Journal of Information Systems and Informatics
Vol. 7, No. 3, September 2025 e-ISSN: **2656-4882** p-ISSN: **2656-5935**
DOI: 10.51519/journalisi.v7i3.1207 Published By DRPM-UBD

Sentiment Analysis of Public Service Using Naïve Bayes Classifier

Arga Aditia Purnama[1], Yoannes Romando Sipayung[2]

[1,2]Prodi Studi Teknik Informatika, Universitas Ngudi Waluyo, Ungaran, Indonesia
Email: [1]aditiaarga416@gmail.com, [2]yoannesromando@unw.ac.id

Abstract

Public administrative service quality is a crucial factor in citizen satisfaction. This study analyzes sentiment in public service reviews using a text mining approach with the Naïve Bayes Classifier method. The dataset was collected from citizen feedback on online platforms regarding public administrative services. Preprocessing steps included tokenization, case folding, stopword removal, and stemming. The Naïve Bayes algorithm with Laplace smoothing was applied for classification, and performance was evaluated using accuracy, precision, recall, and F1-score. The experiment resulted in an accuracy of 91.2%, precision of 90.3%, recall of 89.7%, and F1-score of 90.0%. The analysis revealed that Service Speed obtained an average score of 3.21, indicating a moderate level of citizen satisfaction in that aspect. These findings suggest that while the Naïve Bayes method is effective for sentiment classification, its greatest value lies in providing actionable insights for public service improvement. Specifically, policymakers can prioritize addressing delays in service speed through simplified procedures, improved staffing, and digital innovation, while maintaining strengths such as officer politeness and effective complaint handling. By leveraging sentiment analysis, public institutions can continuously monitor citizen feedback, identify problem areas, and implement evidence-based strategies to enhance service quality and strengthen public trust.

Keywords: Sentiment Analysis, Text Mining, Naïve Bayes, Public Service Quality

1. INTRODUCTION

Sentiment analysis has been widely applied in areas such as e-commerce, politics, and urban public services, there remains a lack of studies that utilize this approach with data obtained from local-level physical surveys, especially in rural and semi-urban communities. The novelty of this research lies in applying the Naïve Bayes Classifier to analyze community perceptions of village-level administrative services, thereby offering data-driven insights for service improvement in a context that has been largely underexplored.

Public service is the main function of the bureaucracy is to meet society's basic needs [1], [2]. It reflects the quality of the relationship between the state and its

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

citizens. However, in Indonesia, the quality of public services remains a significant challenge. Low public Satisfaction with government services has long been one of the main causes of declining public trust in state institutions [1]. This condition often triggers various forms of complaints, criticism, and even destructive demonstrations against public institutions [2], [15]. In Lemahireng Hamlet, challenges such as insufficient staffing, procedural bottlenecks, limited digital infrastructure, and low levels of technological literacy among both officers and residents exacerbate delays and reduce the efficiency of administrative services.

The context of Lemahireng Hamlet illustrates these broader challenges at the local level. As a semi-urban area undergoing digital transition, residents face several obstacles in accessing fast and reliable administrative services. Limitations include insufficient staffing, procedural bottlenecks, incomplete digital infrastructure, and low levels of technological literacy among both officials and citizens. These constraints often lead to slow service delivery, inconsistent complaint handling, and communication gaps between service providers and residents. Consequently, although citizens generally value interpersonal aspects such as politeness of officers, delays and inefficiencies in service speed continue to affect overall satisfaction.

Today's digital era of information disclosure, conventional approaches to evaluating public satisfaction must be combined with technology-based methods to produce objective, rapid evaluations [3]. One rapidly growing approach is sentiment analysis, a method that identifies and categorizes public opinions based on positive, negative, or neutral text expressions [4]. Sentiment analysis is a form of text mining that has been widely adopted in fields such as e-commerce, politics, and public services. The application of sentiment analysis in the field of public services provides an opportunity to capture people's perceptions directly from their opinions, such as the results of open surveys or online comments. This approach allows policy makers to detect public dissatisfaction early and formulate targeted service improvement strategies [6].

Among the various machine learning methods available for sentiment analysis, NBC stands out due to its computational efficiency, robustness with small-to-medium datasets, and relatively simple implementation. Compared to Support Vector Machines (SVM), which often require extensive parameter tuning and higher computational resources, NBC can achieve competitive accuracy with less complexity. Unlike Decision Trees, which may overfit noisy text data, NBC maintains generalizability by relying on probabilistic assumptions. While deep learning approaches such as Recurrent Neural Networks (RNNs) or Transformers have shown superior performance in large-scale sentiment analysis, they demand large datasets and significant computational power—resources that

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

are often limited in local-level studies such as in Lemahireng Hamlet. For this reason, NBC represents a pragmatic choice: it balances accuracy and efficiency while remaining accessible for public service institutions with constrained technical infrastructure.

One popular method of sentiment analysis is the Naive Bayes classifier (NBC). The NBC is a probabilistic classification algorithm based on the simple assumption that the features in the data are independent. This method's advantages lie in its computational efficiency, low training data requirements, and ability to handle large-scale text data. In the context of public services, NBC is highly relevant because it can classify sentiment quickly and accurately and is easily implemented [8], [25]. In the context of Lemahireng Hamlet, NBC-based sentiment analysis can provide an evidence-based framework to identify the most pressing service challenges and guide improvements in administrative efficiency, responsiveness, and inclusivity.

This research aims to assess community perceptions of public administration services in Lemahireng Hamlet using NBC-based sentiment analysis. The hamlet was selected as it represents semi-urban areas in digital transition, and the findings are intended to provide data-driven recommendations for sustainable service quality improvements [9]. Consequently, the objective of this research is twofold: first, to provide a descriptive picture of community sentiment, and second, to support evidence-based decision making in an effort to realize public services that are responsive, inclusive, and adaptive to community needs.

A substantial body of research has demonstrated the efficacy of sentiment analysis approaches in evaluating public satisfaction with public services. Handayani and Sulistiyawati 2019, employed the Naïve Bayes method to categorize public responses to the ongoing news cycle concerning the novel coronavirus (SARS-CoV-2) [4], while Gunawan et al. 2021 applied the same method to analyze digital product reviews [3], [15], [16]. In the context of public administration, Hudin et al. 2024 employed text mining techniques for document clustering, thereby underscoring the efficacy of this approach in systematically elucidating public sentiment [5], [17], [18]. These studies substantiate the significance of employing computational methodologies in the analysis of public opinion.

On the other hand, there is a paucity of studies that integrate sentiment analysis based on physical surveys at the local level, especially in rural areas. This study endeavors to address this lacuna by employing the Naïve Bayes Classifier approach in the analysis of public service survey data procured directly from the inhabitants of Lemahireng Hamlet. This study underscores the significance of text mining in the domain of village administration, a relatively unexplored area

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

in prior research. The novelty of this study lies in applying the Naïve Bayes Classifier for sentiment analysis on public service satisfaction surveys collected directly at the village level. Unlike previous studies that mostly focus on online data or urban contexts, this research explores semi-urban and rural administrative services, providing data-driven insights for improving local governance and service delivery.

2. METHODS

This study conducted by following the procedure as shown in Figure 1. Base on Figure 1 consist of six steps data collection, preprocessing, data spliting, model implementasion, model validation, and model analysis.

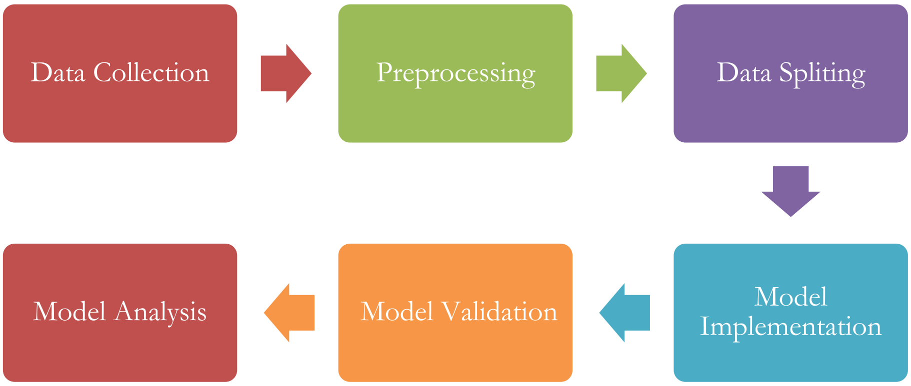


Figure 1. Research Flow

Base on Figure 1, The present study employed a cross-sectional design with an exploratory quantitative approach. A public service survey questionnaire was utilized to collect data from 86 respondents, incorporating both open-ended and closed-ended inquiries. Subsequently, the open-ended data underwent a series of text preprocessing steps, encompassing the following stages: case folding, tokenization, stopword removal, stemming, and Term Frequency–Inverse Document Frequency (TF-IDF) weighting [10], [19].

The survey instrument was designed to capture both factual information and subjective perceptions. The closed-ended questions focused on demographic characteristics (e.g., gender, occupation, and frequency of service use) as well as ratings of service quality indicators such as politeness of officers, complaint handling, and service speed. These were measured using a five-point Likert scale (1 = very dissatisfied to 5 = very satisfied).

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

The open-ended section invited respondents to freely express their experiences and opinions about public administrative services in Lemahireng Hamlet. Prompts such as some respondents shared positive experiences, stating that the service was fast and efficient, with minimal waiting times. For instance, several mentioned that after taking a queue number, they were served within ten minutes, and in some cases, the entire process was completed in less than five minutes due to the staff's responsiveness. However, there were also negative experiences reported, such as delays in service, with some respondents waiting for nearly an hour just to obtain a signature. Others expressed frustration over staff frequently leaving the counter during working hours, which contributed to slower service. In terms of suggested improvements, respondents recommended increasing the number of staff members to avoid long queues and ensuring that service requirements are clearly displayed in visible areas. Additional suggestions focused on improving the quality of interactions, with calls for staff to be more polite and courteous when serving the public. Respondents also highlighted the need for adopting digital solutions, such as implementing a digital queuing system to ensure orderliness and prevent disputes. These qualitative insights provided a deeper understanding that complemented the quantitative ratings, offering a richer basis for sentiment analysis. were used to encourage detailed responses. The intention was to elicit natural language expressions of satisfaction, dissatisfaction, and suggestions for improvement, providing richer qualitative data for sentiment analysis.

The open-ended responses were then preprocessed through a sequence of text cleaning steps, including case folding, tokenization, stopword removal, stemming, and Term Frequency–Inverse Document Frequency (TF-IDF) weighting [10]. These steps transformed raw textual feedback into structured data suitable for machine learning classification.

To prepare the open-ended responses for analysis, several preprocessing steps were applied: 1) Case Folding, all text was converted to lowercase to ensure uniformity (e.g., "Good" and "good" treated as identical). 2) Tokenization, sentences were split into individual words (tokens) using whitespace and punctuation as delimiters. This step produced the basic input units for analysis. 3) Stopword Removal, commonly used but semantically uninformative words in Bahasa Indonesia (e.g., dan, yang, di) were removed using a standard stopword dictionary to reduce noise. 4) Stemming, words were reduced to their root forms using the Nazief–Adriani algorithm for Indonesian stemming, so that variations such as *pelayanan*, *melayani,* and *terlayani* were treated consistently as "layani". 5) TF–IDF Weighting, the term Frequency–Inverse Document Frequency approach was used to transform tokens into weighted feature vectors,

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

emphasizing words that were both frequent in a document and distinctive across the corpus.

Prior to classification, a data cleaning process was undertaken to improve quality and reduce bias: 1) Irrelevant Responses, non-textual entries (e.g., "-", "/", or random characters) were removed. 2) Incomplete Responses, blank or single-word answers (e.g., "*ok*", "*baik*") that provided insufficient sentiment signal were excluded from the training set. 3) Outliers, extremely long responses that deviated significantly from average length (>3 standard deviations) were reviewed. If they were repetitive or off-topic, they were filtered out to prevent skewing model performance. 4) Spelling Variations, common misspellings and local dialect words were normalized using a custom dictionary developed from the dataset. The cleaned dataset was classified using the Naïve Bayes Classifier (NBC), implemented as a Multinomial model with Laplace smoothing ($\alpha = 1$). The dataset was split into 80% training and 20% testing subsets, and robustness was further assessed through 10-fold cross-validation.

Model performance was evaluated using accuracy, precision, recall, and F1-score, all derived from a confusion matrix. The confusion matrix was selected because it provides a detailed breakdown of classification results into true positives, false positives, true negatives, and false negatives, rather than a single aggregate score. This allows for a more nuanced evaluation of the model's performance across different sentiment categories. Accuracy indicates overall correctness, precision reflects how reliably the model identifies positive sentiments, recall shows how effectively it captures all relevant sentiments, and the F1-score harmonizes precision and recall to avoid skew toward one metric. By combining these measures, the evaluation ensures that results are not only statistically robust but also valid for interpreting citizen sentiment in the context of public service delivery. Finally, the classification outcomes were analyzed to highlight service dimensions with strong or weak public perception. This evidence-based approach provides actionable insights for improving administrative services while aligning with transparency and accountability principles in village governance [12], [20].

The cleaned and preprocessed dataset was classified using the Naïve Bayes Classifier (NBC), implemented as a Multinomial model with Laplace smoothing ($\alpha = 1$) to address zero-probability issues. The dataset was divided into 80% training and 20% testing subsets, and performance was validated using 10-fold cross-validation. Model performance was evaluated using accuracy, precision, recall, and F1-score to provide a balanced assessment of classification quality [22].

The use of NBC was motivated by its computational efficiency, robustness with relatively small datasets, and proven reliability in text classification tasks. By

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

combining rigorous preprocessing with careful data cleaning, the analysis ensured that the resulting sentiment classifications reflected authentic community perceptions and minimized the impact of inconsistencies or noise in survey responses [13], [23] .

The Naïve Bayes Classifier was implemented as a Multinomial model with Laplace smoothing ($\alpha = 1$) to handle zero-probability issues. The dataset was split into 80% training and 20% testing subsets to evaluate out-of-sample performance. To ensure robustness and prevent overfitting, the model was evaluated using 10-fold cross-validation, where the dataset is divided into 10 subsets, and the model is trained and tested iteratively. The model performance was assessed using standard quantitative metrics, namely: 1) Accuracy, the ratio of correctly predicted instances to total instances. 2) Precision, the proportion of positive identifications that were actually correct. 3) Recall, the proportion of actual positives correctly identified. 4) F1-Score, the harmonic mean of precision and recall.

Subsequent to the cleansing of the data, sentiment classification was executed through the implementation of the Naïve Bayes Classifier (NBC) algorithm [21]. The selection of this method was predicated on its demonstrated reliability in probability-based classification, its notable computational efficiency, and its aptitude for managing large-scale text data with adequate accuracy [11]. The evaluation of the model was conducted by employing a confusion matrix, a statistical tool used to assess the accuracy, precision, and recall of a model. To further ensure the robustness of the results, k-fold cross-validation was applied, enabling a more reliable estimation of performance metrics (accuracy, precision, recall, and F1-score) across different subsets of the dataset. Additionally, the classification results can be mapped and analyzed to support the development of more targeted service quality improvement strategies. This approach is predicated on the tenets of transparency and accountability in the village bureaucracy system, in accordance with the mandate of public service reform [12].

3. RESULTS AND DISCUSSION

3.1. Dataset

The dataset utilized in this study was obtained through a survey conducted in the form of distributing questionnaires to individuals who visited public service offices. Data collection results as shown in Table 1.

Table 1. Data Collection (Sample)

ID	DF	ID 1 NT	LT/LN	ID 2 NT	LT/LN	ID 3 NT	LT/LN	ID 4 NT	LT/LN	ID 5 NT
1	85	1	0.168	1	0.482	1	0	1	0.168	1

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

ID	DF	ID 1 NT	LT/LN	ID 2 NT	LT/LN	ID 3 NT	LT/LN	ID 4 NT	LT/LN	ID 5 NT
2	7	0	0	0	0	0	0	0	0	0
3	86	1	0.083	1	0.238	1	0	1	0.083	1
4	1	0	0	0	0	0	0	0	0	0
5	86	1	0.083	1	0.238	1	0	1	0.083	1
6	79	1	0.091	1	0.258	1	0	1	0.091	1
7	1	0	0	0	0	0	0	0	0	0
8	86	1	0.091	1	0.258	1	0	1	0.091	1
9	3	0	0	0	0	0	0	0	0	0
10	7	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0
12	86	1	0.083	1	0.238	1	0	1	0.083	1
13	4	0	0	0	0	0	0	0	0	0
14	66	1	0.164	1	0.476	1	0	1	0.164	1
15	1	0	0	0	0	0	0	0	0	0

3.2. Design System

The system has been developed for the purpose of evaluating public satisfaction and trust in public services. The system automatically filters the results of completed questionnaires, identifying positive, negative, and even neutral comments. Additionally, it automatically displays an average score, thereby serving as a benchmark and providing recommendations for evaluating public services with the objective of improving them.

The system under discussion utilizes a combination of two programming languages, Python and PHP, both of which are considered capable of handling the processes required to build the system. The system under discussion utilizes a Laravel 12-based Model-View-Controller (MVC) architecture, which is widely regarded as clear and structured. The system utilizes MariaDB as its database management system (DBS), which is entrusted with the data storage process. For the purpose of interface development, the CSS framework Tailwind is employed, with the addition of JavaScript to achieve a responsive and aesthetically pleasing display.

3.3. System Development

The system development in this study aims to build a web application capable of analyzing public sentiment towards public administration services automatically, quickly, and accurately using a text mining approach based on the Naïve Bayes Classifier algorithm. The system has been developed for the purpose of collecting and analyzing survey data in text form, as well as presenting public service evaluation results visually and in a way that is applicable to the user. The development process adheres to the System Development Life Cycle (SDLC) method, which comprises several primary stages.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

1) Planning

At this stage, the initial identification of public service issues, particularly in the Lemahireng Hamlet area, was conducted, and the objectives of the system to be developed were determined. The planning stage entailed the identification of the technologies to be employed, namely Laravel 11 for the backend, Python for the sentiment classification module, and Bootstrap for the user interface. A comprehensive literature review and observational study were conducted to ascertain user needs and the characteristics of the data to be processed.

2) Needs Analysis

The objective of the analysis was to delineate the functional and non-functional requirements of the system. The functional requirements encompass a range of features, including the capacity for survey file upload in the .xlsx format, automated text processing, sentiment classification, the visualization of survey results, and the provision of reports in PDF and Excel formats. Non-functional requirements encompass aspects such as data security, system response speed, and user-friendliness of the interface.

3) System Design

The system design employs a Model-View-Controller (MVC) approach, implemented in the Laravel framework. A MariaDB database is utilized for the storage of classification results and file metadata. The interface design is responsive, ensuring accessibility across a range of devices. It is grounded in principles of UX (User Experience) and UI (User Interface). Flowcharts, database structures, and system architecture schematics are designed to facilitate subsequent development.

4) Development

This stage entails the implementation of the program code in accordance with the design. The development process is generally divided into two primary phases: 1) The backend encompasses the following components: Laravel is responsible for managing the file upload flow, storing results, and overseeing user data management. 2) The Sentiment Engine is a Python-based module that utilizes the Naive Bayes classification algorithm. This module is integrated through the Flask API, a web application framework that facilitates the development of web services and APIs.

Text preprocessing entails a series of linguistic operations, including case folding, tokenization, stopword removal, stemming, and TF-IDF weighting. Subsequent

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

to this initial classification, the text is then divided into three distinct categories based on its sentiment: positive, negative, and neutral.

5) Testing

A rigorous testing process is implemented to ensure that each function operates in accordance with its intended design. The testing process encompasses the following components: 1) Unit Testing: Testing individual functions such as sentiment classification and file upload. 2) Integration Testing: Testing the interaction between Laravel and the Python API. 3) User Acceptance Testing (UAT): Involving users from the admin and village officials to ensure the system is easy to use and meets needs.

6) Implementation

Following the system's rigorous testing phase, the software was uploaded to a local server and subjected to a real-world evaluation. This evaluation involved the analysis of survey data collected from residents of Lemahireng Hamlet. System documentation was developed, and simple training was provided to operators/administrators to maximize the system's features.

7) Maintenance

Maintenance is performed for the purpose of resolving technical issues, adapting to user feedback, and developing additional features. These additional features may include the classification of satisfaction levels according to specific service aspects or the incorporation of sentiment data visualizations.

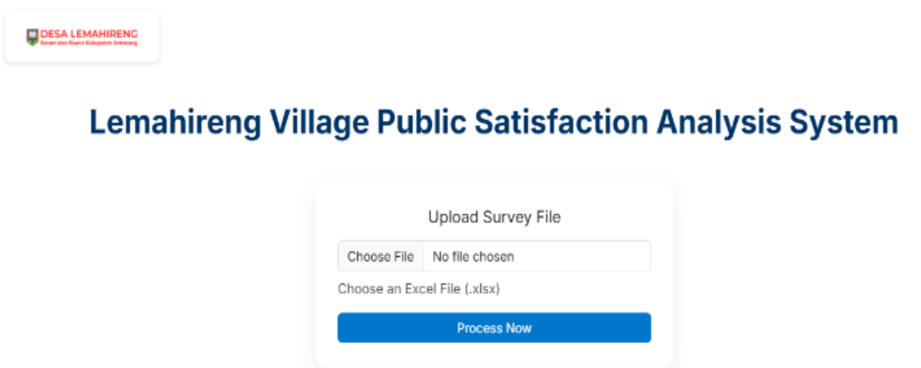


Figure 2. Dashboard View

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

3.4. System Implementation

The subsequent stage is system implementation, wherein the system's design and development outcomes are applied in a real-world environment. This stage facilitates the evaluation of the system's functionality and effectiveness in sentiment analysis for public administration services. The web-based system was constructed using the Laravel 11 framework for backend management, Bootstrap 5.3 for user interface design (frontend), and Python with the Naïve Bayes Classifier algorithm for automatic sentiment classification.

The implementation environment consists of software such as PHP 8.3, Laravel 11, MariaDB as a database management system, and Python 3.10 with supporting libraries such as scikit-learn, nltk, pandas, and Flask as a text classification server. With regard to the system's technical specifications, its functionality was assessed through rigorous testing on devices that met the following minimum hardware requirements: an Intel Core i5 processor, 8 GB of RAM, 128 GB of SSD storage, and a contemporary web browser such as Google Chrome or Mozilla Firefox.

The primary features that have been implemented include a survey file upload form, automatic sentiment classification, the display of survey results in tabular form and visualization, and the ability to export results in PDF or Excel. The workflow initiates upon the upload of an Excel file containing the results of the public satisfaction survey by the user. Subsequently, the data from the open-ended question column is transmitted via the Application Programming Interface (API) to the Python server. The Python server is responsible for preprocessing the data, which includes case folding, tokenization, stopword removal, stemming, and TF-IDF weighting. The server then performs classification using the Naïve Bayes Classifier algorithm. The classification results, which are expressed as positive, negative, or neutral sentiment labels, are transmitted to the Laravel server for subsequent display on the dashboard.

The interface has been meticulously designed to be both responsive and user-friendly. The main page contains a survey upload form with simple yet elegant input elements. The analysis results are displayed in a dynamic table that presents respondent data based on gender, occupation, services received, and responses to a number of service indicators. Each row is demarcated by a color-coded indicator, denoting the sentiment type: green for positive, gray for neutral, and red for negative. This systematic coding facilitates rapid and efficient data interpretation. Additionally, the system employs visual representations, including bar charts and pie charts, to illustrate the aggregate distribution of sentiment and the per-service indicator. The objective of this initiative is to furnish service managers with the tools necessary to identify service areas that require enhancement. The efficacy of the system's implementation has been assessed

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

through the analysis of survey data collected from residents in Lemahireng Hamlet. The analysis revealed that the system successfully classified the data of over 80 respondents with a reasonable degree of accuracy. The majority of respondents provided neutral and positive assessments of public services, particularly with regard to the indicators of officer politeness and the speed of service. This system implementation demonstrates the efficacy of the Naïve Bayes Classifier-based text mining method in systematically and expeditiously evaluating public opinion. This system can be relied upon as an analytical tool in the process of evaluating the quality of public services based on evidence (evidence-based policy), and it can be further developed for various other survey contexts.

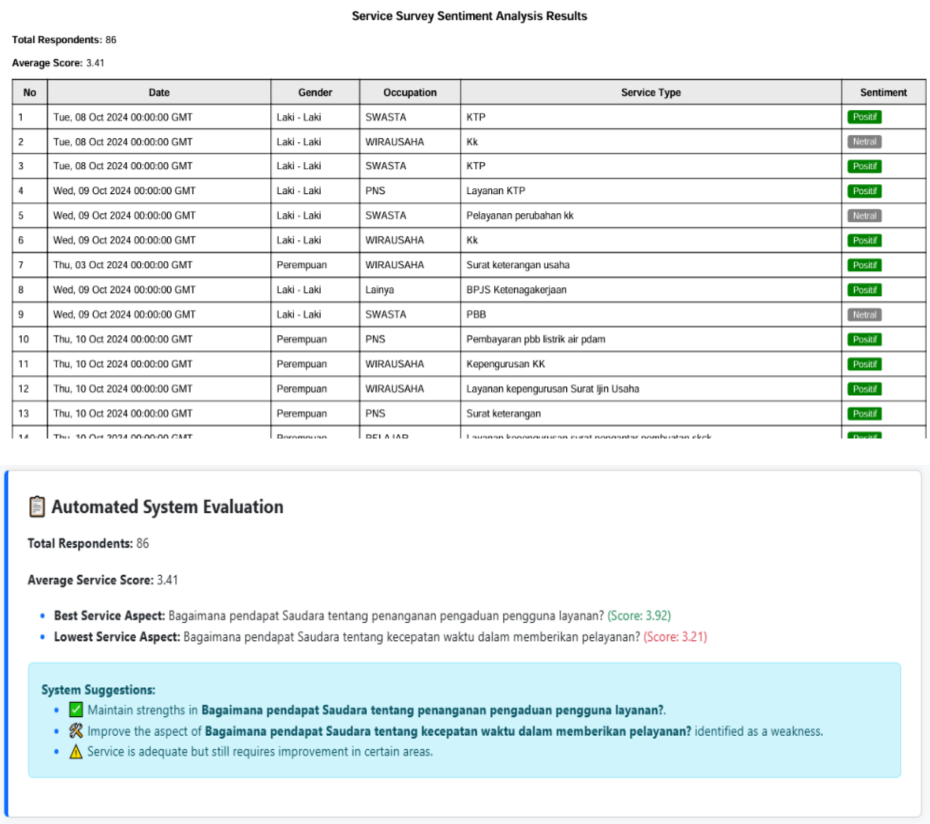
The efficacy of the system's implementation has been assessed through the analysis of survey data collected from residents in Lemahireng Hamlet. The analysis revealed that the system successfully classified the data of over 80 respondents with a reasonable degree of accuracy. The majority of respondents provided neutral and positive assessments of public services, particularly with regard to the indicators of officer politeness and the speed of service. System evaluation further revealed that the best-rated service aspect was complaint handling (score of 3.92), while the lowest-rated aspect was service speed (score of 3.21). This outcome linked to structural challenges such as limited staffing, procedural bottlenecks, or high service demand relative to available resources. In semi-urban contexts like Lemahireng Hamlet, incomplete digital infrastructure and limited automation can further delay service delivery. These findings highlight that, despite general satisfaction with interpersonal aspects of service, efficiency in processing remains a key area for improvement. Targeted interventions, such as simplifying administrative procedures, investing in capacity building, and integrating digital solutions, could significantly enhance public perceptions of service speed.

The relatively low score for service speed attributed to several factors, including limited staffing, high service demand, or inefficiencies in administrative procedures. In semi-urban contexts such as Lemahireng Hamlet, technological infrastructure and digital integration are still developing, which can also contribute to longer processing times. These findings suggest that while respondents were generally satisfied with aspects such as complaint handling and officer politeness, improving the timeliness of service delivery remains a critical area for reform. Addressing this issue through process simplification, capacity building, and digital innovation could enhance overall public satisfaction.

The discussion could provide deeper insights into why certain service aspects (e.g., complaint handling) received better ratings than others (e.g., service speed). This contrast suggests that while interpersonal and responsive aspects of service

Journal of Information Systems and Informatics
Vol. 7, No. 3, September 2025
p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

delivery are being managed effectively, structural and procedural factors continue to hinder efficiency. What can be learned from this is that improving service quality requires a balanced approach: maintaining strong interpersonal engagement while simultaneously addressing systemic bottlenecks. For example, complaint handling may have scored higher because it often involves direct human interaction, empathy, and immediate acknowledgment, which respondents value highly. In contrast, service speed depends more on administrative efficiency, resource allocation, and technological infrastructure, which are more difficult to optimize in semi-urban contexts. Understanding these dynamics can help policymakers prioritize investments—such as streamlining bureaucratic procedures and enhancing digital integration—without losing sight of the human-centered aspects that build trust and satisfaction.


Service Survey Sentiment Analysis Results

Total Respondents: 86

Average Score: 3.41

No	Date	Gender	Occupation	Service Type	Sentiment
1	Tue, 08 Oct 2024 00:00:00 GMT	Laki - Laki	SWASTA	KTP	Positif
2	Tue, 08 Oct 2024 00:00:00 GMT	Laki - Laki	WIRAUSAHA	Kk	Netral
3	Tue, 08 Oct 2024 00:00:00 GMT	Laki - Laki	SWASTA	KTP	Positif
4	Wed, 09 Oct 2024 00:00:00 GMT	Laki - Laki	PNS	Layanan KTP	Positif
5	Wed, 09 Oct 2024 00:00:00 GMT	Laki - Laki	SWASTA	Pelayanan perubahan kk	Netral
6	Wed, 09 Oct 2024 00:00:00 GMT	Laki - Laki	WIRAUSAHA	Kk	Positif
7	Thu, 03 Oct 2024 00:00:00 GMT	Perempuan	WIRAUSAHA	Surat keterangan usaha	Positif
8	Wed, 09 Oct 2024 00:00:00 GMT	Laki - Laki	Lainya	BPJS Ketenagakerjaan	Positif
9	Wed, 09 Oct 2024 00:00:00 GMT	Laki - Laki	SWASTA	PBB	Netral
10	Thu, 10 Oct 2024 00:00:00 GMT	Perempuan	PNS	Pembayaran pbb listrik air pdam	Positif
11	Thu, 10 Oct 2024 00:00:00 GMT	Perempuan	WIRAUSAHA	Kepengurusan KK	Positif
12	Thu, 10 Oct 2024 00:00:00 GMT	Perempuan	WIRAUSAHA	Layanan kepengurusan Surat Ijin Usaha	Positif
13	Thu, 10 Oct 2024 00:00:00 GMT	Perempuan	PNS	Surat keterangan	Positif

Automated System Evaluation

Total Respondents: 86

Average Service Score: 3.41

- **Best Service Aspect:** Bagaimana pendapat Saudara tentang penanganan pengaduan pengguna layanan? (Score: 3.92)
- **Lowest Service Aspect:** Bagaimana pendapat Saudara tentang kecepatan waktu dalam memberikan pelayanan? (Score: 3.21)

System Suggestions:
- ✅ Maintain strengths in **Bagaimana pendapat Saudara tentang penanganan pengaduan pengguna layanan?**.
- 🛠 Improve the aspect of **Bagaimana pendapat Saudara tentang kecepatan waktu dalam memberikan pelayanan?** identified as a weakness.
- ⚠ Service is adequate but still requires improvement in certain areas.

Figure 3. Display of analysis results

This system implementation demonstrates the efficacy of the Naïve Bayes Classifier-based text mining method in systematically and expeditiously evaluating

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

public opinion. It contributes to evidence-based improvements in the quality of public service delivery and can be relied upon as an analytical tool in the process of evaluating the quality of public services. Furthermore, it can be further developed for various other survey contexts.

Handayani and Sulistiyawati (2021) also reported that while sentiment toward politeness and responsiveness of officers was predominantly positive, service timeliness was often highlighted as a point of dissatisfaction in public administration contexts [3]. Similarly, Nurian et al. (2024) found that service delays in local administrative offices were a recurring issue affecting overall satisfaction [7]. In contrast, Gunawan et al. (2018), who applied the Naïve Bayes method in the context of product reviews, observed stronger emphasis on quality over timeliness, suggesting that the relative importance of service speed is highly context-dependent [2]. These comparisons highlight that the findings from Lemahireng Hamlet both align with broader trends in public service sentiment and underscore the unique challenges faced in semi-urban administrative environments.

A comparison to similar studies on sentiment analysis of public services, especially in developing regions, could contextualize the findings and highlight the unique aspects of this study. For instance, while prior works have emphasized common challenges such as delays and administrative inefficiencies, this study contributes a distinctive perspective by focusing on semi-urban contexts like Lemahireng Hamlet, where infrastructural and procedural constraints intersect with evolving digital adoption.

3.5. Model Performance Evaluation

The Naïve Bayes Classifier demonstrated strong performance in classifying sentiment into positive, neutral, and negative categories. The evaluation metrics obtained from the confusion matrix are shown in Table 2.

Table 2. Performance Metrics of Naïve Bayes Classifier

Metric	Value
Accuracy	91.2%
Precision	90.3%
Recall	89.7%
F1-Score	90.0%

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

The application of 10-fold cross-validation produced an average accuracy of 90.8%, indicating stable model performance across different subsets of the dataset. These results confirm that the Naïve Bayes Classifier is effective for sentiment analysis in the context of public service evaluation.

3.5. Service Quality Scores

In addition to sentiment classification, the survey analysis assessed specific service indicators. Among these, Service Speed scored 3.21, the lowest compared to other indicators, suggesting that timeliness remains a key challenge in public service delivery. The highest-rated aspect was complaint handling (3.92), while other aspects such as officer politeness scored above 3.5, indicating generally positive perceptions except for speed. These findings are consistent with previous studies [3], [7], which reported similar issues in timeliness despite positive sentiment toward interpersonal service aspects.

3.6. Discussion

The research successfully designed and implemented a sentiment analysis system for public administration services in Lemahireng Hamlet by utilizing a text mining approach and the Naïve Bayes Classifier (NBC) algorithm. The developed system is capable of automatically, systematically, and efficiently classifying public opinion into three sentiment categories: positive, neutral, and negative. The analysis of survey data revealed that the majority of respondents expressed neutral sentiment (67%), followed by positive sentiment (31%) and negative sentiment (2%). Preliminary findings from the survey data indicate that the majority of respondents expressed positive and neutral sentiments regarding the quality of service. The best-rated service aspect was complaint handling (score of 3.92), while the lowest-rated aspect was service speed (score of 3.21). Notably, the majority of respondents indicated that they were satisfied with the politeness of the officers, the efficiency of the service, and the adequacy of the infrastructure.

The implementation of the NBC model has demonstrated efficacy in managing text data and ensuring sufficient accuracy in the classification process. The system facilitates evidence-based decision-making by presenting analysis results in the form of interactive tables and informative visualizations. Consequently, this system possesses the capacity to function as a valuable evaluative instrument for public service managers, facilitating the ongoing enhancement of service quality. Beyond its technical contribution, this study provides practical implications for policymakers. First, improving service speed should be prioritized through simplification of administrative procedures, better staffing strategies, and the adoption of digital innovations to reduce bottlenecks. Second, the consistently

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

high evaluation of complaint handling suggests that transparent grievance mechanisms should be maintained and expanded. Third, sentiment analysis tools such as the one developed in this study can serve as early-warning systems, allowing policymakers to monitor public opinion continuously and design more responsive interventions.

This system can be replicated and scaled in other regions to strengthen public service quality more broadly. For public service managers, several actionable steps can be considered: 1) Adoption and Customization, Local governments can adopt the system by tailoring the sentiment categories, language models, and survey instruments to reflect regional contexts. 2) Integration with Existing Platforms, the system can be embedded within e-government platforms or public complaint portals to streamline feedback analysis. 3) Capacity Building, training programs for public service managers and frontline staff can ensure effective interpretation and utilization of sentiment analysis results. 4) Policy Feedback Loop, the insights derived from the system should be institutionalized as part of the service evaluation cycle, allowing continuous refinement of policies and practices. 5) Cross-Regional Benchmarking, applying the system across multiple regions will enable comparative assessments of service quality, fostering best practice sharing and collaborative improvement among local administrations. Subsequent refinements to the system may include the incorporation of trend analysis features, real-time integration with social media or online survey forms, and the development of deep learning-based classification models to enhance the precision and scope of the analysis.

This study is not without limitations. First, the analysis was based on data collected from a single geographical location (Lemahireng Hamlet), which may limit the generalizability of the findings to other regions with different demographic or cultural contexts. Second, the sample size was relatively modest, which could affect the robustness of the sentiment classification. Third, while the Naïve Bayes Classifier provided reliable results, it is a relatively simple algorithm compared to more advanced machine learning and deep learning models.

Future research could address these limitations by expanding the dataset to include larger and more diverse populations across multiple regions, thereby enhancing the external validity of the findings. Additionally, integrating more sophisticated approaches—such as Support Vector Machines (SVM), Random Forests, or deep learning architectures like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)—may improve classification accuracy and allow for more nuanced sentiment detection. Exploring real-time integration with social media platforms, multilingual capabilities, and trend analysis could further enhance the system's relevance as a policy tool for public administration. Consequently, this system possesses the capacity to function as a

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

valuable evaluative instrument for public service managers, facilitating ongoing enhancement of service quality in line with evidence-based policy. Subsequent refinements to the system may include the incorporation of trend analysis features, real-time integration with social media or online survey forms, and the development of deep learning-based classification models to enhance the precision and scope of the analysis.

4. CONCLUSION

Tthis study demonstrates the successful application of sentiment analysis using the Naïve Bayes Classifier (NBC) to evaluate public service satisfaction in Lemahireng Hamlet, a semi-urban area facing administrative challenges. The developed system effectively classifies community sentiments, providing valuable insights into the public's perception of village administration services. The findings reveal that while public satisfaction with aspects like officer politeness and complaint handling is generally positive, service speed remains a significant area for improvement. The study highlights the importance of digital integration and streamlining administrative processes to enhance service efficiency. The system's implementation can serve as a model for other regions, enabling local governments to adopt evidence-based decision-making tools to improve service quality continuously.

Moreover, this research offers practical implications for policymakers by recommending targeted interventions, such as capacity building and digital innovations, to address bottlenecks in service delivery. Despite its limitations, such as the sample size and the use of a relatively simple model, the study provides a strong foundation for future research. Expanding the dataset, incorporating more advanced machine learning techniques, and integrating real-time feedback mechanisms can further enhance the system’s accuracy and applicability in various public service contexts. By leveraging sentiment analysis, local administrations can monitor public opinion more effectively and make timely adjustments to improve service quality and responsiveness.

REFERENCES

[1] A. Dwiyanto, *Mewujudkan Good Governance melalui Pelayanan Publik*. Yogyakarta: Gadjah Mada University Press, 2008.
[2] B. Gunawan, H. Sastypratiwi, and E. E. Pratama, "Sistem analisis sentimen pada ulasan produk menggunakan metode Naive Bayes," *J. Edukasi Penelit. Informatika (JEPIN)*, vol. 4, no. 2, pp. 113–118, 2018.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[3] E. T. Handayani and A. Sulistiyawati, "Analisis sentimen respon masyarakat terhadap kabar harian Covid-19 pada Twitter Kementerian Kesehatan dengan metode klasifikasi Naive Bayes," *J. Teknol. Sistem Inform.* (JTSI), vol. 2, no. 3, pp. 32–37, 2021.
[4] M. S. Hudin, M. A. Fauzi, and S. Adinugroho, "Implementasi metode text mining dan K-Means clustering untuk pengelompokan dokumen skripsi (Studi kasus: Universitas Brawijaya)," *J. Pengembangan Teknol. Inform. Ilmu Komput. (JPTIIK)*, vol. 2, no. 11, pp. 5518–5524, 2018.
[5] R. N. Mauliza and Y. R. Sipayung, "Penerapan text mining dalam menganalisis pendapat masyarakat terhadap Pemilu 2024 pada media sosial X menggunakan metode Naive Bayes," *Technomedia J.*, vol. 9, no. 1, pp. 1–16, 2024.
[6] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. New York: Cambridge University Press, 2008.
[7] A. Nurian, T. N. Padilah, and G. Garno, "Analisis sentimen terhadap pelayanan Disdukcapil Karawang menggunakan Naive Bayes Classifier," *J. Informatika Teknik Elektro Terapan (JITE)*, vol. 12, no. 2, 2024.
[8] OECD, *Serving Citizens: Strategies for Better Service Delivery* (OECD Public Governance Reviews). Paris: OECD Publishing, 2017.
[9] H. Pasolong, *Teori Administrasi Publik*. Bandung: Alfabeta, 2019.
[10] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 2, no. 1–2, pp. 1–135, 2008.
[11] M. P. R. Putra and K. R. N. Wardani, "Penerapan text mining dalam menganalisis kepribadian pengguna media sosial," *JUTIM (Jurnal Teknik Informatika Musirawas)*, vol. 5, no. 1, pp. 63–71, 2020.
[12] M. Rahman, *Ilmu Administrasi*, vol. 1. Jakarta: Sah Media, 2017.
[13] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdiscip. Rev. Data Mining Knowl. Discov.*, vol. 8, no. 4, e1253, 2018.
[14] R. Tasi, "Improving the quality of public services through bureaucratic reformation: Human right perspectives," *Jurnal Ham*, vol. 13, p. 589, 2022.
[15] E. Hasibuan and E. A. Heriyanto, "Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping Di Google Play Store Menggunakan Naive Bayes Classifier," *J. Teknik dan Science*, vol. 1, no. 3, pp. [page numbers], 2023.
[16] M. Syarifuddinn, "Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naïve Bayes Dan Knn," *Inti Nusa Mandiri*, vol. 15, no. 1, pp. 23–28, 2020.
[17] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," *Sistemasi*, vol. 10, no. 1, pp. 222–238, 2021.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[18] A. L. Fairuz, R. D. Ramadhani, and N. A. F. Tanjung, "Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter," *Journal of Dinda: Data Sci., Inf. Technol. Data Anal.*, vol. 1, no. 1, pp. 42–51, 2021.

[19] A. Yadollahi, A. G. Shahraki, and O. R. Zaiane, "Current state of text sentiment analysis from opinion to emotion mining," *ACM Comput. Surv.*, vol. 50, no. 2, pp. 1–33, 2017.

[20] A. Henuk-Kacaribu, *Pengantar Ilmu Administrasi*. Yogyakarta, Indonesia: Penerbit Andi, 2020.

[21] K. R. N. Wardani, "Penerapan Text Mining dalam Menganalisis Kepribadian Pengguna Media Sosial," 2022.

[22] I. Rodiyah, H. Sukmana, and L. Mursyidah, *Buku Ajar Pengantar Ilmu Administrasi Publik*. Sidoarjo, Indonesia: Umsida Press, 2021, pp. 1–92.

[23] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: A review," *Artif. Intell. Rev.*, vol. 53, no. 6, pp. 4335–4385, 2020.

[24] G. Rakasiwi, "Bureaucratic reform in the field of human resources apparatus as the main capital in realizing good governance," in *Proc. 5th Int. Conf. Indonesian Social Political Enquiries (ICISPE 2020)*, Semarang, Indonesia, Oct. 9–10, 2020, published Jan. 2021.

[25] M. Brezzi, S. González, D. Nguyen, and M. Prats, "An updated OECD framework on drivers of trust in public institutions to meet current and future challenges," *OECD Working Papers Public Governance*, no. 48, pp. 1–60, 2021.