second model, but this kernel will have three parameters, namely the C, gamma, and degree parameters. These parameters are determined using GridSearchCV to facilitate the research.

GridSearchCV is one of the parameter optimization methods provided by the scikit-learn library. This technique is used to find the best combination of parameters for a machine learning algorithm by performing an exhaustive search of the parameter values that have been predetermined in a grid. By using GridSearchCV, researchers can find the optimal parameter configuration so that the model performs better in terms of accuracy, precision, recall, and F1-score. After the three models are run, the next process is to combine the models to obtain a more accurate and better final prediction result. This final prediction will be the final model for heart disease prediction. The method used in this ensemble is stacking, with model characteristic adjustments. The meta learner also uses the same algorithm, namely SVM.

4.4 Model Evaluation

Model evaluation involves using a confusion matrix consisting of accuracy, precision, recall, and F1-score. By using these metrics, we can assess the effectiveness of the SVM ensemble method in predicting heart disease and ensure that the model built is not only accurate but also reliable in detecting critical health conditions. A confusion matrix is a process in which the performance of an established model is evaluated by displaying the number of correct and incorrect predictions for each class in a matrix format [39]. A confusion matrix serves to assess the effectiveness of the machine learning algorithm used. The confusion matrix is divided into four terms, namely True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). The explanations are as follows.

Table 3 - Confusion Matrix

TP (True Positive)	FN (False Negative)
FP (False Positive)	TN (True Negative)

5. Result and Analysis

In the first training, we utilized an SVM with a linear kernel, which achieved an accuracy of 97% with confusion matrix scores of TP=91, FN=4, FP=2, and TN=103. This model demonstrates strong classification performance and high reliability in identifying positive and negative cases. In the next experiments, we employ two additional SVM models using RBF and polynomial kernels. The RBF kernel achieved an accuracy of 80% with TP=68, FN=27, FP=13, and TN=92, indicating moderate predictive performance. Similarly, the polynomial kernel also reached 80% accuracy with TP=74, FN=21, FP=18, and TN=87, showing comparable effectiveness in detecting heart disease. Detailed comparisons of accuracy, precision, recall, and F1-score for all three models are presented in Table 4.

Table 4 - Details of Linear, RBF and Polynomial Kernel Accuracy

Kernel	**Metrics**			
	Accuracy	**Precision**	**Recall**	**F1-Score**
Linear	97%	96.2%	98.1%	97.1%
RBF	80%	77.3%	87.6%	82.1%
Polynomial	80.5%	80.5%	82.8%	81.6%

At the next stage, we employed GridSearchCV, a hyperparameter optimization method, to provide the best parameters for each kernel. This resulted in the best performance of the three models. The Linear kernel produced the best parameters C = 10. For the RBF kernel, the best parameters were C = 1, gamma = 0.1. Meanwhile, for the polynomial kernel, the best parameters were C = 1, degree = 3, gamma = scale. There is an improvement in accuracy in each SVM model. For complete accuracy results, see the following table.

Table 5 - Accuracy Results with GridSearchCV

Kernel	**Accuracy**
Linear	97.1%
RBF	97.2%
Polynomial	96.3%

At the final stage, we combined the three models using an ensemble learning approach with the stacking method. The base model in this study is the SVM algorithm with the three kernels, namely linear, RBF, and Polynomial. The linear kernel was chosen because it has the highest accuracy compared to the other kernels. By using a stacking ensemble, there was an increase in model accuracy of 97.5% with a detailed confusion matrix of TP = 91, FN = 4, FP = 1, and TN = 104.

Table 6 – Comparison with Other Studies

Method	Accuracy
ANN	73.77%
SVM (Linear)	86.9%
Ensemble Stacking (GNB, KNN, SVM)	97.19%
Ensemble Stacking SVM	97.5%

The final results show that the Ensemble SVM model using the stacking method achieved the highest accuracy of 97.5%, outperforming the individual SVM models and several previous studies. Compared to the study by [10], which used the ANN algorithm with an accuracy of 73.77%, and the study [11], which used an individual SVM with the best kernel achieving 86.92%. The proposed Ensemble SVM model demonstrates a significant improvement in predictive performance. Furthermore, when compared with the study by [12], which obtained an accuracy of 83% using SVM and NN, and the study [13], which achieved 97.19% accuracy using ensemble stacking for anemia prediction, the results of this study indicate that ensemble stacking applied specifically to SVM kernels can produce more stable results. These findings confirm that combining multiple SVM kernels through stacking can effectively enhance classification performance. It can reduce false negatives and improve the reliability of early heart disease detection systems.

6. Conclusion

Based on the results of the research conducted, the implementation of the SVM algorithm with the ensemble stacking approach has proven to provide higher accuracy compared to the use of individual SVM models. This study utilized three types of SVM kernels linear, RBF, and polynomial, with the linear kernel selected as the meta model due to its superior individual performance. Parameter optimization using GridSearchCV further improved the performance of each kernel, resulting in accuracies of 97.1% for the linear kernel, 97.2% for the RBF kernel, and 96.3% for the polynomial kernel.

According to the experimental result, the combination of the three optimized SVM kernels using the stacking ensemble produced the highest accuracy of 97.5%, with a TP=91, FN=4, FP=1, and TN=104. These findings indicate that the ensemble stacking approach substantially improves classification performance and stability compared to using individual SVM models. Compared with previous studies, the proposed Ensemble SVM model achieved superior performance to several existing methods, including ANN, individual SVM kernels, and hybrid SVM–NN. It is believed that applying ensemble stacking with SVM kernels can yield more accurate results for heart disease prediction. By combining the strengths of multiple SVM kernels, the resulting model not only achieves high accuracy but also improves the reliability of early heart disease detection. This research highlights the potential of ensemble-based SVM models as a significant contribution to the medical field, particularly in enhancing diagnostic quality and supporting preventive efforts for cardiovascular diseases in the future.

As future research, it can construct a new Ensemble SVM model using larger and more diverse datasets to evaluate its generalization capability. In addition, integration with deep learning algorithm methods could be explored to further improve accuracy and the model's ability to capture more complex patterns. Future studies may also consider incorporating additional clinical features such as genetic history, lifestyle factors, and laboratory test results to produce more contextually relevant and realistic predictions. Moreover, implementing that could be a practical step to assist medical professionals in performing early heart disease detection more efficiently.

References

[1] Doppala B. P. and Bhattacharyya D., “Cardiovascular Disease Dataset Description,” doi: 10.17632/dzz48mvjht.1.

[2] T. R. Hapsari, L. Sukmarini, and T. Herawati, “Cardiovascular Disease Detection in ECG Signals Based on Deep Learning,” *Journal of Telenursing (JOTING)*, vol. 5, no. 1, pp. 1376–1383, Jun. 2023, doi: 10.31539/joting.v5i1.5178.

[3] E. Erdania, M. Faizal, and R. B. Anggraini, “Factors Associated with the Incidence of Coronary Heart Disease at Dr. (H.C.) Ir. Soekarno Hospital, Bangka Belitung Province, 2022,” *Jurnal Keperawatan*, vol. 12, no. 1, pp. 17–25, Jun. 2023, doi: 10.47560/kep.v12i1.472.

[4] D., S. E., M. K. Maniarti, and S. Notoatmodjo, *Lifestyle of Determinant: Coronary Heart Disease Patients*. PT. RajaGrafindo Persada – Rajawali Pers, 2021.

[5] M. Majid and I. Nawangsih, “Comparison of Ensemble Methods to Improve Machine Learning Algorithm Accuracy in Predicting Breast Cancer,” *Jurnal SAINTIKOM*, vol. 23, no. 1, p. 97, Feb. 2024, doi: 10.53513/jis.v23i1.9563.

[6] P. Prihandoko, R. G. Alam, G. Gunawan, and D. Abdullah, *Understanding Machine Learning Concepts and Implementation*. PT. Sonpedia Publishing Indonesia, 2024.

[7] M. R. S. Alfarizi, M. Z. Al-Farish, M. Taufiqurrahman, G. Ardiansah, and M. Elgar, “Using Python as a Programming Language for Machine Learning and Deep Learning,” 2023.

[8] N. I. Widiastuti, E. Rainarli, and K. E. Dewi, “Summarization and Support Vector Machine in Document Classification,” *JURNAL INFOTEL*, vol. 9, no. 4, p. 416, Nov. 2017, doi: 10.20895/infotel.v9i4.312.

[9] M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, “Effective Heart Disease Prediction Using Machine Learning Techniques,” *Algorithms*, vol. 16, no. 2, p. 88, Feb. 2023, doi: 10.3390/a16020088.

[10] Setiawan, A. Muhammad, and S. H. F. Dewi, “Application of Classification Algorithms for Early Detection of Coronary Heart Disease Based on Clinical Symptoms,” *Teknik: Jurnal Ilmu Teknik dan Informatika*, vol. 5, no. 1, pp. 18–26, May 2025, doi: 10.51903/teknik.v5i1.706.

[11] S. Heristian, “Comparison of Machine Learning Algorithms for Heart Disease Classification,” *Jurnal Infortech*, vol. 6, no. 1, pp. 46–51, Jun. 2024, doi: 10.31294/infortech.v6i1.21888.

[12] S. K. M. T. P. D. D. E. Sondakh, S. W. Taju, M. G. Tene, and A. E. T. Pangaila, “Sentiment Analysis System for Online Shopping Application Reviews Using Ensemble Learning,” *CogITo Smart Journal*, vol. 9, no. 2, pp. 280–291, Dec. 2023, doi: 10.31154/cogito.v9i2.525.280-291.

[13] G. D. Sepbriant and D. W. Utomo, “Ensemble Learning in E-Commerce Product Categorization Using Boosting Techniques,” *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 9, no. 2, pp. 123–133, May 2024, doi: 10.14421/jiska.2024.9.2.123-133.

[14] Pradana, M. L. Alghifari, M. F. Juna, and D. Palaguna, “Heart Disease Classification Using Artificial Neural Networks,” *Indonesian Journal of Data and Science*, vol. 3, no. 2, pp. 55–60, Jul. 2022, doi: 10.56705/ijodas.v3i2.35.

[15] L. N. Farida and S. Bahri, “Heart Failure Classification Using the Support Vector Machine (SVM) Method,” *Komputika: Jurnal Sistem Komputer*, vol. 13, no. 2, pp. 149–156, Oct. 2024, doi: 10.34010/komputika.v13i2.11330.

[16] M. D. F. Tino, H. Hasanah, and T. D. Santosa, “Comparison of Support Vector Machines (SVM) and Neural Networks for Heart Disease Classification,” *INFOTECH Journal*, vol. 9, no. 1, pp. 232–235, May 2023, doi: 10.31949/infotech.v9i1.5432.\

[17] D. Anita, N. Mukhlisah, R. Indira, N. R. Dewi, and Y. Andriani, “Implementation of Stacking Ensemble Techniques for Anemia Classification,” *Journal of Information Engineering and Educational Technology*, vol. 9, no. 1, pp. 51–55, Jul. 2025, doi: 10.26740/jieet.v9n1.p51-55.

[18] H. Yusufi, A. Kharisma, A. D. Adinata, D. F. Ramzy, and M. M. Santoni, “Mortality Risk Prediction in Cardiovascular Disease Patients Using Ensemble Learning,” 2022.

[19] N. B. Binna, T. Rohana, H. Y. Novita, and S. Faisal, “Classification of Tomato Varieties Using K-Nearest Neighbor and Support Vector Machine Algorithms,” *Jurnal Informatika Teknologi dan Sains (Jinteks)*, vol. 7, no. 2, pp. 800–807, May 2025, doi: 10.51401/jinteks.v7i2.5743.

[20] Y. Amrozi, D. Yuliati, A. Susilo, N. Novianto, and R. Ramadhan, “Classification of Banana Types Based on Color Images Using the SVM Method,” *Jurnal Sisfokom*, vol. 11, no. 3, pp. 394–399, Dec. 2022, doi: 10.32736/sisfokom.v11i3.1502.

[21] Mustofa and S. Pradana, “Comparison of Phishing Detection Tests Using SVM with RBF and Linear Kernels,” *SISTEMASI: Jurnal Sistem Informasi*. [Online]. Available: http://sistemasi.ftik.unisi.ac.id