Application of Machine Learning to Analyze Teacher and Student Performance As an Effort to Improve the Quality of Education

Wais Al Qarni[1,*], Trional Novanza, Xosya Salassa[3]

[1]*Sistem Komputer, Fasilkom, Universitas Sriwijaya*
[2]*Sistem Informasi, Fasilkom, Universitas Sriwijaya*
[3]*Sistem Komputer, Fasilkom, Universitas Sriwijaya*
Jl. Raya Palembang – Prabumulih Km. 32 Indralaya, OI, Sumatera Selatan 30662
[*]*wais.alq99@gmail.com*

Abstract

Currently the education sector has produced a lot of data as a result of the teaching and learning process. This data has the potential to improve the quality of education when it processed in depth using machine learning technology, especially it can evaluate the achievement performance of teachers and students in a systemic, intelligent and dynamic manner. Professional teachers and the application of effective teaching methods are the benchmarks for improving the quality of education. however, dynamic teacher performance assessment indicators have not been applied. so that to find out the method used in the lesson is suitable or not still cannot be known accurately. Likewise with students, there are still no accurate indicators that can analyze students' talents and weaknesses based on learning achievement. In this paper not only analyzes the performance of teachers and students, but also designs media to improve the quality of teachers and students. This paper applies a model of both the naïve Bayes algorithm and the decision tree algorithm to classify learning achievement. This paper uses high school student grade recap data that has been processed and sampled by stratified sampling technique. The results of the analysis with the classification using the Naïve Bayes algorithm get an accuracy of 85% while the decision tree algorithm has an accuracy of 80%. Therefore, the chances of success or failure based on the applied education system can be predicted early to improve the quality of education.

Keywords*: education quality, machine learning, performance,*

1 . Introduction

The quality of education greatly impacts to the progress of the nation both in terms of knowledge and character of students[1]. The results of testing by the Program for International Student Assessment (PISA) in 2018, showed a decrease in the quality of Indonesian education in the last 6 ranks with various assessment indicators[2]. Technological innovation in the field of education in Indonesia is absolutely must. the supporting factor is the development of science and technology, especially in terms of data analysis [3].

Many studies using education data have been conducted to improve the quality of education. Research [4] using student score summary data to predict achievement by applying machine learning so that students and teachers can anticipate the learning process early to get maximum learning outcomes. This research applies machine learning that will find useful patterns in data sets, thereby producing new discoveries that cannot be observed with traditional statistical approaches. so that various problems that occur can be predicted and become recommendations for policies to be taken to improve the quality of education[5].

2 . Research Methodology

This research applies a quantitative approach in which the implementation is carried out by experiments involving a number of training data and testing data. The testing data collection technique used stratified sampling technique with the data used consisted of 101 data lines, five attributes and five data labels.

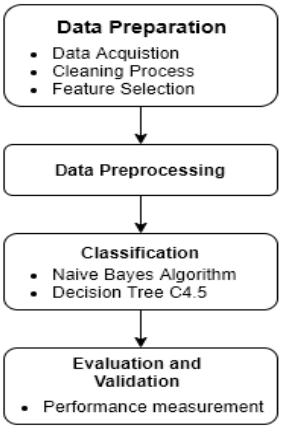
The data used in this study were limited to the students' cognitive and psychomotor scores, which then combined the two categories at the semester level. The affective and presence value variables can be used as supporting variables for the analysis performance.

Variable	Score
Cognitive value	10 – 100
Psychomotor value	10 – 100
Affective value	Very good, good, less
Presence	High, medium, low

(Table 1. Student performance graph variables)

The data analysis process is carried out with a machine learning approach that will classify the talents and weaknesses of students in the subjects and provide accurate department recommendations with the naïve Bayes algorithm and decision tree

algorithm. The naïve Bayes algorithm is a statistical classification algorithm that predicts the probability of membership of a tuple of data that will join a particular class[6]. The stages in this research are:



(Figure 2.1 Research Methodology)

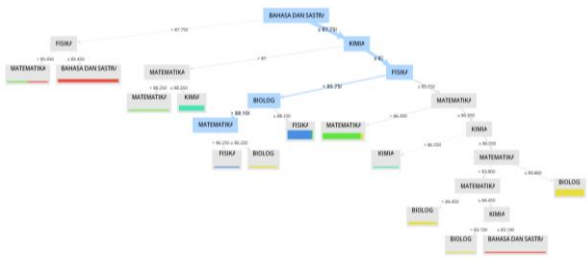
At the data preparation stage the data was taken from the academic website of the SMA Negeri 1 Bantul. Then simplify the header on the dataset so that it is easy in the classification process. After that, feature selection is carried out.

The next stage is preprocessing. At this stage, the data manipulation process is carried out by reducing data that has too high a mode on one of the labels and increasing data mode that is too low. This process is carried out so that there is no imbalanced data that can cause a tendency to one of the variables which can reduce accuracy in classification.

After the preprocessing stage, the data will enter the classification stage. At this stage the data is divided into training data and testing data. Then the classification is carried out using the naïve Bayes algorithm and decision tree. At this stage, the process of labeling data which is the target of evaluation is also carried out, namely the labeling of students' talents and weaknesses in each subject.

The last stage is the evaluation stage. At this stage, an evaluation of the algorithm used will be carried out with indicators including accuracy, precision and recall.

3 . Result and Discussion



(Figure 3.1 Decision Tree Flowchart)

After classifying the data using the naïve Bayes algorithm and decision tree, it is known that there are differences in accuracy results. In the naïve Bayes algorithm, the accuracy obtained is 85.09%, while in the decision tree algorithm the accuracy is 80.09% with the same classification data of 101 data.

	true FISIKA	true KIMIA	true MTK	true BIO	true BHS	class precision
pred. FIS	24	0	1	3	1	82%
pred. KIM	0	16	2	0	0	88%
pred. MTK	0	0	14	2	1	82%
pred. BIO	1	0	2	25	2	83%
pred. BHS	0	0	0	0	7	100%
class recall	96%	100%	73%	83%	63%	

(Table 3.1 Evaluation with confusion matrix)

4. Conslusions

Based on the results of data analysis using two algorithm methods. So it can be seen that the naïve Bayes algorithm and decision tree are suitable for analyzing student performance. Because the accuracy obtained is at 80% with a fairly high recall value and precision. Therefore the process of analyzing teacher and student performance can be carried out dynamically with the aim of evaluating teaching methods by teachers and analyzing students' talents and weaknesses in the subject. With this method, improving the quality of education can be measured and planned.

References

[1] Y. Mujiwati, “Pendidikan karakter dalam pembangunan bangsa,” *Peran. Pendidik. Karakter DALAM Pembang. Karakter BANGSA*, pp. 165–170, 2017.

[2] M. Tohir, “Hasil PISA Indonesia Tahun 2018,” *Pap. Mat.*, vol. 2, no. 1, pp. 1–2, 2019, doi: 10.31219/osf.io/pcjvx.

[3] T. Kadi and R. Awwaliyah, “Inovasi Pendidikan : Upaya Penyelesaian Problematika Pendidikan Di Indonesia,” *J. Islam Nusant.*, vol. 1, no. 2, pp. 144–155, 2017, doi: 10.33852/jurnalin.v1i2.32.

[4] A. Noviriandini and N. Nurajijah, “Analisis Kinerja Algoritma C4.5 Dan Naïve Bayes Untuk Memprediksi Prestasi Siswa Sekolah Menengah Kejuruan,” *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 5, no. 1, pp. 23–28, 2019, doi: 10.33480/jitk.v5i1.607.

[5] A. Saifudin, “Metode Data Mining Untuk Seleksi Calon Mahasiswa,” vol. 10, no. 1, pp. 25–36, 2018.

[6] F. Handayani and S. Pribadi, “Implementasi Algoritma Naive Bayes Classifier dalam Pengklasifikasian Teks Otomatis Pengaduan dan Pelaporan Masyarakat melalui Layanan Call Center 110,” *J. Tek. Elektro*, vol. 7, no. 1, pp. 19–24, 2015, doi: 10.15294/jte.v7i1.8585.