International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 7 Issue 5, May - 2023, Pages: 25-30

Application of the Support Vector Machine Classification Method to Elementary School Accreditation Data in Samarinda City

Marthinus Siahaya[1], Darnah[2], Moh. Nurul Huda[3], Surya Prangga[4] and Meiliani Siringoringo[5]

[1]Department of Mathematics, Mulawarman University, Samarinda 75123,
Indonesia correspondence author: darnah.98@gmail.com

Abstract: *Accreditation is recognition of an educational institution after it is assessed that the institution meets standard requirements or predetermined criteria. This accreditation is usually categorized into 4 categories, namely Very good (A), Good (B), Fair (C), and not accredited. The purpose of accreditation is to determine the eligibility level of a school in providing educational services and obtain an overview of school performance. In this study the method used is the Support Vector Machine (SVM) classification, this method is a function model that can describe and differentiate data into classes by finding the best hyperplane to separate two classes of data using a support vector approach. This study uses two internal kernel functions (SVM), namely the linear kernel function and RBF with the proportion of 70%:30% and 90%:10%. The data used is elementary school accreditation data where there are 131 school data using 8 independent variables namely, content standards, process standards, graduate competency standards, educator and education staff standards, facilities and infrastructure standards, management standards, financing standards, and assessment standards . Based on the research that has been done, it is found that the linear kernel function SVM classification method is the most suitable classification algorithm for classifying data in this study with the highest accuracy value of 97.43% with a proportion of 70%:30%.*

Keywords: Accreditation; Classification; Support Vector Machine.

1. INTRODUCTION

Classification is one of the most important lessons in data mining. Classification is defined as a form of data analysis to extract a model that will be used to predict class labels. Class in classification is the most important data set attribute for independent variables in statistics. Data classification consists of two processes, namely the learning stage and the classification stage. Classification aims to group objects into certain classes based on attribute values related to the object being observed. One method that can be used in classification is the Support Vector Machine (SVM). SVM is a relatively new technique for making predictions, both in the case of classification and regression, which is very popular recently (Santosa, 2007). SVM is a technique for finding hyperplane that can separate two data sets from two different classes. SVM has the basic principle of a linear classifier, namely classification cases that can be separated linearly, but SVM has been developed so that it can work on non-linear problems with the kernel concept in high-dimensional workspaces where hyperplane will be sought in maximizing the distance (margin) between data classes. SVM has advantages including determining distances using support vector so that the computation process becomes fast (Christiani, 2000). The Support Vector Machine (SVM) method is also quite good at solving classification problems using kernel functions, namely linear kernel, Radial Basis Function (RBF), polynomial, and sigmoid (Zhang & Wang, 2010). Government Regulation Number 19 of 2005 Article 2 paragraph 2 concerning National Education Standards states that education quality assurance and control in accordance with National Education Standards (SNP) needs to be carried out in three integrated programs namely evaluation, accreditation and certification. Confirmation of the importance of accreditation can be seen in Law Number 20 of 2003 concerning the National Education System (Tilaar, 2012). The school/madrasah accreditation policy in Indonesia is that every citizen has the right to obtain quality education. To be able to organize quality education, each unit/educational program the scope of school/madrasah accreditation includes Kindergarten (TK)/Raudhatul Artfal (RA), Elementary School (SD)/Madrasah Ibtidaiyah (MI), Middle School (SMP)/Madrasah Tsanawiyah (MTS), High School (SMA)/Madrasah Aliyah (MA), Vocational High School (SMK)/Madrasah Aliyah Vocational (MAK), and Special Schools (SLB) (Hidayat, 2015). Accreditation shows an important role for schools in efforts to increase school productivity and guarantee the quality of an educational unit. In reality on the ground, school accreditation is meant more to gain formal status and recognition, while it true meaning is not widely known and implemented seriously. Therefore, it is necessary to have a system with statistical methods used to classify the accreditation of a school that can be used as an alternative or reference from the government and the community. One way that can be done is to classify school accreditation based on the standards in each school (Zazin, 2011).

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 7 Issue 5, May - 2023, Pages: 25-30

2. RESEARCH METHOD

There are 9 variables used in this study, divided into one dependent variable (Y) and eight independent variables (X). The variables used in this study are presented in Table 1.

Table 1 research variable

Notation	Variable	Scale
Y	Accreditation	Ordinal
X_1	Content standard	Intervals
X_2	Process standard	Intervals
X_3	Graduate competence standard	Intervals
X_4	Standards for educators and education personnel	Intervals
X_5	Facilities and infrastructure standards	Intervals
X_6	Management standard	Intervals
X_7	Financing standard	Intervals
X_8	Rating standards	Intervals

2.1 Research Methods

Data mining is also known as Knowledge Discovery in Database (KDD). Simply put, data mining extracts important information or patterns from existing data in large databases (Siregar & Pusbabhuana, 2017). The descriptive method has the goal of finding patterns, relationships, or data anomalies that are easily understood by humans (Sigit & Yuita, 2018). Classification is an important part of data mining which is used to predict labels or classes from previously studied models. Through a classification technique approach, the data that has been studied will then produce a pattern or rule There are many classification models that can be used such as decision trees, probabilistic classifiers, support vector machines and so on [9].

2.2 *Support Vector Machine* (SVM)

SVM is a machine learning method that works on the principle of Structural Risk Minimization (SRM) which aims to find the best hyperplane to separate two classes of data. Basically SVM has a linear principle, but now SVM has developed so that it can work on non-linear problems. The way SVM works on non-linear problems is by including the kernel concept in high-dimensional space. In this dimensionless space, we will look for a separator or what is often called a hyperplane (Pratiwi, 2017)

The SVM algorithm in general can be written as follows:

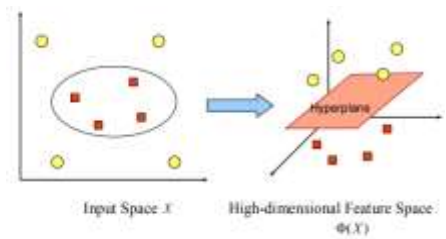
1. Initialize the initial value of the model parameters
2. Determine the limiting equation of each class
3. Determine the quadratic program objective function Quadratic Programing (QP)
4. Solve the optimization problem

For example, the data in the training data set is denoted as $\boldsymbol{x_i} \in \Re^q$ whereas the label of each class is expressed as $y_i \in \{-1, +1\}$ the general linear model used in the SVM method to generate hyperplanes (Fitriyah, Maruddani & Wasito, 2020) namely:

$$y_i = W^T x_i + b, i = 1,2,...,n \quad (1)$$

2.2.1 SVM Non-Linear Separable Data

In SVM there is a kernel function, which is a function used to solve non-linear problems. The kernel function makes it possible to implement a model in a higher dimensional space (feature space). Could be seen in the following illustration.



Picture 1. Hyperplane

Basically the learning process in SVM in finding support vectors only depends on the dot product of the data in the feature space, namely $\Phi(x_i).\Phi(x_j)$. Because the transformation of Φ is unknown and very difficult to understand, the calculation of the dot product can be replaced by the kernel function $K(x_i, x_j)$ which implicitly defines the transformation function Φ. The Kernel Trick formulation is as follows (Cortes & Vapnik, 1995).

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \Phi(\boldsymbol{x}_i).\Phi(\boldsymbol{x}_j) \quad (2)$$

Information:

$K(\boldsymbol{x}_i, \boldsymbol{x}_j)$: Kernel functions

$\Phi(\boldsymbol{x}_i).\Phi(\boldsymbol{x}_j)$: Dot product of data in feature space

Furthermore, it can calculate the hyperplane as a decision boundary efficiently. The use of kernel functions causes the Lagrange equation to be modified to become:

$$L_d = \sum_{i=1}^{n} a_i - \frac{1}{2}\sum_{i,j=1}^{n} a_i a_j y_i y_j \mathrm{K}(\boldsymbol{x_i}^T \boldsymbol{x}_j) \quad (3)$$

The resulting function from training is as follows:

$$Sign(f(\boldsymbol{x})) = \sum_{i,j=1}^{n} a_i y_i \mathrm{K}(\boldsymbol{x}_{i,}, \boldsymbol{x}_j) + b\ ; i, j = 1,2, \dots, n \quad (4)$$

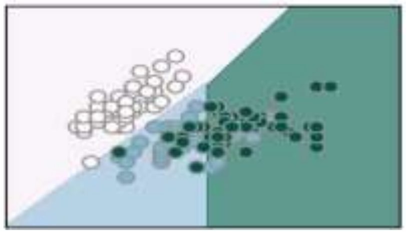
After the solution to the Quadratic Programing (QP) problem has found a value (a_i) (Suyanto, 2017). Then the class of an x testing data can be determined by the following equation:

$$f(\boldsymbol{x}) = \sum_{s=1}^{v} a_s\, y_s \mathrm{K}(\boldsymbol{x}_s . \boldsymbol{x}_t) + b\ ; s = 1,2, \dots, n \quad (5)$$

2.3 Linear Kernel Function

Linear kernel is a kernel that is used in finding hyperplane that are formed as straight lines (Suyanto, 2017). Here's the linear kernel equation:

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x_i}^T \boldsymbol{x_j} \quad (6)$$

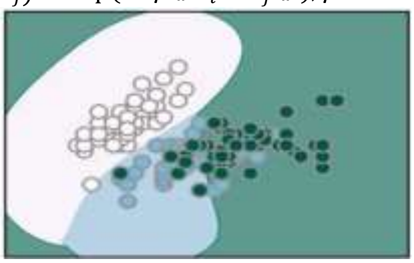


Picture 2. Linear Kernel illustration

2.4 Gausian Kernel

Gaussian Kernel or Radial Basis Function (RBF) is a kernel that is able to recognize objects precisely and accurately. So that the RBF kernel is able to make the hyperplane able to separate data properly (Suyanto, 2017). Here's the Gaussian kernel / RBF equation:

$$(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp(-\gamma \parallel \boldsymbol{x}_i - x_j \parallel^2), \gamma > 0 \quad (7)$$



Picture 3. RBF Kernel illustration

2.5 Classification Method Performance Evaluation

Illustration of the confusion matrix can be seen in Table 2

Table 2. Confusion Matrix

Actual Class	Prediction Class	
	Positive	**Negative**
Positive	TP	FN
Negative	FP	TN

Information:
TP: True Positive
FP: False Positive
FN: False Negative
TN: True Negative
Equation to calculate accuracy:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (8)$$

The method used in this study is the (SVM) method using linear kernels and kernels (RBF) with parameters C=1 for linear kernel functions and γ=0.01 and C=1 for kernel functions (RBF). Flowchart for this study as follows:

Inputs

Pre-processing data

Split data: data *training* and data *testing*

Define kernel functions and their parameters

Determine the amount of penalti C

Build the SVM function from training to get the value of a_i and b

Determine classification results from data testing

Counting accuracy

Choose best accuracy

Finishe

Picture 4. Data Analysis Flowchart

3. RESULT and DISCUSSION

3.1 Descriptive statistics

Before classifying elementary school (SD) accreditation A and B using the linear kernel function (SVM) method and (RBF), a descriptive analysis was carried out first, which aims to provide an overview of the research variables. Descriptive results can be seen in Table 3 and Table 4.

Table 3. Descriptive Statistics of SD Accredited A

Descriptive Statistics	Variable							
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Average	93,31	91,31	90,41	89,40	88,71	90,59	94,19	93,09
Baku Devi	5,74	5,70	6,34	5,07	6,84	5,57	5,04	4,74
Minimum	74	73	75	76	75	76	74	77
Maximum	100	100	100	99	100	100	100	100

Table 4. Descriptive Statistics of SD Accredited B

Descriptive Statistics	Variable							
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Average	82,67	80,07	75,32	80,16	77,52	80,08	87,89	84,81
Baku Devi	8,32	8,78	9,27	6,99	7,35	7,84	8,37	7,46
Minimum	63	57	52	63	62	60	56	64
Maximum	97	94	100	98	94	95	100	99

3.2 Data Standardization

Data standardization is carried out so that all variables are within the same range so that the proportion of influence on the classifier function can be balanced. The calculations for data standardization are as follows:

1. Calculate the average for each variable

$$\bar{X}_1 = \frac{1}{131}(79 + 94 + \cdots + 88)$$
$$= 87{,}38$$
$$\bar{X}_2 = \frac{1}{131}(85 + 93 + \cdots + 86)$$
$$= 85{,}04$$
$$\bar{X}_3 = \frac{1}{131}(88 + 96 + \cdots + 89)$$

$= 82$

$\bar{X}_4 = \frac{1}{131}(78 + 85 + \cdots + 86)$
$= 84{,}21$

$\bar{X}_5 = \frac{1}{131}(94 + 84 + \cdots + 73)$
$= 82{,}47$

$\bar{X}_6 = \frac{1}{131}(91 + 95 + \cdots + 90)$
$= 84{,}73$

$\bar{X}_7 = \frac{1}{131}(96 + 94 + \cdots + 89)$
$= 90{,}67$

$\bar{X}_8 = \frac{1}{131}(98 + 94 + \cdots + 91)$
$= 88{,}47$

2. Calculate the standard deviation value for each variable

$$= \sqrt{\frac{1}{(131-1)} \times ((79-87{,}38)^2 + (94-87{,}38)^2 + \cdots + (88-87{,}38)^2)}$$
$= 8{,}99$

$$s_2 = \sqrt{\frac{1}{(131-1)} \times ((85-85{,}04)^2 + (93-85{,}04)^2 + \cdots + (86-85{,}04)^2)}$$
$= 9{,}40$

$$s_3 = \sqrt{\frac{1}{(131-1)} \times ((88-82)^2 + (96-82)^2 + \cdots + (89-82)^2)}$$
$= 11{,}04$

$$s_4 = \sqrt{\frac{1}{(131-1)} \times ((78-84{,}21)^2 + (85-84{,}21)^2 + \cdots + (86-84{,}21)^2)}$$
$= 7{,}74$

$$s_5 = \sqrt{\frac{1}{(131-1)} \times ((94-82{,}47)^2 + (84-82{,}47)^2 + \cdots + (73-82{,}47)^2)}$$
$= 9{,}03$

$$s_6 = \sqrt{\frac{1}{(131-1)} \times ((91-84{,}73)^2 + (95-84{,}73)^2 + \cdots + (90-84{,}73)^2)}$$
$= 8{,}67$

$$s_7 = \sqrt{\frac{1}{(131-1)} \times ((96-90{,}67)^2 + (94-90{,}67)^2 + \cdots + (89-90{,}67)^2)}$$
$= 7{,}73$

$$s_8 = \sqrt{\frac{1}{(131-1)} \times ((98-88{,}47)^2 + (94-88{,}47)^2 + \cdots + (91-88{,}47)^2)}$$
$= 7{,}59$

3. Calculate the standardized value for each data

$$\hat{x}_{(1,1)} = \frac{79 - 87{,}38}{8{,}99}$$
$= -0{,}93$

$$\hat{x}_{(2,1)} = \frac{94 - 87{,}38}{8{,}99}$$
$= 0{,}73$

$\vdots$

$$\hat{x}_{(131,8)} = \frac{91 - 88{,}47}{7{,}59}$$
$= 0{,}70$

3.3 Distribution of Data Training and Data Testing

Distribution of training data and testing data was carried out after the data randomization process. The distribution of training data and testing data in this study uses 2 proportions, namely 70%: 30% and 90%: 10%

1. Distribution of data with the proportion that is 70%:30%, the calculation is:

$ntraining_{70\%}$ = Proportion data $training \times 131$
$= \frac{70}{100} \times 131$
$= 91{,}7 \approx 92$

$ntesting_{30\%}$ = Proportion data $testing \times 131$
$= \frac{30}{100} \times 131$
$= 39{,}3 \approx 39$

Based on calculations on the proportion of data 70%:30%, 92 training data were obtained, meaning that the first data to the 92nd data after randomization were used as training data and the 93rd to 131st data after randomization were used as testing data.

2. Distribution of data with the proportion that is 70%:30%, the calculation is:

$ntraining_{90\%}$ = Proportion data $training \times 131$
$= \frac{90}{100} \times 131$
$= 117{,}9 \approx 118$

$ntesting_{10\%}$ = Proportion data $testing \times 131$
$= \frac{10}{100} \times 131$
$= 13{,}1 \approx 13$

Based on calculations on the proportion of data 90%:10%, 118 training data were obtained, meaning that the first data to the 118th data after randomization were used as training data and the 119th to 131st data after randomization were used as testing data.

3.4 SVM Clasificassion

SVM classification is carried out based on the proportion of training and testing data 70%:30% and 90%:10% in the linear kernel and RBF kernel.

3.4.1 SVM Classification of Linear Kernel Functions

In the SVM method the linear kernel function uses 92 training data, which is as much as 70% of the 131 overall data will be built based on the SVM function. Form a linear kernel function SVM model with the penalty parameter C=1 as follows:

1. SVM Classification Linear Kernel Function Data Proportion 70%:30%

$$sign(f(\boldsymbol{x})) = \sum_{i,j=1}^{92} a_i y_i \mathrm{K}(\boldsymbol{x_i}^T \boldsymbol{x}_j) + (-0{,}974) \quad (9)$$

Furthermore, it will be used to classify test data:

$$f(\boldsymbol{x}) = \sum_{s=1}^{24} a_s\, y_s \mathrm{K}(\boldsymbol{x}_s . \boldsymbol{x}_t) + (-0{,}974) \quad (10)$$

Table 5. Confusion Matrix of SVM Classification Result Linear Kernel Function
Using 39 testing data (30% testing)

Actual Class	SVM Clasification Prediction Class		Total
	A	**B**	
A	18	*1	19
B	*0	20	20
Total	18	21	39

Next calculate the accuracy value:

$$\text{Accuracy} = \frac{18+20}{18+0+20+1} = 0{,}974 \text{ or } 97{,}43\% \quad (11)$$

2. SVM Classification Linear Kernel Function Data Proportion 90%:10%

$$sign\big(f(\boldsymbol{x})\big) = \sum_{i,j=1}^{118} a_i y_i \mathrm{K}\,(\boldsymbol{x_i}^T \boldsymbol{x}_j) + (-0{,}850) \quad (12)$$

Furthermore, it will be used to classify test data:

$$f(\boldsymbol{x}) = \sum_{s=1}^{25} a_s\, y_s \mathrm{K}(\boldsymbol{x}_s . \boldsymbol{x}_t) + (-0{,}850) \quad (13)$$

Table 6. Confusion Matrix of SVM Classification Result Linear Kernel Function
Using 13 testing data (10% testing)

Actual Class	SVM Clasification Prediction Class		Total
	A	**B**	
A	6	*1	7
B	*0	6	6
Total	6	7	13

Next calculate the accuracy value:

$$\text{Accuracy} = \frac{6+6}{6+0+6+1} = 0{,}923 \text{ or } 92{,}30\% \quad (14)$$

3.4.2 Classification of SVM Kernel RBF Functions

The first step in classifying using the SVM method is the RBF kernel function, namely determining the Kernel RBF parameters and the C penalty parameter. The SVM parameter for the RBF kernel function used is γ=0.01. In addition, the penalty parameter C which in this study uses the penalty parameter C = 1.

1. SVM Classification Kernel Function RBF Data Proportion 70%:30%

$$sign\big(f(\boldsymbol{x})\big) = \sum_{i,j=1}^{92} a_i y_i \mathrm{K}(\exp(-0{,}01 \times \| \boldsymbol{x_i} - \boldsymbol{x}_j \|^2) + (-0{,}301) \quad (15)$$

Furthermore, it will be used to classify test data:

$$f(\boldsymbol{x}) = \sum_{s=1}^{53} a_s\, y_s \mathrm{K}(\exp(0{,}01 \times \| \boldsymbol{x}_s . \boldsymbol{x}_t \|)^2) + (-0{,}301) \quad (16)$$

Table 7. Confusion Matrix of SVM Classification Results Kernel Function RBF
Using 39 testing data (30% testing)

Actual Class	SVM Clasification Prediction Class		Total
	A	**B**	
A	17	*2	19
B	*0	20	20
Total	17	22	39

Next calculate the accuracy value:

$$\text{Accuracy} = \frac{6+6}{6+0+6+1} = 0{,}948 \text{ or } 94{,}87\% \quad (17)$$

2. SVM Classification Kernel Function RBF Data Proportion 90%:10%

$$sign\big(f(\boldsymbol{x})\big) = \sum_{i,j=1}^{118} a_i y_i \mathrm{K}(\exp(-0{,}01 \times \| \boldsymbol{x_i} - \boldsymbol{x}_j \|^2) + (-0{,}350) \quad (18)$$

Furthermore, it will be used to classify test data:

$$f(\boldsymbol{x}) = \sum_{s=1}^{61} a_s\, y_s \mathrm{K}(\exp(0{,}01 \times \| \boldsymbol{x}_s . \boldsymbol{x}_t \|)^2) + (-0{,}350) \quad (19)$$

Table 8. Confusion Matrix of SVM Classification Results Kernel Function RBF
Using 13 testing data (10% testing)

Actual Class	SVM Clasification Prediction Class		Total
	A	**B**	
A	6	*1	7
B	*0	6	6
Total	6	7	13

Next calculate the accuracy value:

$$\text{Accuracy} = \frac{6+6}{6+0+6+1} = 0{,}923 \text{ or } 92{,}30\% \quad (20)$$

3.5 Best Accuracy Selection

Selection of the best accuracy is done by looking at and comparing the accuracy values in the SVM classification of linear kernel functions and RBF based on the proportion of data. The following accuracy values can be seen in Table 9.

Table 9. SVM Accuracy Value

Proportion	Linear SVM ($\boldsymbol{C = 1}$)	SVM RBF ($\boldsymbol{\gamma = 0{,}01, C = 1}$)
70%:30%	97,43%	94,87%
90%:10%	92,30%	92,30%

Based on Table 9 the linear kernel function SVM classification method at the proportion of 70%:30% is a better classification algorithm for classifying elementary school accreditation data in Samarinda City in 2016-2021. It can be

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 7 Issue 5, May - 2023, Pages: 25-30

seen that the highest accuracy value obtained is 97.43%.

4. CONCLUSION

The results of the classification of Samarinda City Elementary School (SD) accreditation for 2016-2021 are based on the best accuracy of the linear kernel function SVM and RBF, namely using the linear kernel function SVM with the proportion of 70%: 30% are 18 schools that are classified correctly and there are no schools that are classified correct. Based on accreditation, out of a total of 18 schools that have accreditation A, there are 18 schools that are correctly classified and one school that is not properly classified. Out of a total of 20 schools that have accreditation category B, there are 20 schools that are not properly classified and there are no schools that are not properly classified. The results of the measurement of accuracy in classifying Samarinda City Elementary School (SD) accreditation in 2016-2021 based on the best accuracy of the linear kernel function SVM method and RBF, namely the linear kernel function SVM with a proportion of 70%:30% in Samarinda City in 2016-2021 are with accuracy of 97.43%.

REFERENCES

[1] Cortes, C., & Vapnik, V. (1995). *Support Vector Machine Networks. Machine Learning*, (20), 273-297.

[2] Cristianini, N. (2000). *An Introduction to Support Vector Machine.* UK: *Cambridge University Press*.

[3] Fitriyah, N., Maruddani, D.A.I., & Warsito, B. (2020). Analisis Senttimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi *Support Vector Machine. Jurnal Gaussian*, 09 (03), 376-390.

[4] Hidayat, A. (2015). *Pengelolaan Pendidikan*. Yogyakarta: Kaukaba.

[5] Pratiwi, Y. R. (2017). Perbandingan Analisis Sentimen Pertalite Melalui Jejaring Sosial Twitter dengan Menggunakan Metode *Support Vector Machine dan Maximum Entropy. Skripsi. Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia.* Yogyakarta.

[6] Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data Untuk Keperluan Bisnis.* Yogyakarta: Graha Ilmu.

[7] Sigit, A., & Yuita, A. S. (2018). Implementasi *Data Mining* Menggunakan WEKA (Cetaka Pe; T. U. Press, Ed.). Malang: UB Press.

[8] Siregar, A. M., & Puspabhuana, A. (2017). *Data Mining:* Pengolahan Data Menjadi Informasi Dengan *RapidMiner*. CV Kekata Group.

[9] Suyanto. (2017). Data Mining untuk Klasifikasi dan Klasterisasi Data. Bandung: Informatika.

[10] Tilaar, H. (2012). Standarisasi Pendidikan Nasional Suatu Tinjauan Kritis. Jakarta: Rineka Cipta.

[11] Zazin, N. (2011). *Gerakan Menata Mutu Pendidikan*. Yogyakarta: Arruzz Media.

[12] Zhang, Y., & Wang, W. (2010). Pattern Clasification of Electroencephalongraphy from the Typical Specialized Student. Education Technology and Computer Science (ETCS), International Workshop, 01, 836-839