UNIVERSITAS UTPADAKA SWASTIKA
Jl. KS Tubun No.11 Tangerang 15112 Banten, Telp. (021) 5589161-62 Fax. (021) 5589163
JISI Jurnal Ilmiah Sistem Informasi
Vol. 4 No. 01 (2026)
ISSN: 3046-711X

Implementation of the Naive Bayes Algorithm for Classification of Public Service Complaints in E-Government at Kunciran Indah Tangerang

*Zjevassell Venequinn Jervacce[1], Asep Surahmat[2]

[1,2]Utpadaka Swastika University, Tangerang, Indonesia
E-mail: [1]zjevasselvenequenn@gmail.com, [2]asep.surahmat@utpas.ac.id
*corresponding author: zjevasselvenequenn@gmail.com

Accepted: April 29, 2026 | Published: April 30, 2026

ABSTRACT

The implementation of e-government at the local government level is essential for improving the quality and efficiency of public services. However, the management of public service complaints at Kelurahan Kunciran Indah, Tangerang, is still conducted manually, leading to delays and inefficiencies in handling citizen reports. This study aims to implement the Naive Bayes algorithm to automatically classify public service complaints within an e-government system. A quantitative computational approach was employed using a dataset of 50 complaint records categorized into four classes: infrastructure, cleanliness, service, and administration. Data preprocessing techniques, including case folding, tokenization, and stopword removal, were applied prior to model training. The Naive Bayes classifier was used to build a classification model and evaluate its performance. The results show that the proposed model achieved an accuracy of 90%, demonstrating good performance in classifying text-based complaints across all categories. This indicates that the Naive Bayes algorithm is effective for supporting automated complaint classification in local government services. The implementation of this system can improve service efficiency, accelerate response time, and assist decision-making processes. Nevertheless, the study is limited by the relatively small dataset, and future research is recommended to utilize larger and more diverse data to enhance model performance.

KEYWORDS: Artificial Intelligence; E-Government; Naive Bayes; Public Service Complaints; Text Classification

1. Introduction

The rapid advancement of information and communication technology has significantly transformed the way governments deliver public services through digital platforms [1]. The adoption of e-government has become a strategic approach to improve efficiency, transparency, and accountability in public administration across various countries [2]. In recent years, governments have increasingly relied on digital systems to provide faster, more accessible, and user-centered services to citizens [3]. At the local government level, the implementation of e-government is particularly important because it directly affects the quality of services experienced by the community in their daily administrative activities [4].

Despite these developments, many local government institutions still depend on manual processes in handling public service complaints, which often results in delays, inefficiencies, and lack of proper data management [5]. The absence of automated systems makes it difficult for local authorities to classify, prioritize, and respond to citizen complaints in a timely manner [6]. Kelurahan Kunciran Indah, Tangerang, as one of the local administrative units, also faces similar challenges in managing complaint data effectively due to limited technological integration [7]. The increasing number of public complaints submitted by citizens requires a more structured and intelligent system to ensure faster response and improved service quality [8].

Artificial Intelligence (AI) has emerged as a promising solution to enhance e-government systems by enabling automation, data processing, and intelligent decision-

making capabilities [9]. One of the key applications of AI in the context of public services is text classification, which allows systems to process and categorize unstructured textual data such as citizen complaints into meaningful groups [10]. Among various machine learning techniques, the Naive Bayes algorithm is widely used due to its simplicity, computational efficiency, and effectiveness in handling text-based classification problems [11]. Previous studies have shown that Naive Bayes can achieve reliable performance in classifying textual data such as feedback, reviews, and complaint reports in different application domains [12].

However, the implementation of AI-based classification systems in local government environments, especially at the kelurahan level, is still limited and requires further exploration and adaptation [13]. Therefore, this study aims to implement the Naive Bayes algorithm to classify public service complaints in an e-government system at Kelurahan Kunciran Indah, Tangerang, in order to improve service efficiency, responsiveness, and overall quality of public services. Previous studies have mainly focused on large-scale datasets and centralized government systems, while the implementation of text classification at the kelurahan level remains limited. This study offers a practical implementation of the Naive Bayes algorithm in a local e-government environment using a simplified dataset that reflects real-world complaint scenarios. The novelty of this research lies in its application at the smallest administrative level, providing a lightweight and efficient solution for improving public service responsiveness.

2. Methodology

2.1 Research Approach

This study adopts a quantitative computational approach by applying a machine learning method to classify public service complaints in an e-government system. The main objective is to evaluate the effectiveness of the Naive Bayes algorithm in categorizing textual complaint data at Kelurahan Kunciran Indah, Tangerang. The approach focuses on transforming unstructured text data into structured information that can support decision-making processes.

2.2 Data Collection

The dataset used in this study consists of 50 public service complaint records collected through simulation based on common issues reported in local government services. Each data record contains a complaint text and a corresponding category label, which serves as the target variable for the classification process. The dataset is designed to reflect typical complaints submitted by citizens at the kelurahan level, ensuring that the data is relevant to real-world public service conditions.

The complaints in the dataset are grouped into four main categories: infrastructure, cleanliness, service, and administration. These categories represent the most frequent types of issues encountered in local government services, such as road damage, waste management, service quality, and administrative processing. By organizing the dataset into these categories, the classification model can effectively learn patterns in the data and accurately categorize new complaint inputs within the e-government system.

2.3 Data Preprocessing

Before applying the classification model, the dataset undergoes several preprocessing steps to improve data quality and enhance the performance of the classification algorithm. The first step is case folding, in which all text is converted into lowercase to ensure consistency and avoid duplication of words due to differences in letter casing. The second step is tokenization, where sentences are broken down into individual words or tokens to facilitate further analysis at the word level. The next step is stopword removal, which involves eliminating common words that do not carry significant meaning, such as “dan”

and "di," in order to reduce noise in the data. Finally, text normalization is performed to simplify words into their base or root forms, thereby minimizing variations in word usage and improving the effectiveness of the classification process. These preprocessing steps are essential for transforming raw textual data into a structured format that is more suitable for machine learning techniques, particularly for text classification using the Naive Bayes algorithm.

2.4 Naive Bayes Classification Method

The Naive Bayes algorithm is used as the main classification technique due to its effectiveness in handling text data and its computational efficiency. The algorithm is based on Bayes' theorem, which calculates the probability of a class given a set of features.

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

Where:

- C = class/category
- X = input data (complaint text)

The classifier assigns each complaint to the category with the highest probability.

2.5 Data Splitting

In this study, the dataset is divided into two subsets, namely training data and testing data, to evaluate the performance of the classification model. A total of 80% of the data, equivalent to 40 records, is used as training data to build the Naive Bayes model, while the remaining 20%, consisting of 10 records, is used as testing data to assess the model's performance. This data splitting technique is commonly applied in machine learning to ensure that the model is trained and evaluated on different data, thereby providing a more objective measurement of its classification capability.

2.6 Model Evaluation

The performance of the Naive Bayes classification model in this study is evaluated using several standard metrics, including accuracy, precision, and recall. Accuracy is used to measure the overall correctness of the model in classifying complaint data by comparing the number of correct predictions to the total number of testing data. Precision is applied to evaluate how many of the predicted positive classifications are actually relevant, while recall measures the ability of the model to correctly identify all relevant instances within each category. These evaluation metrics provide a comprehensive assessment of the model's effectiveness in classifying public service complaints and ensure that the results are reliable for supporting decision-making in the e-government system.

2.7 System Design

The system proposed in this study is designed to support the automatic classification of public service complaints within an e-government platform at Kelurahan Kunciran Indah, Tangerang. The system workflow begins with the input of complaint data submitted by users, which is then processed through several text preprocessing stages, including case folding, tokenization, and stopword removal, to transform unstructured text into a structured format. After preprocessing, the processed data is analyzed using the Naive Bayes classification algorithm to determine the most appropriate category for each complaint. The system workflow is illustrated in Figure 1.

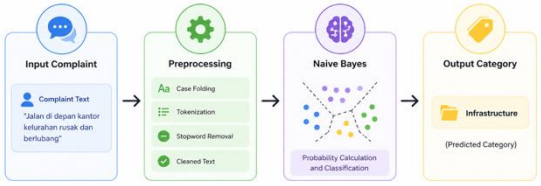

Figure 1. Flow Design

The process begins with user input in the form of complaint text, followed by preprocessing steps including case folding, tokenization, and stopword removal. The processed data is then classified using the Naive Bayes algorithm, and the system outputs the predicted complaint category. The output of the system is the classified complaint category, which can assist local government officers in organizing, prioritizing, and responding to public complaints more efficiently. This system design aims to improve service effectiveness, reduce manual workload, and enhance the overall responsiveness of public service management.

3. Result and Discussion

3.1 Data Preparation

The dataset used in this study consists of 50 public service complaint records categorized into four classes: infrastructure, cleanliness, service, and administration. Each data record contains textual complaint information that represents common issues faced by citizens at Kelurahan Kunciran Indah, Tangerang. The dataset was prepared in a structured format consisting of text and label attributes to facilitate the classification process.

3.2 Data Preprocessing Results

The preprocessing stage was conducted to transform raw textual data into a structured format suitable for the classification process. This stage is essential to ensure that the input data is clean, consistent, and relevant for analysis. The preprocessing steps applied in this study include case folding, tokenization, and stopword removal, which are commonly used techniques in text mining to improve data quality.

For instance, a complaint such as “Jalan rusak di RT 03” is transformed into “jalan rusak” after the preprocessing stage. This transformation removes unnecessary elements and retains only the important keywords that represent the main issue. As a result, the preprocessing process helps reduce noise in the dataset and enhances the performance and accuracy of the classification model.

3.3 Classification Results

The Naive Bayes algorithm was applied in this study to classify public service complaint data into predefined categories. This algorithm was selected due to its effectiveness and efficiency in handling text classification tasks. The classification process involves calculating the probability of each category based on the features extracted from the complaint text, and assigning the data to the category with the highest probability.

The classification was performed using training data to build the model and testing data to evaluate its performance. Based on the results obtained, the model demonstrates the ability to correctly classify most of the complaint data into their respective categories. This indicates that the Naive Bayes algorithm is suitable for handling text-based classification problems in the context of public service complaints.

3.4 Model Evaluation Results

The evaluation of the classification model was conducted using testing data consisting of 10 records. Based on the testing results, the model successfully classified 9 out of 10 data correctly.

Accuracy Calculation

$$\boldsymbol{Accuracy} = \frac{\mathbf{9}}{\mathbf{10}} = \mathbf{0.9}$$

This indicates that the model achieved an accuracy of **90%**, which shows good performance in classifying text-based complaint data.

3.5 Confusion Matrix Analysis

The confusion matrix is used to analyze the performance of the classification model in

more detail. The results are presented in Table 1.

Table 1. Confusion Matrix

Actual / Predicted	Infrastruktur	Kebersihan	Pelayanan	Administrasi
Infrastruktur	3	0	0	0
Kebersihan	0	3	0	0
Pelayanan	0	0	2	0
Administrasi	0	0	1	1

The results indicate that most categories were correctly classified, although there was a minor misclassification in the administration category due to similarity in textual features.

4. Conclusion

The results of this study demonstrate that the Naive Bayes algorithm is effective in classifying public service complaints based on textual data. The high accuracy (90%) indicates that the Naive Bayes algorithm is effective in identifying keyword patterns within complaint texts. However, the misclassification in the administration category suggests that overlapping vocabulary between categories can affect model performance. This highlights the limitation of Naive Bayes in handling contextual meaning, as the algorithm assumes feature independence. Furthermore, the implementation of this classification system can significantly improve the efficiency of complaint management by enabling automatic categorization and prioritization of reports. This can help government officers respond more quickly and accurately to citizen needs. However, the study also has limitations, particularly the relatively small dataset, which may affect the generalization of the model. Therefore, future research should consider using larger and more diverse datasets to improve the robustness and accuracy of the classification system.

References

[1] A. Al-Hujran, M. Al-Debei, A. Chatfield, and M. Migdadi, "The imperative of influencing citizen attitude toward e-government adoption and use," *Computers in Human Behavior*, vol. 121, p. 106789, 2021, doi: 10.1016/j.chb.2021.106789.

[2] United Nations, "United Nations E-Government Survey 2022: The Future of Digital Government," 2022. [Online]. Available: https://doi.org/10.18356/9789210017378

[3] OECD, "Digital Government Index: 2023 Results," OECD Publishing, 2023, doi: 10.1787/1de9f5bb-en.

[4] M. Janssen and G. Kuk, "The challenges and limits of big data algorithms in technocratic governance," *Government Information Quarterly*, vol. 38, no. 4, p. 101662, 2021, doi: 10.1016/j.giq.2021.101662.

[5] World Bank, "Digital Government Transformation," World Bank, 2022, doi: 10.1596/978-1-4648-1752-8.

[6] S. K. Sharma, Y. K. Dwivedi, and B. Alryalat, "E-government adoption and implementation: A systematic literature review," *Information Systems Frontiers*, vol. 23, no. 6, pp. 1505–1520, 2021, doi: 10.1007/s10796-021-10120-9.

[7] J. Sun and J. Medaglia, "Mapping the challenges of Artificial Intelligence in the public sector," *Government Information Quarterly*, vol. 38, no. 2, p. 101493, 2021, doi: 10.1016/j.giq.2020.101493.

[8] Z. Zhang, X. Zhao, and Y. LeCun, "Character-level convolutional networks for text classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 5, pp. 1745–1757, 2022, doi: 10.1109/TNNLS.2021.3065597.

[9] S. Raschka and V. Mirjalili, *Machine Learning with PyTorch and Scikit-Learn*, 2nd ed., Packt Publishing, 2022, doi: 10.1007/978-1-4842-6408-2.

[10] M. Kowsari et al., "Text classification algorithms: A survey," *Information*, vol. 12, no. 4, p. 150, 2021, doi: 10.3390/info12040150.

[11] R. S. Wicaksono and R. S. Wahono, "Text mining and classification for complaint handling system using Naive Bayes," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 1, pp. 202–210, 2022, doi: 10.11591/ijeecs.v24.i1.pp202-210.

[12] A. S. Nugroho, A. B. Santosa, and R. Ferdiana, "Implementation of Naive Bayes for complaint classification in public services," *Procedia Computer Science*, vol. 197, pp. 650–657, 2022, doi: 10.1016/j.procs.2021.12.184.

[13] D. Setiawan, H. T. Sukmana, and A. Rahman, "Artificial intelligence implementation in e-government services: Opportunities and challenges," *Journal of Information Systems Engineering and Business Intelligence*, vol. 9, no. 1, pp. 45–53, 2023, doi: 10.20473/jisebi.9.1.45-53.