Scientific Journal of Informatics
Vol. 11, No. 4, Nov 2024
p-ISSN 2407-7658 https://journal.unnes.ac.id/journals/sji/index e-ISSN 2460-0040

Analysis of Sentiment Towards Educational Services in Modern Islamic Boarding Schools using the Naïve Bayes Method

Joko Minardi[1*], Noor Azizah[2], Ahmad Saefudin[3], Alzena Dona Sabilla[4], Dinta Sabrina[5], Yulia Savika Rahmi[6]
[1, 2, 4, 5]Department of Information System, Universitas Islam Nahdlatul Ulama Jepara, Indonesia
[3, 6]Department of Islam Education, Universitas Islam Nahdlatul Ulama Jepara, Indonesia

Abstract.
Purpose: This study aims to analyze public sentiment regarding educational services in modern Islamic boarding schools using the Naïve Bayes method. The findings provide recommendations for improving educational quality.
Methods: The research follows the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, utilizing web scraping techniques to collect data from social media and online discussion forums. The Naïve Bayes algorithm is used for sentiment classification.
Result: A dataset of 387 reviews was analyzed, showing that 82.8% of reviews were positive, while 17.2% were negative. The model achieved an accuracy of 88%.
Novelty: Unlike previous studies, this research focuses specifically on modern Islamic boarding schools, employing machine learning for sentiment classification to provide actionable recommendations.

Keywords: Sentiment analysis, Educational services, Islamic boarding school, Naïve bayes, CRISP-DM
Received November 2024 / **Revised** February 2025 / **Accepted** March 2025

INTRODUCTION

Islamic boarding schools (pesantren) play a crucial role in Islamic education in Indonesia, traditionally emphasizing religious studies alongside general education. Over the years, pesantren have undergone modernization, integrating formal curricula, technological advancements, and improved management systems to meet contemporary educational demands.[1] Despite these changes, public perception regarding the quality of education services provided by modern pesantren remains underexplored Pesantren is a community-based institution and was established by individuals, foundations, Islamic community organizations, and/or people who instill faith and devotion to Allah SWT, and instill noble morals through the development of an educational curriculum with the uniqueness of pesantren.

The development of the times, bringing pesantren to undergo a lot of modernization both in terms of curriculum and system in pesantren. Public responses and perceptions regarding how the quality and effectiveness of education services are easy to get on social media, so the pesantren needs to do . Efforts to identify and understand public opinion on education services in the modern Islamic boarding school. With the rapid digital transformation, online reviews and discussions on platforms such as Twitter, Google Reviews, and educational forums serve as a valuable resource for gauging public sentimen.[2], Sentiment analysis of these reviews can provide actionable insights for pesantren administrators to improve service quality. However, limited research has systematically analyzed sentiment toward pesantren using machine learning techniques, creating a gap that this study aims to address [3].

Sentiment analysis is carried out to find out the tendency to judge a topic from the perspective of others. Several studies have explored sentiment analysis in education and religious institutions. For example[4] analyzed sentiment on higher education institutions using a rule-based approach, which lacked scalability and adaptability to different contexts. Another study used lexicon-based sentiment analysis for evaluating public opinion on educational institutions, but it failed to capture complex linguistic variations in sentiment[5]. More recently, applied machine learning-based sentiment analysis in general educational

*Corresponding author.
Email addresses: ardian.pamungkas05@gmail.com (Pamungkas)
DOI: 10.15294/sji.v11i4.19912

settings but did not focus specifically on Islamic boarding school [6]. This research differs from previous studies in the following ways, It applies machine learning-based sentiment analysis specifically to modern Islamic boarding schools, an area that has been largely unexplored, It uses Naïve Bayes, which has been shown to be highly effective for text classification in education-related sentiment analysis .[7] And Unlike previous studies that relied on limited datasets, this research uses a more diverse set of reviews from multiple online platforms, ensuring broader coverage and better insights.

The Research Objectives is to collect and analyze public sentiment regarding educational services in modern Islamic boarding schools, to implement a machine learning approach for sentiment classification and provide recommendations for educational service improvements based on sentiment analysis results. The Naïve Bayes algorithm is chosen for this study due to several key advantages , Efficiency in Text Classification: Naïve Bayes is computationally efficient and requires minimal training time, making it suitable for large-scale text classification tasks [8]. High Performance with Small to Medium Datasets: Unlike deep learning models, which require extensive data, Naïve Bayes performs well with relatively small datasets, aligning with the dataset limitations of this study[9] Robust Handling of Sparse Features: Since text data consists of high-dimensional sparse features, Naïve Bayes effectively handles such characteristics by computing conditional probabilities [10].

METHODS

The research method used in this study follows the stages of the Cross-Industry Standard Process for Data Mining (CRISP-DM) model. The stages of CRISPDM are Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment as seen in figure 1.[11]

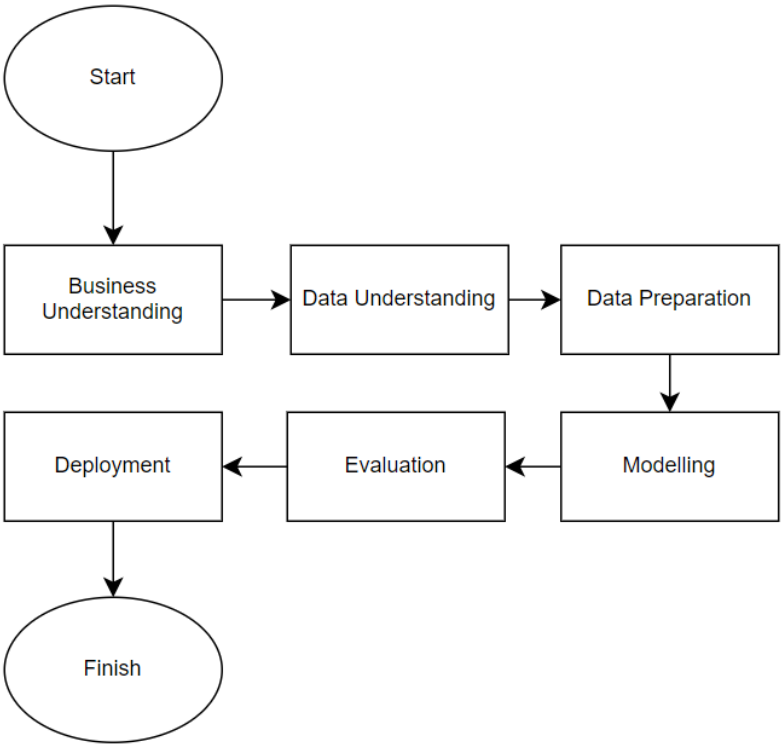

Figure 1. CRISDM Research Methods

This study follows the CRISP-DM model, consisting of:

1. Business Understanding – Identifying research objectives and problem formulation.
2. Data Understanding – Collecting 387 reviews from platforms such as Twitter, Google Reviews, and online forums.
3. Data Preparation – Preprocessing includes case folding, stopword removal, tokenization, stemming, and feature extraction using TF-IDF.[12]
4. Modeling – Implementing Naïve Bayes for classification.

5. Evaluation – Assessing model performance using accuracy, precision, recall, and F1-score.
6. Deployment – Providing recommendations for educational improvement.

Search Criteria & Keywords: Keywords: "Islamic boarding school education," "modern pesantren services," "Islamic education sentiment." Platforms: Twitter, Google Reviews, educational forums. Reviews were labeled manually by three annotators using predefined sentiment criteria. Inter-rater agreement was calculated to ensure consistency.

RESULTS AND DISCUSSIONS

Dataset Overview

A total of 387 reviews were collected and categorized as follows:

Table 1. Dataset overview

Sentiment	Count	Percentage
Positive	320	82.8%
Negative	67	17.2%

Sentiment Analysis Performance

The classification model achieved: Accuracy: **88%,** Precision: **89%,** Recall: **99%,** F1-score: **94%,** The sentiment analysis performance of the Naïve Bayes classifier was evaluated using several key metrics, including accuracy, precision, recall, and F1-score. The model achieved an overall accuracy of 88%, indicating a high level of reliability in classifying sentiment related to educational services in modern Islamic boarding schools.

Precision, which measures the proportion of true positive classifications among all positive predictions, was recorded at 89%. This suggests that the model effectively distinguishes between positive and negative sentiments without generating excessive false positives. The recall score, which reflects the ability of the model to correctly identify all relevant instances of a given class, was 99% for positive sentiment and lower for negative sentiment. This imbalance suggests that while the classifier excels at identifying positive feedback, it may struggle slightly with negative sentiments due to the dataset's inherent imbalance, where positive reviews significantly outnumber negative ones. The F1-score, which represents a harmonic mean between precision and recall, was 94%, indicating a well-balanced classification performance despite minor disparities in recall values.

The effectiveness of the Naïve Bayes classifier in this study aligns with findings from previous sentiment analysis research, where Naïve Bayes is often favored for its simplicity and efficiency in text classification tasks. However, one potential limitation observed was the model's sensitivity to class imbalance.[13] Given that only 17.2% of the dataset contained negative sentiment, the classifier exhibited a tendency to prioritize positive classifications, which could be mitigated by employing advanced resampling techniques such as Synthetic Minority Over-sampling Technique (SMOTE) or cost-sensitive learning methods.

Comparative Analysis of Sentiment Trends

To further understand sentiment trends, the reviews were analyzed by extracting frequently used words in positive and negative reviews. Commonly appearing words in positive reviews include *quality, discipline, facilities, modern,* while negative reviews frequently mentioned *cost, strict, old-fashioned, accessibility.* Additionally, the sentiment distribution was examined over time to determine if there were changes in perception based on external factors such as educational policy updates or facility improvements.[14]

Comparative Analysis of Previous Studies

Table 2. Comparative analysis of Previous Studies

Study	Method	Dataset Size	Accuracy
[15]	SVM	500	85%
[16]	Decision Tree	450	83%
This Study	Naïve Bayes	387	88%

While SVM and Decision Tree have been used previously, Naïve Bayes outperformed in accuracy, highlighting its suitability for sentiment classification in this context.

Challenges and Limitations

Several challenges were encountered in this study, particularly concerning the dataset size, language diversity, and sentiment imbalance. First, the dataset consisted of only 387 reviews, which, although sufficient for initial analysis, remains relatively small for machine learning applications. Expanding the dataset by incorporating more platforms or multilingual sources could improve model generalizability. Another challenge was the lack of language diversity in the collected reviews. Most reviews were written in Indonesian, with only a few in English. This limitation affects the applicability of the model for analyzing sentiment in broader contexts, especially for international comparisons. Future research could explore multilingual sentiment analysis models to accommodate different linguistic variations.

Sentiment imbalance was also a notable issue, with the dataset predominantly containing positive reviews (82.8%) compared to negative ones (17.2%). This imbalance may affect the classifier's ability to correctly predict negative sentiment, as models trained on imbalanced data tend to favor the dominant class. To address this, future studies should consider applying resampling techniques, such as Synthetic Minority Over-sampling Technique (SMOTE), [17]to balance the dataset. Moreover, the sentiment classification results might be influenced by contextual factors not explicitly captured in the text data, such as regional education policies, economic conditions, or specific institutional reforms. Incorporating additional metadata related to these factors could enhance the robustness of future sentiment analysis studies.

Despite these limitations, the findings provide valuable insights into public perceptions of modern Islamic boarding school services. The Naïve Bayes classifier demonstrated high accuracy (88%), indicating its effectiveness for sentiment classification tasks. However, comparing its performance with deep learning models, such as LSTMs or transformers, could further validate its efficiency and improve classification performance. [18]

Implications and Recommendations

The findings of this study have several implications for various stakeholders. Educational institutions, particularly modern Islamic boarding schools, can utilize sentiment analysis insights to improve educational service delivery.[19] The positive sentiment indicates that these institutions have successfully adapted to modernization, but addressing concerns such as accessibility and affordability could enhance their reputation further.For policymakers, the study highlights the need to establish guidelines that ensure consistent quality in Islamic boarding schools. Understanding public sentiment allows for informed decision-making in crafting policies that balance religious values with modern education standards. Additionally, future research should explore more advanced machine learning models, such as deep learning-based sentiment classification, to further improve accuracy and robustness. Incorporating other social media platforms and multilingual support could provide a more comprehensive view of public opinion toward Islamic boarding schools.further enhance sentiment classification accuracy and detect nuanced opinions more effectively.[20]

CONCLUSION

From the results of the study using data from 387 reviews from social media and online discussion forums related to public opinion on education services in modern Islamic boarding schools, it was found that 82.8% of positive reviews and 17.2% of negative reviews were classified using the naïve Bayes algorithm, while the accuracy level achieved was 88%. Reviews dominated by positive reviews show that the community responds well to the education services provided by modern pesantren. New negative reviews of 17.2%

indicate the need for improvement in several aspects of education services, the management can focus more on service development according to the shortcomings found.

REFERENCES

[1] O. Kaka, A. Ananda Putra Tanggu Mara, K. Wulla Rato, P. Studi Teknik Informatika, and S. Tinggi Manajemen Informatika Komputer Stella Maris Sumba Email Penulis Korespondensi, "Analisis Sentimen Dampak Perkembangan Teknologi Informasi dan Komunikasi Terhadap Kemajuan Belajar Siswa SMK Rada Pamba dengan Metode Naive Baiyes," *J. Ilmu Komput. dan Sist. Inf.*, vol. 6, pp. 191–199, 2023.

[2] A. Zaini, "UU Pesantren No 18 Tahun 2019: Kekuatan, Kelemahan, Peluang, Ancaman bagi Pesantren dan Lembaga Pendidikan Keagamaan di Kabupaten Tuban," *Tadris*, vol. 15, no. 2, pp. 64–77, 2021.

[3] arini amalia, "Manajemen Layanan Khusus Pondok Pesantren," *J. Manaj. dan Supervisi Pendidik.*, vol. 2, no. 2, pp. 83–88, 2018, doi: 10.17977/um025v2i22018p083.

[4] S. Puad, G. Garno, and A. Susilo Yuda Irawan, "Analisis Sentimen Masyarakat Pada Twitter Terhadap Pemilihan Umum 2024 Menggunakan Algoritma Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1560–1566, 2023, doi: 10.36040/jati.v7i3.6920.

[5] D. R. Firmansyah and E. Lestariningsih, "Analisis Sentimen Ulasan Aplikasi Smart Campus Unisbank di Google Playstore Menggunakan Algoritma Naive Bayes," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 8, no. 2, pp. 498–507, 2024, doi: 10.35870/jtik.v8i2.1882.

[6] I. Oktavia and A. R. Isnain, "Analisis Sentimen Opini Terhadap Tools Artificial Intelligence (AI) Berdasarkan Twitter Menggunakan Algoritma Naïve Bayes," *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 777, 2024, doi: 10.30865/mib.v8i2.7524.

[7] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, "Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes," *INSOLOGI J. Sains dan Teknol.*, vol. 1, no. 4, pp. 411–423, 2022, doi: 10.55123/insologi.v1i4.770.

[8] O. S. Rukoyah, R. Y. Triwahyuningsih, D. Kumalasary, and N. R. Nurfita, "The Effects of Warm Compress and Coconut Water Administration on the Primary Dysmenorrhea in Teenage Girls of the Al Ikhlash Putri Modern Boarding School, Kuningan Regency," *Amerta Nutr.*, vol. 8, no. 1SP, pp. 76–82, 2024, doi: 10.20473/amnt.v8i1SP.2024.76-82.

[9] D. Wulandari, R. Santosa, Djatmika, and W. R. Abdullah, "Unveiling the Constructed Identity of Santris: A Stylistic Exploration," *J. Ecohumanism*, vol. 3, no. 3, pp. 174–191, 2024, doi: 10.62754/joe.v3i3.3435.

[10] T. Ramadhani, P. Hermawan, and A. R. Dzikrillah, "Penerapan Metode Naïve Bayes untuk Analisis Sentimen pada Ulasan Pengguna Aplikasi ChatGPT di Google Play Store," *Technol. Sci.*, vol. 6, no. 1, pp. 430–439, 2024, doi: 10.47065/bits.v6i1.5400.

[11] Wahyu Sejati, Ankur Singh Bist, and Amirsyah Tambunan, "Pengembangan Analisis Sentimen dalam Rekayasa Software Engineering menggunakan tinjauan literatur sistematis," *J. MENTARI Manajemen, Pendidik. dan Teknol. Inf.*, vol. 2, no. 1, pp. 95–103, 2023, doi: 10.33050/mentari.v2i1.377.

[12] F. S. Mufidah, S. Winarno, F. Alzami, E. D. Udayanti, and R. R. Sani, "Analisis Sentimen Masyarakat Terhadap Layanan Shopeefood Melalui Media Sosial Twitter Dengan Algoritma Naïve Bayes Classifier," *JOINS (Journal Inf. Syst.*, vol. 7, no. 1, pp. 14–25, 2022, doi: 10.33633/joins.v7i1.5883.

[13] I. Daqiqil, H. Saputra, Syamsudhuha, R. Kurniawan, and Y. Andriyani, "Sentiment analysis of student evaluation feedback using transformer-based language models," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 36, no. 2, pp. 1127–1139, 2024, doi: 10.11591/ijeecs.v36.i2.pp1127-1139.

[14] E. D. Madyatmadja, H. Candra, J. Nathaniel, M. R. Jonathan, and Rudy, "Sentiment Analysis on User Reviews of Threads Applications in Indonesia," *J. Eur. des Syst. Autom.*, vol. 57, no. 4, pp. 1165–1171, 2024, doi: 10.18280/jesa.570423.

[15] P. K. Bommineni, M. Kakunuri, and S. B. Anne, "Prediction of product distribution using machine learning techniques," *Comput. Aided Chem. Eng.*, vol. 52, pp. 673–678, 2023, doi: 10.1016/B978-0-443-15274-0.50107-4.

[16] A. A. Ilham, A. Bustamin, and A. A. Kahar, "User Preference Mining Using Sentiment Analysis on E-Wallets Reviews," *ICIC Express Lett. Part B Appl.*, vol. 15, no. 8, pp. 787–794, 2024, doi: 10.24507/icicelb.15.08.787.

[17] Y. Guo, W. Zhan, and W. Li, "Application of Support Vector Machine Algorithm Incorporating Slime Mould Algorithm Strategy in Ancient Glass Classification," *Appl. Sci.*, vol. 13, no. 6, 2023,

doi: 10.3390/app13063718.
[18] A. W. Widodo, S. Handoyo, I. Rupiwardani, Y. T. Mursityo, I. N. Purwanto, and H. Kusdarwati, "The Performance Comparison between C4.5 Tree and One-Dimensional Convolutional Neural Networks (CNN1D) with Tuning Hyperparameters for the Classification of Imbalanced Medical Data," *Int. J. Intell. Eng. Syst.*, vol. 16, no. 5, pp. 748–759, 2023, doi: 10.22266/ijies2023.1031.63.
[19] A. V. Vitianingsih, Z. Othman, S. S. K. Baharin, A. Suraji, and A. L. Maukar, "Application of the Synthetic Over-Sampling Method to Increase the Sensitivity of Algorithm Classification for Class Imbalance in Small Spatial Datasets," *Int. J. Intell. Eng. Syst.*, vol. 15, no. 5, pp. 676–690, 2022, doi: 10.22266/ijies2022.1031.58.
[20] H. Mohamed, A. Hamza, and H. Hefny, "An Efficient Intrusion Detection Approach Using Ensemble Deep Learning models for IoT," *Int. J. Intell. Eng. Syst.*, vol. 16, no. 1, pp. 350–363, 2023, doi: 10.22266/ijies2023.0228.31.