Sentiment Analysis of Electronic Government Services Utilizing the Naive Bayes Algorithm

Winny Purbaratri¹, Hindriyanto Dwi Purnomo², Danny Manongga², Iwan Setyawan², Hendry²

¹Institut Keuangan-Perbankan dan Informatika Asia Perbanas, Jakarta, Indonesia
²Universitas Kristen Satya Wacana, Salatiga, Indonesia

Article Info

Article history:
Received August 10, 2023
Revised January 15, 2024
Accepted March 26, 2024

Keywords:
First Classifier
Google Play store
Nave Bayes
Opinion Mining
Sentiment Analysis

ABSTRACT

E-Government yang melibatkan penggunaan teknologi komunikasi dan informasi untuk menyediakan layanan publik memiliki tiga kendala. Salah satu kendala tersebut adalah penerapan e-Government oleh pemerintah daerah otonom masih dilakukan secara individual. Selain itu, penerapan website daerah juga tidak didukung oleh sistem manajemen dan proses kerja yang efisien, hal ini sebagian besar disebabkan oleh kurangnya penyiapan peraturan, prosedur, dan terbatasnya jumlah sumber daya manusia. Selain itu, banyak pemerintah daerah yang menganggap penerapan e-Government hanya melibatkan pengembangan situs web pemerintah daerah. Konsekuensinya, penerapan e-Government hanya sebatas pada tahap kematangan dan mengabaikan tiga tahap penting lainnya yang perlu diselesaikan. Tujuan dari penelitian ini adalah untuk mengetahui tingkat persetujuan masyarakat terhadap layanan aplikasi pemerintah. Penelitian ini menggunakan pendekatan Naive Bayes Classifier sebagai metodologinya. Sumber data yang digunakan dalam penelitian ini terdiri dari review pengguna dan komentar yang diperoleh dari Google Play Store. Hasil penyelidikan ini menghasilkan tingkat presisi tertinggi yaitu mencapai skor 83%. Selain itu menunjukkan tingkat akurasi sebesar 83%, tingkat kelengkapan sebesar 100%, dan F-measure sebesar 90,7%.

Copyright ©2024 The Authors.
This is an open access article under the CC BY-SA license.

Corresponding Author:
Winny Purbaratri, +6281219203209
Faculty of Information Technology and Doctoral of Computer Science,
Universitas Kristen Satya Wacana, Salatiga, Indonesia
Email: 982022025@student.uksw.edu.

How to Cite:
This is an open access article under the CC BY-SA license (https://creativecommons.org/licenses/by-sa/4.0/)

Journal homepage: https://journal.universitasbumigora.ac.id/index.php/matrik
1. INTRODUCTION

E-government is a supporting tool for realizing clean and good governance because Indonesia is still new in implementing e-government and has not developed rapidly, so clean and good governance in Indonesia has not been achieved [1]. Moreover, E-gov is a government-initiated digital innovation aimed at enhancing the effectiveness and efficiency of business processes to fulfill the public’s needs [2]. This study combines topic modeling and clustering methods to examine public comments on e-government. The primary goal of the suggested approach is to identify the specific subjects that make up written citizen comments in the e-government setting and enhance the quality of e-services [3]. Research to analyze public opinion related to e-government use, predict public opinion related to e-government use, and analyze modeling topics related to e-government use has been carried out by several researchers [4, 3, 5–8, 2, 9] When a software product is used in a certain context, the interaction between humans and the program is related to the software quality (QinU). Additionally, sentiment analysis is used as a method to assess the caliber of application goods [9–18] one of them is e-government application product [7]. The main objective of the proposed sentiment analysis is to find out citizens’ comments about being written in e-government environments and to improve the quality of e-services [3]. Alpukat Mobile Application is a population application used by the provincial government of DKI Jakarta [5].

E-government research has attracted the attention of several researchers including topics that discuss open government, smart cities and analytics. This study investigates the quality of e-government apps by employing sentiment analysis, a methodology commonly employed to evaluate the quality of software products. Prior research has indicated the efficacy of Naive Bayes algorithms in sentiment analysis for government applications, with diverse outcomes in terms of accuracy and methodology. Several research have examined the utilization of the Naive Bayes algorithm in the Mobile Alpukat and Sentuhh Tanahku applications [11], along with the deployment of the Support-vector-machine approach in the Jamsostek application [15]. Sentiment analysis research that uses different methods is Kuwait e-government with Logistic regression and the result Logistic regression [13], M-Health Peduli Lindungi Application [14] with Lexicon-based and Naive Bayes Method and Mobile JKN [10] with Maximum Entropy and Mutual Information Selection. The findings indicate a significant degree of precision in the classification of sentiment. This study distinguishes itself by employing a novel methodology to examine the implementation of the Alpukat application in DKI Jakarta, hence contributing unique insights within this specific context.

This study presents a novel viewpoint on applying sentiment analysis to enhance e-government services through a comparative examination of prior research. This study contributes to the extant body of literature by employing the Naive Bayes method for sentiment analysis in the context of e-government applications. Prior research has presented varying methodologies and degrees of precision regarding sentiment analysis of government applications. This study differentiates it by focusing its research on the Alpukat application, employed by the DKI Jakarta provincial administration, and demonstrating a significantly elevated degree of precision in sentiment classification. This research offers a novel perspective on the efficacy of the Naive Bayes approach in the context of e-government through a comparative analysis with earlier studies. This study situates itself within the e-government framework, specifically examining the Alpukat application employed by the provincial administration of DKI Jakarta. While there have been prior investigations into sentiment analysis in the context of e-government, this study distinguishes itself by employing a distinct methodological approach and providing novel perspectives on this particular application. This study demonstrates a notable degree of accuracy in sentiment categorization using the Naive Bayes approach. This distinguishes it from prior research endeavors that may have employed alternative methodology or focused on disparate applications. This study contributes significantly to the current body of literature by introducing a fresh viewpoint on the effectiveness of the Naive Bayes methodology in e-government applications. Many of QinU’s measurement models are subjectively incoherent and tied to inefficient measurement formulations. To accurately measure QinU consuming software reviews, use the QinUF framework [19]. Quality of service is fundamental to public acceptance and use of e-government websites, although this aspect is frequently overlooked during the design and implementation stages of online public services [10].

This research aims to determine the level of public approval for government application services and contribute to the field of e-Government related to the use of a unique sentiment analysis approach that uses the Naive Bayes algorithm. This methodology is applied to the Avocado application used by the DKI Jakarta Provincial Government. These differences differentiate this research from previous studies which may have used different methods or concentrated on different uses. This research makes a valuable contribution to the field of e-government by demonstrating the effectiveness of the Naive Bayes technique in sentiment classification. This report offers a new perspective in measuring the quality of user interactions with e-Government services, especially in terms of accuracy.
2. RESEARCH METHOD

The study was conducted in July of 2023. The present study employs an experimental design, utilizing data gathered from the Google Play Store as the primary experimental material. The case study platform utilized by the Jakarta Provincial Population Service is Alpukat Apps. There are six steps to take in this study, shown in Figure 1. This involves gathering the data, manually labeling it, preprocessing it, TF-IDF weighing it, separating the data into training and test sets, classifying the data using the Naive Bayes algorithm, and assessing the results. One of these emotional labels Positive, Neutral, or Negative can be the system’s output.

![Research Methodology](image)

In this study, we have identified a research topic, formulated a research question, and established research objectives. Table 1 addresses the study inquiries and substantiates the research aims. The topic under consideration will be examined and analyzed in the next section, dedicated to findings and analysis.

<table>
<thead>
<tr>
<th>Research Problems</th>
<th>Research Question</th>
<th>Research Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citizens may accept applications submitted by the DKI Jakarta provincial administration.</td>
<td>Is the DKI Jakarta Alpukat Application’s categorization system accurate in gathering public opinion?</td>
<td>The DKI Jakarta provincial government developed the Alpukat Application, which uses an algorithm to provide accurate public opinion results at citizen acceptance.</td>
</tr>
</tbody>
</table>

2.1. Data Scraping

Data scraping from the internet is the first step in gathering and entering data into this system. We use Google Colab. The data gathered from the Google Play Store comprised 100 comments. In this study, training data were derived from 100 prior data, and 20% of the test data were derived from 20 test data. After data is gathered, positive or negative labels are manually assigned, as presented in Figure 4.

2.2. Manual Labeling

The gathered reviews will be classified as good, negative, or neutral. The procedure is manual. Positive review data, negative review data, and neutral review data were all extracted from the labeling findings. The importance of positive, negative, and neutral labels follows.

2.3. Preprocessing

Preprocessing, which consists of several stages, is one of the stages to eliminate problems that can interfere with the outcome of data processing. The preprocessing phase includes. Case Folding: This involves several steps, including removing usernames or user mentions (@), hashtags (#), and non-letter characters, converting text to lowercase (case folding), and eliminating URLs or links from each comment. Filtering: In this stage, unnecessary words are removed from the token results. Commonly occurring words considered insignificant, along with stopwords such as "period" and "link," are also filtered out. Stemming: This process simplifies words into their root or base form. Tokenizing: Each word is segmented into tokens based on identified gaps or spaces within the text.

2.4. Transformation

One technique for creating word weighting in the word extraction process by utilizing standard word calculations on returned data is TF-IDF (Term Frequency-Inverse Document Frequency), showing the formulas (3.5.) and (2). The word frequency and the
concept of inverse frequencies are combined in this weighting technique. Frequency refers to the frequency of occurrence of a term in a document.

\[
TF(term) = \frac{\text{the number of occurrences of the term in the document}}{\text{the total amount of data in the document}}
\]

By examining how often a word occurs across the full collection of documents, the Inverse Document Frequency (IDF) method may determine how essential a word is. Words that are seldom used are given additional weight by this IDF component. The following formula is used to calculate IDF. Typically, the formula is as follows.

\[
IDF(term) = \log \frac{\text{total jumlah dokumen dalam koleksi}}{\text{the number of documents containing the term}}
\]

2.5. Separation of Data

The Naive Bayes Classifier is a popular machine learning method for classification tasks, particularly for text classification on sentiment analysis. The Naive Bayes Model is trained to predict model parameters for this text categorization using labeled training data sets. The Naive Bayes algorithm performs probability calculations and other classification-related calculations throughout the training phase. This parameter includes probability and prior probability based on observed characteristics and class labels in the training data. Here is the source code for using the 'Multinomial ('\texttt{Multinomial}') class from the scikit-learn package to train using Naive Bayes. The training data (X\text{\texttt{train}} and Y\text{\texttt{train}}) are used to train this model.

Training data and test data are the two categories of labeled data. The authors split the training and test data into two groups for the following reasons: first, the total amount of training data is less than the total amount of test data, and second, the total amount of exercise data is more than the total amount of test data.

2.6. Confusion Matrix

To determine the accuracy performance outcomes of the Naive Bayes algorithm, which includes precision, precision, and recall, at this level of assessment, we will employ matrix confusion. One technique for evaluating the effectiveness of categorization models is the confusion matrix. Accuracy (3), Precision (4), Recall (5), and F-measure (6) are the formulas used as follows:

\[
Accuracy = \frac{TP + TN}{TP + FP + FN + TN}
\]

\[
Precision = \frac{TP}{TP + FP}
\]

\[
Recall = \frac{TP}{TP + FN}
\]

\[
F - measure = \frac{2 \times precision \times recall}{Precision + recall}
\]

See also:

True Positive (TP) is the amount of data with a positive value correctly predicted as positive.

True Negative (TN) is the amount of data that has a negative value and is correctly predicted to be negative. False Positive (FP) is several data that is negative but predictable to be positive,
False negative (FN) is a data set with a positive but predicted negative value.

This study is categorized using the Naive Bayes method. The Bayes Theorem, derived from Bayesian statistics with strong independent assumptions, forms the method’s foundation.

2.7. Visualization

The term "cloud" is employed to visually represent the most frequently occurring words. Word clouds are a visually captivating technique for presenting textual data, highlighting the frequency of word occurrences in a given text by adjusting the size of each word. Within the scope of this study, the term "word cloud" was utilized to visually depict the words that were most commonly utilized in user reviews of the Alpukat e-Government service application.

3. RESULT AND ANALYSIS

This part shows the outcomes and analyses based on the sentiment analysis performed on user reviews of the Alpukat e-Government service application utilizing the Naive Bayes Classifier. The main data source utilized for this research was a collection of user reviews directly sourced from the Google Play Store. After the data collection process, 100 reviews were obtained, and the labeling process was carried out manually using positive, neutral, and negative labels. As shown in Table 5, the Alpukat Application gets negative user reviews, as shown in Figure 4. The results of the dataset are shown in Table 2.

The dataset underwent a variety of preprocessing processes before the application of the Naive Bayes Classifier. Preparing the text data for analysis included multiple stages, such as eliminating extraneous characters and words, segmenting the text into discrete units of words, and standardizing these units to minimize duplication and enhance the precision of the analysis. Implementing preprocessing techniques had a crucial role in preparing the dataset for sentiment analysis, which allowed the classifier to concentrate on the most significant aspects of the text.

The preprocessed data was subjected to the Naive Bayes Classifier to classify each review into one of three sentiment categories: positive, Neutral, or Negative. The categorization was conducted utilizing the probabilistic model of the Naive Bayes technique, which computes the probability of a given text being assigned to a specific category by considering the frequency and distribution of words.

The investigation’s findings indicated that the users of the Alpukat application expressed a wide spectrum of sentiments. The categorization of reviews into sentiment groups was as follows: The statistical data provided illustrates the distribution of reviews across different categories. The classification accuracy was represented using a confusion matrix, which included the identification of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) instances. Show formula (3).

Furthermore, a word cloud was generated to graphically represent the prevailing words within the various categories of reviews. Utilizing this visualization tool facilitated the identification of the predominant themes and characteristics frequently referenced by users, hence enhancing the comprehension of the underlying sentiments in a more nuanced manner.

<table>
<thead>
<tr>
<th>Sentiment</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>16</td>
</tr>
<tr>
<td>Neutral</td>
<td>1</td>
</tr>
<tr>
<td>Negative</td>
<td>83</td>
</tr>
</tbody>
</table>

3.1. Data collection

In this study, a total of 100 data points comprising 100 training data and 100 test data were sourced from Google Play data. Web scraping was employed as the method for data collection during the user evaluation of the Alpukat application. Google Colab and the provided software code were utilized to obtain the data. The outcomes of the web scraping process are illustrated in Figure 2.
After scraping data with Google Colab, the result of data scrapping is shown in Figure 3.

![Image](image-url)

**Figure 3. Scrapping Data with Google Colab Result**

After scraping data, we have to label positive, neutral, and negative data. The procedure is manual. The outcomes of this labeling are described as data training and data testing. Table 3 shows example labeling data.

<table>
<thead>
<tr>
<th>Label</th>
<th>Teks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Terimakasih bpk/ibu yang sudah membantu selama ini, pengurusan apa pun hanya dengan online sudah selsai semuanya tanpa ribet, tanpa pungut biaya, gratis - gratis.. terimakasih pemerintah jakarta. semoga lebih di tingkatkan lg untuk aplikasi alpukat betawinya.</td>
</tr>
<tr>
<td>Neutral</td>
<td>Gg</td>
</tr>
</tbody>
</table>

---

3.2. Preprocessing

Pre-processing is a preliminary step in text mining that is carried out to weed out unnecessary words from the document and extract valuable information from unstructured data. Case Folding, Filtering, Tokenizing, and Stemming are the four steps in data preparation.

1. Case folding is the data processing stage that makes all letters into small letters. Table 4 shows the results of the case folding stage.

<table>
<thead>
<tr>
<th>Table 4. Case Folding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teks</td>
</tr>
<tr>
<td>terimakasih bpk/ibu yang sudah membantu selama ini, pengurusan apa pun hanya dengan online sudah selsai semuanya tanpa ribet, tanpa pungut biaya, gratis -gratis,.. terimakasih pemerintah jakarta. semoga lebih di tingkatkan lg untuk aplikasi alpukat betawinya.</td>
</tr>
</tbody>
</table>

2. Filtering: Removing words from token results is not important. Readings and stopwords are removed. The stopword is also removed if a sentence contains words that often come out and are considered unimportant, such as time, links, etc. Example filtering data is show in Table 5.

<table>
<thead>
<tr>
<th>Table 5. Filtering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teks</td>
</tr>
<tr>
<td>terimakasih bpk/ibu yang sudah membantu selama ini, pengurusan apa pun hanya dengan online sudah selsai semuanya tanpa ribet, tanpa pungut biaya, gratis -gratis,.. terimakasih pemerintah jakarta. semoga lebih di tingkatkan lg untuk aplikasi alpukat betawinya.</td>
</tr>
</tbody>
</table>

3. Stemming: The term is now reduced to its most fundamental form. Table 7 shows an example of stemming data.

<table>
<thead>
<tr>
<th>Table 6. Stemming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teks</td>
</tr>
<tr>
<td>trimakasi dengan adanya aplikasi ini. Pengalaman yg menyenangkan buat saya. Saat saya butuh potocopy KK dan saat itu saya kehabisan potocopy KK saya. Saya mencoba meminta melalui aplikasi ini.. ternyata respon yg sangat cepat saya terima langsung. Dan sy dikirimkan KK yang saya butuhkan dgn fomat PDF. Trimakasi Dukcapil dan orang yg bernama Adi Prihartono yg dengan senang hati merespon permohonan saya pada aplikasi Alpukat Betawi ini. Tetaplah berkarya demi bangsa.</td>
</tr>
</tbody>
</table>

4. Tokenizing: Each word is now divided based on the discovered gaps. Table 8 shows an example of tokenizing.

<table>
<thead>
<tr>
<th>Table 7. Tokenizing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teks</td>
</tr>
</tbody>
</table>

5. Tokenizing: Each word is now divided based on the discovered gaps. Table 8 shows an example of tokenizing.

Sentiment Analysis of …(First Author)
### 3.3. Transformation

Transforming data during the transformation step into forms that can be processed [20].

### 3.4. Separation of Data

After separation, data is divided into training and test sets. 80 percent of the training data and 20 percent of the test data from 100 data sets successfully scraped from Google Play made up the 8:2 total separation of the data in this research. The separation of data is shown in Table 9.

Table 9. Separation Of Data

<table>
<thead>
<tr>
<th>Data Review</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Training</td>
<td>80</td>
</tr>
<tr>
<td>Data Testing</td>
<td>20</td>
</tr>
</tbody>
</table>

### 3.5. Analysis and Evaluation

Using the Naive Bayes technique, we conducted sentiment analysis on Alpukat app reviews and compared the results across different categories. Table 10 contains the confusion matrix generated from the test data of 20 data.

Table 10. Confusion Matrix

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predictive</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Negative</td>
<td>0</td>
<td>83</td>
</tr>
<tr>
<td>Neutral</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

According to the data in Table 10, there are 16 predictable positive review data, 0 predictable positive review data as negative reviews, 0 predicting negative review data as positive reviews and 83 accurately forecast negative review data out of the 100 data predicted using the Naive Bayes Classifier algorithm. This study uses accuracy testing as a means to assess the performance of the classification approach using equation for Accuracy (3), equation (4) for Precision, equation (5) for Recall, and equation (6) for F-measure Value. Below are the classification calculation’s outcomes based on the confusion matrix:

\[
Accuracy = \frac{83 + 0}{(83 + 17 + 0 + 0)} = \frac{83}{100} = 0.83 \times 100\% = 83\%
\]

\[
Precision = \frac{83}{83 + 17} = \frac{83}{100} = 0.83 \times 100\% = 83\%
\]

\[
Recall = \frac{83}{83 + 0} = \frac{83}{83} = 1 \times 100\% = 100\%
\]

\[ F - \text{measure} = \frac{2 \times 83 \times 100}{83 + 100} = \frac{16600}{183} = 90.7\% \]

The level of accuracy in the Nave Bayes algorithm is depicted in Figure 4 of the chart diagram.

The comparison diagram between positive, negative, and neutral sentiment is shown in Figure 5.
After comparing diagram accuracy, Figure 6. shows value Accuracy, precision, and recall.

![Figure 6. Alpukat application user review diagram](image)

3.6. Visualization

From the data preprocessing results obtained, data visualization is used to see the words often used in reviews by users of the Alpukat application. This data visualization is carried out on the overall data, which is both positive and negative. Visualization of data presented in the Word cloud is shown in Figure 7.

![Figure 7. Word cloud analyses sentiment user reviews of applications](image)

4. CONCLUSION

The present study effectively used the Naive Bayes method for sentiment analysis of e-government services, specifically focusing on the Alpukat application within the Special Capital Region of Jakarta. The analytical findings indicate the efficacy of this methodology in discerning and classifying user sentiment, hence offering significant insights into user contentment. This study validates the significance of sentiment analysis within the e-government domain. It presents prospects for future investigations to enhance the caliber of public services by gaining a deeper comprehension of user input. The findings of this study have significant implications that can serve as a foundation for policymakers in formulating policies to enhance user satisfaction with e-government services.
5. ACKNOWLEDGEMENTS

Thanks to Universitas Kristen Satya Wacana for supporting the research.

6. DECLARATIONS

AUTHOR CONTRIBUTION

The first authors collected data, forecasted data, analyzed data, and preprocessed data. All writers thoroughly revised and analyzed the paper.

FUNDING STATEMENT

Universitas Kristen Satya Wacana provided the funding for this project.

COMPETING INTEREST

The Authors affirm that this research does not have any conflict of interest.

REFERENCES


Sentiment Analysis of . . . (First Author)


