Classification of Graduation Students of the Faculty of Computers and Multimedia, Universitas Islam Kebangsaan Indonesia Using the Naïve Bayes Classifier Algorithm

Mutasar[1], T. Irfan Fajri[2], Novia Hasdyna[3]

[1,2,3]Informatics Study Program, Faculty of Computers and Multimedia, Universitas Islam Kebangsaan Indonesia, Indonesia

Corresponding Author: T. Irfan Fajri

ABSTRACT

The purpose of this study to discuss classification of graduation Students of the Faculty of Computers and Multimedia, Universitas Islam Kebangsaan Indonesia using the naïve bayes classifier algorithm. The methodology used in this study uses a unified modeling language. The results show that the data mining system for classifying graduation Students of the Faculty of Computers and Multimedia, Universitas Islam Kebangsaan Indonesia succeeded in processing datasets, namely training data and sample data to form classes in the classification graduation Students of "Graduated on Time" or "Graduated Late". Processing data mining with the naïve bayes classifier algorithm is able to describe and draw conclusions on large amounts of datasets even though the data is complicated to understand. Research with this web-based system is easy for academics or faculties to use and data processing can be done quickly.

Keywords: Classification, Naïve Bayes Classifier, Algorithm

INTRODUCTION

The Faculty of Computers and Multimedia is one of the faculties within the scope of the Universitas Islam Kebangsaan Indonesia which is one of the private higher education campuses under the auspices of LL13Dikti Aceh which is based in Bireun Regency and Lhokseumawe City. The Faculty of Computers and Multimedia, Universitas Islam Kebangsaan Indonesia has accepted many independent and bidikmisi students so that on the bidikmisi pathway it is highly prioritized for a timely completion of study, namely eight semesters, considering that only during this time period students are freed from the responsibility of studying fees and the need for supporting accreditation.

In addition to this, timely graduation of students is a priority of one of the main indicators of achieving success as measured by the Ministry of Education and Culture, especially during the independent learning campus program. In this study implementing a system to be able to classify or predict student graduation on time with data mining techniques, one of which is by using the naïve bayes classifier algorithm (Dinata et al., 2021).

The naïve bayes classifier algorithm is one of several machine learning techniques that perform probability calculations as well as statistics. This theory was originally coined by a British scientist, namely Thomas Bayes. Thomas Bayes predicts probabilities for the future based on past experience (Bayes et al., 2019). The naïve bayes classifier algorithm is also a probabilistic classification method that can be calculated based on Bayes' theorem using the

assumption that there is no relationship between attributes in the classification process (Annur, 2018; Gandhi, et al. 2021). Classification is the process of finding a model that describes and distinguishes data classes, or how to classify data into one or several predefined classes (Handoko and Neneng, 2021; Indrayuni, 2019). Classification techniques that are widely used include neural networks, rough sets, K-nearest neighbors, bayesian classifiers, and others (Gunawan et al. 2018; Pujianto and Ristanti, 2019).

To be able to draw conclusions from the Student graduation data dataset to be classified based on class whether students "Graduated on Time" or "Graduated Late" so that this information is very useful for faculties to be able to make decisions and policies for teaching and learning development at the Faculty of Computers and Multimedia.

In processing data mining, the classification of student graduation is designed using a web-based program, namely PHP. While the dataset used to draw conclusions is the graduation data of students who have graduated, namely both student data graduating on time or late. This data will become training data to be processed as a dataset so that it can be classified against sample data, namely undergraduate students who will graduate later.

It is hoped that this system will make it easier for the faculty to observe and predict student graduation for policy making and decisions in controlling the progress of teaching and learning.

The purpose of this study to discuss classification of graduation Students of the Faculty of Computers and Multimedia, Universitas Islam Kebangsaan Indonesia using the naïve bayes classifier algorithm.

RESEARCH METHODS

The unified modeling language is a methodology for developing OOP systems and a set of tools to support systems development.

1. Use Case Diagram

Use case diagram describe how the processes are carried out by actors on a system, as shown in Figure 1:

Figure 1. Use Case Diagram
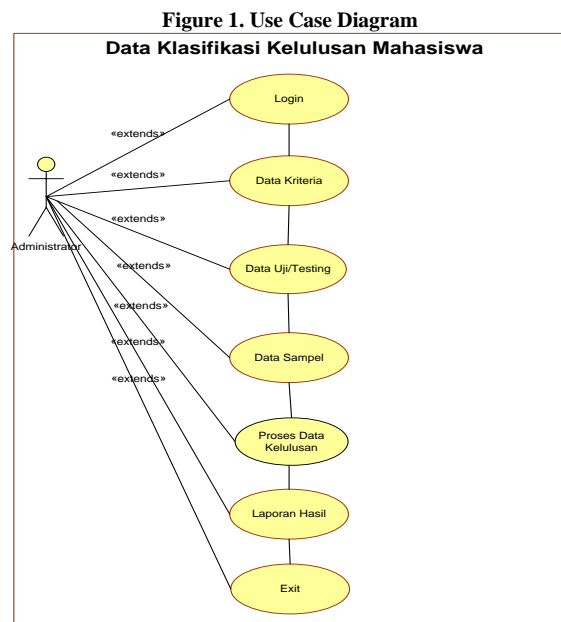
2. Activity Diagram

Activity diagram is diagram that describe the various activity flows in the system being designed, how each flow begins, the results that may occur, and how it ends, as shown in Figure 2.
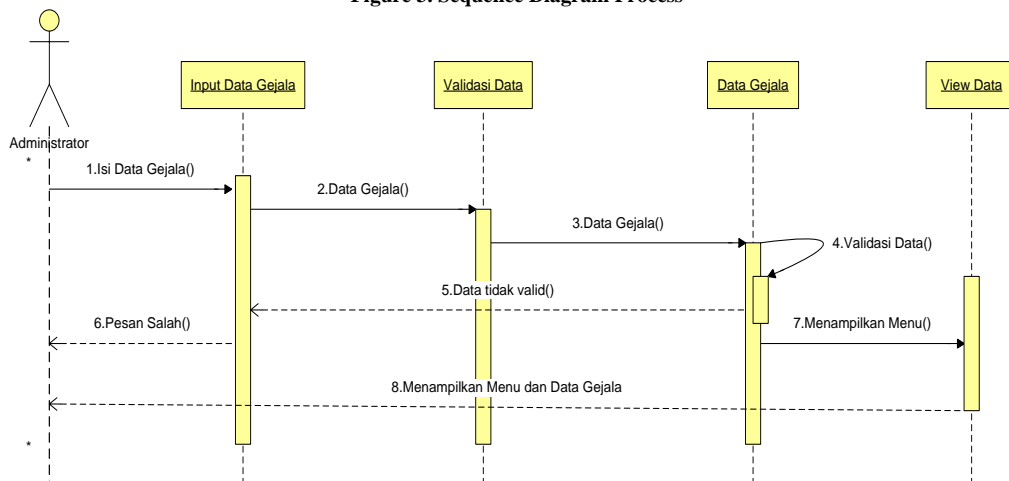
Figure 2. Activity Diagram
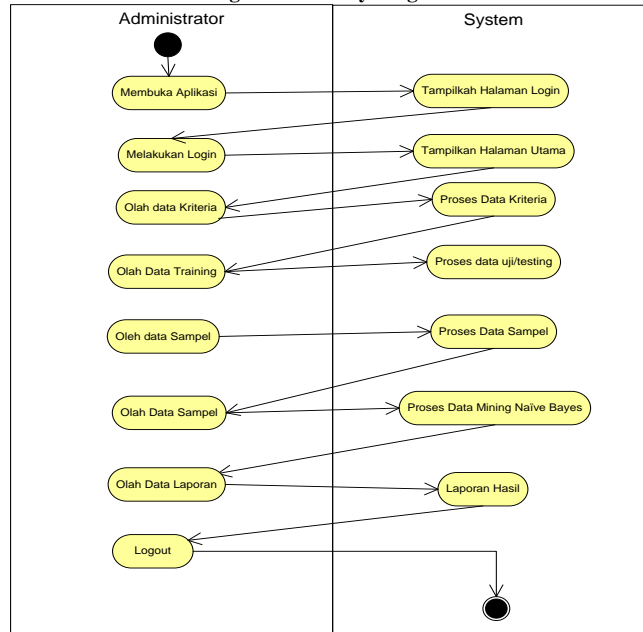
3. Sequence Diagram Process

The design of the data mining sequence diagram naïve Bayes classifier illustrates how the actions of actors on objects and data in using the system in the process of inputting data for calculation data, as shown in Figure 3.

Figure 3. Sequence Diagram Process

RESULT

1. Database Design

Table design in this web application uses a My-SQL database where the author creates a database consisting of 6 Tables, namely:

Table 1. Admin

No.	Field Name	Type Field	Long	Information
1	username	Varchar	20	Username (Primary Key)
2	password	Varchar	20	User Passwords

The admin table is used to store system user admin names.

Table 2. Criteria

No.	Field Name	Type Field	Long	Information
1	Id	Int	11	Criteria id (Primary Key)
2	kdkriteria	Varchar	10	Criteria Code
3	Kriteria	Varchar	200	Criteria Name
4	Subkriteria	Varchar	50	Sub Criteria

The criteria table is used to store criteria data for measuring student achievement.

Table 3. Alternative

No.	Field Name	Type Field	Long	Information
1	kdalternatif	Varchar	5	Alternative Code (Primary Key)
2	alternative	Varchar	100	Alternative Name
3	keterangan	Varchar	200	Description

The alternative table is used to store alternative data, namely students.

Table 4. Criteria Value

No.	Field Name	Type Field	Long	Information
1	idnilai kriteria	int	3	Kd Criterion Value
2	Kdalternatif	Varchar	5	Alternative Code
3	Sub_kriteria	Varchar	50	Sub Criteria
4	Kdkriteria	Varchar	10	Criteria Code
5	Keterangan	Varchar	20	Description

The criteria value table is used to input criteria value data between criteria and alternatives.

Table 5. Sample Value

No.	Field Name	Type Field	Long	Information
1	Idnilai_sampel	Int	3	Id Value (Primary Key)
2	Kdsampel	Varchar	5	Kd Sample
3	Nama_sampel	Varchar	20	Sample Name
4	Sub_kriteria	Varchar	50	Sample Value
5	Kdkriteria	Varchar	10	Criteria Code
6	Keterangan	Varchar	20	Information
7	Klasifikasi	Varchar	20	Achievement Classification

The sample value table is used to store sample value data.

Table 6. Classification Result

No.	Field Name	Type Field	Long	Information
1	Idhasil	Int	4	Classification Result Id (Primary Key)
2	Kdalternatif	Varchar	10	Alternative Code
3	Klasifikasi	Varchar	5	Shortest Distance Rating

The classification result table is used to accommodate the classification result data after the naïve bayes process.

2. Classification Results with the Naïve Bayes Classifier Algorithm

In this study, the process of classifying graduation the Faculty of Computers and Multimedia, Universitas Islam Kebangsaan Indonesia students used the naïve bayes classifier algorithm, which is a web system for processing data mining on training data, namely students who have passed the sample data, namely students who are still undergoing studies to be able to measure abilities and clustering their achievements based on the values obtained are on the criteria that have been determined by campus policy, namely in this case gender, entry point, grade point value per semester 1-5. The system interface display can be seen in the following Figure:

Figure 4. Criteria Data Input Process
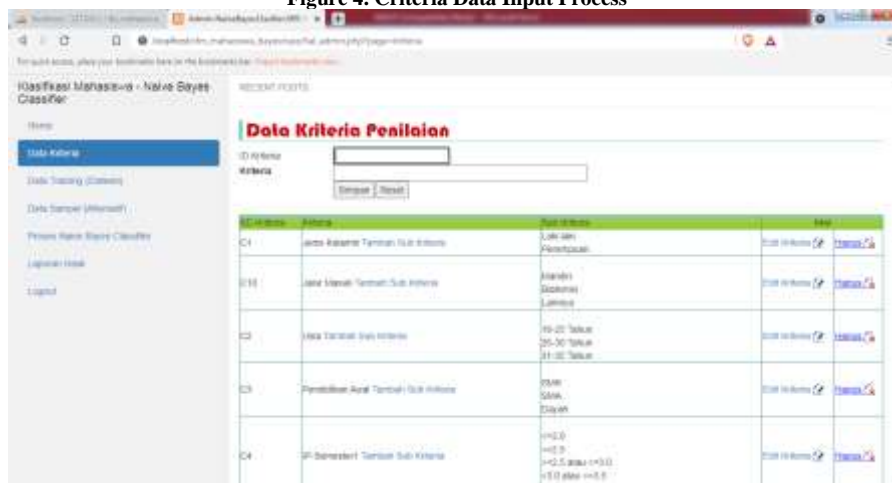
In Figure 4, the assessment criteria data can be inputted by the admin by entering the criteria code, criteria, sub criteria and actions consisting of edit and delete.

Furthermore, in Figure 5 is the process of inputting sample data which displays information on sample value data for each student against the criterion value, in this form there is a button for adding training data or sample data to input sample data. In Figure 6-8, the form of the naïve bayes classifier algorithm process is the main process for analyzing and implementing data mining the naïve bayes classifier algorithm to determine outstanding students. On this page there is a combobox option to determine student sample data (X) for the naïve bayes classifier algorithm calculation process so that it can be concluded that class is related to the sample data.

Figure 5. Data Training Input Process

Figure 6. The Process of Determining Sample Data in the Naïve Bayes Classifier

Figure 7. Class Determination Process in Naïve Bayes Classifier

Figure 8. Student Graduation Classification Results Report

The results show that the data mining system for classifying graduation Students of the Faculty of Computers and Multimedia, Universitas Islam Kebangsaan Indonesia succeeded in processing datasets, namely training data and sample data to form classes in the classification graduation Students of "Graduated on Time" or "Graduated Late". Processing data mining with the naïve bayes classifier algorithm is able to describe and draw conclusions on large amounts of datasets even though the data is complicated to understand. Research with this web-based system is easy for academics or faculties to use and data processing can be done quickly.

CONCLUSION
The results show that the data mining system for classifying graduation Students of the Faculty of Computers and Multimedia, Universitas Islam Kebangsaan Indonesia succeeded in processing datasets, namely training data and sample data to form classes in the classification graduation Students of "Graduated on Time" or "Graduated Late". Processing data mining with the naïve bayes classifier algorithm is able to describe and draw conclusions on large amounts of datasets even though the data is complicated to understand. Research with this web-based system is easy for academics or faculties to use and data processing can be done quickly.

REFERENCES
1. Annur, H. (2018). Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes. *ILKOM Jurnal Ilmiah, 10*(2), 160-165.
2. Bayes, K. B. M. A. N., & Boosting, S. F. D. (2019). *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 3*(2), 227-232.
3. Dinata, R. K., Hasdyna, N., & Alif, M. (2021). Applied of Information Gain Algorithm for Culinary Recommendation System in Lhokseumawe. *Journal Of Informatics And Telecommunication Engineering, 5*(1), 45-52.
4. Gandhi, B. S., Megawaty, D. A., & Alita, D. (2021). Aplikasi Monitoring Dan Penentuan Peringkat Kelas Menggunakan Naïve Bayes Classifier. *Jurnal Informatika Dan Rekayasa Perangkat Lunak, 2(*1), 54-63.
5. Gunawan, B., Sastypratiwi, H., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. *JEPIN (Jurnal Edukasi dan Penelitian Informatika), 4*(2), 113-118.
6. Handoko, M. R., & Neneng, N. (2021). Sistem Pakar Diagnosa Penyakit Selama Kehamilan Menggunakan Metode Naive Bayes Berbasis Web. *Jurnal Teknologi Dan Sistem Informasi, 2*(1), 50-58.
7. Indrayuni, E. (2019). Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan

Algoritma Naive Bayes. *Jurnal Khatulistiwa Informatika, 7*(1).

8. Pujianto, U., & Yuni Ristanti, P. (2019). Perbandingan Kinerja Metode C4. 5 dan Naive Bayes dalam Klasifikasi Artikel Jurnal PGSD Berdasarkan Mata Pelajaran. *TEKNO: Jurnal Teknologi Elektro dan Kejuruan, 29*(1), 50-67.

******