Classification of Red Wine Beverage Quality Using a Support Vector Machine Algorithm Based on Forward Selection as Feature Selection

Muhammad Azhar Luthfi[1], Ajib Susanto[2]
[1,2]Study Program in Informatics Engineering, University of Dian Nuswantoro, Semarang, Indonesia

Article Info:

Keywords:

Classification
Feature selection
Red wine baverage
Support Vector Machine

ABSTRACT

Wine is an alcoholic drink produced from the fermentation process of grapes or can use other types of fruit. Currently, the provision of wine quality certification uses physical tests and chemical tests. In this research, a machine learning technology approach was used to classify the quality of wine drinks. Support Vector Machine (SVM) is a supervised learning type method. Support Vector Machine (SVM) is an algorithm used for the prediction process, both in classification cases. Before the SVM model creation process is carried out, an attribute selection process will be carried out using forward selection techniques to see what attributes influence the formation of wine quality. After that, a model was created using the Support Vector Machine method using the RBF kernel function to produce classification results. There are two processes for creating a machine learning model, namely the first model is an SVM model without parameter optimization and the second model is an SVM model with parameter optimization, and the process of creating both models both use the SVM kernel function, namely the RBF kernel function. In the SVM model without carrying out parameter optimization, we obtained an optimum level of accuracy of 0.78 or 78%, while for the results of the second model, namely the SVM model with parameter optimization, we obtained an optimum level of accuracy of 0.85 or 85% with parameters C = 100 and gamma value = 0.5

Author Correspondence:

Ajib Susanto,
Study Program in Informatics Engineering
Faculty of Computer Science
University of Dian Nuswantoro, Indonesia, 50131
Email: ajib.susanto@dsn.dinus.ac.id

1. INTRODUCTION

Machine learning or Machine Learning is one of the branches of artificial intelligence or (AI) which is related to the development of techniques that can be used as programs and learn from past data. So that from the data that has been collected a pattern can be obtained. In this field it will intersect with the science of probability and statistics and even optimization. Support Vector Machine (SVM) is a method used in Supervised Learning type research [1], [2]. Support Vector Machine (SVM) is a technique for carrying out prediction processes, both in classification and regression cases. Support Vector Machine has the basic principles of linear and non-linear classifiers by including the kernel concept.

Wine is an alcoholic drink made from the fermentation process of grapes or other types of fruit. Currently, wine is still a luxurious and classy drink. To make the wine industry grow even more in the countries that enjoy it. Enables the use of machine learning technology to improve manufacturing and sales processes. Providing certification or assessing the quality of drinks is the core of this research. Certification aims to prevent illegal wine counterfeiting. Wine certification is generally carried out using physical and chemical tests [3], [4]. Laboratory tests used to label the quality of wine drinks are based on variables such as density, alcohol level and Ph, while sensory tests rely on humans. It should be emphasized that taste is most difficult for the human senses to understand and its accuracy may still not be complete. Thus, wine classification is a difficult

task. Where, the relationship between physicochemical and sensory analysis is very complex and still not fully understood.

In this study, researchers used a dataset assessing the quality of wine from Portugal "Vinho Verde" in 2009, the contents of the dataset were the results of laboratory tests with 11 independent variables and produced output in the form of wine quality. Researchers will use machine learning to produce a model that will be used to predict and classify the quality of red wine drinks.

2. METHOD

2.1. Machine Learning

Machine learning or in Indonesian called machine learning is part of AI (Artificial Intelligence) which gives computers the task of studying previous data. Machine Learning can be interpreted as a computer that has the ability to learn from the experiences received on tasks that have been completed to improve performance. Machine learning is defined as a distinctive artificial intelligence, which gives computers the ability to learn data without having to follow explicitly programmed instructions [5]. The way the machine learning process works is by following the way humans learn, namely learning patterns from an example being analyzed. It is true that not everything can be solved with machine learning, but in machine learning there are complex algorithm systems that can be used in everyday life. Examples include: Person Face Detection Tool, Disease Diagnostic Tool, Weather Detector, Film Recommendations, and so on.

The first process that is carried out when you want to create a machine learning system is to collect a dataset. After the dataset is obtained, the next stage is to choose what method you want to use. In machine learning, there are many methods that can be used, such as linear regression, logistic regression, support vector machines, artificial neural networks and others. After selecting the method, the next step is processing the data so that it can be processed using the selected method. Like cleaning data, checking whether the data contains empty or null values, making it easier to train on test data or training datasets. Machine learning has two general types of techniques, namely Supervised Learning and Unsupervised Learning. The supervised learning process is divided into two parts, namely classification and regression. Here is a detailed explanation of the types and types of machine learning [6]:

1. Directed learning (Supervised Learning). Directed Learning (Supervised Learning) where the classes in each data are clear or the information is known or owned. So that the processing of this data is directed or clear. Therefore, this method requires an input mode and an output mode to identify information. Several algorithms applied for supervised learning techniques include Naïve Bayes, Backpropagation, Support Vector Machine and others.
2. Undirected Learning (Unsupervised Learning). Unsupervised Learning is a learning method whose classes are not yet known. The concept of unsupervised learning is very different from the supervised learning method. The main goal of this method is to identify objects that can be combined into a single unit based on groups that have similar values. This method is very suitable for researching or looking for classification patterns where many objects do not exactly match each object. Some examples of methods applied for undirected learning techniques are K-means, KNN and others.

2.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a technology used for classification and prediction as well as regression. Support Vector Machine was originally developed by Boser, Guyon, and Vapnik. Support Vector Machine was first introduced to the public at the annual computational learning theory seminar held in 1992. The basic concept of Support Vector Machine is a combination of existing computational theories. previously, such as the margin hyperplane introduced by someone named Duda and Hart in 1973 and the kernel by Aronszajn in 1950. However, there was no attempt to assemble or combine these components until 1992 [7].

This method is different from the strategy used by neural networks, where artificial neural networks try to find separate hyperlanes in each class, whereas the Support Vector Machine (SVM) method uses the method of finding the best hyperplane in the input space [8]. In Support Vector Machine the basic principle is a linear classifier which has been further developed, so that it can process non-linear problems by incorporating kernel concepts in a high-dimensional workspace.

The data used is given the notation $\overrightarrow{X_i} \in \Re^d$ and each label for each data is given the notation $yi \in \{-1, +1\}$ for $i = 1,2, \dots, n$. Where n is the amount of data. It can be assumed that the -1 and +1 categories can be completely separated from the d-dimensional hyperlane, which can be denoted:

$$\overrightarrow{W}.\vec{X} + b = 0 \quad (1)$$

For the pattern $\overrightarrow{X_i}$ which is included in category -1 or can be said to be a negative value sample, it can be expressed as a pattern that satisfies the inequality:

$$\overrightarrow{W}.\vec{X} + b \leq -1 \quad (2)$$

And for the pattern $\overrightarrow{X_i}$ or is in the +1 category or can be said to be a positive sample, the inequality is:

$$\overrightarrow{W}.\vec{X} + b \geq +1 \quad (3)$$

For the process of finding and maximizing the distance value on the hyperplane from the closest point at the largest margin, it is $\frac{1}{\|\vec{w}\|}$. In this case it can be formulated as a Quadratic Programming Problem or in Indonesian it can be interpreted as quadratic programming, the process carried out is to look for the minimum point or lowest point by paying attention to the constraint values:

$$\min_{\vec{W}} \tau(w) = \frac{1}{2} \| \overrightarrow{w} \|^2 \quad (4)$$

$$y_i\left(\overrightarrow{X_i W} + b\right) - 1 \geq 0, \forall_i \quad (5)$$

So this problem or problem can be solved using a technique called Lagrange Multiplier.

$$L(w, b, a) = \frac{1}{2} \| \vec{w} \|^2 - \sum_{i=0}^{n} a_i \left(y_i\left((x_i . w_i + b) - 1\right)\right) \quad (6)$$

$(i = 1,2,3, \ldots, n)$

a_i is Lagrange Multipliers, where a_i has a value of zero or is positive ($a_i \geq 0$). The optimal value produced by equation (6) can be calculated or processed by minimizing the L value for w $\vec{}$ and b, and maximizing the L value relative to the a_i value. Remembering that the best point for the gradient is the value L = 0, therefore equation (6) can be modified to maximize a problem that only contains the value a_i as the equation shown as follows:

Maximization:

$$\sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} a_i a_j y_i y_j \overrightarrow{X_i X_j} \quad (7)$$

With the Constraint process:

$$a_i \geq 0 (i = 1,2,3, \ldots, n) \sum_{i-1}^{n} a_i y_i = 0 \quad (8)$$

Most of the a_i obtained from the calculation results are positive. Data associated with positive values of a_i are called support vectors. In the description above, it is only assumed that the two categories or classes can be perfectly separated via hyperlane (linear separable). However, what generally happens is that the two classes or categories contained in the input space cannot be perfectly separated (non-linear separable). Therefore, this has the effect of constraints such as equation (8) which cannot be fulfilled, and results in the inability to carry out the optimization process. To overcome this problem, Support Vector Machine (SVM) introduces a solution, where the solution is a technique called softmargin. In this softmargin technique, modification is made to equation (5) by entering the slack variable ξ(ξ>0) as below:

$$y_i(X_i . w + b) \geq 1 - \xi_i \forall_i \quad (9)$$

So equation (4) is changed to:

$$\min_{\vec{W}} \tau(w, \xi) = \frac{1}{2} \| \overrightarrow{w} \|^2 + C \sum_{i=0}^{n} \xi_i \quad (10)$$

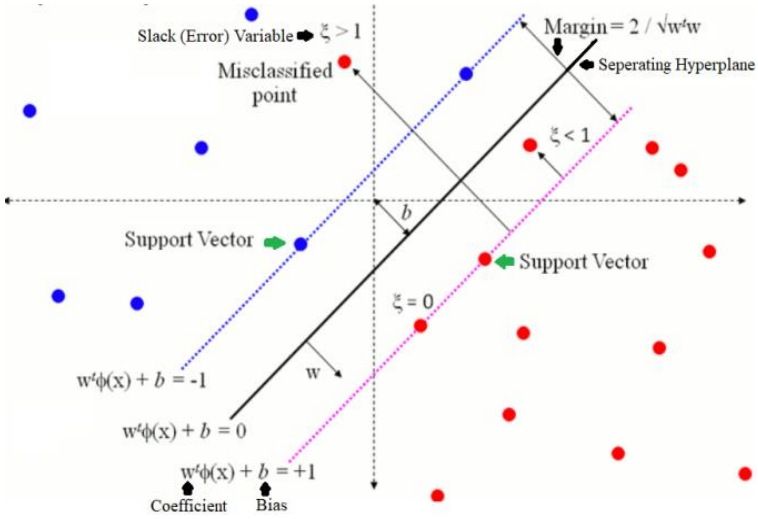

Figure 1. SVM without kernel

The parameter C is used to control the trade-off between margin and classification error ξ. So for a large value of C, it will have a greater penalty effect for misclassification. The characteristics found in Support Vector Machines are generally as follows:

- In principle, Support Vector Machine is a linear classifier.
- Pattern recognition is carried out by changing the data in the input space into a higher dimensional space (feature space) and optimization is carried out in the new vector space.
- Implement a structural risk minimization strategy.
- Basically, the working principle of SVM can only handle the classification of two categories or classes, but it has been developed for the classification of more than two classes with pattern recognition.

2.3. Confusion Matrix

Confusion Matrix is a matrix that has the function of displaying the performance of the classification algorithm. The aim is to compare the results of the classification process carried out by the model system with the actual classification.

Table 1. Confusion matrix

		True Values	
		True	*False*
Prediction	*True*	TP *Correct result*	FN *Unexpected result*
	False	FP *Missing result*	TN *Correct absence of resut*

Information :
TP: True Positive (The number of correct predictions in the positive class)
FP: False Positive (Number of incorrect predictions in the positive class)
FN: False Negative (Number of incorrect predictions in the negative class)
TN: True Negative (The number of correct predictions in the negative class)

In Table 1, the True Positive (TP) value and True Negative (TN) value show the accuracy of the classification process. The process of calculating the percentage of accuracy, precision and recall can be done using the formula below:

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP} \times 100\% \quad (11)$$

$$Precison = \frac{TP}{TP+FP} \times 100\% \quad (12)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (13)$$

Classification accuracy is a measure or range of performance accuracy values of a classification algorithm. The higher the level of accuracy value indicates that the performance of the classification is good. The levels of accuracy values are divided into several as in Table 2.

Table 2. Accuracy Score

No	*Score Accuracy*	*Classification Level*
1	0.90 - 1.00	Excellent
2	0.80 - 0.90	Good
3	0.70 - 0.80	Fair
4	0.60 - 0.70	Poor
5	0.50 - 0.60	Failure

2.4. Cross-validation

Cross-validation is a method in statistics that can be used to carry out the evaluation process of a model or algorithm, where the data will be separated, namely data for the learning process (training data) and data for validation or evaluation (testing data). The model will be trained or the data training process will be carried out by the learning subset and will be validated by the validation subset. In the next process, selecting the type of CV can be based on the size of the dataset. K-fold Cross-validation is one method of cross-validation that is popularly used. In K-Fold, there is a K value. Where the K value is the total fold you want to make. For more details, look at Figure 2.

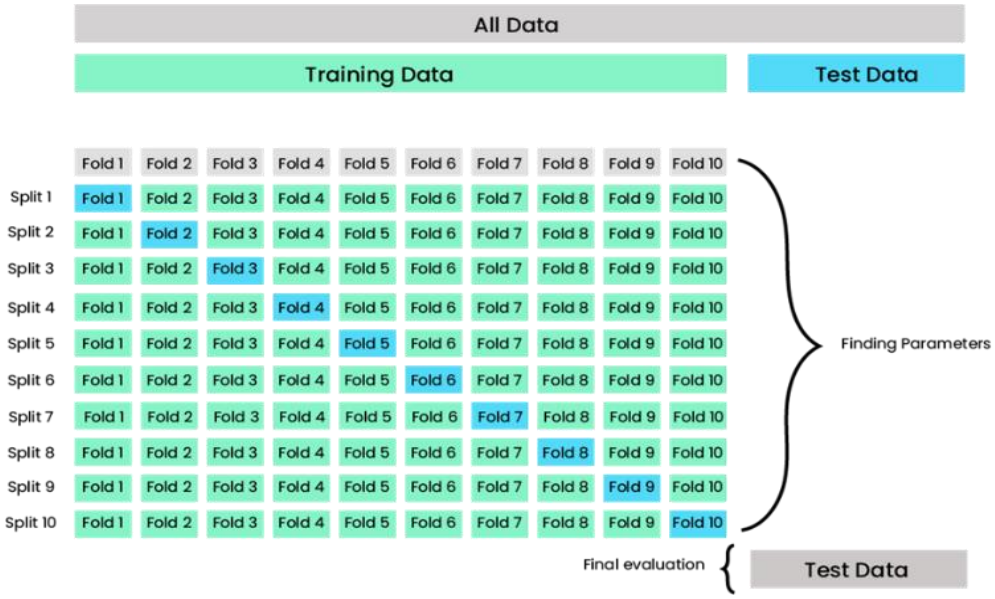

Figure 2. K-fold cross validation dengan nilai k = 10

2.5. Wine

Wine is an alcoholic drink made from the fermentation process of grapes or other types of fruit. Wine is a drink whose history can be traced back to 6000 BC, and originated in the Mesopotamia region and then spread throughout the world. Because in the composition of wine there is a natural chemical balance, the grapes in wine can ferment without the addition of sugar, acid, enzymes, water or other nutritional content [9]. In the process of making wine, a fermentation process is carried out. In the fermentation process, yeast will be consumed by the sugar content of the grapes and convert it into ethanol and carbon dioxide. Variations in wine types can occur when in the wine making process different types of grapes are used [10], [11].

There are several types of wine, namely, white wine, red wine, rose wine, sparkling wine, sweet wine. White wine is a wine made from green grapes (white grapes). Red wine is a type of wine made from red grapes [12]. Rose wine is pink or pink wine made from red grape varieties but with a shorter color extraction process compared to the process of making red wine. Sparkling wine or what is mistakenly called Champagne is the name of a region producing sparkling wine in France. This type of wine contains a lot of CO2 gas or carbon dioxide bubbles in it. Sweet wine is wine that still contains a lot of residual sugar from fermentation so that it tastes sweet.

2.6. Feature Selection

Feature selection is the most important stage during the pre-processing process. This technique reduces the number of irrelevant features. The main goal of feature selection is to select the best features from a data set. Feature selection can be classified into wrapper methods and filter methods [13]. The application of feature selection methods can usually improve the performance of the models created, especially models that have high dimensions.

Forward selection is carried out by entering predictors in stages, these predictors are based on the largest partial correlation value. In the forward selection method, the predictor variables included in the model cannot be removed again. When used, the forward selection method aims to increase or remove independent variables gradually so as to reduce the complexity of an algorithm and the level of accuracy of the algorithm. To carry out the forward selection process there are several steps including the following:

1. Create a model by regressing the Y variable with each predictor. After that, sort out the models that have a very large R2 value.
2. Regressing the response variable Y, with the predictor X_a, plus each predicator not only X_a as well as other predicators. After that, the model with the highest R2 value is selected, for example containing the additional predicator X_b, namely $\hat{Y} = b_0 + b_a X_a + b_b X_b$. The selected predicator X_b means it has a very high Fsquential. The $F_{squential}$ formula for X_b is $F_{seq} = R(\beta_b|\beta_0, \beta_a)$ / MSE / db. The $F_{squential}$ value for X_b can also be obtained by squaring the statistical value of the pre-dicator T test X_b.
3. The process is repeated until we get $F_{squential} > F_{in}$. Nilai $F_{in} = F(1, v, \alpha_{in})$, so the best model selected is a model that does not have a predictor with $F_{squential} < F_{in}$.

2.7. Django

Django is a web framework that uses the Python programming language. Django is also a full stack framework. This means that Django can be used for the frontend as well as the backend. Django makes it easier and faster to build web applications with less syntax or code as in Figure 3. One of the advantages of Django is that this framework uses Object Relationship Mapping (ORM). Object Relationship Mapping (ORM) is a method or technology from a programming language whose purpose is to convert data from an object-oriented programming (OOP) language environment to a relational database environment. ORM acts as a connector and at the same time makes it easier for us to create relational databases to complete application work more efficiently.

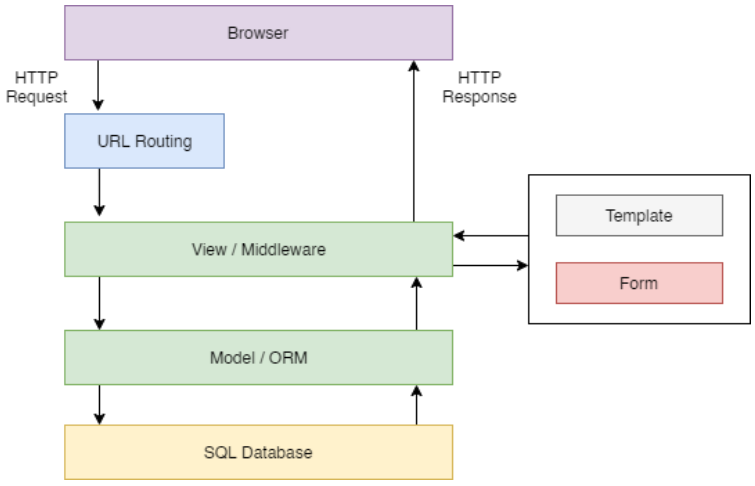

Figure 3. Django architecture framework

2.8. Data Collection

Dataset taken from research on the quality of Portuguese red wine "Vinho Verde" by Paulo Cortez, PhD from the University of Minho, then the data from the research results were published in the University of California at Irvine (UCI) Machine Learning Repository. The amount of data used was 1600 with 12 variables. For more complete information about the variables that will be used in this research process, see Table 3.

The dataset has been divided into 2 parts, namely training data and test data. This training data is data that will be used to train the model using algorithms or methods used for the machine learning process. Meanwhile, test data or testing data is used to calculate the level of accuracy of the classification model created using training data. According to the testing process it can be seen whether the model produces good

performance. There are no special rules in the percentage process for dividing training data and testing data, only that the distribution of test data or testing data cannot be greater than the test data or training data used, so that the resulting performance is good. In general, 3/4 of the total data is used as training data and the rest is used as testing data.

Table 3. List of variables

Symbol	Variables
X_1	*Fixed acidity (tartaric acid - g/dm3)*
X_2	*Volatile acidity (acetic acid - g/dm3)*
X_3	*Citric acid (g/dm3)*
X_4	*Residual sugar (g/dm3)*
X_5	*Chlorides (sodium chloride - g/dm3)*
X_6	*Free sulfur dioxide (mg/dm3)*
X_7	*Total sulfur dioxide (mg/dm3)*
X_8	*Density (g/cm3)*
X_9	*pH*
X_{10}	*Sulphates (potassium sulphate - g/dm3*
X_{11}	*Alcohol (% by volume)*
Y	*Quality*

Information in Table 3 ;

- Fixed acidity is the sour taste contained in wine.
- Volatile acidity is the amount of acetic acid in wine.
- Citric acid is citric acid which can add a fresh taste to wine.
- Residual sugar is the amount of sugar remaining after the wine fermentation process is complete.
- Chlorides are the amount of salt in wine.
- Free sulfur dioxide is the amount of SO2 that is not bound to other molecules.
- Total sulfur dioxide is a measure of free and bound SO2, excessive amounts of SO2 can inhibit the fermentation process.
- Density is the density of water which depends on the percentage of alcohol and sugar content.
- pH is used to describe how acidic a wine drink is.
- Sulphates are substances that can contribute to levels of sulfur dioxide gas.
- Alcohol is the percentage of alcohol content in wine.
- Quality is the quality of the wine.

Data balancing is a process carried out to handle unbalanced data to make it balanced. Data is said to be unbalanced if in the training data or training data on the target variable, the number of classes or categories is disproportionate and has quite large differences between classes or categories. Most types of machine learning algorithms or methods do not work well with imbalanced datasets. There are many ways that can be used to handle imbalanced data, including oversampling and undersampling. Oversampling is a data balancing technique, by taking a minority class or smaller class and then carrying out the process of adding or duplicating the minority class. In the process of adding classes, this can be done based on the closest distance from the minority class. Meanwhile, the undersampling technique is a technique for balancing data by randomly reducing or removing samples from the majority class. So that the composition of each class is balanced with the minority class.

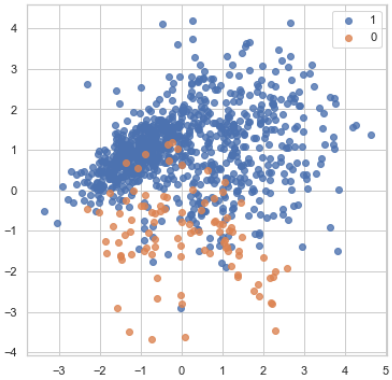

Figure 4. Example of an unbalanced dataset

As shown in Figure 4, where in the class of the determining variable in the data, the number of class percentages is not proportional. It is necessary to carry out data resampling techniques, in order to make the percentage of the minority class not too far from the majority class. For the resampling process on unbalanced data, you can use the technique described previously, namely using the Oversampling technique and the Undersampling technique.

3. RESULTS AND DISCUSSION

In this section, we will explain the implementation of SVM using the Python programming language. The Python programming language itself supports the field of artificial intelligence. There are many libraries provided to carry out data science processes, such as Sklearn, Python, TensorFlow, Keras, and so on. In this research, the Sklearn library was used. The editor code used for the machine learning model creation process is Jupyter Lab with version 3.0.5 and for the Jupyter Lab kernel version it uses the Python kernel 3.8.6. With Jupyter Lab, the machine learning model design process will be carried out by carrying out mathematical computational calculations. The application of SVM which is based on complex mathematical calculations can be implemented in the program to make it easier. The existing red wine quality data has been created for a classification model using the SVM method and using the SVM kernel function. The kernel used is the Radian Basis Function (RBF) Kernel.

The following are the results of the SVM model process using test data or red wine quality testing data on the RBF kernel function. By using test data of 400, the results obtained are as shown in Table 4. Based on table 4.7, the results of the model process created on the original red wine quality data, the resulting level of accuracy in the model has a good accuracy value using the RBF kernel function in the SVM. From the SVM model, the performance of the model can be measured using a confusion matrix.

Table 4. Confusion matrix dor SVM without kernel

Performance			
Accuracy	Precision	Recall	RMSE
0.78	0.31	0.82	0.47

Based on Figure 5, 400 test data or testing data on red wine quality data, it was found that the prediction results from the SVM model were correct for good red wine quality as many as 37 and for incorrect predictions as many as 8. Furthermore, for correct predictions the quality of red wine was bad as many as 273 and predictions wrong as many as 82. The results obtained from 400 testing data from the training model using the SVM method, the accuracy value obtained was 0.78 or 78%. The precision value obtained was 0.31 or 31%, the recall value was 0.82 or 82%. And from this model the Root Mean Squared Error (RMSE) value is 0.47.

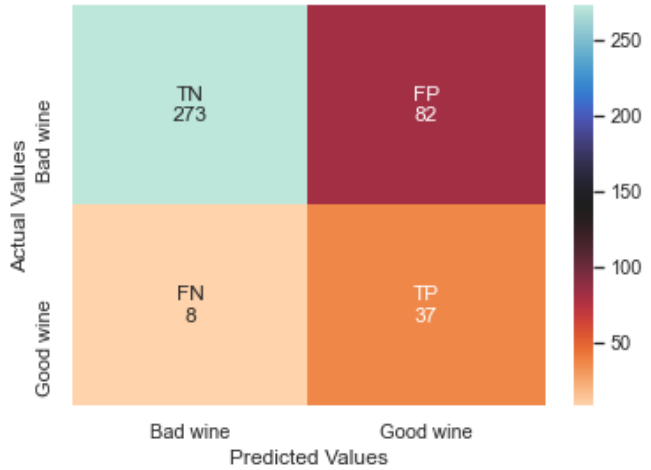

Figure 5. SVM Confusion matrix without kernel

In the parameter optimization process in the SVM model, researchers have determined the parameters that will be used in the optimization process using GridSearchCV. The SVM parameters to be determined are the C and gamma parameters. We have determined the parameters that will be used in the optimization process

using GridSearchCV. The SVM parameters to be determined are the C and gamma parameters. For the specified C parameter values, see Table 5. There are 4 parameter values, namely, C = 1, C = 10, C = 100, C = 1000. There are 5 values for the gamma parameter (γ), namely, γ=0.1, γ =0.2, γ=0.3, γ=0.4, γ=0.5. The two parameters C and gamma (γ) will be used for the optimization process of the red wine quality classification model. With GridSearchCV, the previous model optimization process will be carried out, with the aim of obtaining optimal results based on parameters that have been determined by previous researchers. From the results of the GridSearchCV process, the best parameters were obtained to be used to obtain maximum results in the SVM model that had previously been carried out.

The best parameters have been produced using GridSearchCV, to get optimal results on the previous SVM model. From the previously determined parameters C and gamma (γ), based on the results of the model optimization process with GridSearchCV. The best parameters were obtained for the value of C = 100 and the value of gamma (γ) = 0.5. The following are the results of testing the SVM model based on the best parameters resulting from the optimization process using GridSearchCV. Based on 400 test data or testing data on red wine quality data, it was found that the prediction results from the SVM optimization model were correct for good red wine quality by 35 and incorrect predictions were 10. Furthermore, correct predictions for bad red wine quality were 305 and incorrect predictions were 50. The results obtained from 400 testing data from the training model using the SVM method, the accuracy value obtained was 0.85 or 85%. The precision value obtained was 0.41 or 41%, the recall value was 0.78 or 78%. And from this model the Root Mean Squared Error (RMSE) value is 0.39.

After carrying out the process of creating a machine learning model using the SVM method. Namely, the SVM model uses the RBF kernel function without optimization and the SVM model uses the RBF kernel function with optimization. To measure the performance of the two models, an accuracy test process will be carried out using K-Fold cross validation. The accuracy performance testing process will be repeated 10 times. Table 5 shows that from the two models that have been created, the results of the accuracy of red wine quality classification using the SVM method with the RBF kernel function have good accuracy values. From the accuracy testing process using K-Fold cross validation, the largest average results were obtained in the SVM model with optimization with an average value of 0.91 or 91%. Meanwhile, the SVM model without optimization gets an average value of 0.85 or 85%.

Table 5. K Fold Validation result

Perulangan ke -	Nilai Akurasi	
	Model SVM tanpa optimasi	Model SVM dengan optimasi
1	0.79	0.87
2	0.79	0.85
3	0.83	0.87
4	0.83	0.94
5	0.84	0.91
6	0.83	0.93
7	0.89	0.93
8	0.92	0.94
9	0.91	0.93
10	0.89	0.92
Rata-rata	**0.85**	**0.91**

After carrying out the process of creating a red wine quality classification model using the Support Vector Machine (SVM) method. The best model is selected which will be deployed into the system to be created. The model chosen is a red wine quality classification model using the Support Vector Machine (SVM) method with optimization. The application used was created with the Python framework, namely Django.

(a)

No	Volatile acidity	Total sulfur	Sulphates	Alcohol	Classification
1	1.0	23.0	55.0	5.0	Good quality
2	3.0	0.1	1.0	5.0	Good quality
3	3.0	0.1	1.0	5.0	Good quality
4	3.0	0.1	1.0	5.0	Good quality
5	5.0	23.0	12.0	12.0	Good quality
6	5.0	23.0	12.0	12.0	Good quality
7	2.0	1.8	1.4	1.2	Bad quality
8	0.0	0.0	0.0	0.0	Bad quality

(b)

Figure 6. (a) Sample of Classification form, (b) Classification result

4. CONCLUSION

The data used is data on quality red wine drinks from Portugal "Vinho Verde" totaling 1600. The red wine quality classification model was successfully created using the Support Vector Machine (SVM) method with several input data variables that can classify the quality of red wine. In the attribute selection process using the forward selection technique, from the 11 predictor variables in the dataset, the 4 largest predictor variables were obtained based on the largest R2 score test value, namely the variables alcohol, sulphates, total sulfur dioxide, and volatile acidity. The process of creating a red wine quality classification model using the Support Vector Machine method using the RBF kernel function is divided into two, namely a red wine quality classification model using the Support Vector Machine method without optimization and a red wine quality classification model using the Support Vector Machine method with parameter optimization. The accuracy testing results of the two classification models, namely the red wine quality classification model without parameter optimization using the SVM method was 0.78 or 78%, while the red wine quality classification model with optimal parameter optimization was C = 100, gamma (γ) = 0.5, the accuracy level increased to 0.85 or 85%. Apart from accuracy testing, precision and recall values were also obtained from the two models. From the SVM model without parameter optimization, the precision value obtained was 0.31 or 31% and the recall value was 0.82 or 82%, whereas in the SVM model where the parameter optimization process was carried out, the precision value obtained increased to 0.41 or 41%, but the recall value decreased to of 0.78 or 78%. This research produces the best method for classifying red wine quality drinks using the Support Vector Machine

(SVM) method with optimization parameters C = 100 and gamma (γ) = 0.5. In future research, it is hoped that researchers will be able to deepen the analysis by knowing what problems can determine the quality of red wine.

REFERENCES

[1] D. K. Jana, P. Bhunia, S. Das Adhikary, and A. Mishra, "Analyzing of salient features and classification of wine type based on quality through various neural network and support vector machine classifiers," *Results in Control and Optimization*, vol. 11, p. 100219, Jun. 2023, doi: 10.1016/j.rico.2023.100219.

[2] N. L. Costa, L. A. G. Llobodanin, I. A. Castro, and R. Barbosa, "Using Support Vector Machines and neural networks to classify Merlot wines from South America," *Information Processing in Agriculture*, vol. 6, no. 2, pp. 265–278, Jun. 2019, doi: 10.1016/j.inpa.2018.10.003.

[3] S. Aich, A. A. Al-Absi, K. Lee Hui, and M. Sain, "Prediction of Quality for Different Type of Wine based on Different Feature Sets Using Supervised Machine Learning Techniques," in *2019 21st International Conference on Advanced Communication Technology (ICACT)*, IEEE, Feb. 2019, pp. 1122–1127. doi: 10.23919/ICACT.2019.8702017.

[4] F. Liu and Y. He, "Use of visible and near infrared spectroscopy and least squares-support vector machine to determine soluble solids content and pH of cola beverage.," *J Agric Food Chem*, vol. 55, no. 22, pp. 8883–8, Oct. 2007, doi: 10.1021/jf072057b.

[5] H. Yu, H. Lin, H. Xu, Y. Ying, B. Li, and X. Pan, "Prediction of enological parameters and discrimination of rice wine age using least-squares support vector machines and near infrared spectroscopy.," *J Agric Food Chem*, vol. 56, no. 2, pp. 307–13, Jan. 2008, doi: 10.1021/jf0725575.

[6] S. Aich, A. A. Al-Absi, K. L. Hui, J. T. Lee, and M. Sain, "A classification approach with different feature sets to predict the quality of different types of wine using machine learning techniques," in *2018 20th International Conference on Advanced Communication Technology (ICACT)*, IEEE, Feb. 2018, pp. 139–143. doi: 10.23919/ICACT.2018.8323674.

[7] B. Chen, C. Tawiah, J. Palmer, and R. Erol, "Multi-class wine grades predictions with hierarchical support vector machines," in *2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, IEEE, Jul. 2017, pp. 111–115. doi: 10.1109/FSKD.2017.8392918.

[8] S. Dey, S. Wasif, D. S. Tonmoy, S. Sultana, J. Sarkar, and M. Dey, "A Comparative Study of Support Vector Machine and Naive Bayes Classifier for Sentiment Analysis on Amazon Product Reviews," in *2020 International Conference on Contemporary Computing and Applications (IC3A)*, IEEE, Feb. 2020, pp. 217–220. doi: 10.1109/IC3A48958.2020.233300.

[9] S. Chhikara, P. Bansal, and K. Malik, "Wine Quality Prediction Using Machine Learning Techniques," in *Lecture Notes in Networks and Systems*, vol. 645 LNNS, Springer Science and Business Media Deutschland GmbH, 2023, pp. 137–148. doi: 10.1007/978-981-99-0769-4_14.

[10] F. J. Acevedo, J. Jiménez, S. Maldonado, E. Domínguez, and A. Narváez, "Classification of Wines Produced in Specific Regions by UV−Visible Spectroscopy Combined with Support Vector Machines," *J Agric Food Chem*, vol. 55, no. 17, pp. 6842–6849, Aug. 2007, doi: 10.1021/jf070634q.

[11] Basvaraj. S. Anami, K. Mainalli, S. Kallur, and V. Patil, "A Machine Learning Based Approach for Wine Quality Prediction," in *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)*, IEEE, Aug. 2022, pp. 1–6. doi: 10.1109/ASIANCON55314.2022.9908870.

[12] S. Kumar, K. Agrawal, and N. Mandan, "Red Wine Quality Prediction Using Machine Learning Techniques," in *2020 International Conference on Computer Communication and Informatics (ICCCI)*, IEEE, Jan. 2020, pp. 1–6. doi: 10.1109/ICCCI48352.2020.9104095.

[13] K. R. Dahal, J. N. Dahal, H. Banjade, and S. Gaire, "Prediction of Wine Quality Using Machine Learning Algorithms," *Open J Stat*, vol. 11, no. 02, pp. 278–289, 2021, doi: 10.4236/ojs.2021.112015.