Informatics and Software Engineering
p-ISSN/e-ISSN: xxxx-xxxx/xxxx-xxxx
Homepage: https://sanscientific.com/journal/index.php/ise
1(1) 30-34 (2023)
DOI: 10.58777/ise.v1i1.61

Application of Naive Bayes in Assessing Student Satisfaction with Facilities and Services at SMK Dastamaco

Deny Novianti
Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika

Received: June 16, 2023; Accepted: June 25, 2023

Abstract
Education is one of the important factors that are highly considered. Starting from the selection of quality schools and also the background of good teachers. However, sometimes to get good facilities, it is also balanced with many costs. Only parents who can afford it will not be a problem, but not so with parents of students whose economic conditions are middle to lower. SMK Dastamaco strives to provide the best service and quality for students so that parents feel safe and comfortable having entrusted their children to school. Usually, it is rare for schools to evaluate students' satisfaction with the service and quality of the existing school. Therefore, the school also needs more information about which side needs to be improved so that the school can be even better later. Based on the above problems, it is necessary to hold a differentiating classification to find out how many students are satisfied with the service and quality of SMK Dastamaco. Moreover, to classify it, the Naive Bayes method can be used, where several criteria are later used to determine whether students are satisfied or dissatisfied with the service and quality of the school. The test results were conducted with 191 data and as many as 27 as test data. An accuracy pattern of 0.556 stated Satisfied, and 0.222 expressed dissatisfaction and dissatisfaction.

Keywords: Naive Bayes, Classification, SMK, Facilities, Services

Abstrak
Pendidikan merupakan salah satu faktor penting yang sangat difikirkan. Mulai dari pemilihan sekolah yang berkualitas dan juga latar belakang pengajarnya yang baik. Namun, terkadang agar mendapatkan fasilitas yang bagus, tentu pula diimbangi dengan biaya yang tidak sedikit. Mungkin baru para orang tua murid yang mampu tidak akan menjadi masalah, tapi tidak begitu dengan orang tua murid yang kondisi perekonomiannya menengah ke bawah. SMK Dastamaco terus berupaya untuk memberikan pelayanan dan kualitas yang terbaik bagi para siswa, agar para orang tuanya merasa aman dan nyaman telah menitipkan anaknya untuk bersekolah. Biasanya jarang ada sekolah yang melakukan evaluasi terhadap kepuasan para siswa terhadap pelayanan dan kualitas sekolah yang ada. Maka dari itu, pihak sekolah juga kurang mendapatkan informasi tentang sisi manakah yang perlu diperbaiki agar sekolah bisa jadi lebih baik lagi nantinya. Berdasarkan masalah di atas perlu diadakannya klasifikasi pembeda untuk mengetahui seberapa banyak siswa yang sudah merasa puas terhadap pelayanan dan kualitas dari SMK Dastamaco. Dan untuk mengklasifikasikannya dapat menggunakan metode Naive Bayes, dimana nantinya digunakan beberapa kriteria untuk dapat mengetahui apakah siswa merasa puas atau tidak puas terhadap pelayanan dan kualitas sekolah. Dari hasil pengujian yang dilakukan dengan data sebanyak 191, dan sebanyak 27 sebagai data tes. Didapatkan pola akurasi sebesar 0, 556 yang menyatakan Puas, serta 0, 222 yang menyatakan kurang puas dan juga tidak puas.

Kata Kunci: Naive Bayes, Klasifikasi, SMK, Fasilitas, Pelayanan

*Corresponding author: Deny Novianti (denynov.dov@bsi.ac.id)

1. Introduction

With the development of the times, education is one of the important factors that are highly considered. Starting from the selection of quality schools and also the background of good teachers. The school environment is also one of the supporting factors for the school to be considered good. Therefore, both the government and the school sides continue to strive to create the best atmosphere for students to study comfortably. However, sometimes to get good facilities, it is also balanced with many costs. Usually, it is rare for schools to evaluate students' satisfaction with the service and quality of the existing school. Therefore, the school also needs more information about which side needs to be improved so that the school can be even better later. Based on the above problems, it is necessary to hold a differentiating classification to find out how many students are satisfied with the service and quality of SMK Dastamaco, and the Naive Bayes Method is one method that can be used. The Bayesian Naive Method is one of the methods of classification and branching of artificial intelligence; these activities will be in the form of a Ranking, namely Satisfied, Dissatisfied, and Dissatisfied, so that the school can find out whether the students are satisfied with the facilities and services in the school.

Furthermore, from research that has been done previously using the Naïve Bayes algorithm method, it is stated that this method can be used in this case. According to research (Muslehatin & Ibnu, 2017), the test results showed an accuracy of 66.67%, i.e., 16 respondents were at risk for moderate obesity. Sixty-nine respondents had normal nutritional levels, and three respondents were malnourished. This obesity must be considered because students who are obese have an 80% chance of being obese as adults or old age. Meanwhile, according to (Alfa Saleh, 2015), the naïve Bayes method is expected to predict the amount of electricity used in each household to make it easier to manage electricity usage. Of the 60 household electricity usage data tested by the Naïve Bayes method, a percentage result of 78.3333% was obtained for prediction accuracy, whereas from 60 household electricity usage data tested, 47 household electricity usage data were successfully classified correctly.

2. Methods

There are two types of data types; the first is Primary Data, namely data obtained directly from the object of research (Andi Prastowo, 2012). Secondary data is obtained from literature, reference books, and internet browsing (Andi Prastowo, 2012).

Data mining, often also called Knowledge Discovery In Database (KDD), is an activity that includes the collection and use of historical data to find patterns or relationships in large data sets. The output of this data mining can be used to improve decision-making in the future. So the term Pattern Recognition is rarely used because it is part of data mining (Santosa, 2007).

Classification is the process of finding a model or function that explains or differentiates a concept or class of data, intending to be able to estimate the class of an object whose label is unknown. In achieving this goal, the classification process forms a model that can distinguish data into different classes based on certain rules or functions. The model can be an "if-then" rule, a decision tree, or a mathematical formula (Bustami, 2014).

Figure 1. Classification Model Diagram

Bayes decision theory is a fundamental statistical approach to pattern recognition based on quantifying trade-offs between classification decisions using the probabilities and costs incurred in those decisions (Santosa, 2007). In addition, Bayesian classification can also predict a class's membership probability. Bayes' theorem has similar classification capabilities to decision trees and neural networks. Bayesian classification is proven to have high accuracy and speed when applied to databases with large data (Jananto, 2013). Bayes' theorem has the following general form:

$$P(H \mid X) = \frac{P(X \mid H)P(H)}{P(X)} \quad \text{.....(1)}$$

Figure 2. Bayes' Theorem

Information:
X = Data with an unknown class
H = Hypothetical data X is a specific class
P(H| X) = Probability of hypothesis H based on condition x (posterior prob.) P(H) = Probability of hypothesis H (prior prob.)
P(X|H) = Probability X based on the condition
P(X) = Probability of X (Jananto, 2013)

The methods used in this study are:
- The data collected from the questionnaire results were 191, and those used for the calculation process were 27.
- The method used for the calculation is the Naive Bayes Algorithm

Rapid Miner is a tool used in techniques that are in the environment of machine learning, data mining, text mining, and predictive analytics (Nugroho, 2014).

3. Results and Discussion

3.1. Determination of Criteria

a. Class: X, XI, XII
b. Type: Asset, Pasiva, Vario, Beat
c. Department: Accounting, Motorcycle Engineering
d. Gender: Male,
Woman
e. Classification: Satisfied, Dissatisfied, and Dissatisfied

3.2. Naive Bayes Calculation Analysis

Here is a table of the result set data questionnaire at SMK Dastamaco. The total consists of 191 data, but Calculations only used as much as 27 data.

Table 1. Data Set Results of Student Satisfaction Assessment Questionnaire on Facilities and Services in SMK Dastamaco

No	Nama	Kela	Tip	Jurusan	Jenis Kelami	Klasifikas
1	Azril Fajar A.	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
2	Dicky Sugara	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	KP
3	Alfian Putra Wangsa	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
4	Tita Putra Wiyana	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	TP
5	Ikbal Rangga Pratama	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
6	Ahmad Rayyan Muhiffal	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	KP
7	Zulhan Efendi P.	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
8	Erwin S.	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	KP
9	Ahmad Zidan	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	TP
10	Ahmad Ozahman	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
11	Aliyasah	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
12	Ari Hidayat	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
13	Islamy Ikhsan R.	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
14	Mochammad Afryan H.	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	TP
15	Wahyu Prasetyo	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
16	Febi Trio	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	KP
17	Dendy Zuhri	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
18	Oky Kurniawan	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	KP
19	Didik Arifudin	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	TP
20	Karno Lesman	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
21	Jules Tambunan	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
22	Nurkholis	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	TP
23	Rizki Wahyudi	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
24	M. Syaiful Arief	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	TP
25	Putra Maulana	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	KP
26	Rifan Mananta	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P
27	Wardan Christian	XI TSM	Vario	Teknik Sepeda Motor	Laki-Laki	P

Naive Bayes Calculation:

a. Calculating the Number of Classes/Labels (Ranking)
   P (Satisfied) = 15/27
   P (Dissatisfied) = 6/27
   P (Dissatisfied) = 6/27

b. Count the same number of cases with the same class (by type)
   P = (Type = Vario | Y = Satisfied) = 15/27
   P = (Type = Vario | Y = Less Satisfied) = 6/27
   P = (Type = Vario | y = dissatisfied) = 6/27

c. Variable Profit Share Satisfied, Less Satisfied, and Dissatisfied

- P (TypeSatisfied)
  = 15/27
  = 0.556
- P (TypeLess Satisfied)
  = 6/27
  = 0.222
- P (TypeDissatisfied)
  = 6/27
  = 0.222

From the calculation results above, the Bayes classification for Ranking from the 27 data obtained is "Satisfied" with a value of 0.556 compared to Less Satisfied and Dissatisfied, which both have a value of 0.222.

3.3. Calculation Using Rapid Miner

Here is the data processing using the Rapid Miner Application.

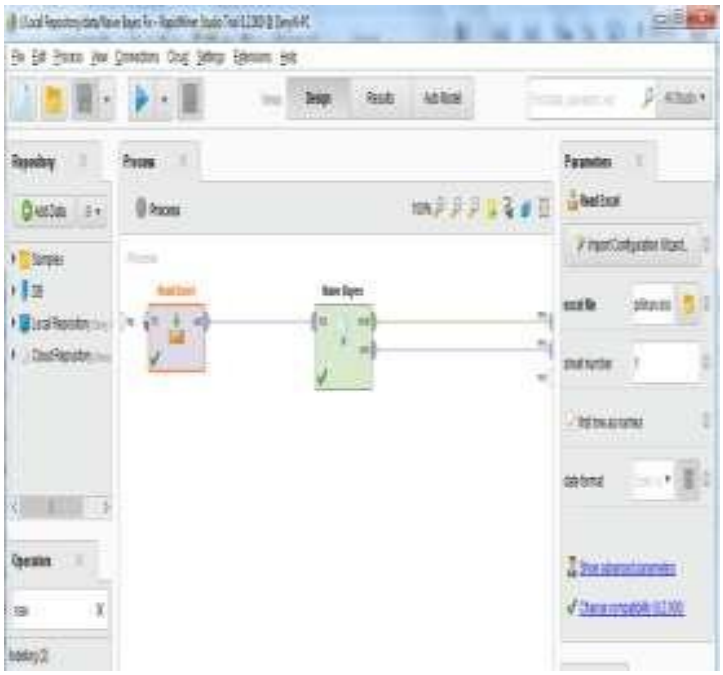

Figure 2. Process Data

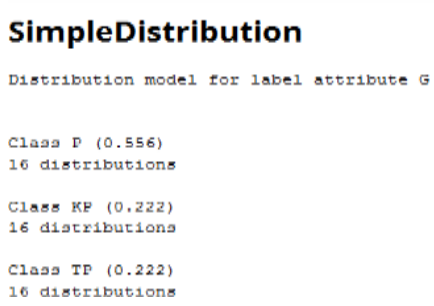

Figure 3. Simple Distribution

Based on the above data, using the Riped Miner Application can show the same results as the calculation of the formula manually. That is the value of "Satisfied" of 0.556, then the value for "Less Satisfied" and "Not Satisfied" is 0.222.

4. Conclusion

The results of the classification calculation show that the students have been satisfied with the facilities and services at SMK Dastamaco, with a satisfied score of 0.556. Manual calculations using Ms. Excel with calculations using the Riped Miner Application show the same results. We get maximum results, and it is expected that the determination of the criteria can be more varied. It is expected to be developed in an application using a programming language so that its use becomes more effective. Different algorithms can also be used for the calculation of this case in order to compare results better.

References

Alfa Saleh. (2015). Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Citec Journal*, *2*(3), 207–217. https://doi.org/10.20895/inista.v1i2.73

Andi Prastowo. (2012). *Metode Penelitian Kualitatif dalam Perspektif Rancangan Penelitian*. Yogyakarta.

Bustami. (2014). Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi. *Techsi: Jurnal Penelitian Teknik Informatika*, *8*(1), 127–145. https://repository.nusamandiri.ac.id/index.php/repo/viewitem/13890

Jananto, A. (2013). Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa. *Teknologi Informasi*, *18*(1), 9–16.

Muslehatin, W., & Ibnu, M. (2017). Penerapan Naïve Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi UIN Suska Riau. *Seminar Nasional Teknologi Informasi, Komunikasi Dan Industri (SNTIKI)*, 2579–5406.

Nugroho, Y. S. (2014). *Data Mining Menggunakan Algoritma Naïve Bayes Untuk Klasifikasi Kelulusan Mahasiswa Universitas Dian Nuswantoro. (Studi Kasus: Fakultas Ilmu Komputer Angkatan 2009).*

Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*.