Received May 5th, 2025; accepted Juni 18th, 2025. Date of publication June 30th, 2025
Digital Object Identifier: https://doi.org/10.25047/jtit.v12i1.446

Utilizing Data Mining Approach For Hypertension Diagnosis Classification

PUDJI WIDODO[1], HERIBERTUS ARY SETYADI[2], HARTATI DYAH WAHYUNINGSIH[3], SUNDARI[4]

[1,2]Universitas Bina Sarana Informatika, Jakarta, Indonesia
[3]Universitas Dharma AUB, Surakarta, Indonesia
[4]Universitas Duta Bangsa, Surakarta, Indonesia

CORESPONDING AUTHOR: PUDJI WIDODO (email:pudji.piw@bsi.ac.id)

ABSTRACT Hypertension is one of the factors contributing to the highest death rates from non-communicable diseases in various countries. Every year, the number of hypertension sufferers increases significantly. It is estimated that in 2025, the number of hypertension sufferers will reach 1.5 billion individuals. Data mining aims to identify patterns that can help in decision making, classification, and prediction. One of the well-known algorithms or methods for classification is the Support Vector Machine (SVM). The SVM method aims to find the best hyperplane or decision boundary function that can separate two or more classes of data in the input space. This research purpose is to determine the classification results and accuracy of the diagnosis of hypertension using the SVM method. Eleven attributes used include age, smoking habits, physical activity, sugar consumption, salt consumption, fat consumption, alcohol consumption, lack of fruit and vegetable consumption, systolic and diastolic blood pressure. This research will utilize Jupyter Notebook tools and Python programming language as research tools. The SVM method was trained with various kernel attributes and hyperparameters to produce the best model. From the results it is known that the RBF kernel used with parameters $C = 100$ and $\gamma = 0.1$ produces an accuracy of 97.5% which is the best model in classifying hypertension. From these results it can be concluded that the SVM method is able to produce a very good classification of hypertension diagnosis and can provide a diagnosis to detect hypertension early.

KEYWORDS: classification, data mining, hypertention, SVM

I.INTRODUCTION

Hypertension or high blood pressure, often referred to as "The Silent Killer" because it is often without complaints [1]. Hypertension or high blood pressure is a disorder of the blood vessels that causes the supply of oxygen and nutrients, carried by the blood, to be blocked to the body's tissues that need it. The incidence of hypertension increases with age [2]. Hypertension can also be interpreted as a condition when a person has systolic blood pressure above 120 mm and diastolic blood pressure above 80 Hg or can be written above 120/80 mmHg [3].

Hypertension is still a concern for all levels of society in all corners of the world. The impacts of this disease can be felt in the short and long term [4]. Hypertension is one of the factors contributing to the highest death rates from non-communicable diseases in various countries [5]. The condition of increased blood pressure will cause the heart to work harder to channel blood throughout the body through the blood vessels. So the higher a person's blood pressure, the higher the person's chances of suffering from heart disease, kidney failure and stroke [6]. Every year, the number of hypertension sufferers increases significantly. It is estimated that in 2025, the number of hypertension sufferers will reach 1.5 billion individuals [7]. Apart from that, it is estimated that every year around 9.5 million people die due to hypertension and the complications it causes [8].

Data mining aims to identify patterns that can help in decision making, classification, and prediction. Data mining can be used to extract information from large amounts of data into a useful and understandable form [9]. Classification is a technique of grouping data mining into certain classes or categories based on their characteristics [10]. One of the well-known algorithms or methods

for classification is the Support Vector Machine (SVM). SVM is an algorithm that can be used for both linear and non-linear data classification [11]. This algorithm is a supervised learning algorithm, which means this algorithm can find out data from certain label models in the dataset used [12]. The SVM method aims to find the best hyperplane or decision boundary function that can separate two or more classes of data in the input space [13]. The SVM method was chosen in this research because it has a clear, systematic and consistent concept when compared with other classification algorithms. The SVM method can produce good accuracy values even though the data analyzed is unbalanced [14]. The SVM method can also reduce classification errors for unknown samples and increase generalization ability compared to Artificial Neural Network methods and sensitivity to outliers [15].

Globally, the World Health Organization (WHO) estimates that the prevalence of hypertension will reach 33% in 2023 and 67% of them are in poor and developing countries. Total people with hypertension is expected to continue to increase, reaching 1.5 billion people worldwide by 2025 [16]. In 2024, the prevalence of hypertension in Indonesia reached 36%. In the 45-54 age group, hypertension showed the highest prevalence rate, which was 45.3% [17]. Therefore, early detection and treatment of hypertension is very important. One way to improve the accuracy and efficiency of medical experts in classifying a disease is to consider a classification system developed using data mining. This research aims to provide a diagnosis classification for hypertension and determine the level of accuracy of using the SVM method from the classification results.

There are several previous researches that also classify hypertension. One of the research results published in a journal form was developed a web-based application for diagnosing hypertension using the Naïve Bayes method. Dataset used came from the Framingham Heart Study (FHS) with 4,434 participants from the National Institutes of Health. Research can produce binary classifications of prediction classes by adjusting small configurations. The test results accuracy for binary classification before discretization were 77.15% and after discretization were 76.4%. The developed model accuracy successfully predicted hypertension classification by 84.28% [18].

Other research developed an Internet of Things (IoT) based system to measure blood pressure and heart rate using the K-Nearest Neighbors (KNN) method. The resulting system can measure blood pressure and detect systolic, diastolic blood pressure, and heart rate in real-time. Data processing using KNN method produces status classification into optimal hypertension, normal, prehypertension, grade 1 hypertension, and grade 2 hypertension. Using the KNN method produces consistent and accurate values, so it can support the system for identifying and producing classification of hypertension conditions [19].

Further research uses the K-Nearest Neighbor (KNN) algorithm, Principal Component Analysis (PCA), and tDistributed Stochastic Neighbor Embedding (t-SNE) for hypertension conditions classification. The classification system produced consists of Normal, Hypertension, Stage 1 Hypertension, and Stage 2 Hypertension. The secondary data used consists of 7,794 samples taken from Labuang Baji Regional Hospital, Makassar. The attributes used by the system include age, weight, and systolic and diastolic blood pressure. The accuracy test results resulted in using the KNN method being 99%, combination of KNN and PCA producing 100% accuracy, and combination accuracy of KNN with t-SNE was 99%. From the accuracy test results, it can be concluded that KNN, especially when combined with t-SNE, can produce a classification of non-linear data structures very accurately and effectively [20].

In previous research that have been explained, there has been no research that has classified hypertension using SVM method. In this research, the SVM method was chosen to classify the diagnosis of hypertension. The research purpose is to determine the hypertension classification results disease diagnosis and the classification accuracy level from hypertension disease diagnosis using the SVM method results. This research will utilize the Jupyter Notebook tools and the Python programming language as research tools. This research purpose is to determine the classification results and accuracy of the diagnosis of hypertension using the SVM method. Eleven attributes used include age, smoking habits, physical activity, sugar consumption, salt consumption, fat consumption, alcohol consumption, lack of fruit and vegetable consumption, systolic and diastolic blood pressure.

II.METHOD

The type of research used is quantitative research. This research uses quantitative data to test 1373 secondary data, which is hypertension data at the Jaten 2 Karanganyar Health Center. Table 1 presents the attributes used in this research.

TABLE 1. Research Attributes

Attribute	Explanation
Age	Patient age
Gender	male/female
Less physical activity (min. 30 minutes/day)	yes/no
Excess sugar consumption (>4 tablespoons/day)	yes/no
Excess salt consumption (>1 teaspoon/day)	yes/no
Excess fat consumption (>4 pieces of fried food/day)	yes/no
Eat less fruit and vegetables (<5 portions/day)	yes/no
Alcohol consumption	yes/no
Systolic blood pressure	Blood pressure at pumps

Diastolic blood pressure	Blood pressure at rest

The research stages in the form of a flowchart are presented in Figure 1.

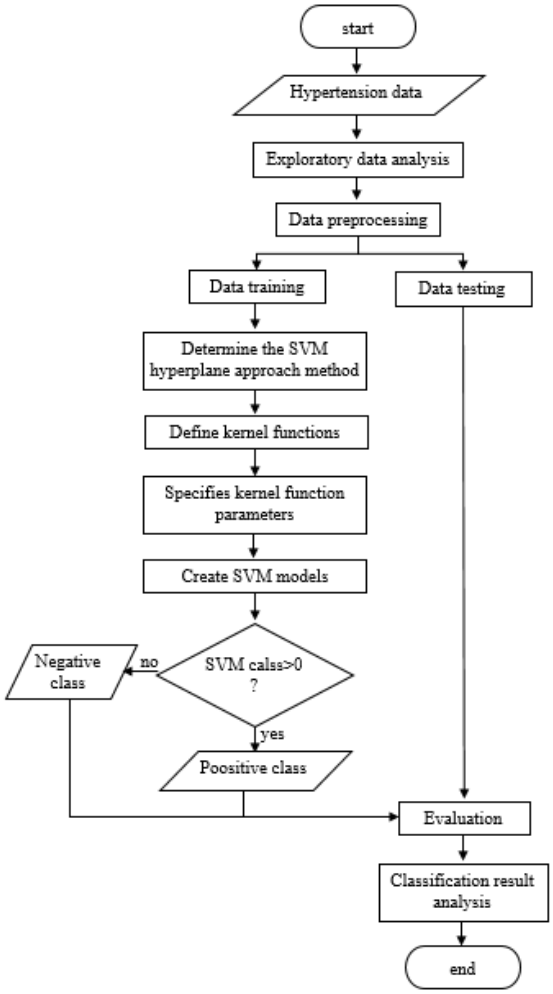

FIGURE 1. Research Flowchart

The stages undertaken in this research include:

1. Diagnosis Classification of Hypertension Using SVM
   a. Collecting hypertension data consisting of dependent and independent variables.
   b. Carrying out Exploratory Data Analysis (EDA) to find out existing information and insights.
   c. Conducting data preprocessing: Data cleaning, Data Transformation, and Data Oversampling.
   d. Dividing the data into two parts, namely training data and testing data with a ratio of 80:20.
   e. Conducting the classification process using the SVM method: choosing a method to determine the SVM hyperplane, choosing a kernel function, determining parameter values, and forming an SVM model using a kernel function.

   To find kernel Gaussian Radial Basis Function (RBF) it can be calculated using the equation:

$$K(x_i, x_j) = \exp\left(-y\|x_i - x_j\|^2\right) \quad (1)$$

[21]

General equation for the non-linear support vector case is:

$$L(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} y_i y_j \alpha_i \alpha_j \emptyset(x_i).\emptyset(x_j) \quad (2)$$

To get the weighting vector value w in Hillbert Space, use equation:

$$w_i = \sum_{i=1}^{N} \alpha_i y_i \emptyset(x_i) \quad (3)$$

[22]

for the bias parameter b can be calculated using equation:

$$b = y_i - \sum_{i,j=1}^{N} \alpha_i y_i K(x_i, x_j) \quad (4)$$

By substituting equation 4, general equation for calculating the hyperplane is obtained:

$$f(\emptyset(x)) = sign\left(\sum_{i,j=1}^{N} \alpha_i y_j K(x_i).(x_j) + b\right) \quad (5)$$

[23]

2. Calculation of Accuracy Value of SVM Method in Classifying Potential Hypertension Disease
   a. Evaluation the SVM method classification model accuracy using k-fold cross validation. K-Fold Cross Validation works by dividing all data into k parts randomly.
   b. Analyzing the SVM method classification model performance to determine the hypertension diagnosis classification prediction results using confusion matrix. There are 4 equations in testing using the confusion matrix:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} x100\% \quad (6)$$

$$Precision = \frac{TP}{TP + FP} x100\% \quad (7)$$

$$Recall = \frac{TP}{TP + FN} x100\% \quad (8)$$

$$FI\ Score = 2\,X \frac{Precision\ x\ Recall}{Precision + Recall} x100\% \quad (9)$$

TP = True Positive, TN = True Negative
FP = False Positive, FN = False Negative

[24]

The attributes that have numeric data types are age, systolic and diastolic blood pressure. The following are the numeric attributes descriptive statistics in table 2.

TABLE 2. Descriptive Statistics On Numeric Attributes

Attribute	Average	Standard Deviation	Min	Max
Age	46.27	14.36	5	98
Systolic	131.88	21.94	79	238
Diastolic	81.19	12.11	12	161

III.RESULT AND DISCUSSION

1. Exploratory Data Analysis (EDA)

Attributes that have categorical data types are smoking habits, lack of physical activity, excessive sugar consumption, excessive salt consumption, excessive fat consumption, lack of fruit and vegetable consumption, and alcohol consumption. Attributes with categorical data types are subjected to bivariate analysis using bar charts to determine the hypertension patients frequency with dependent variables. The relationship between attributes that have categorical data types and hypertension attributes frequency is presented in Figure 2.

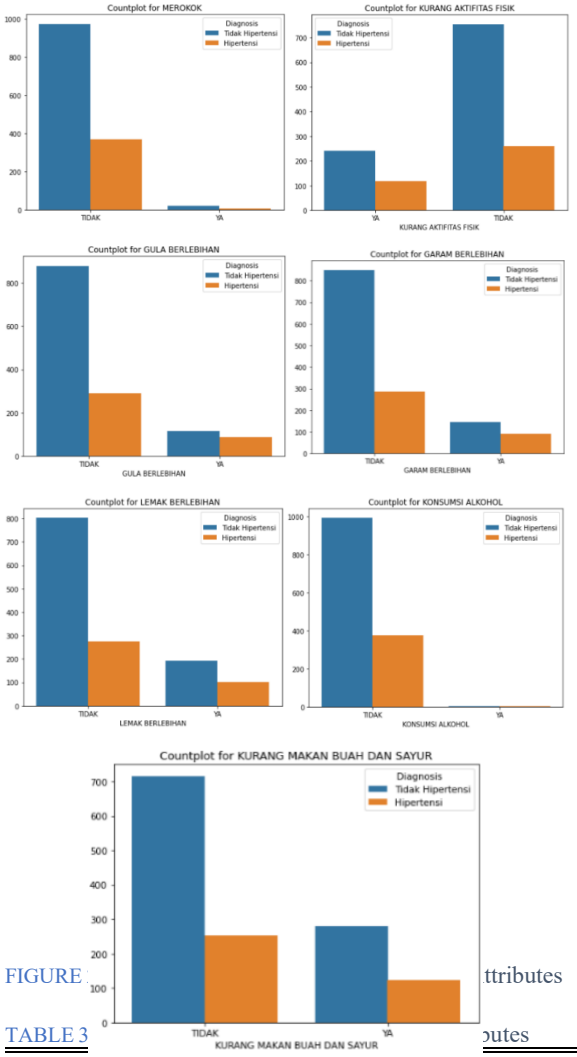

FIGURE ttributes

Based on Figure 2, it can be interpreted as follows: people who don't smoke have a lower risk of suffering from hypertension, people who are physically active often have a lower risk of suffering from hypertension, people who do not have a history of excess sugar are less likely to suffer from hypertension, people who don't consume a lot of salt have a lower risk of suffering from hypertension, people who don't consume much fat have a lower risk of suffering from hypertension, people who eat less fruit and vegetables have a higher risk of suffering from hypertension, and people who do not consume alcohol have a very low risk of developing hypertension.

2. Data Preprocessing

a. Data Transformation

To change the data in the dataset according to the format that can be processed by the software used. At this stage, there is a process of transforming categorical attribute data into numeric. The answer "yes" is converted to number "1" and the answer "no" is converted to number "0".

After all attributes are converted into numeric, then the correlation between variables can be searched using Pearson correlation. The correlation result will be a value ranging from -1 to +1, with the following interpretation:

1) If r = -1 then the correlation is perfect negative (the strongest opposite direction relationship).
2) If r = 1 then the correlation is perfect positive (the strongest directional relationship).
3) If the correlation coefficient shows 0, then the two variables most likely do not influence each other.
4) If the value of r is between -1 and 0 or 0 and +1 then the relationship is linear with varying strength (the closer to -1 or +1, the stronger relationship).

The pearson correlation results between attributes used in this study are presented in Table 3. A1: smoking habits, A2: lack of physical activity, A3: excessive sugar consumption, A4: excessive salt consumption, A5: excessive fat consumption, A6: lack of fruit and vegetable consumption, A7: alcohol consumption, A8: systolic blood pressure, A9: diastolic blood pressure, A10: Diagnosis.

TABLE 3 utes

Correlation	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
A1	1	-0.02399	0.014997	0.004989	0.027817	0.057886	0.164968	0.025479	-0.01798	-0.01597
A2	-0.02399	1	0.045112	0.153985	0.082979	0.251102	0.018998	0.091998	0.06198	0.046001
A3	0.014997	0.045112	1	0.203998	0.179897	0.155978	-0.02499	0.209989	0.08597	0.140121
A4	0.004989	0.153985	0.203998	1	0.189889	0.170102	0.035979	0.210102	0.11142	0.120103
A5	0.027817	0.082979	0.179897	0.189889	1	0.141013	0.026988	0.139968	0.07501	0.063010
A6	0.057886	0.251102	0.155978	0.170102	0.141013	1	0.013989	0.082898	0.02502	0.059011
A7	0.164968	0.018998	-0.02499	0.035979	0.026988	0.013989	1	0.033341	0.01675	0.022012
A8	0.025479	0.091998	0.209989	0.210102	0.139968	0.082898	0.033341	1	0.53176	0.546989

A9	-0.01798	0.06198	0.08597	0.11142	0.07501	0.02502	0.01675	0.53176	1	0.14110
A10	-0.01597	0.06601	0.136998	0.11201	0.081997	0.041979	0.016989	0.566001	0.670111	1

Table 3 shows that the strongest positive correlation value is the diastolic blood pressure attribute with a diagnosis of hypertension of 0.670111 and the strongest negative correlation value is the relationship between the excessive sugar attribute and alcohol consumption of -0.02499.

b. Oversampling Data

Equalizes data samples of both values by increasing the minority data to match the majority value. SMOTE selects samples from the minority class in the dataset. Then, for each selected minority sample, SMOTE finds its nearest neighbors in the feature space, usually using a method such as K-Nearest Neighbors (KNN). SMOTE creates new synthetic samples between the minority sample and its nearest neighbors. These synthetic samples are on the line connecting the two samples in the feature space. Finally, these new synthetic samples are added to the dataset, increasing the number of samples in the minority class until it reaches a desired level or balances with the majority class. The label depiction of the data class before and after the oversampling process is presented in Figure 3.

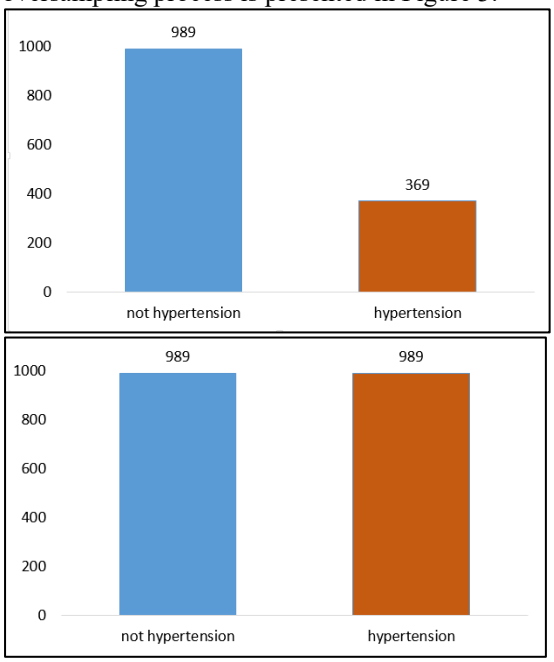

FIGURE 3. Data Before and After Oversampling Display

From Figure 3, it can be seen that before data oversampling, there were 989 non-hypertension data and 369 hypertension data. After the data oversampling process, both classes had the same amount of data, namely 989 data for each class.

c. Classification using SVM Method

The classification process begins by dividing the data into two, namely training data and testing with a ratio of 80:20. This ratio was chosen based on the results of trials that have been carried out, where it was proven that models trained with this ratio achieved a high level of accuracy. The training data consisting of several variables and target classes is used as sample data to form the SVM model. Meanwhile, testing data is used to evaluate the accuracy of the classification model based on the confusion matrix formed.

In manual calculations, SVM classification is carried out using the RBF kernel defined in Equation (1) with parameter γ (gamma) = 1. In the manual calculation example, 4 data are used which are presented in Table 4.

TABLE 4. Sample Data For Manual Calculation

X_1	X_2	Y
1	0	0
0	1	1
0	0	1
0	1	0

The first step in calculating SVM classification is calculating the kernel K matrix value. Kernel K can be calculated with dimensions m x m, where m is the amount of data.

$$K(x_1, x_1) = \exp\left(-1\left(\sqrt{(1-1)^2 + (0-0)^2}\right)^2\right)$$
$$\exp(0) = 1$$
$$K(x_1, x_2) = \exp\left(-1\left(\sqrt{(1-0)^2 + (0-1)^2}\right)^2\right)$$
$$\exp(-2) = 0.1353$$
$$K(x_1, x_3) = \exp\left(-1\left(\sqrt{(1-0)^2 + (0-0)^2}\right)^2\right)$$
$$\exp(-3) = 0.3679$$
$$K(x_1, x_4) = \exp\left(-1\left(\sqrt{(1-0)^2 + (0-1)^2}\right)^2\right)$$
$$\exp(-2) = 0.1353$$

Those calculation is carried out to the values x_4,x_4 so that it will produce a kernel matrix K:

$$K = \begin{bmatrix} 1 & 0.1353 & 0.3679 & 0.1353 \\ 0.1353 & 1 & 0.3679 & 1 \\ 0.3679 & 0.3679 & 1 & 0.3679 \\ 0.1353 & 1 & 0.3679 & 1 \end{bmatrix}$$

By using kernel K as a substitute for dot-product $\phi(x_i).\phi(x_j)$ in the lagrange multiplier duality equation in Equation (2):

$$L(\alpha) = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}(\alpha_1\alpha_2 + {\alpha_2}^2 + 0.3679\alpha_2\alpha_3 + {\alpha_3}^2)$$

with conditions $-\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 = 0$ and $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \geq 0$

In the objective function, the second term has been multiplied by $y_i y_j$. This equation meets Quadratic Programming standards so that it can be solved using a commercial solver for Quadratic Programming (QP). By using additional tools, you can get the following results: $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$. This results show that all data in this example are support vectors, because $a \neq 0$. Then, the w (weight) value can be searched using Equation (3) to obtain the results shown in Table 5.

TABLE 5. Weight Value

w_i	Weight
w_1	-1.639
w_2	2.49
w_3	2.08
w_4	-2.49

To get the b (bias) value, a calculation is carried out using Equation (4). In this research, the value of b = -0.115 was obtained. After getting the values of w (weight) and b (bias) then an SVM model can be formed which is used in the classification process using Equation (6) as follows:

$$f(\emptyset(x)) = sign\left(\sum_{i,j=1}^{N} \alpha_i y_j K(x_i).(x_j) - 0.115\right)$$

To classify the first data in Table 4, calculations can be carried out as follows:

f(ϕ(1)) = sign((1 x -1 x 1 - 0.115) + (1 x -1 x 1.353 - 0.115) + (1 x -1 x 0.3679 – 0.115) + (1 x -1 x 0.1353 – 0.115))

f(ϕ(1)) = sign(-2.11) = -1

Based on the calculation above, the first data is classified as a negative class. The calculation process above will be applied to all training data to obtain the optimal SVM model in hypertension classification. The kernel function used for SVM classification is the RBF kernel with parameters C = 0.01, 0.1, 1, 10, and 100 and parameters γ (gamma) = 0.01, 0.1, 1, 10, and 100. The experiment was carried out once using 25 RBF kernel parameters. After the SVM model is formed using training data, model accuracy values are obtained using testing data on RBF kernel parameters as follows in Table 6.

TABLE 6. RBF SVM Kernel Model Parameter Accuracy

Parameter	Accuracy			
	y = 0.01	y = 0.1	y = 1	y = 10
C = 0.01	75.75%	90.25%	70.25%	50.25%
C = 0.1	91.00%	92.75%	90.75%	70.75%
C = 1	93.75%	94.75%	96.25%	90.00%
C = 10	95.00%	96.25%	97.00%	91.25%
C = 100	95.75%	97.75%	97.25%	91.25%

From Table 5 and Table 6, the highest accuracy is obtained in the RBF kernel with parameters C = 100 and γ ($gamma$) = 0.1 with an accuracy value of 97.75%. From this model, the number of support vectors is 133 points.

The formed SVM classification model can be visualized in the form of a plot. The plot is formed by a combination of two independent variables. From the combination of two independent variables, the overall plots are obtained as follows in Figure 4.

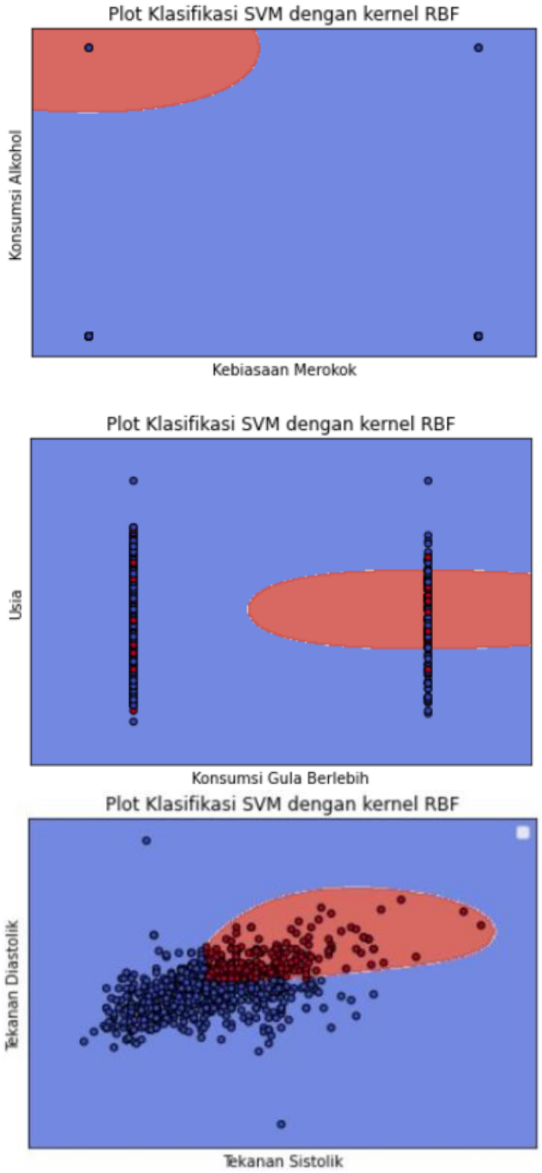

FIGURE 4. SVM Kernel RBF Classification Plot with $\boldsymbol{\gamma}$ = **0.1** and $\boldsymbol{C}$ = 100

d. Evaluation

Evaluation of the classification model results accuracy can be determined using K-Fold Cross Validation with k of 5 fold. Each piece of data will be used as test data once, while the other data will be used as training data. Thus, the classification model will be trained 5 times and tested 5 times. Following are the accuracy, precision and recall results obtained using 5-Fold Cross Validation in Table 7.

TABLE 7. Evaluation Result Using 5-Fold Cross Validation

k paramater	Precision	Recall	Accuracy
k = 1	0.960101	0.938997	0.969988
k = 2	0.971102	0.960110	0.899878
k = 3	0.969898	0.945987	0.977496
k = 4	1.000000	0.936574	0.944997
k = 5	1.000000	0.964988	0.964989
Average	0.980220	0.949331	0.951470

Based on Table 4.8, the average precision is 0.980220, the average recall is 0.949331, and the average accuracy is 0.951470. These results show that the SVM model with the RBF kernel has good performance in classifying data.

e. Classification Model Performance Analysis

Accuracy calculation of the hypertension classification disease diagnosis using the SVM kernel RBF method with parameters C = 100 and γ = 0.1 can be analyzed using the confusion matrix in Table 8.

TABLE 8. Confusion Matrix Research Results

Data		Actual	
		Non hypertension	Hypertension
prediction	Non hypertension	195	8
	Hypertension	2	203

Based on Table 8 of the confusion matrix, it shows that 195 non-hypertensive patients were correctly predicted as non-hypertensive, 8 hypertensive patients were incorrectly predicted as non-hypertensive, 203 hypertensive patients were correctly predicted as hypertensive, and 2 non-hypertensive patient was predicted incorrectly as hypertensive. Then the accuracy, precision, recall and F1-Score values are calculated as follows:

$$Accuracy = \frac{195 + 203}{195 + 203 + 8 + 2} x100\% = 97.5\%$$

$$Precision = \frac{195}{195 + 8} x100\% = 96\%$$

$$Recall = \frac{195}{195 + 2} x100\% = 98.9\%$$

$$FI\ Score = 2\ X \frac{96\%\ x\ 98.9\%}{96\% +\ 98.9\%} x100\% = 97.4\%$$

Based on the calculation results, it can be concluded that the classification of hypertension diagnosis using SVM has an accuracy of 97.5%. This proves that the SVM method is able to classify hypertension diagnosis very well. The precision was 96%, which means that of the 201 non-hypertensive patients predicted by the SVM model, 195 of them were truly non-hypertensive. The recall value was 98.9%, which means that of the 197 non-hypertensive patients, 195 patients were identified correctly by the SVM model. The F1-Score value obtained was 97.4%, which means that the SVM model is able to balance precision and recall.

IV.CONCLUSION

Based on the results of the research that has been conducted, it is proven that the SVM method can classify hypertension with an accuracy level of 97.5%. The SVM classification results method in hypertension classification diagnosis show that the best model uses the RBF kernel. Parameter values used are C = 100 and γ (gamma) = 0.1. The classification results analysis using confusion matrix showed that 195 non-hypertensive patients were correctly predicted as non-hypertensive, 8 hypertensive patients were incorrectly predicted as non-hypertensive, 203 hypertensive patients were correctly predicted as hypertensive, and 2 hypertensive patient was incorrectly predicted as non-hypertensive. The accuracy level obtained from hypertension classification diagnosis using SVM kernel RBF method with parameters C = 100 and γ (gamma) = 0.1 was 97.5%. This shows that the SVM method is able to classify hypertension diagnoses very well. Based on the results of this study, the SVM method can be an alternative classification method that can be used to classify diagnoses of hypertension.

ACKNOWLEDGMENT

We would like thank to Bina Sarana Informatika University especially the LPPM team, Dharma AUB University, and Duta Bangsa University for support in this research.

REFERENCE

[1] S. Ratwatte and D. S. Celermajer, “The latest definition and classification of pulmonary hypertension,” *Int. J. Cardiol. Congenit. Hear. Dis.*, vol. 17, no. July, p. 100534, 2024, doi: 10.1016/j.ijcchd.2024.100534.

[2] N. Benaired, M. H. Meghraoui, and Z. A. Benselama, “Traitement du Signal Hypertension Management via Photoplethysmography : An Ensemble Learning-Based Approach for Classification of Blood Pressure Using Fourier Synchrosqueezed Transform,” *Trait. du Signal*, vol. 41, no. 5, pp. 2263–2278, 2024, doi: https://doi.org/10.18280/ts.410504.

[3] E. M. S. Rochman *et al.*, “Classification of hypertension disease using Artificial Neural Network (ANN) backpropagation method case study in mitigating health risk: UPT Modopuro Mojokerto Health Center,” *BIO Web Conf.*, vol. 146, pp. 1–8, 2024, doi: 10.1051/bioconf/202414601083.

[4] J. Li *et al.*, “Analysis of Related Factors Influencing Hypertension Classification among Centenarians in Hainan, China,” *Rev. Cardiovasc. Med.*, vol. 25, no. 7, pp. 1–10, 2024, doi: 10.31083/j.rcm2507235.

[5] C. Xu *et al.*, “Etiological Diagnosis and Personalized Therapy for Hypertension: A Hypothesis of the REASOH Classification,” *J. Pers. Med.*, vol. 13, no. 2, 2023, doi: 10.3390/jpm13020261.

[6] F. J. Charchar *et al.*, “Lifestyle management of hypertension: International Society of Hypertension position paper endorsed by the World Hypertension League and European

Society of Hypertension," *J. Hypertens.*, vol. 42, no. 1, pp. 23–49, 2024, doi: 10.1097/HJH.0000000000003563.
[7] Tri Sutanti Puji Hartati and Emyr Reisha Isaura, "Three Body Mass Index Classification Comparison In Predicting Hypertension Among Middle-Aged Indonesians," *Media Gizi Indones.*, vol. 18, no. 1, pp. 38–48, 2023, doi: 10.20473/mgi.v18i1.38-48.
[8] R. Kreutz *et al.*, "2024 European Society of Hypertension clinical practice guidelines for the management of arterial hypertension," *Eur. J. Intern. Med.*, vol. 126, no. May, pp. 1–15, 2024, doi: 10.1016/j.ejim.2024.05.033.
[9] M. Dohan, R. B. Mohammed, W. H. Gwad, M. Khalaf, and K. M. Z. Othman, "Predicting Vehicle Driver Preference from the Analysis of In-Vehicle Coupon Recommendation Data," *J. Soft Comput. Data Min.*, vol. 5, no. 2, pp. 274–282, 2024, doi: 10.30880/jscdm.2024.05.02.020.
[10] F. Fang, "A Study on the Application of Data Mining Techniques in the Management of Sustainable Education for Employment," *Data Sci. J.*, vol. 22, no. 1, pp. 1–13, 2023, doi: 10.5334/dsj-2023-023.
[11] A. Amriana, A. A. Ilham, A. Achmad, and Y. Yusran, "Optimization of Herbal Plant Classification Using Hybrid Method Particle Swarm Optimization with Support Vector Machine," *Int. J. Informatics Vis.*, vol. 9, no. 1, pp. 396–406, 2025, doi: 10.62527/joiv.9.1.2576.
[12] A. Biswas and M. S. Islam, "A Hybrid Deep CNN-SVM Approach for Brain Tumor Classification," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 9, no. 1, pp. 1–15, 2023, doi: 10.20473/jisebi.9.1.1-15.
[13] Jumanto *et al.*, "Optimizing Support Vector Machine Performance for Parkinson's Disease Diagnosis Using GridSearchCV and PCA-Based Feature Extraction," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 10, no. 1, pp. 38–50, 2024, doi: 10.20473/jisebi.10.1.38-50.
[14] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, "An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review," *Inf.*, vol. 15, no. 4, pp. 2–36, 2024, doi: 10.3390/info15040235.
[15] H. Nurrani, Andi Kurniawan Nugroho, and Sri Heranurweni, "Image Classification of Vegetable Quality using Support Vector Machine based on Convolutional Neural Network," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 1, pp. 168–178, 2023, doi: 10.29207/resti.v7i1.4715.
[16] R. Kurniawan *et al.*, "Hypertension prediction using machine learning algorithm among Indonesian adults," *IAES Int. J. Artif. Intell.*, vol. 12, no. 2, pp. 776–784, 2023, doi: 10.11591/ijai.v12.i2.pp776-784.
[17] M. Z. Ardiansyah and E. Widowati, "Hubungan Kebisingan dan Karakteristik Individu dengan Kejadian Hipertensi pada Pekerja Rigid Packaging," *HIGEIA (Journal Public Heal. Res. Dev.*, vol. 8, no. 1, pp. 141–151, 2024, doi: 10.15294/higeia.v8i1.75362.
[18] Y. Munali and Armansyah, "Classification of Hypertension Using Naïve Bayes Method with Data Discretization Approach Risk Factors," *J. Sist. Cerdas*, vol. 7, no. 1, pp. 1–12, 2024, doi: 10.37396/jsc.v7i1.381.
[19] S. S. Nurhadiva, "Blood Pressure and Heart Rate Measurement for Hypertension Classification Using the K-Nearest Neighbors Method Based on IoT," *Piksel*, vol. 12, no. 225, pp. 373–382, 2024, doi: 10.33558/piksel.v12i2.9824.
[20] A. A. C. Resky, J. C. Lapendy, A. A. N. Risal, D. F. Surianto, and A. Wahid, "PCA and t-SNE Implementation for KNN Hypertension Classification," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 9, no. 1, pp. 175–184, 2025, doi: https://doi.org/10.29207/resti.v9i1.6208.
[21] Z. Jun, "The Development and Application of Support Vector Machine," *J. Phys. Conf. Ser.*, vol. 1748, no. 5, pp. 1–6, 2021, doi: 10.1088/1742-6596/1748/5/052006.
[22] F. O. Awalullaili, D. Ispriyanti, and T. Widiharih, "Klasifikasi Penyakit Hipertensi Menggunakan Metode Svm Grid Search Dan Svm Genetic Algorithm (GA)," *J. Gaussian*, vol. 11, no. 4, pp. 488–498, 2023, doi: 10.14710/j.gauss.11.4.488-498.
[23] A. K. Darmawan, M. W. Al Wajieh, M. B. Setyawan, T. Yandi, and H. Hoiriyah, "Hoax News Analysis for the Indonesian National Capital Relocation Public Policy with the Support Vector Machine and Random Forest Algorithms," *J. Inf. Syst. Informatics*, vol. 5, no. 1, pp. 150–173, 2023, doi: 10.51519/journalisi.v5i1.438.
[24] E. Rahmawati, G. S. Nurohim, C. Agustina, D. Irawan, and Z. Muttaqin, "Development of Machine Learning Model to Predict Hotel Room Reservation Cancellations," *J. Teknol. Inf. Dan Terap.*, vol. 11, no. 2, pp. 58–65, 2024, doi: https://doi.org/10/25047/jtit.v11i2.5440.

PUDJI WIDODO. Born in Temanggung, July 22, 1973, graduated with a Bachelor's degree in Information Systems and a Master's degree in Computer Science from STMIK Nusa Mandiri. Currently working as a Lecturer at Bina Sarana Informatika University, Banyumas Campus. Active in various activities such as training, research, community service, seminars and writing computer books. Scientific articles published in journal focus on the development of information systems, management information systems, computer networks and machine learning.

HERIBERTUS ARY SETYADI holds a Master of Information Systems degree from Diponegoro University, Semarang, Indonesia. Received Bachelor of Engineering degree from the Akprind Yogyakarta Institute of Science and Technology, Indonesia. Joined as a lecturer at Bina Sarana Informatika University, Surakarta Campus until now. Practitioners in developing information systems as systems analysis. Some published scientific research concentrates on decision support systems, information systems development, expert systems and machine learning.

HARTATI DYAH WAHYUNINGSIH
Since 2002 until now, he has been a lecturer at Dharma University , Computer Science Faculty, Information Systems study program. Alumni of Pancasila Money & Bank Academy Surakarta, , continuing his Bachelor's degree at Slamet Riyadi University Surakarta, and Magister at Muhammadiyah University Surakarta. Research focuses on management information systems, management, and entrepreneurship.

SUNDARI.
Alumni of the Undergraduate and Postgraduate Management study program at Dharma University AUB Surakarta. Currently holds the status of a permanent lecturer at the Informatics Engineering Study Program, Faculty of Computer Science, Duta Bangsa University Surakarta since August 2014. Research focuses on management information systems, management, and entrepreneurship.