THE TITLE OF THE MANUSCRIPT DATA MINING FOR DETERMINING QUALIFICATION LEVEL OF THE HIGH SCHOOL INSTITUTIONS

Megawati Busrah[1*], Zahir Zainuddin[2], Hazriani[3]
[1*,3] Universitas Handayani, Makassar, Indonesia
[2] Univrsitas Hasanuddin, Makassar, Indonesia
egha.Palopo@gmail.com

Abstract

This study focuses on the use of data mining methods in educational classification. The heterogeneity within a group of schools can influence the achievement level of individual students and the overall evaluation of the school. One approach to improving the average school performance is by clustering schools with similar characteristics based on criteria such as accreditation, facilities, and performance. The resulting clusters are expected to minimize significant gaps, ensuring that the members of each group have relatively equal potential. By utilizing data mining clustering, a set of student groups with comparable abilities can be identified, allowing for targeted interventions. The results of this study will provide each school with information about their ranking. This method is used to classify and rank the data of schools. The ranking classification is performed using the SAW (Simple Additive Weighting) algorithm, with the highest performance scores.

Keywords: Data mining, Clustering, Classification, Cluster, Simple Additive Weighting

INTRODUCTION

Information systems, as modern tools, can be applied to support efforts in improving the performance of educational institutions, specifically in terms of the average National Examination scores. In addition to direct operational techniques such as providing extra study time, practice materials, and various forms of tryouts, it is also important to consider the conditions and capabilities of each student. A number of students trapped in unhealthy environments may experience a decline in their learning ability. This can be understood if students psychologically perceive a significant gap in their environment. Classification of student groups is believed to contribute to improving school performance by grouping students with similar potential and abilities, thus fostering self-confidence in each individual.

Data mining is used to cluster schools based on the quality and educational criteria they possess. This research will focus on reviewing the weight or quality of education in several Senior High Schools (SMA). The results of this review will then be used for clustering to determine the average weight score. This is considered to provide a projection of strategic learning steps aimed at achieving maximum results that can be felt by all students and schools in general. In conventional systems, the determination of sub-classes can easily be made by improving the quality of the results. However, this is considered less representative, as sometimes parents and class teachers only focus on the current period, which does not reflect the true achievements of the students. This approach is not ideal, as it may create the impression of "instant" students, which can be misleading.

Based on the explanation above, this research will process the qualification data of senior high schools using the SAW (Simple Additive Weighting) algorithm and the Naive Bayes algorithm to determine the weight of school scores and generate school rankings. The result of this research will be a web-based application that can be used by students to obtain recommendations on schools that can be chosen according to their preferences and capabilities.

METHODS

Data mining refers to data that is stored over a long period of time, and when this data is collected, it is found to have certain patterns after analysis. Data mining typically involves statistics, economics, weather forecasting, and various other types of data related to work. Data mining is the study of what has happened in the past and will be applied in the future to achieve better results. Specifically, the collection of methods known as data mining provides methodologies and technical solutions to address medical data analysis and construct predictive models.

A. Research Stages

The research method used is the experimental research method, with the following research stages in the research framework:

1. Data Gathering Method
   This research uses school data to predict the school level.
2. Data Pre-processing Method
   The acquired data is transformed to obtain relevant attributes that match the input format of the Naive Bayes algorithm.
3. Proposed Method
   The proposed model/method involves comparing the accuracy levels of the Naive Bayes algorithm, which can be used for prediction.
4. Method/Model Testing and Experimentation
   Testing is conducted by using part of the school data for training and the other part for testing.

B. System Description

The stages required in developing a program include analyzing the existing system. Typically, teachers and school administrative staff handle the administrative tasks, with the admin section responsible for recording data in a register book as an archive. Based on the analysis of the current system, the author provides a solution to the problem by designing and implementing the SAW (Simple Additive Weighting) algorithm to determine the classification of schools in Palopo.
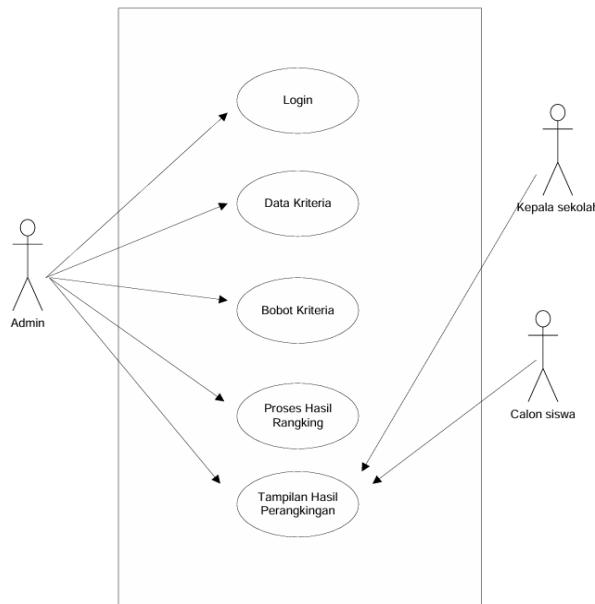
Figure 1 Use Case Diagram

The Use Case diagram above illustrates the workflow of the system to be developed, where there is one user, the Program Manager, and the admin staff. The Program Manager inputs the school data to be processed using the K-means algorithm. The system will then display the classification and ranking of the schools.

C. Implementation of the SAW (Simple Additive Weighting) Algorithm

The implementation process in the program is carried out by applying the SAW (Simple Additive Weighting) method to solve the problem of determining the highest ranking. The SAW method is also known as the weighted summation method. The basic concept of the SAW method is to find the weighted sum of the performance ratings for each alternative across all attributes. The SAW method requires the normalization of the decision matrix (X) to a scale that can be compared with all the existing alternative ratings.

1. Determining the criteria as the sum of the data specified as new objects. The training data is as follows

Atribut	Kriteria		
Atribut	K.1	K.2	K.3
A1	70	90	70
A2	70	70	80
A3	70	70	70
A4	70	80	90
A5	70	90	90
A6	70	90	90
A7	90	100	90
A8	120	100	90
A9	120	100	90
A10	90	70	90
A11	120	100	90
A12	60	100	90
A13	60	100	90

2. Calculating the weight of the new matrix data against all the data that has been trained.

```
$sql="SELECT kode, nama_sekolah1, p_fasilitas, p_pembayaran, p
_akreditasi,
    b_fasilitas, b_pembayaran, b_akreditasi
                FROM kriteria, bobot ORDER BY p_fasilitas D
ESC, p_pembayaran DESC, p_akreditasi DESC";
        $result=mysqli_query($konek_db,$sql) or die(mysql_error(
)); //row melihat dari sql
        $rangking =0;
```

RESULT AND DISCUSSION

A. Implementation of the Naive Bayes Algorithm
   Algoritma ini menggambarkan tentang bagaimana cara menghubungkan data ke algoritma Naive Bayes
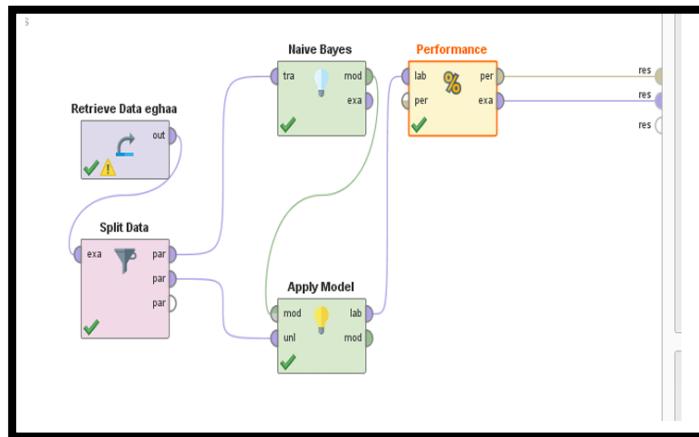
Figure 2 Naive Bayes Algorithm Model

The image above explains how to determine evaluation criteria to measure the performance of an algorithm

Figure 3 Implementation of the Naive Bayes Algorithm

The image shows the results of implementing the Naive Bayes algorithm using RapidMiner, where the processed data displays classification predictions for each entry based on predefined criteria such as facilities, accreditation, and payment. Each prediction is accompanied by a confidence level, indicating the model's certainty about the classification result. In the image, the confidence values show the level of certainty the model has regarding the predicted classifications, with some predictions showing very high confidence, reaching 1.000, indicating that the model is very confident in its predictions. These results are used to assess the qualification of schools based on data analyzed using the Naive Bayes algorithm.

B. System implementation
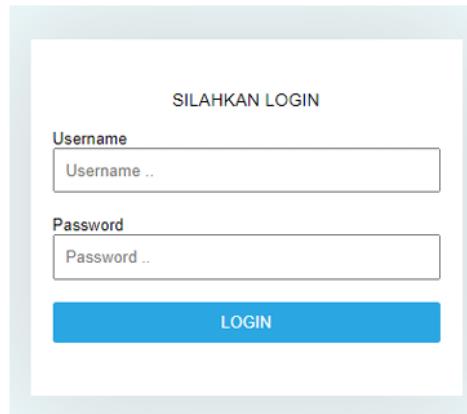
1. Login Screen

Figure 4 Login Screen

The image above shows the login form screen, which functions to display the login form. The admin must input the username and password in order to log in and access the main menu page

2. Admin Main Menu Screen

Figure 5 Admin Main Menu Screen

The image above shows the main menu screen, which functions to display the admin's main menu page after a successful login. On this main page, there are menus for Input Alternative Data, Input Points Data, Input Weights Data, View Alternative Data, View Weights Data,

CONCLUSION

Based on the research results, the implementation of the SAW (Simple Additive Weighting) algorithm for school ranking can be used to predict school improvements. It can be concluded that: 1. The SAW classification method for new objects works by determining the improvement in the level of high schools using an approach to calculate the proximity between new and old cases based on training data, and then ranking the new cases according to the old ones. 2. The use of the SAW-based classification application for ranking schools can assist the program administrators in ranking high schools in Palopo. 3. Based on the implementation of the SAW (Simple Additive Weighting) algorithm, it can be used to predict rankings specifically for high schools.

REFERENCES

Ahmad Setiadi[*1] , Yunita[2] , Anisa Ratna Ningsih[3] [1,2]Program Studi Manajemen Informatika, Penerapan Metode Simple Additive Weighting(SAW)

Bastian, A., Sujadi, H., Febrianto, G., Studi, P., Informatika, T., Majalengka, U., … No, M. (2009). Penerapan Algoritma. Jsi.Cs.Ui.Ac.Id, 14(1), 26–32. Retrieved from http://jsi.cs.ui.ac.id/index.php/jsi/article/view/566

Diana Laily Fithri, Eko Darmanto (2014) sistem pendukung keputusan untuk memprediksi kelulusan mahasiswa menggunakan metode naïve bayes Program Studi Sistem Informasi, Fakultas Teknik, Universitas Muria Kudus

Efendi, M. S., & Wibawa, H. A. (2018). Prediksi Penyakit Diabetes Menggunakan Algoritma ID3 dengan Pemilihan Atribut Terbaik (Diabetes Prediction using ID3 Algorithm with Best Attribute Selection). Juita, VI, 29–35.

Hartama, A. A. K. (2017). Klasifikasi Penyakit Hipertensi Menggunakan Algoritma C4.5 Studi Kasus RSU Provinsi NTB.1–154.

Maliha, D. Z., Santoso, E., & Furqon, M. T. (2019). Penerapan Metode Neighbor Weighted K-Nearest Neighbor Dalam Klasifikasi Diabetes Mellitus. 3(3), 2910–2915.

Melisa Elistri, Jusuf Wahyudi, Reno Supardi Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Dehasen Bengkulu. penerapan metode saw dalam sistem pendukung keputusan pemilihan jurusan pada sekolah menengah atas. (2014)