

APPLICATION OF DATA MINING USING THE NAÏVE BAYES CLASSIFIER METHOD TO ANALYZE THE LEVEL OF CUSTOMER SATISFACTION IN ICE CREAM MIXUE

Anita^{*1}, Deslin Fitri Y. Lumban Gaol², Meta Doner Septia Sipayung³

^{1,2,3}Faculty of Science and Technology, Prima University of Indonesia Jl. Sampul No. 4, Medan Petisah, Medan City, North Sumatra

E-mail : anitayakub_pichan@yahoo.com

ABSTRACT- Ice Cream Mixue is a company engaged in the production of beverages and food. They provide a wide range of ice cream flavors with premium quality and natural ingredients, attracting many customers' attention. This study uses a quantitative method because the data sources are apparent. The research method is a scientific way to obtain data with specific purposes and uses. The technique used is the Naïve Bayes Classifier Algorithm with rapidminer results. 356 customers answered the questionnaire results satisfied, 64 customers who answered the questionnaire results were neutral, and 80 customers responded to the questionnaire results and were not happy. In this case, the Naïve Bayes Classifier Method can predict the level of customer satisfaction for ice cream mixes using the Rapidminer Application with Satisfied Words. This analysis obtained Class Precision satisfaction results of 94.72%, pred. neutral 2.0%, and pred. They are dissatisfied by 4%.

Keywords: Ice Cream Mixue, Naïve Bayes, Rapidminer

1. INTRODUCTION

Ice Cream Mixue is a company engaged in producing beverages and food. They provide a wide range of ice cream flavors with premium quality and natural ingredients, attracting many customers' attention. As a customer-oriented company, a deep understanding of the level of customer satisfaction is very important. Therefore, the application of data mining techniques becomes relevant to analyze and understand the factors that influence the level of customer satisfaction in Ice Cream Mixue. Ice cream is a processed drink made from fresh cow's milk and skim milk that has undergone a freezing process.[1][2]

As the business grew and the popularity of Ice Cream Mixue, so did the amount of data generated by the company. The data includes customers' taste preferences, purchase locations, frequency, and customer feedback on products and services. Manual analysis of this amount of data needs to be more realistic and tends to overlook more complex and hidden patterns.

The rapid development of data mining (DM) cannot be separated from the development of information technology, which allows large amounts of data.[3] *Data mining* is finding hidden patterns from a set of unknown data.[4][5] The Naïve Bayes Classifier method is a statistical classifier that can predict class membership probabilities. Naïve Bayes Classifier is a basic theorem with classification capabilities similar to decision trees and neural networks. This classification is proven to have accuracy and speed when applied to databases[6].

Through data mining using the Naïve Bayes Classifier method, Ice Cream Mixue is expected better to understand its customers' preferences and satisfaction levels.[7] With this deeper understanding, companies can optimize business strategies, improve product and service quality, and maintain positive customer relationships for long-term growth and success.

The Naïve Bayes classifier is a classification method based on probability and the Bayesian Theorem, assuming that each X variable is independent or independent and has nothing to do with other variables.[8]

2. RESEARCH CONTENT

This study uses a quantitative method because the data sources are apparent. The research method is a scientific way to obtain data with specific purposes and uses.[9]

This study uses a questionnaire that has been collected to collect information about customer satisfaction with ice cream Mixue. This research aims to improve customer service quality by collecting information using the Naïve Bayes Classifier data mining method.[10]

2.1 Data Identification

This study uses a dataset of mixed customer satisfaction, assessed objectively based on the dataset where each piece of information has a total of 500 rows of data with 15 columns.

Table 2.1 Mixue Questionnaire Question Data

No	Questionnaire	Total Questionnaire
1	Is the appearance of the ice cream mixue waitress neat and clean	701
2	How is the ice cream service served in a friendly and timely manner?	1008
3	Building is neat and clean	1019
4	Yanti Clean, tidy	958
5	Meta the place where ice cream mixue is provided is comfortable and clean	684
6	Mixue ice-cream products have affordable prices for all groups.	927
7	Mixue ice cream products	991
8	Does it provide information for details on purchasing menu types?	966
9	Does the officer respond to customer complaints quickly and precisely	1121
10	The serving utensil used to serve ice cream mixue is safe to use (serving knife is not damaged)	1071
11	Timely serving of dishes (no long waits)	1305
12	Clean serving utensils (free of dirt, dust, etc.)	1151
13	Waiters give a polite attitude when customers want to make payments or refunds	996
14	Ice cream mixue provides good quality remote ordering (GoFood)	896
15	Are you satisfied with the quality of the ice cream mixue product	1023

Table 2.1 Display of Question Results

JK	P1	P2	P3
23	1	2	2
21	4	3	1
32	3	6	4
22	3	3	3
28	1	2	1

2.2 Research Design

This study was designed with observational data collection and calculations. After the data is entered into Microsoft Excel, Rapid miner follows the steps of the Naive Bayes Method. The company with the same name develops data science software called Rapid miner. The platform offers a unified environment for machine learning, text mining, and predictive analysis (predictive analytics).[11][12]. The research design is as follows.

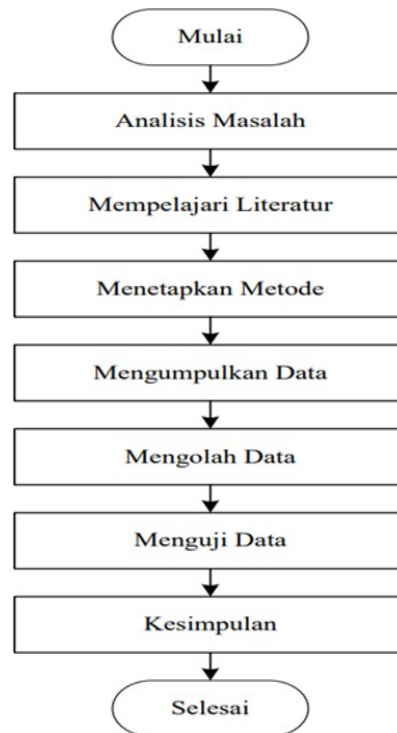


Figure 2.2 Research design

Information:

- 1) Problem Analysis is analyzing a problem related to mixed customer satisfaction, which is assessed objectively based on the dataset. The method used is the Naïve Bayes Classifier Algorithm which can be applied to predict the level of customer satisfaction in Ice Cream Mixue. Studying the Literature, This research must be based on several references that can be used to obtain information in research.
- 2) Establishing Methods Establishing methods for solving problems. This study uses the Naïve Bayes Classifier method
- 3) Collecting data on a list of proposed customer names to predict customer satisfaction with ice cream mixture was obtained from question-and- answer interviews by employees with customers.
- 4) Data processing using data mining in the Naïve Bayes Classifier Algorithm method.
- 5) Testing Data Data testing is carried out using RapidMiner.5.3 tools.
- 6) The conclusions obtained in classifying service results on Ice Cream Mixue.

2.3 Data Mining

Extracting hidden information, patterns, and essential relationships from large and complex data sets is known as data mining.[13]In data mining, one of the Naive Bayes classification methods often used is the analysis of customer satisfaction or sentiment. This method is based on Bayes' theory which states name develops data science software called that each data feature is independent of one another.[14] *Data mining* is a data processing method to find hidden patterns in the data; data mining methods can be used to make decisions in the future.[15]

2.3.1 Data Mining Process

- 1) Business Understanding: The first step in the data mining process is understanding the problem or question to be answered and achieving the business goal through data analysis.
- 2) Data Understanding: This stage involves collecting relevant data for analysis and understanding the characteristics of the data, including the type, size, and structure of the data. Understanding data helps identify limitations and constraints that may exist in the data mining process.
- 3) Data Cleaning: At this stage, invalid, duplicate, or irrelevant data is cleaned and deleted as data is often incomplete, defective, or missing. The main goal is to ensure that the data used for analysis is high quality.
- 4) Data Integration: If there are several different data sources, this step involves merging and integrating those data sources into one consistent data source.
- 5) Data Transformation: This process involves normalizing, aggregating, or coding data to conform to the format and schema required for further analysis.

2.4 Naïve Bayes Algorithm

A machine teaching technique that uses probability and statistical calculations invented by British scientist Thomas Bayes, the Naïve Bayes Classifier predicts future probabilities based on past experiences.[3][16] One of the main assumptions of the Naïve Bayes method is that all observed features are independent of one another. However, there is the possibility that some of the features are related. The Naïve Bayes classifier is a classification method based on probability and the Bayesian Theorem, assuming that each X variable is independent or independent and has nothing to do with other variables.

2.5 Results of the Naïve Bayes Algorithm Using Rapidminer

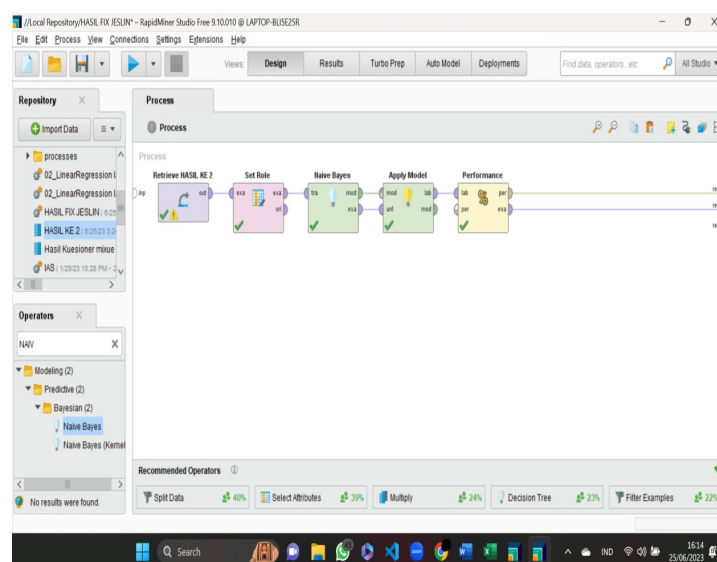


Figure 2.5 The Naïve Bayes Algorithm results using Rapidminer

After the sales data is linked and the parameters have been set in Figure 2.5 above, then by clicking the Run button on Rapidminer, you can process and display the results of the analysis. This software can work in standalone and network environments; Rapidminer can interact with data mining, text mining, machine learning, predictive analysis, and business analysis.[17]-[18]

Table 2.5 Results of Naive Bayes Rapidmine

NAME	K1	K2	K3	K4	K5	..	K15	Results
Yanti	SS	TS	TS	N	SS	..	SS	P
Anita	SS	N	SS	SS	SS	..	SS	P
Meta	SS	TS	TS	TS	SS	..	st	P
Rosita	SS	N	N	SS	N	..	N	P
Reza	SS	N	TS	SS	TS	..	SS	P

Information :

- SS (Strongly Agree)
- TS (Disagree)
- N (Neutral)
- P (Satisfied)

There are 15 questionnaires consisting of questions regarding the physical appearance of the mixture venue, empathy, constraints (reliability), responsiveness (responsiveness), and certainty (assurance) with customer data that answered 500 questionnaires. Mixue Customer Questionnaire Data. Then the results of the rapid miner application with the algorithm method can be seen below.

2.6 Cluster Data Results

Data clustering is grouping data into groups or clusters with similar characteristics. This analysis obtained a satisfied Class Precision result of 94.72%, pred. neutral 2.0%, and pred. Dissatisfied by 4%.

Table 2.5 Cluster Data Results from the RapidminerApplication

	True Satisfied	True Dissatisfied	True Neutral	Class Precision
pred. Satisfied	354	0	1	92.72%
pred. Not satisfied	0	81	0	4.00%
pred. Neutral	1	0	62	2.00%
Class recall	99.72%	100.00%	98.41%	

Segregation is a term for clustering. In some instances, this method uses groups based on similar characteristics. Clustering separates several data groups based on their respective characteristics. Group objects can be people, events, or other things.[19]Data clustering is used to find hidden patterns or structures in data that may be difficult or invisible to identify directly. Finding and grouping data with similar characteristics is known as clustering. The main goal of this technique is to group large amounts of data or objects into sets or groups so that each set can contain the most similar data.[20]

2.7 Rapid Miner Result Data Display

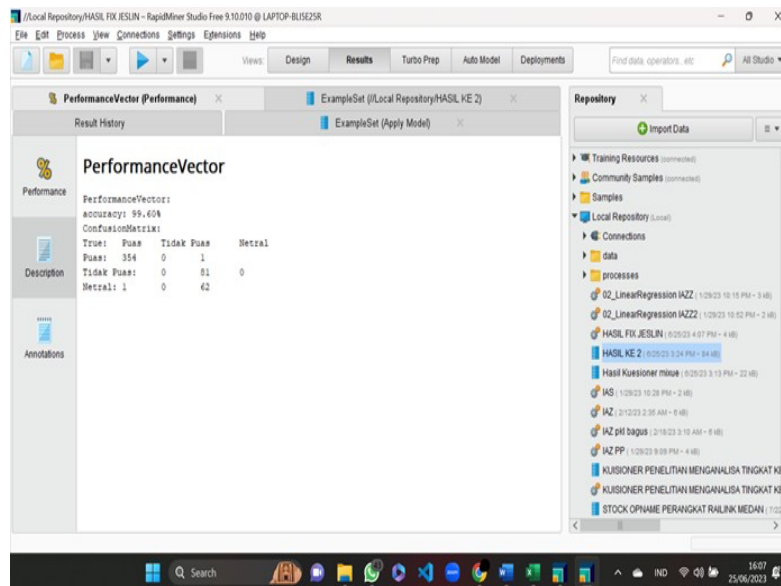


Figure 2.7 Rapid Miner Results Data

Is the result of the rapid miner application, then the above is mixed customer satisfaction which is assessed objectively based on the dataset. The method used is the Naïve Bayes Classifier Algorithm with rapidminer results; 356 customers answered the questionnaire results satisfied, 64 customers who answered the results of the questionnaire were neutral to happy customers, and 80 customers who answered the questionnaire results were dissatisfied with customers who feel dissatisfied or disagree. In this case, the Naïve Bayes Classifier Method can predict customer satisfaction for ice cream mixes that use the Rapidminer Application with Satisfaction.

3. CONCLUSION

Based on rapid miner results 356 customers answered the questionnaire results satisfied, 64 customers who answered the results of the neutral questionnaire, and 80 customers who answered the questionnaire results were not satisfied for the results of this analysis obtained Class Precision results that were satisfied by 94.72%, pred. neutral 2.0% and pred. Dissatisfied by 4%.

4. CLOSING

Based on the results of the naïve Bayes classifier method that has been designed, it is expected that the naïve Bayes classifier method, in analyzing the level of customer satisfaction, can find out how satisfied customers are with the ice cream mixture product. So that this research contributes to the company so that the results of this research provide benefits and benchmarks to further improve what must be evaluated in order to maintain the level of customer satisfaction.

BIBLIOGRAPHY

- [1] Wicaksana and T. Rachman, “濟無No Title No Title No Title,” *Angew. Chemie Int. Ed.* 6(11), 951–952., vols. 3, no. 1, pp. 10–27, 2018, [Online]. Available: <https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf>
- [2] N. Widayanti, S. Mai Murni, DS Mulyadi, NDP Ratnasari, S. Aisyah, and N. Syahria, “Innovation of Using Vegetables to Become Smoothies and Vegetable Ice Cream,” *J. Abadimas Adi Buana*, vol. 2, no. 1, pp. 47–53, 2018, doi:10.36456/abadimas.v2.i1.a1635.
- [3] U. Optimization and S. Marketing, “Journal of Information and Communication Technology STMIK Subng, April 2018 ISSN: 2252-4517,” no. April, pp. 84–95, 2018.
- [4] R. Takdirillah, "Application of Data Mining Using the Apriori Algorithm for Transaction Data to Support Sales Strategy Information," *Edumatic J. Educator. inform.*, vol. 4, no. 1, pp. 37–46, 2020, doi:10.29408/edumatic.v4i1.2081.
- [5] PMS Tarigan, JT Hardinata, H. Qurniawan, M. Safii, and R. Winanjaya, "Implementation of Data Mining Using the Apriori Algorithm in Determining Goods Inventory," *J. Janitra Inform. and Sist. inf.*, vol. 2, no. 1, pp. 9–19, 2022, doi: 10.25008/janitra.v2i1.142.
- [6] HF Putro, RT Vlandari, and WLY Saptomo, "Application of the Naive Bayes Method for Customer Classification," *J. Technol. inf. and Commun.*, vol. 8, no. 2, 2020, doi: 10.30646/tikomsin.v8i2.500.
- [7] Y. Siska, "Application of Data Mining With Customer's Naive Bayes Algorithm for Car Service Services (Case Study: Katamso Service)," *Maj. Science. CORE*, vol. 6, p. 354, 2019.
- [8] P. Postgraduate, U. Amikom, North JR, M. Naive, and B. Classifier, "Expert System for Refining Acute Respiratory Diseases Using the Web-Based Naive Bayes Classifier Method," pp. 127–138.
- [9] N. Darna *et al.*, “Choosing the right research method: for management science research,” vol. 5, no. April, pp. 287–292, 2018.
- [10] S. Wirma, "Data Mining Using the Naïves Bayes Classifier Method in Predicting the Level of Satisfaction with Population Document Services," *J. Inform. Econ. Business*, vol. 4, no. 3, pp. 156–160, 2022, doi: 10.37034/infec.v4i3.155.
- [11] R. Nofitri and N. Irawati, "Data Analysis of Profit Results Using Rapidminer Software," *JURTEKSI (Journal of Technology and Information Systems)*, vol. 5, no. 2, pp. 199–204, 2019, doi: 10.33330/jurteksi.v5i2.365. GH Wastito, "Chapter II Basic Theory," *J. Chem. inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2018.
- [12] Yuli Mardi, "Data Mining: Classification Using the C4 Algorithm. 5 Data mining is part of the Knowledge Discovery in Database (KDD) process stage . Journal of Informatics Education," *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2019.
- [13] HF Putro, RT Vlandari, and WLY Saptomo, "Application of the Naive Bayes Method for Customer Classification," *J. Technol. inf. and Commun.*, vol. 8, no. 2, Oct. 2020, doi: 10.30646/tikomsin.v8i2.500.
- [14] BD Meilani, N. Susanti, JT Informatics, FT Information, I. Technology, and A. Tama, "Accuracy of Data Mining to Generate Student Graduation Patterns with the NAÏVE BAYES Method," *J. Sist. inf. Univ. Suryadarma*, vol. 3, no. 2, pp. 182–189, 2014, doi: 10.35968/jsi.v3i2.66.
- [15] N. Chatrina, Siregar, R. Ruli, A. Siregar, and M. Yoga, Distra, Sudirman, "Implementation of the Naive Bayes Classifier (NBC) Method on School Community Comments Regarding the Implementation of Distance Learning (PJJ)," *J. Technol.*, vol. 34, no. 1, pp. 102–110, 2020, [Online]. Available: <https://aperti-e-journal.id/technology/article/view/67>
- [16] February, *The application for measuring the effectiveness of finished steel damage uses*

the Naïve Bayes method. 2019.

- [17] B. Lutkevich, *Social Media*, vol. 58, no. 10. 2010. [Online]. Available:
<https://www.unodc.org/southasia/en/topics/frontpage/2009/trafficking-in-persons-and-smuggling-of-migrants.html>
- [18] Yulia and M. Silalahi, "Application of Data Mining Clustering in Grouping Books Using the K-Means Method," *Indonesia. J. Comput. sci.*, vol. 10, no. 1, 2021, doi: 10.33022/ijcs.v10i1.3008.
- [19] M. Anjelita, AP Windarto, A. Wanto, and I. Jadiri, "Development of Clustering Datamining in Cases of Environmental Pollution," *Monday. Nas. Technol. computer. Science*, pp. 309–313, 2020.

