Sentiment Analysis of Mobile Legends Play Store Reviews Using Support Vector Machine and Naive Bayes

Mona Alkhoze[1,*], Miada Almasre[2]

[1,2]Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

ABSTRACT

This study applies sentiment analysis to Mobile Legends Play Store reviews to classify user feedback as positive, negative, or neutral, offering insights into the factors influencing user satisfaction. Utilizing machine learning models—Naive Bayes and Support Vector Machine (SVM)—user sentiment is evaluated, and key themes in user feedback are identified. Both models demonstrate high accuracy, with SVM slightly outperforming Naive Bayes. Specifically, the SVM model records an accuracy of 84.95%, a precision of 81.76%, and an F1-score of 83.31%, while Naive Bayes achieves an accuracy of 84.10%, a precision of 82.09%, and an F1-score of 82.57%. This classification highlights a predominance of positive reviews, revealing players' appreciation for the game's graphics and gameplay. In contrast, negative reviews expose common frustrations related to lag and technical issues, indicating areas for potential improvement. The analysis also uncovers the challenge of accurately classifying neutral sentiments due to the informal language and slang found in reviews written in Bahasa Indonesia. Future studies could address this by incorporating advanced NLP techniques, such as word embeddings or deep learning models, to better capture linguistic nuances. Overall, this research provides actionable insights for game developers, enabling them to prioritize updates and feature enhancements that align with player preferences and feedback trends.

Keywords Sentiment Analysis, Mobile Legends Reviews, Support Vector Machine, Naive Bayes, User Feedback

Introduction

The rapid expansion of mobile gaming, particularly within the Multiplayer Online Battle Arena (MOBA) genre, has significantly reshaped the global gaming industry. Among these titles, *Mobile Legends: Bang Bang* stands out as one of the most popular and enduring games, especially across Southeast Asia. This rise in popularity is largely attributed to the increasing accessibility of mobile devices, which offer immersive gaming experiences comparable to those on traditional consoles and PCs. The MOBA genre, known for its real-time strategy and team-based gameplay, has become a cornerstone of competitive mobile e-sports. Within this landscape, *Mobile Legends* has not only cultivated a massive player base but has also established itself as a major title in the global competitive gaming scene.

Accessibility has been a driving factor in this growth. The widespread use of smartphones enables players to enjoy games anywhere and anytime, fostering both convenience and long-term engagement. Studies have shown that mobile gaming attracts a diverse audience, from casual users to highly competitive players, due to its simplicity and availability [1]. The competitive structure of *Mobile Legends* motivates players to enhance their skills and improve their

Submitted 1 May 2025
Accepted 2 June 2025
Published 13 November 2025

Corresponding author
Mona Alkhoze,
malkhozae@kau.edu.sa

Additional Information and Declarations can be found on page 386

DOI: 10.47738/jdmdc.v2i4.44

How to cite this article: M. Alkhoze, M. Almasre, "Sentiment Analysis of Mobile Legends Play Store Reviews Using Support Vector Machine and Naive Bayes," *J. Digit. Mark. Digit. Curr.*, vol. 2, no. 4, pp. 368-389, 2025.

rankings, deepening their emotional connection with the game. As emphasized by [2], the combination of accessibility and competitive challenge has been crucial in sustaining player interest, generating both short-term excitement and long-term loyalty.

Equally important to *Mobile Legends'* appeal is its social dimension. MOBA games require cooperation and communication between players, fostering teamwork and a shared sense of purpose. This aspect creates strong social bonds that enhance player retention and satisfaction. Research indicates that multiplayer games meet psychological needs for competence, autonomy, and relatedness [3], [4]. The ability to form teams, build friendships, and compete globally transforms *Mobile Legends* from a simple gaming platform into a social ecosystem where interaction and collaboration are central to the experience.

In Indonesia, *Mobile Legends* has become a dominant force in the gaming market and a leading e-sports title. Millions of Indonesian players actively participate in tournaments and community events [5]. The game's ranking systems, reward structures, and regular updates encourage consistent engagement, motivating players to continually improve [6]. These features sustain long-term participation by rewarding achievement and competitiveness, reinforcing *Mobile Legends'* position as both a recreational and professional e-sports game.

User reviews from platforms such as the Google Play Store provide a valuable reflection of player satisfaction and engagement. Reviews not only serve as feedback for developers but also influence how potential users perceive a game's quality. As observed by [7], the volume and tone of user reviews significantly affect public perception and app retention rates. Players often comment on gameplay, graphics, controls, and performance, making reviews a vital data source for understanding user experience and expectations in real time.

Moreover, reviews reveal issues that require developer attention, such as bugs, crashes, or balance problems. Negative reviews can harm reputation and player retention, while positive ones reinforce trust and appreciation for improvements. Studies show that unaddressed negative feedback can lead to substantial drops in user engagement [8]. In contrast, integrating user suggestions and acknowledging feedback helps maintain satisfaction and strengthens community loyalty [9], [10]. By analyzing these reviews systematically, developers can identify critical areas for refinement and maintain strong player relationships.

Beyond individual comments, aggregated reviews reveal broader sentiment trends that may not be evident through quantitative metrics alone. Sentiment analysis enables developers to extract patterns of satisfaction, frustration, or enthusiasm across large datasets [11]. These insights are essential for enhancing player experience and optimizing future updates. Social aspects—such as cooperative gameplay, chat functionality, and community events—play an equally crucial role, as they sustain engagement by fostering a sense of belonging [3].

The emergence of sentiment analysis and machine learning has made it possible to analyze user feedback more efficiently. Techniques such as Support Vector Machines (SVM) and TF-IDF have been successfully used in various domains, from analyzing public opinion on policies to classifying product reviews [12], [13], [14]. Similarly, in digital finance and cryptocurrency research, sentiment analysis has been applied to study user behavior and engagement dynamics [15], [16], [17], [18]. Drawing from these methodologies, this study applies SVM and Naive Bayes algorithms to classify *Mobile Legends* reviews, aiming to uncover patterns that reflect player satisfaction, frustration, and engagement.

However, analyzing large volumes of multilingual reviews presents considerable challenges. *Mobile Legends* attracts users from diverse linguistic backgrounds, including Bahasa Indonesia, which introduces variability in syntax, slang, and cultural context [19], [22], [23]. Informal language use and spelling variations complicate manual analysis, making automated NLP approaches essential [24]. Automated sentiment analysis thus offers a scalable solution for classifying reviews, identifying patterns, and generating insights efficiently. Given the growing volume of user-generated content, such tools are crucial for understanding and improving player experience in a dynamic gaming environment.

In light of these challenges and opportunities, this study aims to conduct sentiment analysis on *Mobile Legends* reviews written in Bahasa Indonesia from the Google Play Store. Using Support Vector Machine (SVM) and Naive Bayes classifiers, the research categorizes reviews into positive, negative, and neutral sentiments [25]. These models have proven effective in text classification tasks, particularly with large datasets [24]. The results are expected to offer valuable insights for developers and researchers alike. For developers, sentiment analysis can identify recurring issues and highlight aspects that influence player satisfaction. Academically, this study contributes to the field of NLP by advancing sentiment classification for underrepresented languages such as Bahasa Indonesia [26]. Ultimately, it bridges the gap between user feedback and practical game design, providing data-driven recommendations to enhance player engagement and satisfaction.

Literature Review

Sentiment Analysis

Sentiment analysis, or opinion mining, is a method in Natural Language Processing (NLP) used to extract subjective information from textual data. It determines the polarity of text by categorizing it as positive, negative, or neutral based on the emotional tone conveyed by the author. The primary goal is to identify and understand the emotions and opinions expressed in written communication such as social media posts, product reviews, and online feedback. This technique has become fundamental in understanding digital communication, where unstructured textual data holds valuable insights into public perception and behavior [27]. Its importance lies in the ability to process vast amounts of textual information and convert them into actionable insights. Organizations frequently use sentiment analysis to assess consumer attitudes toward products, services, or brands, which supports marketing, customer experience, and product development strategies [28], [29]. By quantifying emotions, sentiment analysis enables decision-makers to monitor satisfaction levels, track public sentiment trends, and detect shifts in opinions [28].

In the context of artificial intelligence (AI), sentiment analysis bridges computational modeling and human cognition, allowing algorithms to interpret emotions and opinions through mathematical and statistical methods [30]. This capability is essential in applications such as social media monitoring, where real-time sentiment tracking helps identify public opinion and emerging issues [31], [32]. Beyond business contexts, sentiment analysis has proven valuable in sectors like finance, healthcare, and politics, where understanding public sentiment can influence strategic decisions and policy-making [33], [34]. In the gaming industry, sentiment analysis serves as a critical tool for developers and publishers to systematically analyze player feedback from sources such as app store reviews, social media platforms, and gaming forums. By classifying sentiments into positive, neutral, or negative categories, it helps identify elements that enhance or hinder player satisfaction, providing data-driven insights that guide design improvements and player engagement strategies [35], [36].

Recent studies have highlighted the expanding role of sentiment analysis in Bahasa Indonesia contexts, emphasizing the need for language-specific preprocessing and modeling techniques [37]. Researchers have explored machine learning and deep learning approaches tailored to the linguistic characteristics of Bahasa Indonesia to enhance classification accuracy. For example, Rozaq et al. conducted sentiment analysis on the “Kampus Mengajar 2” initiative using Naive Bayes and Euclidean distance methods, finding that Naive Bayes performed effectively in classifying Indonesian-language sentiment [38]. These findings illustrate the adaptability of sentiment analysis across various domains, including education, tourism, politics, and gaming. Furthermore, localized model development and aspect-based sentiment analysis enable deeper insights into user preferences and emotions, helping developers and organizations make better data-driven decisions aligned with Indonesian linguistic and cultural nuances.

Natural Language Processing (NLP) for Sentiment Analysis

Tokenization is a fundamental process in NLP in which text is divided into smaller units called tokens. Tokens typically represent individual words but can also be subwords or characters depending on the language and application [39]. Tokenization enables efficient text analysis by breaking complex sentences into manageable elements. In sentiment analysis, this process is essential for transforming unstructured text into a structured format that can be processed by machine learning algorithms [40]. Tokenization allows the analysis of vast textual data from reviews, social media posts, and other sources.

Several approaches exist for tokenization, with word-based tokenization being the most common. In this method, spaces and punctuation marks act as delimiters that separate sentences into words [41]. For example, the sentence "Mobile Legends is fun" becomes "Mobile," "Legends," "is," and "fun." In languages with complex morphology, such as Bahasa Indonesia, subword tokenization methods like Byte Pair Encoding (BPE) or WordPiece are often more effective, as they can manage rare or out-of-vocabulary words by dividing them into smaller, meaningful units [42]. Effective tokenization is therefore crucial for ensuring that NLP models can capture linguistic patterns and contextual relationships within text data.

Stopword removal is another essential step in NLP preprocessing, especially for sentiment analysis. Stopwords are common words such as “and” or “the” that carry little semantic meaning. In Bahasa Indonesia, stopwords include “dan” (and) and “yang” (that), which serve grammatical functions but do not contribute to sentiment [43]. The goal of stopword removal is to eliminate linguistic noise, allowing models to focus on words that convey sentiment, such as adjectives and verbs expressing emotion [44]. Researchers have developed specialized stopword lists tailored to Bahasa Indonesia to account for its grammatical and linguistic features [45]. Removing stopwords enhances model performance by increasing the signal-to-noise ratio in text data [46]. When combined with techniques such as stemming and lemmatization, stopword removal produces cleaner and more focused datasets for sentiment analysis.

Stemming and lemmatization are techniques used to reduce words to their base or root forms, allowing different variations of a word to be analyzed as a single entity. Stemming removes prefixes or suffixes to produce a root form, even if it is not a valid word. For example, “running” and “runner” can both be reduced to “run” [47]. Although efficient, stemming may result in inaccuracies since the truncated form does not always preserve meaning [48]. Lemmatization, by contrast, ensures that the reduced form is a valid word by considering the grammatical context and part of speech [49]. For instance, “better” would be lemmatized to “good,” preserving semantic accuracy [50]. Both methods are vital in NLP, and their use depends on the trade-off between processing speed and linguistic precision.

Feature Extraction for Sentiment Analysis

The Bag of Words (BoW) model is a foundational NLP technique that represents text as a collection of words, ignoring grammar and word order but preserving word frequency [51]. Each document is transformed into a high-dimensional vector, where each feature represents a word and its frequency. Despite its simplicity and usefulness, BoW cannot capture context or meaning, as it treats all words independently. For instance, “not good” and “good” may appear similar in BoW representation. Additionally, BoW often results in high-dimensional data, which may cause overfitting. To address this, dimensionality reduction techniques such as Singular Value Decomposition (SVD) and Principal Component Analysis (PCA) are applied [53].

Term Frequency-Inverse Document Frequency (TF-IDF) extends the BoW model by weighting words according to their frequency within a document and their rarity across the corpus [54]. The TF component measures how often a word appears, while the IDF component down-weights common words and emphasizes rare but informative ones. This balance highlights terms that are both relevant and distinctive. TF-IDF has proven effective for text classification and keyword extraction tasks. In sentiment analysis, TF-IDF helps models focus on sentiment-bearing words such as “excellent” or “disappointed” [55]. Its adaptability across specialized domains, including legal and biomedical contexts, underscores its versatility in extracting meaningful text features [56].

Machine Learning Algorithms

Naive Bayes is a probabilistic classifier based on Bayes’ Theorem, which calculates the probability of a class given certain features. It assumes conditional independence between features, meaning that the presence of one feature does not affect another. Despite this assumption, Naive Bayes performs well in text classification tasks, including sentiment analysis [57]. Bayes' Theorem can be expressed as:

$$P(C \mid X) = \frac{P(X \mid C) \cdot P(C)}{P(X)} \tag{1}$$

Naive Bayes simplifies computation and performs efficiently in high-dimensional text data. Variants such as Multinomial Naive Bayes handle word counts effectively, while Bernoulli Naive Bayes manages binary features representing word presence or absence [58], [59]. Its simplicity and scalability make it a popular choice for applications such as spam detection, review analysis, and customer feedback classification [60].

Support Vector Machines (SVM) are powerful classification models that identify an optimal hyperplane separating data into distinct classes [61], [62]. SVM maximizes the margin between classes, ensuring good generalization to new data. In sentiment analysis, SVM excels in high-dimensional spaces, focusing only on support vectors, which are the most critical data points near the decision boundary [63], [64]. This approach minimizes overfitting and improves robustness. Studies have shown that SVM often outperforms Naive Bayes in handling complex, nonlinear sentiment data [65]. Its adaptability through different kernel functions, such as linear and radial basis function (RBF) kernels, allows it to manage both linear and nonlinear relationships effectively [61], [66], [67]. The accuracy and flexibility of SVM solidify its position as one of the most effective machine learning algorithms in NLP.

Method

The research method for this study consists of several steps to ensure a comprehensive and accurate analysis. The flowchart in figure 1 outlines the detailed steps of the research method.

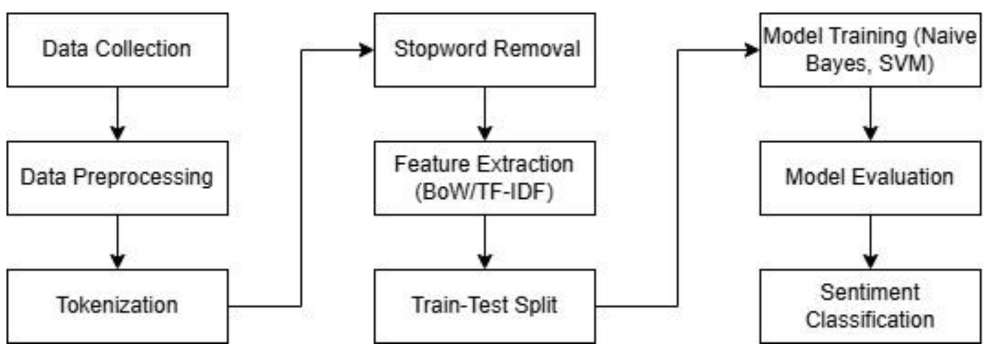

Figure 1 Research Method Flowchart

Dataset Description

The dataset for this study was collected from user reviews of the Mobile Legends: Bang Bang application on the Google Play Store. Each review entry includes various attributes that provide insights into user feedback and engagement with the game. Key columns in the dataset include `content`, which contains the text of the user's review; `score`, representing the rating assigned by the user on a scale from 1 to 5; and `thumbsUpCount`, which indicates the number of other users who found the review helpful. Additional fields such as `reviewId`, `userName`, `userImage`, and `reviewCreatedVersion` provide supplementary metadata about each review, allowing for a deeper understanding of the context and frequency of user feedback.

Before conducting any analysis, it was essential to preprocess the data to enhance its quality and usability for sentiment analysis. The first step in data preprocessing involved text cleaning, where all reviews in the `content` column were converted to lowercase to ensure uniformity. Non-alphabetical characters such as punctuation marks, numbers, and special symbols were removed to eliminate noise from the data. This cleaning process was crucial for standardizing the textual data, enabling a more accurate comparison of words and phrases across reviews. Additionally, any missing or irrelevant data entries were identified and removed to maintain the dataset's integrity.

Following text cleaning, tokenization was applied to split each review into individual words or tokens. This process facilitated further analysis by breaking down the text into manageable units. Given that the reviews were written in Bahasa Indonesia, a custom stopword list for the language was used to remove common words like “yang” (that) and “dan” (and), which do not contribute meaningful information to the sentiment analysis. Stopword removal helped to focus the analysis on more impactful terms that directly reflect users' sentiments and opinions. Afterward, stemming was performed using the Sastrawi stemmer, which is specifically designed for Bahasa Indonesia. This step reduced words to their base forms, ensuring that different variations of a word were treated as a single term, thereby enhancing the consistency and interpretability of the data.

The final output of the preprocessing stage was a cleaned and tokenized dataset, ready for feature extraction and model training. Each review was represented by a processed text string in the `cleaned_content` column, which served as the primary input for subsequent machine learning analyses. Additionally, the sentiment of each review was classified based on its `score`: reviews with scores of 4 or 5 were labeled as positive, scores of 3 as neutral, and scores of 1 or 2 as negative. This classification allowed for a clear distinction between different sentiment classes, laying the foundation for training and evaluating the Naive Bayes and Support Vector Machine models. The preprocessing steps were essential in transforming raw user reviews into a structured format suitable for in-depth sentiment analysis, aimed at uncovering user perceptions of Mobile Legends.

Exploratory Data Analysis (EDA)

The initial stage of the EDA involved examining the distribution of sentiments within the dataset, categorized as positive, neutral, or negative based on user ratings. Reviews with scores of 4 and 5 were labeled as positive, scores of 3 as neutral, and scores of 1 and 2 as negative. Visualizations, including a bar chart and a pie chart, were created to illustrate the proportion of each sentiment class within the dataset. The bar chart (figure 2) displayed the frequency of reviews for each sentiment category, while the pie chart provided a percentage-based breakdown, offering a quick view of the sentiment landscape. These visualizations revealed that positive reviews comprised the majority, followed by negative reviews, with neutral reviews being the least frequent. This distribution provided a foundational understanding of user sentiment toward the Mobile Legends application.

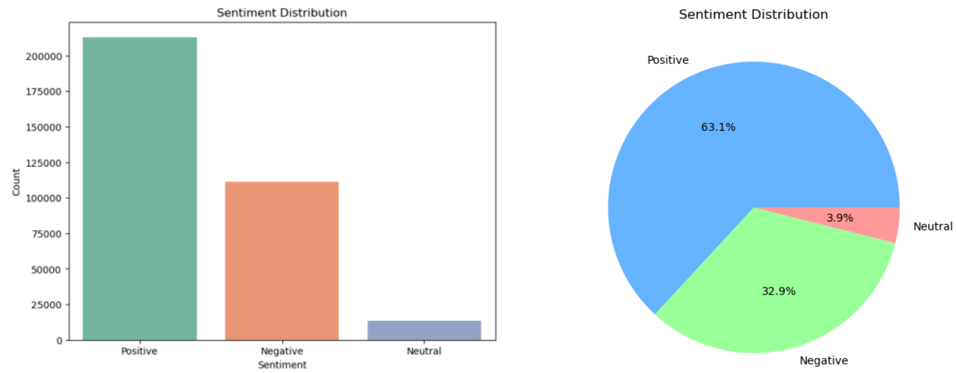

Figure 2 **Sentiment Distribution**

The bar chart displays the sentiment distribution across user reviews for the Mobile Legends Play Store app. The three sentiment categories are Positive, Negative, and Neutral. According to the chart, the majority of reviews are Positive, indicating a largely favorable response from users. Specifically, Positive reviews exceed 200,000, suggesting that a significant portion of users had a satisfactory experience with the app. Negative reviews form the second-largest group, totaling slightly over 100,000. This volume of Negative feedback indicates notable areas of user dissatisfaction, which might point to specific issues with gameplay, performance, or features within the app that require further investigation. Lastly, the Neutral category is the smallest, with relatively few reviews. This low count of Neutral reviews implies that users typically have strong opinions about the app, either positive or negative, rather than neutral or indifferent views.

The pie chart provides a visual representation of the sentiment distribution in Mobile Legends Play Store reviews, categorized into Positive, Negative, and Neutral sentiments. The largest portion, at 63.1%, consists of Positive reviews, indicating that a significant majority of users have a favorable opinion of the app. This dominant positive sentiment suggests that most users are satisfied with the features and performance of Mobile Legends. Negative reviews account for 32.9% of the total, highlighting areas where users have expressed dissatisfaction. This substantial percentage of negative feedback underscores the presence of notable issues or features that may require attention from the developers to improve user experience. The Neutral category comprises only 3.9% of the reviews, indicating that few users have mixed or indifferent feelings about the app. The relatively small proportion of Neutral reviews suggests that user experiences are polarized, with most users feeling strongly either positively or negatively about the app. Overall, while the majority sentiment is positive, the sizable negative feedback points to areas for potential improvement.

The sentiment distribution analysis highlighted a significant proportion of both positive and negative reviews, indicating diverse user experiences with the game. While positive reviews suggested areas of the game that users enjoyed, the presence of negative reviews underscored aspects of the game that may require improvement. Analyzing this distribution allowed for a quantitative grasp of user satisfaction and dissatisfaction, setting the stage for a deeper qualitative analysis of specific topics that users mentioned frequently. This sentiment analysis laid the groundwork for identifying key themes within each sentiment category through further textual analysis.

To complement the sentiment distribution, the next step in the EDA involved identifying the most frequently mentioned words in positive (figure 3) and negative reviews (figure 4).

Figure 3 Most Frequent Words in Positive Reviews

The word cloud for the most frequent words in positive reviews reveals several key terms that highlight what users appreciate about Mobile Legends. The largest words, such as "game," "bagus" (good), "seru" (fun), and "nya" (a suffix indicating possession or emphasis), indicate a general positive sentiment toward the game itself. The repeated presence of "game" and "bagus" suggests that users commonly describe the app as a good or enjoyable gaming experience. Additionally, terms like "gratis" (free) and "skin" appear prominently, which likely reflect user appreciation for free skins or content available within the game. This indicates that users value free features and are likely satisfied with the rewards or content provided by the game. Other words, such as "mantap" (awesome), "keren" (cool), and "sistem" (system), point to specific elements of the game or its mechanics that players find engaging or impressive. The inclusion of "dark" and "sistem" could relate to specific game features or events, suggesting that users are particularly engaged with certain themes or mechanics. Overall, the word cloud suggests that users have a generally positive perception of the game, enjoy its features, and appreciate both the gameplay experience and the available free content.

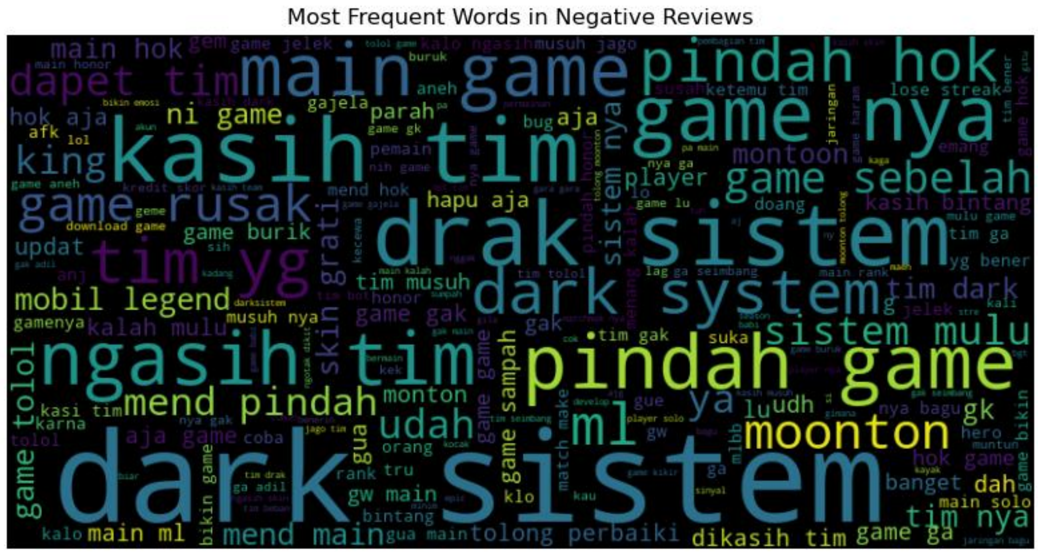

Figure 4 **Most Frequent Words in Negative Reviews**

The word cloud for the most frequent words in negative reviews highlights several issues that users commonly complain about regarding Mobile Legends. Prominent words such as "dark," "sistem" (system), and "tim" (team) suggest that players have frustrations related to game mechanics, likely involving team dynamics and system issues. The term "dark" paired with "sistem" could indicate dissatisfaction with the matchmaking system or issues perceived as unfair in the game's mechanics.nWords like "pindah" (move) and "ngasih" (give) imply complaints about the game’s ability to assign or transfer players between teams, possibly hinting at frustrations with random matchmaking or perceived imbalances in team formation. The recurrence of terms like "kasih" (give), "hok" (luck), and "tim" reflects users' dissatisfaction with how the game assigns teammates, suggesting that users feel their gaming experience is overly dependent on luck or random assignment. Additionally, terms such as "rusak" (broken), "burik" (ugly), and "parah" (severe) indicate that players have encountered technical issues or are unhappy with aspects of the game’s quality and performance. The presence of terms like "moonton" (the developer) and "main" (play) implies that users are directly addressing the game developers, possibly demanding fixes or improvements. Overall, the negative sentiment revolves around frustrations with team management, perceived unfairness, and technical problems, highlighting areas where players desire improvements.

Together, the sentiment distribution and word frequency analyses in the EDA stage provided a well-rounded view of user experiences with Mobile Legends, helping to inform the subsequent topic modeling and sentiment analysis.

Data Preprocessing

The data preprocessing stage began with comprehensive text cleaning to prepare the Mobile Legends user reviews for analysis. All text in the content column was converted to lowercase to maintain uniformity, while punctuation, special characters, and numbers were removed as they provided no semantic value for sentiment classification. Extra whitespace was also stripped to ensure a consistent text structure. Following this, tokenization was performed to split each review into individual words or tokens, allowing for more granular analysis. Since the dataset was written in Bahasa Indonesia, a stopword removal process using a predefined list of common Indonesian words (e.g., “yang,” “dan”) was applied to eliminate terms that did not contribute meaningful information to sentiment detection.

Next, stemming was implemented to reduce words to their root forms, consolidating variations of the same word into a single representation. For example, “menggunakan” (using) and “digunakan” (used) were stemmed to “guna.” This step was crucial for improving text consistency and reducing feature dimensionality, particularly given the morphological richness of Bahasa Indonesia. The Sastrawi stemmer, a tool specifically designed for the Indonesian language, was employed to perform this task accurately and efficiently, ensuring that all words were represented in their base form while retaining their semantic meaning.

Finally, the preprocessed text was transformed into numerical features using the Term Frequency–Inverse Document Frequency (TF-IDF) technique to prepare it for machine learning models. TF-IDF assigns a weight to each term based on its importance, calculated as:

$$\text{TF-IDF}(t,d) = \text{TF}(t,d) \times \text{IDF}(t) \tag{2}$$

$$\text{TF}(t,d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}} \text{ and } \text{IDF}(t) = \log\left(\frac{N}{n_t}\right). \tag{3}$$

This process ensures that frequent but less informative words receive lower weights, while rare and sentiment-bearing words gain higher significance. The resulting TF-IDF vectors formed the input for the Support Vector Machine (SVM) and Naive Bayes classifiers, providing a strong and reliable foundation for effective sentiment analysis.

Feature Extraction

The feature extraction stage aimed to convert the preprocessed text data into numerical representations suitable for machine learning algorithms through the Term Frequency–Inverse Document Frequency (TF-IDF) technique. TF-IDF is widely used in text mining because it evaluates the importance of each word within a document relative to its occurrence across all documents in the dataset. In this study, each review was represented as a vector in a high-dimensional space, where each dimension corresponded to a unique word. Words that appeared frequently within a review but rarely across the entire corpus, such as “lag” or “seru” (exciting), were given higher importance, while common terms like “game” or “grafik” (graphics) were assigned lower weights. This weighting mechanism helped emphasize sentiment-bearing words that were most relevant for sentiment classification, allowing the analysis to focus on expressions that truly reflected user opinions.

The application of TF-IDF resulted in a sparse matrix, where each row represented an individual review and each column represented a unique word from the dataset. Due to the vast vocabulary and varying word frequencies, most entries in the matrix were zero, indicating that many words appeared infrequently. To improve efficiency and prevent model overfitting, the vocabulary size was limited to the 5,000 most informative words based on their TF-IDF scores. This dimensionality reduction preserved essential textual features while maintaining computational feasibility. The resulting numerical vectors served as inputs for the Naive Bayes and Support Vector Machine (SVM) classifiers, enabling them to effectively identify and classify sentiment patterns in Mobile Legends user reviews. Through this process, TF-IDF provided a robust foundation for converting unstructured text into structured features that captured both the frequency and contextual importance of words.

Model Training

To classify the sentiment of Mobile Legends Play Store reviews, two machine learning algorithms — Multinomial Naive Bayes (MNB) and Support Vector Machine (SVM) — were trained using the TF-IDF feature representations. The Naive Bayes classifier assumes conditional independence among features, making it efficient for high-dimensional text data. Given the multinomial distribution of term frequencies, the MNB variant was selected. The model was trained by learning the conditional probability of each word given a sentiment class based on labeled TF-IDF vectors. Using k-fold cross-validation, the dataset was divided into $k$partitions to ensure robust generalization. During training, Laplace smoothing was applied to handle zero-frequency terms, and the optimal smoothing parameter (α) was determined through experimentation. The trained model could then predict sentiment classes (positive, neutral, or negative) for unseen reviews based on the posterior probabilities learned during training.

Algorithm 1 Multinomial Naive Bayes & SVM Model Training

Input:
$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$where x_i= TF-IDF features, y_i= sentiment label
Output: Trained Naive Bayes and SVM models

--- Multinomial Naive Bayes Training ---
For each class $c \in C$:
$P(c) = \frac{N_c}{N}$
$P(w_i \mid c) = \frac{N_{w_i,c}+\alpha}{\sum_j N_{w_j,c}+\alpha|V|}$

For each document x:
$\hat{y} = \arg\max_{c \in C} P(c) \prod_{i=1}^{m} P(w_i \mid c)$

--- SVM Training ---
Given: Training set D, regularization parameter C

Objective:
$\min_{\mathbf{w},b} \frac{1}{2} \| \mathbf{w} \|^2 + C \sum_{i=1}^{n} \max(0, 1 - y_i(\mathbf{w}^\top x_i + b))$

Decision Function:
$f(x) = \text{sign}(\mathbf{w}^\top x + b)$

--- Cross-Validation ---
Split Dinto $k$folds
For each fold $j = 1, \dots, k$:

Train model on $D_{\text{train}}^{(j)}$
Validate model on $D_{\text{test}}^{(j)}$
Compute average performance across folds

In parallel, a Support Vector Machine (SVM) classifier was trained using the same TF-IDF features to complement the probabilistic nature of Naive Bayes with a margin-based approach. The SVM seeks to find an optimal hyperplane that maximizes the separation margin between sentiment classes in the feature space. A linear kernel was selected due to its effectiveness with text data, and the regularization parameter (C) was tuned via grid search to balance bias and variance. Cross-validation was again employed to validate model stability and minimize overfitting. The optimization objective for SVM can be described as finding the weight vector w and bias b that minimize:

$$\min_{\mathrm{w},b} \quad \frac{1}{2} \parallel \mathrm{w} \parallel^2 + C \sum_{i=1}^{n} \max\left(0, 1 - y_i(\mathrm{w}^\top x_i + b)\right) \tag{4}$$

The resulting models — Naive Bayes for its interpretability and SVM for its discriminative power — provided a strong comparative framework for sentiment classification. Together, they enabled a comprehensive evaluation of how different algorithmic approaches could effectively capture and predict user sentiment toward the Mobile Legends game.

Model Evaluation

The performance of the Naive Bayes and Support Vector Machine (SVM) classifiers was evaluated using several key metrics, including accuracy, precision, recall, and F1-score. Accuracy, which measures the overall correctness of the model by calculating the proportion of correctly classified instances, served as an initial indicator of each model's performance. However, to gain a more nuanced understanding of the classifiers' effectiveness, precision and recall were also considered. Precision reflects the proportion of true positive predictions among all instances predicted as positive, providing insight into the model's reliability in identifying specific sentiment classes. Recall, on the other hand, measures the model's ability to capture all relevant instances of a particular class, making it particularly useful for evaluating the model’s sensitivity in detecting positive, neutral, and negative sentiments.

In addition to these metrics, the F1-score was calculated to provide a balanced assessment that considers both precision and recall. The F1-score, defined as the harmonic mean of precision and recall, offered a single metric that captured the trade-off between these two aspects. This metric was particularly valuable in cases where the dataset was imbalanced, as it accounted for the model's ability to both correctly identify and comprehensively capture instances of each sentiment class. To visualize the performance of the classifiers, confusion matrices were generated, depicting the distribution of true positive, false positive, true negative, and false negative predictions for each sentiment class. These matrices enabled a detailed examination of misclassifications, highlighting areas where each model was prone to errors and revealing patterns in incorrect sentiment predictions.

Cross-validation was employed to assess the robustness of both classifiers and to ensure that the evaluation results were not overly dependent on a single split of the dataset. In this study, a k-fold cross-validation technique was used, where the dataset was divided into k subsets, and the models were iteratively trained on k-1 subsets while tested on the remaining subset. This process was repeated for each subset, and the performance metrics were averaged across all iterations to provide a comprehensive view of the models' generalization capabilities. Cross-validation mitigated the risk of overfitting by ensuring that the models were evaluated on various portions of the data, thereby enhancing the reliability and validity of the evaluation results.

The combined use of accuracy, precision, recall, F1-score, and confusion matrices, along with cross-validation, provided a robust framework for evaluating the Naive Bayes and SVM classifiers. This comprehensive evaluation approach allowed for a thorough comparison of the two models, highlighting their strengths and weaknesses in classifying the sentiment of Mobile Legends reviews. By leveraging these metrics and validation techniques, the study effectively assessed the classifiers' abilities to accurately and consistently interpret user sentiment, paving the way for actionable insights based on user feedback.

Result and Discussion

Sentiment Distribution

The sentiment distribution of Mobile Legends Play Store reviews was visualized using a bar chart, as shown in Figure 1. This chart provides a clear breakdown of user sentiments categorized into positive, negative, and neutral classes. Positive reviews, which were the most numerous, indicated a generally favorable perception of the game, with users expressing satisfaction regarding various aspects of gameplay and features. Negative reviews represented a significant portion of the feedback, suggesting areas where users encountered issues or dissatisfaction. Neutral reviews, while the least frequent, provided a more balanced view of the game, indicating mixed feelings or ambivalence among some users.

In addition to the visual representation, a summary table was created to quantify the sentiment distribution across the dataset. Out of the total reviews, the majority were classified as positive, totaling over 200,000 entries. Negative reviews accounted for approximately 125,000 entries, highlighting areas that could benefit from improvement. Neutral reviews were far fewer, comprising a small fraction of the total dataset. This distribution reflects a substantial amount of user engagement, with a large number of users willing to share their experiences and opinions about the game. The high number of positive reviews points to a strong user base that appreciates certain aspects of the game, while the substantial number of negative reviews signals issues that may need to be addressed.

The sentiment distribution analysis provides valuable insights into user perceptions of Mobile Legends. The predominance of positive reviews suggests that the game is generally well-received; however, the significant number of negative reviews emphasizes the importance of addressing user concerns. This feedback serves as an essential resource for game developers, enabling them to identify and prioritize potential areas for improvement. By understanding the balance between positive and negative sentiments, developers can make informed decisions aimed at enhancing user satisfaction and retaining a loyal player base.

Model Performance Comparison

To evaluate the performance of the Naive Bayes and Support Vector Machine (SVM) classifiers on the sentiment analysis task, confusion matrices were generated for both models, providing a visual representation of the classification outcomes for each sentiment category. The Naive Bayes model demonstrated effective classification for the positive and negative classes, with a significant number of true positives in both categories. However, the model struggled with the neutral class, resulting in a high number of misclassifications. Similarly, the SVM model showed strong classification performance for positive and negative reviews but faced challenges with the neutral class, often mislabeling neutral reviews as either positive or negative. These confusion matrices highlighted areas where both models succeeded and areas where they encountered difficulties, particularly in distinguishing neutral sentiments.

In addition to the confusion matrices, a detailed metric table was created to summarize the accuracy, precision, recall, and F1-score for both models. The Naive Bayes classifier achieved an accuracy of 0.8410, with a precision of 0.8209, recall of 0.8410, and an F1-score of 0.8257. These metrics indicated that the Naive Bayes model performed well in accurately classifying positive and negative reviews, though its performance on neutral reviews was limited, as reflected in the low precision and recall for this class. The SVM classifier outperformed Naive Bayes slightly, achieving an accuracy of 0.8495, with a precision of 0.8176, recall of 0.8495, and an F1-score of 0.8331. This performance improvement was mainly observed in the positive and negative classes, where SVM consistently achieved higher recall and F1-scores, indicating a stronger capability in identifying true positive cases within these categories.

The comparison of weighted averages further underscored the relative strengths and weaknesses of each model. For both Naive Bayes and SVM, the weighted averages for precision, recall, and F1-score indicated that positive and negative reviews were classified with reasonable accuracy, while the neutral class posed a challenge. Specifically, the weighted F1-scores were 0.83 for both models, reflecting balanced overall performance. However, the macro averages, which provided a less biased view across all classes, highlighted the difficulty both models faced with neutral sentiment classification. The Naive Bayes model achieved a macro F1-score of 0.56, while the SVM model achieved a similar macro F1-score of 0.57, indicating room for improvement in handling the neutral category.

In conclusion, both the Naive Bayes and SVM classifiers demonstrated strong performance in classifying positive and negative sentiments in the Mobile Legends Play Store reviews, with SVM slightly outperforming Naive Bayes in terms of overall accuracy and F1-score. Despite these successes, both models faced challenges with the neutral sentiment class, which impacted the macro average scores. These findings suggest that while Naive Bayes and SVM are effective for binary sentiment classification tasks, additional refinements or alternative approaches may be required to improve the classification of neutral reviews in a multi-class sentiment analysis setting.

Discussion

The comparative analysis of the Naive Bayes and Support Vector Machine (SVM) classifiers revealed distinct strengths and limitations in predicting the sentiment of Mobile Legends Play Store reviews. Both models showed robust performance in identifying positive and negative sentiments, with SVM achieving slightly higher accuracy and F1-scores overall. This marginal difference suggested that the SVM classifier's margin-based approach enabled it to capture the subtle distinctions between positive and negative sentiments more effectively than the probabilistic Naive Bayes model. However, both models struggled with neutral sentiment classification, often misclassifying neutral reviews as either positive or negative. This challenge was evident from the low recall and precision scores for the neutral class, reflecting the difficulty of distinguishing nuanced or mixed sentiments in the dataset.

An exploration of sentiment trends in the positive and negative reviews uncovered distinct themes associated with each sentiment. Positive reviews frequently contained terms such as “grafik” (graphics), “seru” (exciting), and “menarik” (interesting), which highlighted aspects of the game that users enjoyed. These terms underscored the importance of engaging gameplay and appealing visual design as factors contributing to user satisfaction. Conversely, negative reviews often featured words like “lag,” “bug,” and “error,” reflecting user frustration with technical performance issues. These recurring terms emphasized the impact of performance-related challenges on user dissatisfaction and provided actionable insights for developers seeking to enhance the user experience by addressing these concerns. The frequency of these terms across sentiment categories pointed to the significance of game performance and stability in shaping overall user sentiment.

The study encountered several challenges in analyzing the reviews, particularly due to the linguistic characteristics of user-generated content in Bahasa Indonesia. Many reviews included informal language, slang, and regional dialects, which posed challenges for tokenization, stemming, and accurate sentiment classification. The presence of imbalanced sentiment classes, with a smaller proportion of neutral reviews, further complicated the analysis, as it affected the models' ability to generalize across all sentiment categories. These challenges highlighted the need for additional preprocessing techniques tailored to the nuances of Bahasa Indonesia and underscored the importance of exploring alternative approaches, such as incorporating word embeddings or deep learning methods, to better handle informal language and imbalanced data. Despite these obstacles, the results provided valuable insights into user perceptions of Mobile Legends and identified areas for potential service improvement.

Conclusion

This study demonstrated the effectiveness of sentiment analysis in extracting meaningful insights from user feedback on the Mobile Legends Play Store reviews. By applying machine learning models such as Naive Bayes and Support Vector Machine (SVM), the analysis was able to categorize user sentiments as positive, negative, or neutral, providing a comprehensive overview of player perceptions. Both models performed well in predicting positive and negative sentiments, with SVM slightly outperforming Naive Bayes in terms of accuracy and F1-score. The SVM model's margin-based classification approach proved advantageous in capturing nuanced distinctions between positive and negative reviews, making it the preferred choice for this sentiment analysis task.

The findings of this study offer valuable contributions to the field of game development, particularly for developers seeking to enhance the user experience in Mobile Legends. By identifying common themes in positive and negative reviews, the analysis highlighted specific areas of satisfaction, such as gameplay graphics and excitement, as well as areas of dissatisfaction, including technical issues like lag and bugs. These insights enable developers to prioritize feature improvements and address player concerns more effectively. Understanding user sentiment also allows developers to tailor their updates and customer service strategies to align with the expressed needs and preferences of their player base, ultimately fostering a more engaged and satisfied community.

Despite its contributions, this study faced several limitations. The Naive Bayes and SVM models, while effective for basic sentiment classification, were unable to capture deeper emotional nuances, such as sarcasm or complex sentiment expressions. Furthermore, the models struggled with the neutral sentiment class, indicating the need for more sophisticated techniques to improve classification accuracy for ambiguous sentiments. Future research could explore the use of more complex models, such as deep learning approaches or transformer-based models, to enhance the understanding of nuanced language and capture specific emotional states beyond basic sentiment categories.

In addition to model advancements, future research could extend the scope of sentiment analysis by implementing multi-label classification to detect a broader range of emotions, such as frustration, joy, or anger, in user reviews. This approach would enable a more granular understanding of user feedback and facilitate targeted improvements based on specific emotional responses. By addressing these limitations and expanding the analytical framework, future studies can build on the insights presented here and continue to support game developers in their efforts to optimize user experiences and maintain a competitive edge in the mobile gaming industry.

Declarations

Author Contributions

Author Contributions: Conceptualization, M.A. and M.A.; Methodology, M.A. and M.A.; Software, M.A. and M.A.; Validation, M.A. and M.A.; Formal Analysis, M.A.; Investigation, M.A. and M.A.; Resources, M.A. and M.A.; Data Curation, M.A.; Writing—Original Draft Preparation, M.A.; Writing—Review and Editing, M.A.; Visualization, M.A. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or

publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Z. Huang, "The Mediating Effect of Engagement in the Relationship Between Self-Efficacy and Perceived Learning in the Online Mathematics Environment Among Chinese Students," *Res. Square Preprint*, vol. 2024, no. Apr., pp. 1–14, Apr. 2024.

[2] A. C. Lusar, Ú. Oberst, R. Cladellas, and H. Fuster, "Effect of the Frustration of Psychological Needs on Addictive Behaviors in Mobile Videogamers—The Mediating Role of Use Expectancies and Time Spent Gaming," *Int. J. Environ. Res. Public Health*, vol. 17, no. 17, pp. 6429–6441, Sep. 2020.

[3] S. T. T'ng, K. H. Ho, and K. Pau, "Need Frustration, Gaming Motives, and Internet Gaming Disorder in Mobile Multiplayer Online Battle Arena (MOBA) Games: Through the Lens of Self-Determination Theory," *Int. J. Ment. Health Addict.*, vol. 21, no. 6, pp. 3821–3841, Dec. 2022.

[4] J. W. Bonny and L. M. Castaneda, "To Triumph or to Socialize? The Role of Gaming Motivations in Multiplayer Online Battle Arena Gameplay Preferences," *Simul. Gaming*, vol. 53, no. 2, pp. 157–174, Apr. 2022.

[5] M. N. H. Nurdin, "Effect of Online Game's Intensity on Verbal Aggression Behavior in Students of SMP Negeri 10 Makassar," *Arrus J. Soc. Sci. Humanit.*, vol. 4, no. 3, pp. 319–328, Jul. 2024.

[6] H. Wijaya, "Impact of Digital Marketing and Intellectual Capital on Business Performance (Case Study of SMEs in Depok City, West Java)," *Best J. Adm. Manag.*, vol. 2, no. 4, pp. 183–189, Apr. 2024.

[7] W. H. Martin, F. Sarro, Y. Jia, Y. Zhang, and M. Harman, "A Survey of App Store Analysis for Software Engineering," *IEEE Trans. Softw. Eng.*, vol. 43, no. 9, pp. 817–847, Sep. 2017.

[8] N. Herlinawati, Y. Yuliani, S. Faizah, W. Gata, and S. Samudi, "Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes and Support Vector Machine," *CESS (J. Comput. Eng. Syst. Sci.)*, vol. 5, no. 2, pp. 293–301, May 2020.

[9] İ. Göksu, A. Aslan, and Y. E. Turgut, "Evaluation of Mobile Games in the Context of Content: What Do Children Face When Playing Mobile Games?," *E-Learn. Digit. Media*, vol. 17, no. 5, pp. 388–407, Oct. 2020.

[10] B. Patzer, B. S. Chaparro, and J. R. Keebler, "Developing a Model of Video Game Play: Motivations, Satisfactions, and Continuance Intentions," *Simul. Gaming*, vol. 51, no. 3, pp. 287–309, Jun. 2020.

[11] T. Lapaš and T. Orehovački, "Evaluation of User Experience in Interaction With Computer Games," in *Proc. Central European Conf. Information and Intelligent*

Systems (CECIIS 2015), vol. 2015, no. Sep., pp. 271–282, 2015.

[12] M. Liebenlito, N. Inayah, E. Choerunnisa, and A. Pradana, "Active Learning on Indonesian Twitter Sentiment Analysis Using Uncertainty Sampling," *J. Appl. Data Sci.*, vol. 2024, no. Oct., pp. 1–10, Oct. 2024.

[13] J. Batmetan and T. Hariguna, "Sentiment Unleashed: Electric Vehicle Incentives Under the Lens of Support Vector Machine and TF-IDF Analysis," *J. Appl. Data Sci.*, vol. 2024, no. Oct., pp. 1–12, Oct. 2024.

[14] Q. Siddique, "Comparative Analysis of Sentiment Classification Techniques on Flipkart Product Reviews: A Study Using Logistic Regression, SVC, Random Forest, and Gradient Boosting," *J. Digit. Mark. Digit. Curr.*, vol. 2024, no. Oct., pp. 1–12, Oct. 2024.

[15] T. Wahyuningsih and S. Chen, "Analyzing Sentiment Trends and Patterns in Bitcoin-Related Tweets Using TF-IDF Vectorization and K-Means Clustering," *J. Curr. Res. Blockchain*, vol. 2024, no. Oct., pp. 1–12, Oct. 2024.

[16] A. Wahid, "Time Series Analysis of Bitcoin Prices Using ARIMA and LSTM for Trend Prediction," *J. Digit. Mark. Digit. Curr.*, vol. 2024, no. Oct., pp. 1–12, Oct. 2024.

[17] A. Mu'amar, "Economic Decentralization through Blockchain: Opportunities, Challenges, and New Business Models," *J. Curr. Res. Blockchain*, vol. 2024, no. Oct., pp. 1–14, Oct. 2024.

[18] C. Izumi and T. Hariguna, "In-Depth Analysis of Web3 Job Market: Insights from Blockchain and Cryptocurrency Employment Landscape," *Int. J. Res. Metaverse*, vol. 1, no. 1, pp. 1–14, Jun. 2024.

[19] Z. Yejian, "Review Classification Based on Machine Learning: Classifying Game User Reviews," *IEEE Access*, vol. 11, no. Dec., pp. 142447–142463, Dec. 2023.

[20] M. F. Pramono, K. Renalda, D. P. Kristiadi, H. L. H. S. Warnars, and W. Kusakunniran, "Matchmaking Problems in MOBA Games," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 11, no. 3, pp. 908–917, Sep. 2018.

[21] J. Trischler, T. Dietrich, and S. Rundle-Thiele, "Co-Design: From Expert- to User-Driven Ideas in Public Service Design," *Public Manag. Rev.*, vol. 21, no. 11, pp. 1595–1619, Nov. 2019.

[22] C. Stanik, M. Haering, and W. Maalej, "Classifying Multilingual User Feedback Using Traditional Machine Learning and Deep Learning," in *Proc. IEEE Int. Conf. Requirements Engineering Workshops (REW)*, vol. 2019, no. Sep., pp. 169–175, Sep. 2019.

[23] R. A. Fischer, R. M. Walczuch, and E. Guzmán, "Does Culture Matter? Impact of Individualism and Uncertainty Avoidance on App Reviews," in *Proc. IEEE/ACM Int. Conf. Software Engineering: Software Engineering in Society (ICSE-SEIS)*, vol. 2021, no. May, pp. 65–71, May 2021.

[24] M. A. Hadi and F. H. Fard, "Evaluating Pre-Trained Models for User Feedback Analysis in Software Engineering: A Study on Classification of App Reviews," *arXiv Preprint*, vol. 2021, no. Apr., pp. 1–10, Apr. 2021.

[25] F. Rustam, A. Mehmood, M. Ahmad, S. Ullah, D. M. Khan, and G. S. Choi, "Classification of Shopify App User Reviews Using Novel Multi Text Features," *IEEE Access*, vol. 8, no. Feb., pp. 30234–30244, Feb. 2020.

[26] A. Rusli, A. Suryadibrata, S. B. Nusantara, and J. C. Young, "A Comparison of Traditional Machine Learning Approaches for Supervised Feedback Classification in Bahasa Indonesia," *Int. J. New Media Technol. (IJNMT)*, vol. 7, no. 1, pp. 28–32,

Jun. 2020.

[27] I. Atoum and A. F. Otoom, "Mining Software Quality From Software Reviews: Research Trends and Open Issues," *Int. J. Comput. Trends Technol.*, vol. 31, no. 2, pp. 74–83, May 2016.

[28] Z. Nawaz, C. Zhao, F. Nawaz, A. A. Safeer, and W. Irshad, "Role of Artificial Neural Networks Techniques in Development of Market Intelligence: A Study of Sentiment Analysis of eWOM of a Women's Clothing Company," *J. Theor. Appl. Electron. Commer. Res.*, vol. 16, no. 5, pp. 1862–1876, Oct. 2021.

[29] Y. M. M. Seethalakshmi, "A Survey on Feature Extraction Techniques, Classification Methods and Applications of Sentiment Analysis," *Braz. Arch. Biol. Technol.*, vol. 66, no. Mar., pp. 1–14, Mar. 2023.

[30] J. Redzimska, "Selected Aspects of Sentiment Analysis in the Context of Human Communication," *Philol. Int. J. Linguist. Lit. Stud. Engl. Lang. Teach.*, vol. 19, no. 1, pp. 193–212, Jan. 2022.

[31] H. Karayiğit, A. Akdağlı, and Ç. İ. Acı, "BERT-based Transfer Learning Model for COVID-19 Sentiment Analysis on Turkish Instagram Comments," *Inf. Technol. Control*, vol. 51, no. 3, pp. 409–428, Sep. 2022.

[32] K. Schouten and F. Frăsincar, "Survey on Aspect-Level Sentiment Analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 3, pp. 813–830, Mar. 2016.

[33] M. Selimi and A. Besimi, "A Proposed Model for Stock Price Prediction Based on Financial News," *SSRN Electron. J.*, vol. 2019, no. Nov., pp. 1–10, Nov. 2019.

[34] W. Che, Y. Zhao, H. Guo, Z. Su, and T. Liu, "Sentence Compression for Aspect-Based Sentiment Analysis," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 23, no. 12, pp. 2111–2124, Dec. 2015.

[35] L. Britto and L. D. S. Pacífico, "A Step-by-Step Approach for User Acceptance Evaluation in Games Based on Sentiment Analysis and Machine Learning," in *Proc. KDMiLe 2022 – Brazilian Workshop on Knowledge Discovery, Modeling and Machine Learning*, vol. 2022, no. Nov., pp. 1–8, Nov. 2022.

[36] Y. Yu, "Mining Insights From Esports Game Reviews With an Aspect-Based Sentiment Analysis Framework," *IEEE Access*, vol. 11, no. Jun., pp. 61161–61172, Jun. 2023.

[37] S. Hansun, A. Suryadibrata, R. Nurhasanah, and J. Fitra, "Tweets Sentiment on PPKM Policy as a COVID-19 Response in Indonesia," *Indian J. Comput. Sci. Eng.*, vol. 13, no. 1, pp. 51–58, Jan. 2022.

[38] A. Rozaq, Y. Yunitasari, K. Sussolaikah, and E. R. N. Sari, "Sentiment Analysis of Kampus Mengajar 2 Toward the Implementation of Merdeka Belajar Kampus Merdeka Using Naïve Bayes and Euclidean Distance Methods," *Int. J. Adv. Data Inf. Syst.*, vol. 3, no. 1, pp. 1–8, Mar. 2022.

[39] D. Alomari, "Exploring Character Trigrams for Robust Arabic Text Classification: A Comparative Analysis in the Face of Vocabulary Expansion and Misspelled Words," *IEEE Access*, vol. 12, no. Apr., pp. 57103–57116, Apr. 2024.

[40] T. Hiraoka, S. Takase, K. Uchiumi, A. Keyaki, and N. Okazaki, "Joint Optimization of Tokenization and Downstream Model," *Findings of the Association for Computational Linguistics: ACL 2021*, vol. 2021, no. Aug., pp. 1–10, Aug. 2021.

[41] J. Petrus, E. Ermatita, Sukemi, and Erwin, "A Novel Approach: Tokenization Framework Based on Sentence Structure in Indonesian Language," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 2, pp. 1–8, Feb. 2023.

[42] Ç. Toraman, E. H. Yılmaz, F. Şahïnuç, and O. Özçelik, "Impact of Tokenization on

Language Models: An Analysis for Turkish," *arXiv Preprint*, vol. 2022, no. Apr., pp. 1–9, Apr. 2022.

[43] A. W. Pradana and M. Hayaty, "The Effect of Stemming and Removal of Stopwords on the Accuracy of Sentiment Analysis on Indonesian-Language Texts," *Kinet. Game Technol. Inf. Syst. Comput. Netw. Comput. Electron. Control*, vol. 4, no. 4, pp. 375–380, Oct. 2019.

[44] K. K. Agustiningsih, E. Utami, and O. M. A. Alsyaibani, "Sentiment Analysis and Topic Modelling of the COVID-19 Vaccine in Indonesia on Twitter Social Media Using Word Embedding," *J. Ilm. Tek. Elektro Komput. And Inform.*, vol. 8, no. 1, pp. 64–73, Feb. 2022.

[45] F. Rahutomo and A. R. T. H. Ririd, "Evaluasi Daftar Stopword Bahasa Indonesia," *J. Teknol. Inf. And Ilmu Komput.*, vol. 6, no. 1, pp. 41–48, Jan. 2019.

[46] A. Miyajiwala, "On Sensitivity of Deep Learning Based Text Classification Algorithms to Practical Input Perturbations," *arXiv Preprint*, vol. 2022, no. Jan., pp. 1–12, Jan. 2022.

[47] M. Alhawarat, H. A. M. Abdeljaber, and A. M. Hilal, "Effect of Stemming on Text Similarity for Arabic Language at Sentence Level," *PeerJ Comput. Sci.*, vol. 7, no. Jun., p. e530, Jun. 2021.

[48] D. C. Comeau, H. Liu, R. I. Dogan, and W. J. Wilbur, "Natural Language Processing Pipelines to Annotate BioC Collections With an Application to the NCBI Disease Corpus," *Database*, vol. 2014, no. Jun., pp. bau056–bau056, Jun. 2014.

[49] F. Barteld, I. Schröder, and H. Zinsmeister, "Dealing With Word-Internal Modification and Spelling Variation in Data-Driven Lemmatization," in *Proc. Workshop on Computational Linguistics for Linguistic Complexity (CL4LC)*, vol. 2016, no. Jun., pp. 45–54, Jun. 2016.

[50] N. A. Razmi, M. Z. Zamri, S. S. S. Ghazalli, and N. Seman, "Visualizing Stemming Techniques on Online News Articles Text Analytics," *Bull. Electr. Eng. Inform.*, vol. 10, no. 1, pp. 365–373, Mar. 2021.

[51] A. Nugumanova, Y. Baiburin, and I. Bessmertny, "Enhancement of Binary Text Classification Using Automatic Relation Extraction," in *Proc. Int. Conf. Computer Science and Information Technologies (CSIT 2015)*, vol. 2015, no. Sep., pp. 686–693, Sep. 2015.

[52] D. Yan, K. Li, S. Gu, and L. Yang, "Network-Based Bag-of-Words Model for Text Classification," *IEEE Access*, vol. 8, no. May, pp. 82641–82652, May 2020.

[53] A. Karami, "Taming Wild High Dimensional Text Data With a Fuzzy Lash," in *Proc. IEEE Int. Conf. Data Mining Workshops (ICDMW)*, vol. 2017, no. Nov., pp. 457–464, Nov. 2017.

[54] H. Naderi, S. Madani, B. Kiani, and K. Etminani, "Similarity of Medical Concepts in Question and Answering of Health Communities," *Health Informatics J.*, vol. 26, no. 2, pp. 1443–1454, Jun. 2019.

[55] Y. Zhou, "Prediction on Bitcoin Price Trends Based on Machine Learning Algorithms," *BCP Bus. Manag.*, vol. 34, no. Sep., pp. 21–29, Sep. 2022.

[56] M. Hao and K. Fan, "A Method for Calculating the Similarity of TF–IDF Texts for Synonyms in Biomedical Domains," in *Proc. 7th Int. Conf. Fuzzy Management and Statistical Methods (FMSMT 2017)*, vol. 2017, no. Dec., pp. 117–123, Dec. 2017.

[57] A. Falasari and M. A. Muslim, "Optimize Naïve Bayes Classifier Using Chi Square and Term Frequency Inverse Document Frequency for Amazon Review Sentiment Analysis," *J. Soft Comput. Explor.*, vol. 3, no. 1, pp. 31–36, Apr. 2022.

[58] W. Xu, L. Sun, and X. M. Liu, “An Improved Weighted Mixed Bayes Classification Model,” *Adv. Mater. Res.*, vol. 179–180, no. Feb., pp. 1260–1265, Feb. 2011.

[59] K.-M. Schneider, “A Comparison of Event Models for Naive Bayes Anti-Spam E-Mail Filtering,” in *Proc. 10th Conf. European Chapter of the Association for Computational Linguistics (EACL)*, vol. 2003, no. Apr., pp. 307–314, Apr. 2003.

[60] N. Umar and M. Nur, “Application of Naïve Bayes Algorithm Variations on Indonesian General Analysis Dataset for Sentiment Analysis,” *J. RESTI (Rekayasa Sist. and Teknol. Inf.)*, vol. 6, no. 4, pp. 585–590, Dec. 2022.

[61] Y. Yao, J. Wang, X. Liu, H. Liang, and C. Wang, “K-SVM: An Effective SVM Algorithm Based on K-Means Clustering,” *J. Comput.*, vol. 8, no. 10, pp. 2632–2639, Oct. 2013.

[62] K. Chong and N. Shah, “Comparison of Naive Bayes and SVM Classification in Grid-Search Hyperparameter Tuned and Non-Hyperparameter Tuned Healthcare Stock Market Sentiment Analysis,” *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 13, no. 12, pp. 1–8, Dec. 2022.

[63] S. H. Imanuddin, “Sentiment Analysis on Satusehat Application Using Support Vector Machine Method,” *J. Electron. Electromed. Eng. Med. Inform. (JEEMI)*, vol. 5, no. 3, pp. 143–149, Sep. 2023.

[64] E. A. S. Roja, “Performance of Machine Learning Models in Predicting Sentiments of Post-COVID Patients,” *Int. J. Recent Innov. Trends Comput. Commun. (IJRITCC)*, vol. 11, no. 10, pp. 2324–2329, Oct. 2023.

[65] M. B. Sitepu, I. R. Munthe, and S. Z. Harahap, “Implementation of Support Vector Machine Algorithm for Shopee Customer Sentiment Analysis,” *Sinkron*, vol. 7, no. 2, pp. 619–627, Apr. 2022.

[66] M. Ahmad and S. Aftab, “Analyzing the Performance of SVM for Polarity Detection With Different Datasets,” *Int. J. Mod. Educ. Comput. Sci. (IJMECS)*, vol. 9, no. 10, pp. 29–36, Oct. 2017.

[67] R. Umar, “Comparing the Performance of Data Mining Algorithms in Predicting Sentiments on Twitter,” *J. RESTI (Rekayasa Sist. and Teknol. Inf.)*, vol. 7, no. 4, pp. 817–823, Dec. 2023.