International Journal on Advanced Science Engineering Information Technology

Vol.13 (2023) No. 5
ISSN: 2088-5334

Monitoring the Quality of *PeduliLindungi* Application based on Customer Reviews on Google Play Using Hybrid Naïve Bayes -*Laney p'* Attribute Control Chart

Nia Triamalia Apsari [a], Muhammad Ahsan [a,*], Muhammad Hisyam Lee [b]

[a] *Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*
[b] *Department of Mathematical Sciences, Universiti Teknologi Malaysia, Johor Bahru, Malaysia*
*Corresponding author: *muh.ahsan@its.ac.id*

***Abstract*—Indonesia is battling the COVID-19 pandemic. One of the government's strategies to break the virus's transmission chain is to track digital contacts in Indonesia using the *PeduliLindungi* application. The Google Play comment section is where users can express their opinions about the app. User opinions discovered on Google Play can be used to perform sentiment analysis and quality evaluation. The Naïve Bayes classification can be used to identify how user opinions contain positive, neutral, or negative sentiments in user reviews of the *PeduliLindungi* app on Google Play. The *p* and *Laney p'* charts can be used for quality evaluation. *Laney p'* control chart is an attribute chart used to monitor the proportion of defects with large and varied sample sizes. The data used in this study is from April 1, 2020, to March 31, 2022. According to the sentiment analysis results of user reviews of the *PeduliLindungi* app on Google Play, there are more negative reviews than positive classes. The classification accuracy has an Area Under Curve (AUC) value of 89.05%. This result shows that the test data has good classification. The monitoring results using *p* and *Laney p'* charts based on ratings and user reviews of the *PeduliLindungi* app show that the processes are still not statistically controlled. These findings indicate that the app developer still needs to make improvements.**

***Keywords*— Sentiment analysis; *Laney p'* control chart; *p* control chart; naïve bayes classifier; *PeduliLindungi*.**

Manuscript received 21 Aug. 2022; revised 21 May 2023; accepted 28 Jun. 2023. Date of publication 31 Oct. 2023.

I. INTRODUCTION

A novel coronavirus, SARS-CoV-2, was first detected in Wuhan, China, in December 2019. COVID-19, the disease caused by SARS-CoV-2, attacks the respiratory system. The virus can spread from person to person through respiratory droplets produced when an infected person coughs or sneezes or through contact with contaminated surfaces. The World Health Organization (WHO) declared COVID-19 a pandemic on March 11, 2020, due to the rapid spread of the virus and its ability to cause severe illness in some people[1].

President Joko Widodo announced Indonesia's first positive case of COVID-19 on March 2, 2020. The number of cases in Indonesia continued to rise rapidly until February 2022, with three waves of the pandemic. The Indonesian government has implemented many policies to combat the pandemic, including contact tracing. *PeduliLindungi* is an application developed to assist government agencies in contact tracing. The application relies on community participation to share location data, which is used to track contact history with COVID-19 cases. The application is available for download on Google Play, the App Store, and the App-Gallery.

Applications have both positive and negative aspects, which can lead to a variety of user responses. One place where users can express their opinions about an application is in the comments section on Google Play. The reviews of *PeduliLindungi* users found on Google Play can be used to evaluate the application's sentiment and quality. In everyday life, businesses and organizations want to know what consumers or the public think about their products and services. Sentiment analysis is a field of study that analyzes people's opinions, judgments, attitudes, and emotions about products, services, events, and attributes.[2].

The large number of customer reviews in the comments section makes it difficult and time-consuming to read them all. Therefore, a classification method is used to categorize existing opinions according to their sentiment class.

Sentiment classes are divided into positive, neutral, and negative. Several previous researchers have conducted research on sentiment analysis using various algorithms, including: the Naïve Bayes algorithm, Support Vector Machine (SVM) [3]–[6], K-Nearest Neighbor (K-NN) [7]–[10], Decision Tree [11]–[13], and Neural Network [14]–[17].

In this study, the classification method used is Naïve Bayes. Naïve Bayes classification is a simple method that applies Bayes' theorem to classify data, assuming that the predictors are independent. [18]. In simple terms, Naïve Bayes classification assumes that the presence of a feature in a class is independent of the presence of any other feature [19]. The advantage of Naïve Bayes classification for text classification is that it can be adapted to each data set's specific nature and needs. Naïve Bayes classification is often used in text classification because it is a simple and effective method.

Statistical process control (SPC) is a statistical method that can be used to monitor and control a process. SPC uses data to identify and correct problems in a process before they cause defects in the output. SPC can be used to improve the quality of products and services, reduce costs, and increase productivity. [20]. SPC aims to improve and maintain the quality of the process by reducing variability [21]. Statistical process control (SPC) is a statistical method that uses simple graphical methods to solve problems. One of the seven tools of SPC is the control chart. Control charts are used to monitor a process over time and identify changes that may indicate a problem. Control charts can be divided into the variable [22]–[27] and attribute [28]–[31] based on the type of variable inspected.

The *p* control chart is an attribute control chart to monitor the proportion of defects with a different sample size for each observation [20]. However, the *p* control chart is sensitive to samples that have large sizes and variations. Therefore, Laney introduced the *Laney p*' control chart for research with large and varied sample sizes. Ahsan et al. conducted a study to investigate the suitability of the *Laney p'* control chart for use with a variable number of samples with a large sample size. The results of the study showed that the *Laney p'* control chart is a more robust tool than the traditional *p* chart for monitoring processes with a variable number of samples and a large sample size. This is because the control limits on the *Laney p'* chart are wider, allowing for more data variation without triggering a false alarm [32].

User reviews of the *PeduliLindungi* application on Google Play can be used to identify the sentiment of user opinions with the Naïve Bayes classification algorithm. The defects or negative comments distribution can be identified with p and Laney p' control charts. The research will be conducted from April 1, 2020, to March 31, 2022. The results of this study are expected to be used as evaluation material regarding the quality of the *PeduliLindungi* application and to identify the obstacles experienced by application users.

Based on this description, The authors are interested in conducting research on monitoring the quality of the *PeduliLindungi* application based on customer reviews on Google Play using the *Laney p'* attribute control chart. The *Laney p'* attribute control chart is a statistical tool that can be used to monitor the quality of a product or service by tracking the number of defects. The authors believe this research can help improve the quality of the *PeduliLindungi* application by identifying and addressing any potential problems.

II. Material and Methods

A. Descriptive Statistics

Descriptive statistics are a set of methods used to summarize and describe the main features of a dataset. Descriptive statistics can be used to present data concisely and attractively and identify patterns and relationships in the data. The descriptive statistics used in this study were bar charts. Bar charts are a type of graph that can be used to represent qualitative data. The horizontal axis of a bar chart represents the different values of the qualitative data, and the vertical axis represents the frequency of each value. The height of each bar in a bar chart is proportional to the frequency of the value that it represents.

B. Text Mining

Text mining is a process of extracting knowledge from unstructured textual data. This knowledge can be in the form of patterns, trends, or relationships. Text mining can extract information from various sources, including documents, emails, and social media posts [33]. Textual data include online media articles, status or Twitter tweets, user reviews, etc. The stages in doing text mining include Knowledge Discovery Goal, Data Preparation, Data Pre-processing, Data Modeling, Evaluation, and Knowledge and Result. Text mining has almost the same properties as data mining. However, there are more stages in the pre-processing stage of text mining, as follows [34].

- Case Folding is the process of converting all capital letters into lowercase letters.
- Tokenizing is the process of breaking down what was originally a sentence into words based on each word that composes it. The tokenizing process also includes the process of removing numbers, removing punctuation marks, and removing spaces.
- Filtering is a process used to retrieve important words from the tokenizing process.
- Stemming is the process of changing words into basic words by removing prefixes, suffixes, and insertions.

C. Sentiment Analysis

Sentiment analysis is a natural language processing (NLP) technique that identifies, extracts and quantifies opinions, sentiments, evaluations, judgments, attitudes, and emotions towards entities such as products, services, organizations, problems, events, topics, and their attributes. Sentiment analysis is used to obtain information from the process of understanding, extracting, and processing textual data. Companies use sentiment analysis to monitor social media, monitor brands, customer feedback, customer service, and market research.Sentiment analysis can be used to detect production or service problems, identify areas for improvement, and track customer satisfaction. It can also be used to identify emerging trends and opportunities. By understanding customer sentiment, companies can make better decisions about product development, marketing, and customer service.

D. Holdout Validation

Holdout validation is a resampling method for evaluating the performance of a machine learning model. The data is split into two sets: training and test sets. The model is trained on the training set and then evaluated on the test set. The holdout validation method is particularly useful for large datasets, as it allows for a more accurate assessment of the model's performance on unseen data.[35]. Holdout stratification is a technique used to ensure that each class is represented proportionally in the training and testing data. This is important because it helps ensure the model is not biased towards any class. The holdout stratification process involves first creating a stratified sample of the data, where each class is represented in proportion to its actual size in the population. The data is then randomly divided into training and testing sets, with each set containing the same proportion of each class. This ensures that the model is trained on a representative sample of the data and that the testing data is representative of the population as a whole [36].

E. Naïve Bayes Classification

Naïve Bayes is a supervised learning algorithm that uses Bayes' theorem to classify new data points. The algorithm assumes that the features of a data point are independent of each other, given the class label. This assumption is called "naive" because it is often not true in real-world data. However, the Naïve Bayes algorithm can still be effective in many cases, especially when the number of features is small. [37]. Naïve Bayes classification is often used in text classification.

Naïve Bayes classification is a supervised learning algorithm that uses Bayes' theorem to classify new data points. The algorithm first trains on a set of labeled data points, which are used to create a classifier model. The classifier model is then used to predict the class of new data points. The accuracy of the classifier can be measured by comparing its predictions to the actual classes of the new data points. Naïve Bayes classification is a simple and effective algorithm that can be used to classify a variety of data types. Bayesian theory is stated in equation (1)

$$P\,(c_i|x) = \frac{P\,(x|c_i)\,P\,(c_i)}{P_x}, i = 1,2,\ldots,k \quad (1)$$

The value of P_x is a constant that has a constant value, so it can be omitted because if the probability values are sorted by class, then the value of P_x will always be the same [38]. Then equation (1) can be simplified into equation (2).

$$P\,(c_i|x) = P\,(x|c_i)\,P\,(c_i) \quad (1)$$

F. Classification Accuracy

Classification measurements are used to evaluate the performance of a classification model. A confusion matrix is a table used to summarize a classification model's results. The rows of the confusion matrix represent the actual classes of the data points, and the columns represent the predicted classes of the data points. Table 1 contains a 3 × 3 Confusion matrix.

The accuracy of a classification model can be assessed using two metrics: accuracy and the area under the curve (AUC). Accuracy is the percentage of correctly classified documents, while AUC measures the model's ability to distinguish between the two classes. A higher accuracy and AUC indicate a better performing model. The formula for calculating accuracy is found in equation (3).

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \quad (2)$$

TABLE I
CONFUSION MATRIX 3X3

Actual Class	Prediction Class		
	Class 1	Class 2	Class 3
Class 1	X_{11}	X_{12}	X_{13}
Class 2	X_{21}	X_{22}	X_{23}
Class 3	X_{31}	X_{32}	X_{33}

Area Under Curve (AUC) is an indicator of ROC (Receiver Operating Characteristic) curve performance that can summarize the performance of a classifier into one value. The AUC value is an opportunity for the classification method to give a higher random positive test case rating than the random negative test. The AUC value lies in the interval 0 to 1, so the closer to 1 the AUC value is, the better [39]. The scale and interpretation used to explain the AUC values are described in Table 2.

TABLE II
INTERPRETATION AUC VALUE

AUC value	Interpretation
0.90 – 1	Excellent Classification
0.80 – 0.89	Good Classification
0.70 – 0.79	Fair Classification
0.60 – 0.69	Poor Classification
0.50 – 0.59	Failure

G. Pareto Chart

A Pareto chart is a bar graph that depicts a simple frequency distribution of data attributes arranged categorically. The problems that occur the most are shown by the highest first bar and are placed on the far left, and so on, until the least problems are shown by the last lowest bar graph and are placed on the far right, as shown in Figure 1. Pareto charts are often used to measure and analyze DMAIC (Define, Measure, Analyze, Improve, and Control). Pareto charts cannot be determined automatically, but most often [20].

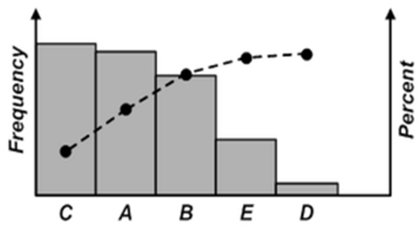

Fig. 1 Pareto Chart

H. p control chart

The *p* control chart is one of the attribute control charts used to control the product's non-conforming (defective) parts. The observations used in making the *p* control chart can have the same or different numbers. The control chart *p* is made by using the size of the defect in the form of the proportion of defective products in each sample taken. The following is the proportion of defective units for each sub-group.

$$p_j = \frac{D_j}{n_j} \quad (4)$$

This control chart's statistical principle is based on binomial distribution. The following is the formula for determining the control limit and the center line in the *p* control chart.

$$\begin{aligned} UCL &= \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_j}} \\ CL &= \bar{p} \\ LCL &= \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_j}} \end{aligned} \quad (5)$$

The average proportion is as follows.

$$\bar{p} = \frac{\sum_{j=1}^{m} D_j}{\sum_{j=1}^{m} n_j}$$

I. Laney p' control chart

This chart was developed by combining the principles of the Z control chart and Donald Wheeler's concepts. The *Laney p'* control chart is used to overcome the sampling problem on the p control chart; when the sample size is very large, the control limit becomes narrower, so there are many observations from the control limit [32]. The formula for the proportion of defective units for each sub-group is given in equation (4). The control limits and the center line on the *Laney p'* control chart are as follows.

$$\begin{aligned} UCL &= \bar{p} + 3\sigma_{p_j}\sigma_z \\ CL &= \bar{p} \\ LCL &= \bar{p} - 3\sigma_{p_j}\sigma_z \end{aligned} \quad (3)$$

With σ_z is the sigma for the individual chart. Standardization is carried out with the following formula.

$$z_j = \frac{p_j - \bar{p}}{\sigma_{p_j}}$$

Then calculate the vector R_j in equation (7) and the average vector R_j in equation (8).

$$R_j = |z_j - z_{j-1}|, j = 1,2,\dots,m \quad (4)$$

$$\bar{R_j} = \frac{1}{m-1}\sum_{j=2}^{m} R_j \quad (5)$$

Therefore σ_z can be calculated by the following formula.

$$\sigma_z = \frac{\bar{R_j}}{1.128}$$

$$\sigma_{p_j} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_j}}$$

According to research conducted by Ahsan et al [32] for the same number of subgroup samples and varying sizes <100 (small sample size), the performance of the p and *Laney p'* control charts is almost the same. In research with a sample size of 200-500 (medium sample size) for the same number of subgroup samples, the performance of the p and *Laney p'* control charts is almost the same, but for the varying number of subgroup samples, the p control chart is more sensitive than the *Laney p'* control chart. Meanwhile, for research with a sample size of 5,000-10,000 (big sample size) for the number of samples for each subgroup, the control chart p becomes more sensitive than the *Laney p'* control chart for the number of samples for each subgroup varies, the p control chart becomes oversensitive than the *Laney p'* control chart. it is causing many observations to go out of control. It can be concluded that the *Laney p'* control chart is suitable for use for varying sample sizes with large sample sizes because the control limit on the *Laney p'* control chart is more comprehensive so the results are more rational [32].

J. PeduliLindungi Application

PeduliLindungi is a digital contact tracing application that uses Bluetooth Low Energy (BLE) technology to collect and store contact information from users who have been in close proximity to each other. Government agencies can then use this information to track the spread of COVID-19 and identify potential cases [40]. Users of the *PeduliLindungi* application will be notified if they are in a crowded area or in a red zone, which is an area where there have been confirmed cases of COVID-19 or where people are under surveillance. The application was made mandatory for all modes of transportation in Indonesia on August 28, 2021. The *PeduliLindungi* application can also help to reduce physical contact by eliminating the need to carry paper documents such as COVID-19 test results or vaccination certificates. This can help to prevent the spread of COVID-19 by reducing the risk of transmission through contact with contaminated surfaces.

K. Data Source

This research uses secondary data obtained through the Google Play website scraping method of the *PeduliLindungi* application from April 1, 2020, to March 31, 2022, with the URL: *https://play.google.com/store/apps/details?id=com.telkom.tracencare&hl=en&gl=US* accessed on April 3, 2022. The data is divided into two phases of process control: phase I data is from April 1 to December 31, 2020, and phase II data is from January 1, 2021, until March 31, 2022.

L. Research Variable

In this study, a naïve Bayes classifier was used to classify user reviews of the *PeduliLindungi* application into positive, neutral, and negative sentiment classes. The predictor variables were the basic words in each user review, and the response variable was the classification of user sentiment. Variables used for *p* and *Laney p'* control chart in this research are negative class sentiment from the customer reviews of the *PeduliLindungi* application with the following types of defects (see Table 3).

TABLE III
TYPES OF DEFECTS PEDULILINDUNGI APPLICATION

Category	Types of Defects
1	Slow application
2	Applications use large memory
3	Application crashes frequently
4	The application uses a lot of battery
5	COVID-19 test results do not show up
6	Personal data leak

Category	Types of Defects
7	Difficulty logging in or registering
8	Difficulty scanning the QR code
9	Difficulty entering the date of birth
10	The vaccination certificate does not show
11	Application displays hard to understand
12	The location zone on the app is not accurate

M. Data Structure

PeduliLindungi application user reviews data is divided into training and testing data using the holdout validation method. The data structure of the Naïve Bayes classification used in this study is presented in Table 4.

TABLE IV
NAÏVE BAYES DATA STRUCTURE

No.	Username	Review $(\boldsymbol{u})$	Keyword				Sentiment class $(\boldsymbol{y})$
			$\boldsymbol{w_1}$	$\boldsymbol{w_2}$	$\cdots$	$\boldsymbol{w_k}$	
1	$user_1$	u_1	$w_{1,1}$	$w_{1,2}$	$\cdots$	$w_{1,k}$	y_1
2	$user_2$	u_2	$w_{2,1}$	$w_{2,2}$	$\cdots$	$w_{2,k}$	y_2
3	$user_3$	u_3	$w_{3,1}$	$w_{3,2}$	$\cdots$	$w_{3,k}$	y_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
n	$user_n$	u_n	$w_{n,1}$	$w_{n,2}$	$\cdots$	$w_{n,k}$	y_n

The data structure of the p and *Laney* p' control chart used in this study is contained in Table 5.

TABLE V
NAÏVE BAYES DATA STRUCTURE

Day	Number of Reviews	Type of Defect					Defect (Negative Class)
		$\boldsymbol{x_1}$	$\boldsymbol{x_2}$	$\boldsymbol{x_3}$	$\cdots$	$\boldsymbol{x_{12}}$	
1	n_1	$n_{1,1}$	$n_{1,2}$	$n_{1,3}$	$\cdots$	$n_{1,12}$	D_1
2	n_2	$n_{2,1}$	$n_{2,2}$	$n_{2,3}$	$\cdots$	$n_{2,12}$	D_2
3	n_3	$n_{3,1}$	$n_{3,2}$	$n_{3,3}$	$\cdots$	$n_{3,12}$	D_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
j	n_j	$n_{j,1}$	$n_{j,2}$	$n_{j,3}$	$\cdots$	$n_{j,12}$	D_j

N. Research Flow

The analytical method used in this study is the p and *Laney* p' control chart. The research flow in this research is as follows.

1. Formulate the problem and determine the research objectives.
2. Conduct a literature study.
3. Collect data on the Google Play website with the title *PeduliLindungi* page using the scraping method through the Python application.
   a. Download and install the Google Play Scraper Package.
   b. Import the required packages.
   c. Copying the App Id (*com. telkom.tracencare.*)
   d. Scraping user reviews.
   e. Converts review data to Pandas Data Frame.
4. Labeling the sentiment class.
5. Preprocessing the text of the *PeduliLindungi* application user review data using the RStudio application, which includes case folding, cleaning punctuation, cleaning number, stemming, stop word, and tokenizing.
6. Create and interpret the Pareto chart of user constraint data (negative comments) on user reviews of the *PeduliLindungi* application for p and *Laney* p' control charts.
7. Distribute training and testing data using holdout validation with a ratio of 80%:20%.
8. Performing Nave Bayes Classification on customer review data of *PeduliLindungi* application users by dividing them into positive, neutral, and negative classes.
9. Calculating the value of accuracy and AUC of the formed Naïve Bayes Classification model. The results of the Naïve Bayes classification will be used for p and *Laney* p' control charts.
10. Analyze the data on the number of user constraints (negative comments) on user reviews of the *PeduliLindungi* application using the p and *Laney* p' control chart for phase I.
11. Analyzing user constraint data (negative comments) on user reviews of the *PeduliLindungi* application using p and *Laney* p' control chart phase II using the average proportion of defects in phase I.
12. Comparison of the control chart analysis that has been done.
13. Draw conclusions and suggestions.

III. RESULT AND DISCUSSION

A. Data Characteristics

This study used user review data for the *PeduliLindungi* application on Google Play, which was obtained by scraping the Google Play website. The data was analyzed using data visualization techniques, namely bar charts and Pareto charts.

1) Bar Chart

Bar charts are a type of data visualization that can be used to show the frequency of data points in a categorical variable. In this study, a bar chart was used to show the number of negative reviews and reviews each month for the *PeduliLindungi* application on Google Play. The following is a bar chart of the number of user reviews of the *PeduliLindungi* application on Google Play.

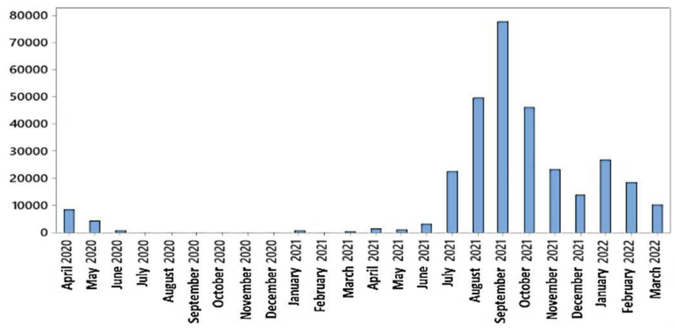

Fig. 2 Number of User Reviews Caring for App on Google Play

Figure 2 shows the number of user reviews of the *PeduliLindungi* application on Google Play from April 2020 to March 31, 2022. The number of reviews varies significantly, with a highest of 78,028 reviews in September 2021 and a lowest of 46 reviews in November 2020. The number of reviews is likely affected by several factors, including app updates and government policies. The total number of reviews over the two-year period was 313,367.

2) *Pareto Chart*

The Pareto chart is a type of data visualization that can be used to identify the most important problems in a set of data. The Pareto chart for the *PeduliLindungi* application on Google Play from April 1, 2020, to March 31, 2022, shows that the most common complaints are about the application's performance, its features, and its user interface. The following is a Pareto chart of the number of complaints from users of the *PeduliLindungi* application on Google Play.

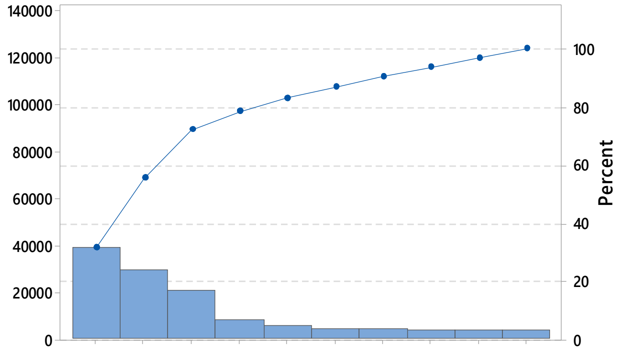

Fig. 3 Pareto Chart of *PeduliLindungi* Application Constraints

The Pareto chart in Figure 3 shows the most common problems PeduliLindungi application users face from April 1, 2020, to March 31, 2022. The most common problem was that vaccination certificates did not appear, accounting for 31.4% of all problems. The second most common problem was difficulty logging in or registering, accounting for 23.9% of all problems. The third most common problem was that the application crashed frequently, accounting for 16.7% of all problems. The category descriptions are provided in Table 3.

B. PeduliLindungi Users Review on Google Play Classification Using Naïve Bayes Classification

The user review data for the *PeduliLindungi* application on Google Play was used to train a classifier. The classifier was able to categorize the reviews into two classes: positive and negative. The training data comprised 250,694 user reviews, and the testing data comprised 62,673 user reviews. The classifier correctly classified 90% of the reviews in the testing data.

TABLE VI
CLASSIFICATION PROBABILITY RESULT

Class	Probability
Negative	0.461
Neutral	0.113
Positive	0.427

TABLE VII
CONFUSION MATRIX TRAINING DATA

Actual	Prediction		
	Negative	Neutral	Positive
Negative	94,282	5,518	6,253
Neutral	20,669	16,631	64,872
Positive	547	6,115	35,807

The classification results of the *PeduliLindungi* application user reviews are shown in Table 6. The table shows the probability of each review being classified as negative, neutral, or positive. The classifier correctly classified 46.1% of the negative reviews, 11.3% of the neutral reviews, and 42.7% of the positive reviews. The performance of the classifier on the training data is summarized in Table 7.

The performance of the Naïve Bayes classifier on the training data is summarized in the confusion matrix. The confusion matrix shows that the classifier correctly classified 35,807 positive reviews, 16,631 neutral reviews, and 94,282 negative reviews. The accuracy of the model on the training data was 58.53%, and the AUC value was 77.91%, indicating fair classification. The model was then applied to the testing data, which consisted of 62,673 user reviews. The confusion matrix for the testing data is shown in Table 8.

The performance of the Naïve Bayes classifier on the testing data is summarized in the confusion matrix in Table 8. The confusion matrix shows that the classifier correctly classified 18,306 positive reviews, 5,263 neutral reviews, and 27,526 negative reviews.

TABLE VIII
CONFUSION MATRIX TESTING DATA

Actual	Prediction		
	Negative	Neutral	Positive
Negative	27,526	990	948
Neutral	1,226	5,263	7,479
Positive	122	813	18,306

The Naïve Bayes classifier was able to classify 81.53% of the testing data correctly, and the AUC value was 89.05%, indicating good classification. The results of the classifier will be used to analyze the data using *p* and *Laney p'* control charts.

C. p Control Chart

The *p* control chart analysis in phases I and II was conducted using a prediction class derived from the Naïve Bayes classification. The results are as follows.

1) *Phase I of p Control Chart*

The *p* control chart in phase I was used to analyze the review data for the *PeduliLindungi* application on Google Play from April 1, 2020, to March 31, 2022. The chart was run four times to obtain a statistically controlled chart with an average defect proportion of 0.300. The average defect proportion from phase I iteration 4 will be used to monitor the data in phase II of the *PeduliLindungi* application rating data on Google Play. Figure 4 shows the control chart for phase I iteration 4 that has been statistically controlled.

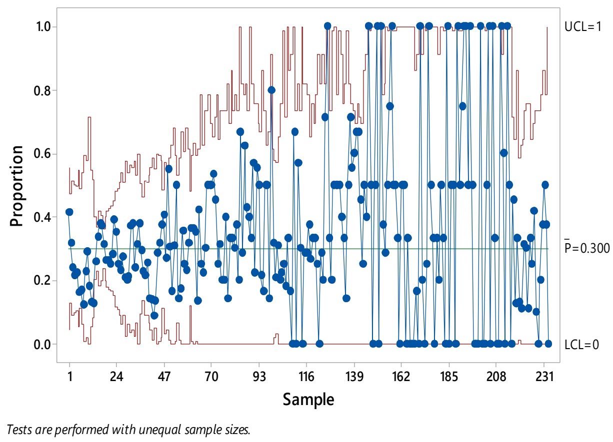

Fig. 4 Statistically Controlled Phase I of *p* Control Chart

2) Phase II of p Control Chart

The *p* chart was used to monitor the proportion of defects in *PeduliLindungi* application reviews on Google Play from January 2021 to March 2022. The average proportion of defects in phase I, which was statistically controlled, was 0.300. A monthly analysis was conducted for phase II of the *p* chart. The results showed that several phase II control charts were not statistically controlled. The number of out-of-control observation points and their causes (see Table 3 for a description of the categories) are presented in Table 9.

TABLE IX
NUMBER OF OUT OF CONTROL OBSERVATION PHASE II OF *P* CONTROL CHART

Number	Month	Out of Control	Out of Control Causes
1	January 2021	4	Category 7
2	February 2021	-	-
3	March 2021	6	Category 10
4	April 2021	3	Category 10
5	May 2021	8	Category 10
6	June 2021	21	Category 10
7	July 2021	29	Category 10
8	August 2021	29	Category 10
9	September 2021	30	Category 10
10	October 2021	31	Category 10
11	November 2021	30	Category 7
12	December 2021	31	Category 10
13	January 2022	23	Category 10
14	February 2022	14	Category 10
15	March 2022	3	Category 10

D. Laney p' Control Chart

The *Laney p'* chart is a modification of the traditional p-chart that is designed to address the problem of overdispersion, which can occur when the sample size is large. The following is the analysis results of the *Laney p'* control diagram for phase I and phase II.

1) Phase I of Laney p' Control Chart

A *Laney p'* control chart was used to monitor the proportion of defects in *PeduliLindungi* application reviews on Google Play from April to December 2020. Four iterations were conducted to obtain a statistically controlled chart. The average proportion of defects was 0.301, which was used to monitor data in phase II. Figure 5 presents the *Laney p'* control chart that has been statistically controlled.

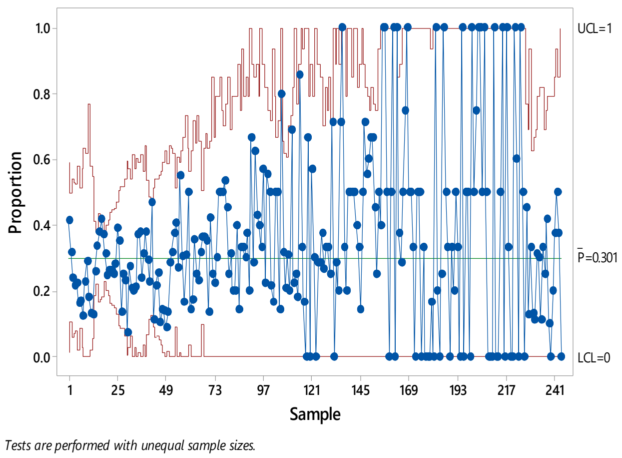

Fig. 5 Statistically Controlled Phase I *Laney p'* Control Chart

2) Phase II of Laney p' Control Chart

The *Laney p'* control chart was used to monitor the proportion of defects in *PeduliLindungi* application reviews on Google Play from January to March 2022. The average proportion of defects in phase I, which was statistically controlled, was 0.301. A monthly analysis of the Laney p' control chart was conducted for phase II. The results showed that several phase II control charts were not statistically controlled. The number of out-of-control observation points and their causes (see Table 3 for a description of the categories) are presented in Table 10.

TABLE X
NUMBER OF OUT OF CONTROL OBSERVATION PHASE II OF *LANEY P'* CONTROL CHART

Number	Month	Out of Control	Out of Control Causes
1	January 2021	1	Category 7
2	February 2021	1	Category 10
3	March 2021	1	Category 10
4	April 2021	2	Category 10
5	May 2021	4	Category 10
6	June 2021	20	Category 10
7	July 2021	22	Category 10
8	August 2021	26	Category 10
9	September 2021	24	Category 10
10	October 2021	31	Category 10
11	November 2021	27	Category 7
12	December 2021	11	Category 10
13	January 2022	5	Category 10
14	February 2022	2	Category 10
15	March 2022	4	Category 10

E. Comparison of the p and Laney p' Control Chart for Phase I and II

A comparison of the *p* and *Laney p'* control charts for phases I and II was conducted to determine the best control chart. The following is a comparison of the phase I control chart.

TABLE XI
COMPARISON OF PHASE I CONTROL CHART ANALYSIS

Control Chart Type	Iteration	Number of Out-of-Control Observations	$\bar{p}$ *in-control*
p Control Chart	0	25	0.300
	1	3	
	2	2	
	3	2	
	4	-	
Laney p' Control Chart	0	17	0.301
	1	2	
	2	1	
	3	1	
	4	-	

Table 11 shows four iterations were conducted to obtain a statistically controlled *p* control chart. The remaining 233 observations yielded an average proportion of defects of 0.300. Four iterations were also conducted to obtain a statistically controlled *Laney p'* control chart, with 241 observations yielding an average proportion of defects of 0.301. Table 12 compares the phase II control charts.

In phase II control chart analysis, it is known that the *p* control chart is more sensitive to out-of-control observations than the *Laney p'* control chart. This is because the *p* control chart is more sensitive to large sample sizes. Most of the out-of-control observations occurred in June-December 2021, which coincides with a significant increase in COVID-19 cases in Indonesia. This could be due to the start of the mass vaccination program in June 2021, which made vaccination certificates accessible through the *PeduliLindungi* application. In July 2021, the Electronic Health Alert (e-HAC) feature, a mandatory requirement for travel by public transportation, was also integrated into the *PeduliLindungi* application.

The e-HAC feature in the application displays vaccination certificates and negative COVID-19 test results as a travel condition. This has led to an increase in the demand for the *PeduliLindungi* application. However, users of the *PeduliLindungi* application still experience several problems, such as not receiving a certificate, being unable to log in or register to the *PeduliLindungi* application, the application frequently crashing or closing unexpectedly, and having difficulty checking in due to network constraints. Some of these obstacles resulted in a high number of negative reviews during that time. In January 2022, the *PeduliLindungi* application was updated with an offline check-in feature that allows users to check in at a location without an internet connection, which solves signal problems. In addition, several other obstacles have also been resolved, resulting in users beginning to give positive reviews of the application.

TABLE XII
COMPARISON OF PHASE II CONTROL CHART ANALYSIS

Month	Number of Out of Control Observations	
	P **Control Chart**	***Laney p'*** **Control Chart**
January 2021	4	1
February 2021	-	1
March 2021	6	1
April 2021	3	2
May 2021	8	4
June 2021	21	20
July 2021	29	22
August 2021	29	26
September 2021	30	24
October 2021	31	31
November 2021	30	27
December 2021	31	11
January 2022	23	5
February 2022	14	2
March 2022	3	4

IV. CONCLUSION

The sentiment analysis of user review data for the *PeduliLindungi* application on Google Play from April 1, 2020, to March 31, 2022, found that 42.7% of the reviews were positive, 11.3% were neutral, and 46.1% were negative. This indicates that there were more negative reviews than positive reviews. The accuracy of the testing data classification was 89.05%, which is considered to be good classification.

The results of the analysis of the *p* and Laney p' control charts are based on the sentiment analysis of user review data for the *PeduliLindungi* application on Google Play.

- In *p* control chart analysis of user reviews of the *PeduliLindungi* application in phase I showed that the average defect proportion was 0.300. The chart was statistically controlled in the 4th iteration. In phase II, the chart was statistically controlled only in February 2021.
- In *Laney p'* control chart analysis of user reviews of the *PeduliLindungi* application in phase I showed that the chart was statistically controlled in the 4th iteration with an average defect proportion of 0.301. In phase II, the chart was not statistically controlled.

The *p* control chart is more sensitive to detecting out-of-control observations than the Laney p' control chart. This is because the *p* control chart is based on the proportion of non-conforming units in a sample, while the *Laney p'* control chart is based on the number of non-conforming units in a sample. The larger the sample size, the more sensitive the *p* control chart will be to detecting out-of-control observations. Results of control chart analysis that have not been statistically controlled indicate that the *PeduliLindungi* application developer still needs to improve the quality of the application. Based on a Pareto chart of user reviews of the *PeduliLindungi* application on Google Play from April 1, 2020, to March 31, 2022, the most common obstacle is that the vaccination certificate does not appear in the application.

ACKNOWLEDGMENT

The authors gratefully acknowledge financial support from the Institut Teknologi Sepuluh Nopember for this work, under the project scheme of the Publication Writing and IPR Incentive Program (PPHKI) 2023.

REFERENCES

[1] W. H. Organization, "WHO Director-General's opening remarks at the media briefing on COVID-19-11 March 2020." Geneva, Switzerland, 2020.

[2] G. D'Aniello, M. Gaeta, and I. La Rocca, "KnowMIS-ABSA: an overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5543–5574, 2022.

[3] C. P. D. Cyril, J. R. Beulah, N. Subramani, P. Mohan, A. Harshavardhan, and D. Sivabalaselvamani, "An automated learning model for sentiment analysis and data classification of Twitter data using balanced CA-SVM," *Concurr. Eng.*, vol. 29, no. 4, pp. 386–395, 2021.

[4] S. Styawati, A. Nurkholis, A. A. Aldino, S. Samsugi, E. Suryati, and R. P. Cahyono, "Sentiment Analysis on Online Transportation Reviews Using Word2Vec Text Embedding Model Feature Extraction and Support Vector Machine (SVM) Algorithm," in *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*, 2022, pp. 163–167, doi: 10.1109/ISMODE53584.2022.9742906.

[5] K.-X. Han, W. Chien, C.-C. Chiu, and Y.-T. Cheng, "Application of support vector machine (SVM) in the sentiment analysis of twitter dataset," *Appl. Sci.*, vol. 10, no. 3, p. 1125, 2020.

[6] B. AlBadani, R. Shi, and J. Dong, "A novel machine learning approach for sentiment analysis on twitter incorporating the universal language model fine-tuning and SVM," *Appl. Syst. Innov.*, vol. 5, no. 1, p. 13, 2022.

[7] C. Villavicencio, J. J. Macrohon, X. A. Inbaraj, J.-H. Jeng, and J.-G. Hsieh, "Twitter sentiment analysis towards covid-19 vaccines in the Philippines using naïve bayes," *Information*, vol. 12, no. 5, p. 204, 2021.

[8] A. R. Lubis, S. Prayudani, M. Lubis, and O. Nugroho, "Sentiment Analysis on Online Learning During the Covid-19 Pandemic Based on Opinions on Twitter using KNN Method," in *2022 1st International Conference on Information System & Information Technology (ICISIT)*, 2022, pp. 106–111.

[9] F. H. Rachman and B. S. Rintyarna, "Sentiment analysis of Madura tourism in new normal era using text blob and KNN with hyperparameter tuning," in *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*, 2022, pp. 23–27.

[10] Q. Zhao, "Social emotion classification of Japanese text information based on SVM and KNN," *J. Ambient Intell. Humaniz. Comput.*, pp. 1–12, 2021.

[11] M. K. M. K. Anwar, M. Yusoff, and M. Kassim, "Decision Tree and Naïve Bayes for Sentiment Analysis in Smoking Perception," in *2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, 2022, pp. 294–299.

[12] R. A. Annisa and E. B. Setiawan, "Aspect Based Sentiment Analysis on Twitter Using Word2Vec Feature Expansion Method and Gradient Boosting Decision Tree Classification Method," in *2022 1st International Conference on Software Engineering and Information Technology (ICoSEIT)*, 2022, pp. 273–278.

[13] M. Aufar, R. Andreswari, and D. Pramesti, "Sentiment analysis on youtube social media using decision tree and random forest algorithm: A case study," in *2020 International Conference on Data Science and Its Applications (ICoDSA)*, 2020, pp. 1–7.

[14] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks," *Concurr. Comput. Pract. Exp.*, vol. 33, no. 23, p. e5909, 2021.

[15] S. O. Alhumoud and A. A. Al Wazrah, "Arabic sentiment analysis using recurrent neural networks: a review," *Artif. Intell. Rev.*, vol. 55, no. 1, pp. 707–748, 2022.

[16] H. T. Phan, N. T. Nguyen, and D. Hwang, "Convolutional attention neural network over graph structures for improving the performance of aspect-level sentiment analysis," *Inf. Sci. (Ny).*, vol. 589, pp. 416–439, 2022.

[17] M. Usama, B. Ahmad, E. Song, M. S. Hossain, M. Alrashoud, and G. Muhammad, "Attention-based sentiment analysis using convolutional and recurrent neural network," *Futur. Gener. Comput. Syst.*, vol. 113, pp. 571–578, 2020.

[18] K. Rrmoku, B. Selimi, and L. Ahmedi, "Application of trust in recommender systems—Utilizing naive Bayes classifier," *Computation*, vol. 10, no. 1, p. 6, 2022.

[19] J. Han, J. Pei, and H. Tong, *Data mining: concepts and techniques*. Morgan kaufmann, 2022.

[20] D. C. Montgomery, *Introduction to statistical quality control*. John Wiley & Sons, 2020.

[21] R. Sanchez-Marquez and J. M. J. Vivas, "Multivariate SPC methods for controlling manufacturing processes using predictive models–A case study in the automotive sector," *Comput. Ind.*, vol. 123, p. 103307, 2020.

[22] M. Ahsan, M. Mashuri, H. Kuswanto, D. D. Prastyo, and H. Khusna, "T2 Control Chart based on Successive Difference Covariance Matrix for Intrusion Detection System," in *Journal of Physics: Conference Series*, 2018, vol. 1028, no. 1, p. 12220.

[23] M. Mashuri, M. Ahsan, M. H. Lee, and D. D. Prastyo, "PCA-based Hotelling's T2 chart with Fast Minimum Covariance Determinant (FMCD) Estimator and Kernel Density Estimation (KDE) for Network Intrusion Detection," *Comput. Ind. Eng.*, p. 107447, 2021.

[24] M. Ahsan, M. Mashuri, H. Kuswanto, D. D. Prastyo, and H. Khusna, "Outlier detection using PCA mix based T 2 control chart for continuous and categorical data," *Commun. Stat. Comput.*, vol. 50, no. 5, pp. 1496–1523, 2021.

[25] M. Ahsan, M. Mashuri, H. Khusna, and M. H. Lee, "Multivariate Control Chart Based on Kernel PCA for Monitoring Mixed Variable and Attribute Quality Characteristics," *Symmetry (Basel).*, vol. 12, no. 11, p. 1838, 2020.

[26] M. Ahsan, M. Mashuri, and H. Khusna, "Kernel principal component analysis (PCA) control chart for monitoring mixed non-linear variable and attribute quality characteristics," *Heliyon*, vol. 8, no. 6, p. e09590, 2022.

[27] M. Imran, J. Sun, X. Hu, F. S. Zaidi, and A. Tang, "Investigating zero-state and steady-state performance of MEWMA-CoDa control chart using variable sampling interval," *J. Appl. Stat.*, pp. 1–22, 2023.

[28] A. Zaka, M. Naveed, and R. Jabeen, "Performance of attribute control charts for monitoring the shape parameter of modified power function distribution in the presence of measurement error," *Qual. Reliab. Eng. Int.*, vol. 38, no. 2, pp. 1060–1073, 2022.

[29] A. Johannssen, N. Chukhrova, and P. Castagliola, "The performance of the hypergeometric np chart with estimated parameter," *Eur. J. Oper. Res.*, vol. 296, no. 3, pp. 873–899, 2022.

[30] A. A. Aly, N. A. Saleh, and M. A. Mahmoud, "An adaptive EWMA control chart for monitoring zero-inflated Poisson processes," *Commun. Stat. Comput.*, vol. 51, no. 4, pp. 1564–1577, 2022.

[31] T. Mahmood, A. Iqbal, S. A. Abbasi, and M. Amin, "Efficient GLM-based control charts for Poisson processes," *Qual. Reliab. Eng. Int.*, vol. 38, no. 1, pp. 389–404, 2022.

[32] M. Ahsan, M. Mashuri, and H. Khusna, "Evaluation of Laney p' Chart Performance," *Int. J. Appl. Eng. Res.*, vol. 12, no. 24, pp. 14208–14217, 2017.

[33] H. Hassani, C. Beneki, S. Unger, M. T. Mazinani, and M. R. Yeganegi, "Text mining in big data analytics," *Big Data Cogn. Comput.*, vol. 4, no. 1, p. 1, 2020.

[34] L. Hickman, S. Thapa, L. Tay, M. Cao, and P. Srinivasan, "Text preprocessing for text mining in organizational research: Review and recommendations," *Organ. Res. Methods*, vol. 25, no. 1, pp. 114–146, 2022.

[35] S. Vijayarani, M. J. Ilamathi, and M. Nithya, "Preprocessing techniques for text mining-an overview," *Int. J. Comput. Sci. Commun. Networks*, vol. 5, no. 1, pp. 7–16, 2015.

[36] A. A. Freitas, *Data mining and knowledge discovery with evolutionary algorithms*. Springer Science & Business Media, 2002.

[37] G. Singh, B. Kumar, L. Gaur, and A. Tyagi, "Comparison between multinomial and Bernoulli naïve Bayes for text classification," in *2019 International Conference on Automation, Computational and Technology Management (ICACTM)*, 2019, pp. 593–596.

[38] S. Kharya, S. Agrawal, and S. Soni, "Naive Bayes classifiers: a probabilistic detection model for breast cancer," *Int. J. Comput. Appl*, vol. 92, no. 10, pp. 975–8887, 2014.

[39] M. J. Zaki, W. Meira Jr, and W. Meira, *Data mining and analysis: fundamental concepts and algorithms*. Cambridge University Press, 2014.

[40] A. Mustopa, E. B. Pratama, A. Hendini, and D. Risdiansyah, "Analysis of user reviews for the pedulilindungi application on google play using the support vector machine and naive bayes algorithm based on particle swarm optimization," in *2020 Fifth International Conference on Informatics and Computing (ICIC)*, 2020, pp. 1–7.