Analysis of Algorithms Support Vector Machine with Naive Bayes Kernel in Data Classification

Juanto Simangunsong[1], Muhammad Zarlis[2] and Tulus[3]
[1]*Graduate Program of Computer Science, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia*
[2] *Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia*
[3]*Department of Computer Science, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia*

Keywords: Support Vector Machine, Naïve Bayes, Accuracy, Classification.

Abstract: This research is about SVM and Naive Bayes in data mining. Many researchers carry out and develop methods to improve the accuracy and classification of data in good results. This research was carried out by conducting experiments on the types of flowers. In this study, it was concluded that the performance of Naïve Bayes was better than Support Vector Machine, Naïve Bayes had excellent results that promised to help classify the best values to get data grouping. This research is better than SVM. The training process has a difference of 28% and a testing process of 0.83% with the accuracy.

1 INTRODUCTION

According to (Lee, 2001) SVM utilizes optimization with quadratic programming, so for high-dimensional data and large amounts of SVM data to be less efficient. Support Vector Machine (SSVM). When compared with SSVM, SVM has a longer running time and smaller accuracy than SSVM. According to Rachman (2011), Huang, (2003) and Byvatov, (2003), Support Vector Machine has a better level of accuracy compared to the logistic regression method, ANN, Naive Bayes, and CART. SVM learning based method that is very promising to be developed because it has high performance and can be widely applied for classification and estimation. According to the study (Honakan, 2018), the classification with the machine vector support process has high accuracy with a combination of stopword, tokenizing, term frequency & chi-square 47.43%. While research (Pratama, 2018) Support Vector Machine (SVM) classifies data into 2 classes using Gaussian kernel RBF with a combination of parameter values $\lambda = 0.5$, constant $\gamma = 0.01$, and ε (epsilon) = 0.001 itermax = 100, c = 1 using training data for 170 datasets. This research resulted in an average accuracy of 80.55%. So that the determination of training data influences the percentage of precision, recall, and accuracy. (Ridwan, 2013).

The application of naïve Bayes method is expected to be able to predict the amount of electricity usage per household so that it is easier to regulate electricity usage. From 60 data on electricity usage in homes has been tested using the Naive Bayes algorithm and its percentage results were 78.3333% for the accuracy of predictions, of which 60 electricity usage in homes has been tested using the Naive Bayes algorithm were 47 data on electricity usage in homes that were successfully classified correctly. (Saleh, 2015).

2 RESEARCH METHODS

2.1 Naïve Bayes

Naive Bayes Classifier (NBC) is an algorithm that performs data mining techniques by applying the Naive Bayes method in classifying data. The theory of naive Bayes characteristic in doing pattern recognition. Naive Bayes has an independent attribute value when output values are used. Output probability through individual probability. NBC is done by entering equation 1 and equation 2. (Santosa, 2002).

Simangunsong, J., Zarlis, M. and Tulus, .
Analysis of Algorithms Support Vector Machine with Naive Bayes Kernel in Data Classification.
DOI: 10.5220/0008553302870291
In *Proceedings of the International Conference on Natural Resources and Technology (ICONART 2019)*, pages 287-291
ISBN: 978-989-758-404-6

The probability of a simple classification that performs calculations with several probabilities by performing several frequencies and mixing values from the dataset used is a naïve Bayes method of testing. Naïve Bayes assumes independent attributes or not interdependence on the variable values of each class. (Patil, 2013). Other researchers say that the naïve Bayes method is made by English people, Thomas Bayes, who classifies probabilities and statistics by predicting the future by doing it using previous experiences (Bustami, 2013).

Naïve bayes is not given an output value and simplifies the independent atriut values conditionally. In other words, by observing the probability of the product's probability. (Ridwan, 2013). The good that is produced by using the method n is that this method only requires little data in conducting data training and in determining estimates of the parameters used in the data classification process. Naïve Bayes works very well in real global life as expected.

Theorem bayes is used to calculate the number of probabilities to events that influence the results of observations. In Bayesian, the parameter is used as a random variable while in the former statistical world, the parameters must always be corrected. Pastor Thomas Bayes is the name of the theorem Bayes, which is described as the relationship between the opportunities of events A and Z, which are explained in the following formula (Kundu, 2011):

$$P(H \mid x) = \frac{P(x \mid H) P(H)}{P(x)}$$

In the X data sample class whose label is not yet known, and H is the hypothesis, the sample data x is transferred to the special class c. P (H / x) is a probability that explains data about research data x. P (H / x) is a posterior probability that resembles trust in predictions after x is given. Conversely, P (H) is the probability H before the sample is used, before the sample is formed. Posterior probability P (H / x) is based on a lot of information from the priori probability P (H). Bayes theory has a way of calculating posterior probability P (H / x) using probabilities P (H), P (X) and P (H / x).

The Bayes method is a statistical approach for induction inference on classification problems. First discussed first about the basic concepts and definitions in the theory of Bayes, then using this theorem to classify in Data Mining. The Bayes method uses conditional proportions as the basis.

2.1.1 Principles of The Bayes Method

The Bayes method has an easy way to add outside information to the data analysis process. The process is done by distributing existing data with approved datasets (Albert, 2009). this method is done with opportunities that have requirements..

2.1.2 Principles of The Bayes Method

Some classification techniques are used (Albert, 2009): Decision tree classifier, Rule based classifier, Neural network, Naive Bayes.

Each technique uses a learning algorithm to identify the model that provides the most appropriate relationship. An example of the Bayesian theory is the case of patients who have difficulty breathing. Decisions taken are between cases of patients suffering from asthma or patients suffering from lung cancer (Bolstad, 2007).

a. Decision 1: states that someone has lung cancer despite the actual symptoms of asthma (cost: high enough, so that it scares patients and makes patients undergo unnecessary examinations).
b. Decision 2: declare someone asthma even though it is actually lung cancer (cost: very high that makes the patient lose the opportunity to treat cancer at the initial or final stage).

2.1.3 Principles of The Bayes Method

Disadvantages of the Bayes Method include is The Bayes method can only be used for classification problems with supervised learning and categorical data, The Bayes method requires initial knowledge to be able to make a decision. The success rate of this method depends on the initial knowledge given.

The advantages of the Bayes Method include is Interpolation: The Bayes method has choices about how much time and effort is made by humans vs. computers; Language: The bayes method has its own language for determining the things prior and posterior; Intuition: Involving priors and integration, two broadly useful activities.

Bayesian probability is the best theory in dealing with estimation problems and drawing conclusions. Bayesian methods can be used to draw conclusions in cases with multiple sources of measurement that cannot be handled by other methods such as complex hierarchical models (Bolstad, 2007).

2.2 Smooth Support Vector Machine

SVM was created by Vapnik in 1992 using a series of superior concepts that are good in the field of

pattern recognition. SVM is still considered young in pattern recognition methods. However, its ability in various applications is often used as an art in making a pattern. SVM is also a learning machine method that works based on the principle of Structural Risk Minimization (SRM), which functions as the best hyperplane that separates an input.

The SVM concept is a right combination of computer theory that has been used for several years, such as the hyperplane margin (Duda & Hart in 1973, Cover in 1965, Vapnik 1964, etc.), Kernel introduced by Aronszajn in 1950, and also with other supporting plans. However, until 1992, there had never been an attempt to assemble these components.

SVM has multi-dimensional features that have plots as data points in classifying by defining the boundary between data points from the surface. SVM aims to make the bottom line or called the hyperlink with the same direction of data partition with each other. In this way, SVM learning combines aspects of the learning of the closest neighbors. For binary classification problems, SVM is very suitable. For example, in Figure 1, below.

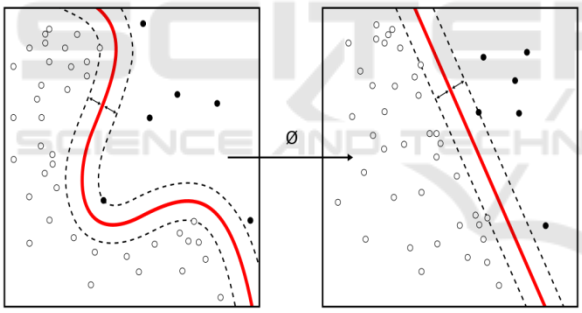

Figure 1: Support Vector Machine

3 IDENTIFICATION OF PROBLEM

From the background of the problem that has been described, then the authors take the formulation of the problem for the data processing required several methods to get better and optimal results. By making a comparison of the methods used it is very necessary for good data processing to analyze the performance of algorithms through a comparison of the Support Vector Machine and Naive Bayes Kernel algorithms with different class classifications from the perspective of precision, recall & accuracy and F-Measure..

4 RESULT AND DISCUSSION

The results of the analysis that we did on the methods in this study, I experimented with a student data set. All of these experiments were carried out using a personal computer, which used a 1.66 GHz Intel Atom Core 2 Duo CPU processor. The computer runs using the Windows 7 operating system, by installing Rapid Miner 8.1.

At this stage, the performance testing of the Naïve Bayes algorithm and SVM is carried out. Images show graphs of classification results measured based on accuracy values obtained from the Naïve Bayes algorithm and SVM

Table 1: Test Results Table

	NAÏVE BAYES	SVM
class precision	95.83%	95.00%
class recall	96.00%	38.00%
accuracy	96.00%	68.00%

In the Naïve Bayes and SVM algorithms there are graphs of different shapes, which can be seen in the following figure:

4.1 Naïve Bayes Graph

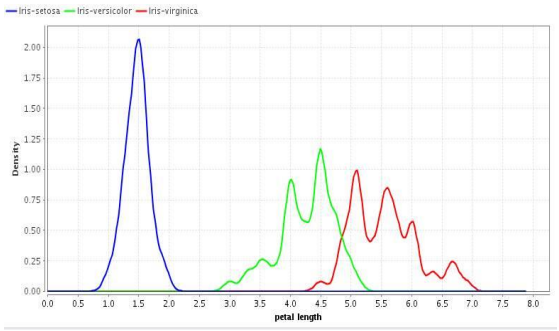

Figure 2: Petal Length Graph

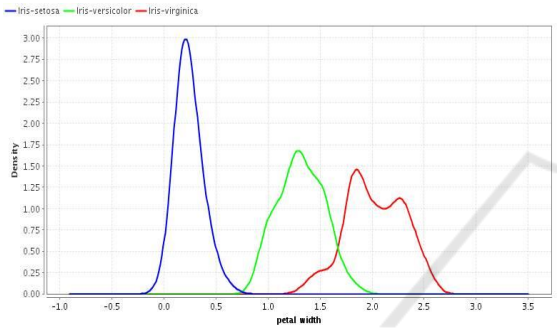

Figure 3: Petal Width Graph

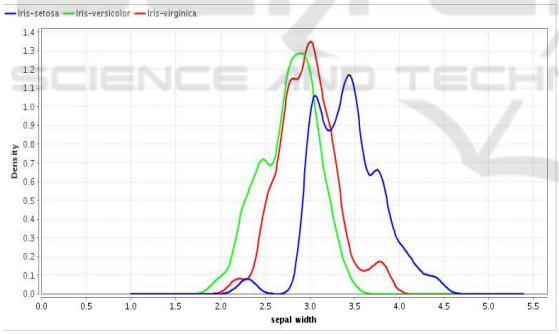

Figure 4: Sepal Width Graph

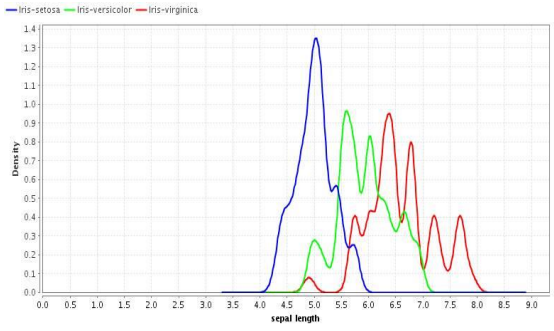

Figure 5: Sepal Length Graph

4.2 Naïve Bayes Graph

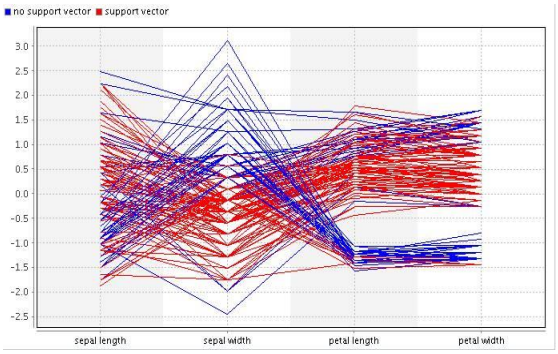

Figure 6: SVM Graph

5 CONCLUSIONS

In conclusion, the results can be seen as follows: Naïve Bayes performance is better than SVM, Naïve Bayes gets impressive results that promise to help find excellent data accuracy. The accuracy of the Naïve Bayes classification in this study is better than SVM, which experiences a 30% difference in accuracy. Therefore, Naive Bayes should be used in testing data classification in a larger form because it can classify data with useful data.

REFERENCES

Albert, J. 2009. *Bayesian Computation with R*. Springer : New York.

Bustami. 2013. *Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi.* TECHSI : *Jurnal Penelitian Teknik Informatika*, 3 (2), 127-146.

Bolstad, W. M. 2007. *Introduction to Bayesian Statistics*. John Wiley and Sons : New Jersey.

Bramer, M. 2007. *Principles of Data Mining*. London: Springer.

Gorunescu, F. 2011. *Data Mining: Concepts, Models and Techniques.* Berlin: Springer-Verlag.

Han, J., Kamber, M. 2001. *Data Mining Concepts and Techniques*, Morgan Kaufman Pub., USA.

Han, J. and Kamber, M. 2006. *Data Mining Concepts and Techniques.* San Francisco, 2nd Edition.

Han, J., Kamber, M., Pei, J. 2011. *Data Mining: Concepts and Techniques.*.San Francisco, 3rd edition.

Patil, T. R., Sherekar, M. S., 2013. Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification. *International Journal of Computer Science and Applications*, 6 (2), 256-261.

Pattekari, S. A., Parveen, A. 2012. Prediction System for Heart Disease Using Naive Bayes. *International*

Journal of Advanced Computer and Mathematical Sciences, 3 (3), 290-294.

Ridwan, M., Suyono, H., Sarosa, M. 2013. Penerapan Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Jurnal EECCIS*, 1(7), 59-64.

Santosa, B. 2007. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis.* Graha Ilmu. Yogyakarta.

Zarlis, M., Sitompul, O. S., Sawaluddin, Effendi, S., Sihombing, P., Nababan, E. B. 2015. *Pedoman Penulisan Tesis.* FasilkomTI. Universitas Sumatera Utara.

Zhang, H., Wang. Z. 2011. A Normal Distributions-Based Over-Sampling Approach to Imbalanced Data Classification. Advanced Data Mining and Applications. In *7th International Conference*, 83-96. Beijing, Springer.