Sentiment Analysis of Reviews from Google Play: Azur Lane, Genshin Impact, Arknights

Kardandi Alfarizi Siregar[1*], Bhagaskara Cahyadi[2], Legiman Samosir[3], Alfani Azhard[4], Supiyandi[5,6]

[1,2,3,4,6] Faculty of Science and Technology, Department of Computer Science, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia
[5]Faculty of Computational Science and Digital Intelligence, Information Technology, Universitas Pembangunan Panca Budi, Medan, Indonesia
E-mail: [1]*kardandi83@gmail.com, [2]bhagaskaracahyadi.07@gmail.com, [3]legiman0701221018@uinsu.ac.id
[4]alfaniazhard02@gmail.com, [5]supiyandi.mkom@gmail.com
*E-mail Corresponding Author: kardandi83@gmail.com

Abstract

Abstract: Sentiment analysis of user reviews for mobile games is essential for understanding player perceptions of a game. This study focuses on sentiment analysis of user reviews for three popular mobile games: Azur Lane, Genshin Impact, and Arknights, using the Support Vector Machine (SVM) algorithm. The objective of this research is to classify reviews into three sentiment categories: Positive, Negative, and Neutral. Data was collected through web scraping from the Google Play Store, with a total of 6,000 reviews analyzed. The data preprocessing steps included cleaning, tokenization, stopword removal, and stemming, followed by TF-IDF feature extraction. The results show that the SVM model achieved an accuracy of 79.50%, with the best performance for Positive and Negative sentiments, but struggled with Neutral sentiment classification. The sentiment distribution revealed that Azur Lane had a higher proportion of Negative reviews compared to Genshin Impact and Arknights, which received predominantly Positive feedback. This study provides insights into the potential of using SVM for sentiment analysis in mobile games, and highlights areas for improvement, such as better handling of Neutral sentiment through more advanced models or balanced datasets.

Keywords: *Sentiment Analysis; Mobile Games; Support Vector Machine; Google Play Store; User Reviews.*

I. INTRODUCTION

Mobile games are a genre of video games specifically designed to be played on mobile devices such as smartphones and tablets. These games can be downloaded and installed through application stores available on mobile platforms, including the App Store for Apple iOS devices, Google Play Store for Android devices, and other app stores for various mobile platforms (Puspitasari, 2023).

The global market for gacha games, a subgenre of mobile games that integrates chance-based mechanics with strategic gameplay, was valued at USD 19.4 billion in 2023 and is expected to grow significantly, reaching USD 43.2 billion by 2032, at a compound annual growth rate (CAGR) of 9.3%. This growth is primarily driven by the increasing penetration of smartphones, widespread internet connectivity, and the rising popularity of mobile gaming. Both casual and hardcore gamers are attracted to gacha games, which combine elements of luck and strategy to create engaging experiences (Patel, 2025).

Sentiment analysis is a technique used to identify and extract opinions, feelings, or emotions from textual data. This technique combines Natural Language Processing (NLP) and machine learning to classify text into sentiment categories, such as positive, negative, or neutral (Effendi & Ernawati, 2025). By utilizing advancements in NLP, sentiment analysis has become a powerful tool for exploring opinions and emotions conveyed in text, offering valuable insights for decision-making and in-depth understanding of user feedback (Fadhilah & Utomo, 2024).

The Google Play Store is the official Android application platform developed by Google. Besides offering app downloads, the Play Store also features a review system where users can share their experiences with apps. These reviews often include ratings as well as detailed opinions on the strengths and weaknesses of applications, providing future users with valuable insights before downloading an app (Rahman et al., 2025). Online reviews also serve as a critical source of information for potential customers when making purchasing decisions, as they are generally free from the bias of the producers (Febrianta et al., 2021).

User sentiment plays a critical role in determining the success or failure of these games. With the increasing volume of user reviews on platforms like the Google Play Store, developers are increasingly turning to sentiment analysis to better understand user opinions and improve game features (Ramadhan & Mutiara, 2023). However, the large volume and diverse nature of reviews make manual analysis inefficient, and the application of automated sentiment analysis methods becomes essential.

Genshin Impact is a widely popular mobile game developed by miHoYo, also available on platforms like PC, PlayStation, and Nintendo Switch. The game features an open-world action role-playing game (RPG) format with stunning landscapes and compelling narratives. Released in 2020, Genshin Impact quickly became a global phenomenon, including in Indonesia, thanks to its immersive world and gameplay that appeal to both casual and hardcore gamers (Arsi et al., 2023).

Arknights is a mobile gacha game developed by Yostar, where players assume the role of “The Doctor,” a character with amnesia, rescued by the Rhodes Island organization. Set in a dystopian world ravaged by natural disasters, the game revolves around Originium, a rare yet dangerous resource that grants magical abilities and causes infection. The game’s tower defense mechanics, along with its rich storyline, character development, and unique gameplay, make it an engaging experience for players (Jensen, 2020).

Azur Lane is a side-scrolling shooter game that reimagines World War II battleships as anthropomorphic waifu characters. Developed by Shanghai Manjuu and Xiamen Yongshi, Azur Lane offers a unique blend of strategy, action, and gacha elements. Originally released in 2017 in Japan and China, it has since expanded internationally. Unlike other games in the genre, Azur Lane provides a more interactive experience with fans and communities, featuring gacha systems that are more enjoyable for players. Its success in the genre is attributed to its engaging gameplay, strategic elements, and beloved waifu characters (Aryani, 2019).

II. RESEARCH METHODOLOGY

The data for this research was collected through web scraping from the Google Play Store, which serves as the primary platform for mobile game reviews. The games chosen for this study are Azur Lane, Genshin Impact, and Arknights, which have large user bases and varying types of feedback. Web scraping is used to automatically extract user reviews, which include both textual feedback and rating information. The extraction process involves utilizing Python and a web scraping package called google-play-scraper, which helps efficiently gather reviews from the Google Play Store for the specified games. Reviews were extracted based on the rating system (1 to 5 stars) to capture a wide range of opinions. A total of 6,000 reviews were collected for the analysis, which were subsequently stored in CSV format for further processing.

The collected data is raw and requires several preprocessing steps before being used for sentiment analysis. The steps include:

- Data Cleaning: Irrelevant characters, such as special symbols, emojis, HTML tags, and unnecessary whitespace, are removed. This step ensures that the text is clean and free from noise that could affect the analysis.
- Case Folding: All text is converted to lowercase to ensure consistency in word representation. This eliminates any discrepancies caused by capitalization and improves the effectiveness of the analysis.
- Tokenization: The text data is split into individual words or tokens, which are the smallest units of meaning. Tokenization helps in breaking down sentences into manageable pieces for sentiment analysis.
- Stopword Removal: Common words such as "and", "the", "in", and "is" that do not contribute significant meaning are removed. This reduces the dimensionality of the dataset and enhances the focus on the important words.
- Stemming: Words are reduced to their root forms. For example, "playing" becomes "play" and "happier" becomes "happy". This process helps group similar words together, improving the model’s ability to generalize.

These preprocessing steps are crucial in making the dataset ready for feature extraction and sentiment classification.

To prepare the data for analysis, the TF-IDF (Term Frequency-Inverse Document Frequency) technique is applied. TF-IDF is a statistical method used to evaluate the importance of a word in a document relative to a collection of documents (the corpus). This method helps identify significant terms that carry important contextual meaning, which are used as features for the sentiment classification model.

The process works in two steps, Term Frequency (TF) calculates how often a word appears in a document. Inverse Document Frequency (IDF) measures the importance of the word across the entire dataset. Words that appear frequently across many reviews are down-weighted, while rare words with strong significance are highlighted. By combining TF and IDF, the model can focus on the most informative features (words) within the reviews, which enhances the accuracy of sentiment classification.

For sentiment analysis, the Support Vector Machine (SVM) algorithm is employed due to its effectiveness in text classification tasks. SVM is a supervised machine learning algorithm that works by finding an optimal hyperplane to separate different classes with maximum margin. In this study, the reviews are categorized into three sentiment classes: Positive, Negative, and Neutral. This classification process helps in determining the overall sentiment of the text, which can then be used for various applications such as understanding customer feedback or analyzing public opinion.

The classification process follows a systematic series of steps to ensure the model is trained and

E-ISSN: xxxx-xxxx
DOI: https://doi.org/10.64803/cessmuds.v1i1.33

evaluated correctly. First, the training data is prepared by splitting the dataset into three subsets: training, validation, and test sets with a 70:20:10 ratio. The training set is used to train the SVM model, while the validation set helps in tuning the hyperparameters of the model. The test set is used to evaluate the final performance of the model after it has been trained. This approach ensures that the model generalizes well to unseen data and does not overfit to the training set.

Model training is carried out using the SVM algorithm with a linear kernel, which is known to perform well for high-dimensional data such as text. The model learns to differentiate between positive, negative, and neutral reviews based on features extracted using Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF is used to weigh the importance of words in a document, allowing the model to focus on the most significant terms for classification. Once the model is trained, it undergoes hyperparameter tuning, where parameters like C (the regularization parameter) and the choice of kernel type are adjusted using cross-validation techniques to optimize performance.

The performance of the trained SVM model is evaluated using several metrics to ensure its accuracy and reliability. Key metrics include accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the model, while precision evaluates the accuracy of positive predictions. Recall assesses how well the model identifies all relevant cases, and F1-score is the harmonic mean of precision and recall, providing a balanced evaluation. These metrics help determine how well the model can correctly classify sentiments and whether it is effective in distinguishing between different sentiment categories.

To ensure the robustness of the results, the model's performance is further validated using the confusion matrix. This matrix shows the true positives, true negatives, false positives, and false negatives, providing a comprehensive view of the model's ability to distinguish between sentiment classes. By analyzing the confusion matrix along with precision, recall, and F1-score, a deeper understanding of the model's strengths and weaknesses can be gained, ensuring its practical application in sentiment classification tasks.

III. RESULTS AND DISCUSSION

In this section, we present the results from the sentiment analysis performed on the reviews of three mobile games: Azur Lane, Genshin Impact, and Arknights. We provide a breakdown of the model's performance, the sentiment distribution among the games, and the accuracy of the classification.

Accuracy of the SVM model: 79.50%

	precision	recall	f1-score	support
Negative	0.76	0.78	0.77	439
Neutral	0.00	0.00	0.00	81
Positive	0.81	0.90	0.85	680
accuracy			0.80	1200
macro avg	0.53	0.56	0.54	1200
weighted avg	0.74	0.80	0.77	1200

Figure 1. SVM Model Performance

The SVM (Support Vector Machine) model achieved an accuracy of 79.50%. As seen in the classification report, the Positive reviews were classified most accurately, with a precision of 0.81, recall of 0.90, and an F1-score of 0.85. However, the model struggled significantly with Neutral reviews, scoring a precision and recall of 0.00. This indicates that the neutral sentiment class was poorly predicted, which could be due to its ambiguous nature or insufficient training data.

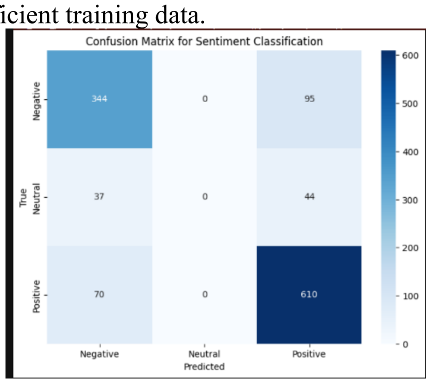

Figure 2. Confusion Matrix for Sentiment Classification

The confusion matrix (Figure 2) provides a detailed breakdown of the SVM model's performance. The matrix shows how well the model predicted the three sentiment classes: Negative, Neutral, and Positive.

- The Negative class was predicted with the highest accuracy (344 true negatives).
- A significant issue arises with Neutral reviews, where the model predicts them as Positive or Negative.
- The Positive class was predicted with good precision and recall, showing that the model was able to correctly classify positive feedback well.

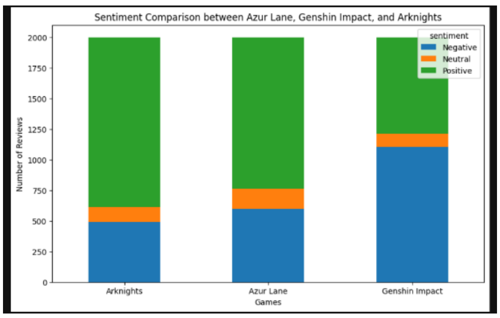

Figure 3. Sentiment Distribution of Reviews for Azur Lane, Genshin Impact, and Arknights

The sentiment distribution bar chart (Figure 3) provides a visual comparison of the sentiments expressed in the reviews for Azur Lane, Genshin Impact, and Arknights. The chart reveals several key observations:

- Positive sentiments dominate across all three games, as indicated by the large green sections of the stacked bars. This suggests that users generally have a favorable opinion of these games.
- Neutral sentiments are relatively low in comparison, as represented by the orange portions of the bars.
- Negative sentiments are more prominent in Azur Lane compared to Genshin Impact and Arknights, as seen in the larger blue portion of the Azur Lane bar. This indicates that Azur Lane received a higher proportion of dissatisfied or critical feedback.

The SVM model's confusion matrix (Figure 2) also provides insights into how the model performed in classifying the sentiments correctly. The matrix indicates that the model performed well in identifying Negative and Positive sentiments, as evidenced by the high values in the diagonal cells (344 true negatives and 610 true positives). However, it did not predict the Neutral class accurately, as shown by the zero values in the neutral category. This is an area of improvement, as neutral reviews were predominantly misclassified into other categories.

The SVM model worked well for positive and negative sentiments but faced challenges with neutral sentiments. This issue can be addressed by fine-tuning the model with more neutral sentiment data or exploring more sophisticated models like BERT that handle nuanced language better. The issues with neutral sentiment classification are similar to those encountered by Kurniawan et al. (2022), who also found that BERT struggles with neutral sentiments in game reviews. However, SVM showed better accuracy than models like Naïve Bayes used in previous studies. The issues with neutral sentiment classification are similar to those encountered by Kurniawan et al. (2022), who also found that BERT struggles with neutral sentiments in game reviews. However, SVM showed better accuracy than models like Naïve Bayes used in previous studies.

IV. CONCLUSION

The sentiment analysis performed on the reviews of Azur Lane, Genshin Impact, and Arknights using Support Vector Machine (SVM) yielded valuable insights into the sentiment distribution of user feedback across these games. The model achieved an accuracy of 79.50%, demonstrating its effectiveness in classifying Positive and Negative reviews. The Positive sentiment was predicted with high precision and recall, indicating that the model was particularly strong in identifying satisfied users. However, the model struggled significantly with Neutral sentiments, showing a precision and recall of 0.00. This poor performance with neutral reviews highlights a key weakness in the model, suggesting that neutral feedback is more complex and difficult to classify using the current approach.

The results also revealed that Azur Lane had a higher proportion of Negative reviews compared to Genshin Impact and Arknights, which received predominantly Positive feedback. This sentiment distribution was visually represented in a stacked bar chart, providing clear insights into how users feel about each game. The confusion matrix further demonstrated that the model was effective in classifying Negative and Positive sentiments, but the lack of accuracy in predicting Neutral sentiments needs to be addressed.

While the SVM model demonstrated strength in certain areas, its inability to accurately classify neutral reviews presents a significant weakness. One possible reason for this could be an imbalance in the data or the complexity of neutral sentiment, which is often nuanced and harder to categorize. To improve performance, future work could involve enhancing the model by incorporating more advanced techniques such as BERT, which has shown better results in capturing complex sentiments. Additionally, balancing the dataset with more Neutral sentiment data or exploring deep learning models and ensemble methods could help improve classification accuracy. Overall, while the research provided meaningful insights, further development is necessary to refine the model and achieve better results in classifying neutral sentiments, ultimately making sentiment analysis a more effective tool for understanding user feedback in mobile games.

V. RECOMMENDATIONS

To improve sentiment analysis in future research, several key steps can be taken. First, enhancing the classification of Neutral sentiments is crucial, as the current model struggled with this category. This could be achieved by gathering more neutral data and exploring advanced models like BERT or other transformer-based approaches. Additionally, balancing the dataset with more Neutral reviews through data augmentation could help improve the model's accuracy. Future studies could also consider using deep learning models such as RNNs or LSTMs, which are better suited for

.

complex textual data. Furthermore, breaking down sentiment categories into more granular groups, such as Very Positive or Mixed, would provide more detailed insights. Finally, expanding the research to include cross-domain sentiment analysis could offer a broader understanding of user feedback across different types of applications. Addressing these recommendations will enhance the effectiveness and reliability of sentiment analysis models in future studies.

VI. REFERENCES

Arsi, P., Subarkah, P., & Kusuma, B. A. (2023). Analisis Sentimen Game Genshin Impact pada Play Store Menggunakan Naïve Bayes Clasifier. *JURNAL ILMIAH TEKNIK MESIN, ELEKTRO DAN KOMPUTER (JURITEK)*, *3*(1), 161–170.

Aryani, T. (2019). *Ini Kantai Collection Tapi Lebih Baik (Ringkasan & Panduan Azur Lane)*. https://medium.com/@aryaniryan44/its-kantai-collection-but-better-an-azur-lane-rundown-guide-26f7efa1f96a

Effendi, P. A., & Ernawati, T. (2025). ANALISIS SENTIMEN ULASAN APLIKASI GAME HAY DAY MENGGUNAKAN ALGORITMA RANDOM FOREST. *Jurnal Informatika Dan Teknik Elektro Terapan*, *13*(3S1). https://doi.org/10.23960/jitet.v13i3S1.6772

Fadhilah, S. N., & Utomo, F. S. (2024). Algoritma Naïve Bayes untuk Analisis Sentiment Review Blibli.com di Google Play Store Naïve Bayes Algorithm for Sentiment Analysis of Blibli.com Review on Google Play Store. *SISTEMASI: Jurnal Sistem Informasi*, *13*(2), 831–840. http://sistemasi.ftik.unisi.ac.id

Febrianta, M. Y., Widiyanesti, S., & Ramadhan, S. R. (2021). Analisis Ulasan Indie Video Game Lokal pada Steam Menggunakan Analisis Sentimen dan Pemodelan Topik Berbasis Latent Dirichlet Allocation. *Journal of Animation & Games Studies*, *7*(2), 117–144.

Jensen, B. (2020). *Arknights Review – Tower Defense Waifus*. https://noisypixel.net/arknights-review-ios/

Patel, D. (2025). *Gacha Games Market*. https://dataintelo.com/report/global-gacha-games-market

Puspitasari, O. (2023). *Mengenal Game Mobile*. https://www.kompasiana.com/oktavianipuspitasari2770/6513d7d54addee576168cab2/mengenal-game-mobile

Rahman, R. N., Rahim, A., & Pranoto, W. J. (2025). Analisis Sentimen Ulasan Game eFootball 2024 Pada Playstore menggunakan Algoritma Naïve Bayes. *JURNAL ILMIAH INFORMATIKA (JIF)*, *13*(1), 38–44.

Ramadhan, Z. F., & Mutiara, A. B. (2023). Sentiment Analysis of Honkai: Star Rail Indonesian Language Reviews on Google Play Store Using Bidirectional Encoder Representations from Transformers Method. *International Journal of Engineering, Science and Information Technology*, *3*(3), 1–6. https://doi.org/10.52088/ijesty.v3i3.462

.