Comparison of Classification Algorithms with Bag of Words Feature in Sentiment Analysis

Fenilinas Adi Artanto

Faculty of Engineering and Computer Science, Informatic
University of Muhammadiyah Pekajangan Pekalongan
Pekalongan, Indonesia
fenilinasadi@gmail.com

Abstract— **The rapid growth of digital culture, especially on social media platforms, has led to the emergence of unique viral phenomena characterized by unconventional humor and illogical logic such as the Italian brainroot anomaly. Although there have been many studies on sentiment analysis, there is still a lack of studies focusing on cultural sentiment such as humor in the Italian brainroot anomaly. This study provides an overview of user sentiment analysis of the game "Hantu Tung Tung Tung Sahur 3D," a culturally viral application anomaly italian brainroot among young people on the Google Play Store during the month of Ramadan. User reviews were collected through web scraping, and data preprocessing involved tokenization, stopword removal, lowercase, stemming, and filtering to prepare the text for analysis. Feature extraction was performed using the Bag of Words method. This study compares the performance of four widely used classification algorithms—Support Vector Machine (SVM), Naïve Bayes, Decision Tree (C4.5), and Random Forest—implemented through Orange Data Mining software, with evaluation based on K-Fold Cross Validation. The novelty of this study lies in its focus on sentiment analysis in a unique and culturally viral digital context, as well as a comparative evaluation of classification algorithms specifically on this dataset. The results show that the Random Forest algorithm achieves the highest Area Under the Curve (AUC) score of 0.529, outperforming Naïve Bayes (0.504), SVM (0.503), and Decision Tree (0.498). These findings provide new insights into the suitability of ensemble methods such as Random Forest for sentiment analysis in specific digital phenomena, highlighting its potential for more reliable sentiment classification in similar contexts.**

Keywords— *Analysis Sentiment, Anomaly Tren, Bag of Words, Google Play Store.*

I. INTRODUCTION

The development of technology and the internet, especially social media, has created a unique and dynamic digital culture ecosystem. One of the emerging phenomena is the anomaly trend, a form of absurd humor and illogical narrative that is popular among children and teenagers. One of the viral anomaly trends is "Tung Tung Tung Sahur", which has its roots in the Italian Brainrot internet subculture which refers to content with broken logic, chaotic visuals, and dramatic sound effects that are deliberately made illogical but entertaining [1]. This trend quickly spread through TikTok, Youtube Shorts, and Instagram Reels, attracting the attention of especially Generation Z who love fast and absurd entertainment [2].

The phenomenon was then adopted into the form of a mobile game "Hantu Tung Tung Tung Sahur 3D" developed by LemauDev. This game presents a unique gameplay where players must avoid the ghost anomaly that chases them because they don't wake up for sahur. With scary but funny visualizations with an inspired story from Brainrot logic, this game went viral in Google Play Store downloads, especially during the month of Ramadan [3].

The popularity of this game not only reflects the developer's success in responding to viral trends, but also generates thousands of user reviews. These reviews reflect users' perceptions, satisfaction, criticism, and participation in the viral culture ecosystem. Therefore, analyzing user reviews is important to understand how the public responds emotionally to this application, whether they feel entertained, disappointed, scared, or even give negative sentiments towards the viral phenomenon.[4].

Method for understanding reviews from application users is through sentiment analysis, which is a method for classifying text data based on its emotional polarity (positive, neutral, or negative) [5]. Text classification in sentiment analysis is included in the data mining method, which is a concept where we look for information from the large amount of data that has been obtained [6]. In data mining there is a method of classification where data is mapped into predetermined groups or classes. The data mining method is a concept where we look for information from the large amount of data that has been obtained [7]. Some techniques or methods in classification are decision trees, support vector machines, neural networks, naïve Bayes, random forests and rule-based classifiers to find information from the large amount of data that has been obtained [8]. Review data in the form of text focuses on the interaction of opinions or views in sentence form, so that it will facilitate the sentiment analysis process [9].

As in Artanto's research (2024) where sentiment analysis on Twitter social media about KPPS members was carried out using Support Vector Machine (SVM), where the SVM algorithm obtained 70% accuracy results and showed that on Twitter social media when discussing KPPS members, it showed that 99% of the public gave positive opinions [10].

In the research of Sari & Wibowo, (2019) the method used in sentiment analysis is the Naïve Bayes algorithm on JD.ID

online store customers, where the results obtained show that Naïve Bayes has an accuracy of 96.44% where the people who gave positive sentiment were 300 reviews and those who gave negative sentiment were 294, while those who gave neutral sentiment were 288 [11].

Meanwhile, in the research of Fatkhudin et al., (2024), the Decision Tree algorithm was used to analyze Twitter sentiment about the use of Artificial Intelligence for Theses. The results obtained showed that the Decision Tree had an accuracy value of only 66% and gave negative sentiment results of 84.4% and 15.6% positive sentiment [12].

In Artanto's research, (2024) in sentiment analysis, the Random Forest algorithm was used to classify Twitter reviews about E-Stamps. The results obtained showed that Random Forest had an accuracy value of 70.1% with 86% neutral sentiment, 9% negative sentiment and 5% positive sentiment [13].

In order to improve text feature extraction in sentiment analysis, the Bag of Words (BoW) approach can be used, where the BoW model is widely used with good results for predicting language modeling and document classification because BoW is simple and flexible for specific text data [14].

In this context, this study aims to:

- Analyze user sentiment towards the "Hantu Tung Tung Tung Sahur 3D" application which represents anomalous and viral digital trends.
- Compare the performance of four popular classification algorithms (Support Vector Machine, Naïve Bayes, Decision Tree C4.5, and Random Forest) in processing review data containing typical absurd and emotional narratives.
- Examine the uniqueness of the review patterns of this application compared to other applications, especially related to language style, emotional intensity, and rapid viral dynamics.

The selection of this case study was criticized and selected because the application is a concrete representation of the viral digital culture phenomenon that has not been widely studied scientifically through text mining and sentiment analysis approaches. By using the Bag of Words method for feature extraction and evaluation using K-Fold Cross Validation, this study not only provides technical contributions in comparing classification algorithms, but also broadens the understanding of how absurd digital culture can be scientifically analyzed using artificial intelligence and data mining.

II. RESEARCH METHODOLOGY

This study uses a quantitative approach by comparing the performance of the accuracy of the classification algorithm in the sentiment analysis of application reviews with Bag of Word feature selection. The research process is divided into several stages as shown in Figure 1:
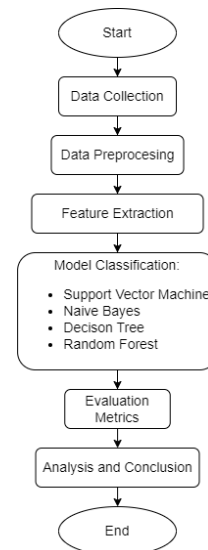

Figure 1. Research Flow

A. Data Collection

The data used in this study were obtained from user reviews of the "Hantu Tung Tung Tung Sahur 3D" application available on the Google Play Store. Data retrieval was carried out using the web scraping technique using the Playstore Scraper tool using the Phython Library on Google Collabs with the syntax:

```
!conda install -y gdown

!pip install google-play-scraper

!pip install PySastrawi

import nltk

tungtung, continuation_token = reviews('net.lemau.hantutungtungtungtung3D', lang='id', country='id', sort=Sort.NEWEST, count=1905)
```

The data used in this study are user reviews of the "hantu Run Tun Tun Saur 3d" application on the Google Play Store. The initial data collected amounted to 1,905 reviews, taken with the parameters lang='id' and country='id' to ensure that all reviews are in Indonesian. The data obtained will later be used in orange datamining to test algorithms for sentiment analysis.

B. Data Preprocessing

Before classification is carried out, the review data will go through a preprocessing process to make it suitable for processing, this stage includes:

- Tokenization
  Tokenization is the stage where review data in the form of sentences is broken down into words [15].
- Stopword Removal
  This stage involves removing stop words, which are common words that are considered not to provide valuable information in text analysis due to their frequent occurrence [16].

- Lowercasing
  This step changes all letters to lowercase [17].
- Stemming/Lemmatization
  This process converts words into their basic or root form. This helps in reducing the variation of words that have the same root, thus making the analysis easier [18].
- Filtering
  This step removes symbols, numbers and punctuation [19].
- Normalization
  In particular, normalization of non-standard words, slang, and typical Brainrot terms (e.g.: "tung tung", "sahur 3D", "brainrot", "ngakak", "random banget") was carried out so that they remain represented in the analysis. This list of terms is arranged based on the frequency of occurrence and cultural relevance in the dataset [20].

C. Data Labeling

The data labeling process is done automatically using a multilingual sentiment analysis model that supports Bahasa Indonesia, so it can capture the nuances of sentiment in the context of local culture and language. This model classifies each review into three sentiment classes: positive, negative, based on the content and tone of the sentence.

D. Feature Extraction

After the data is cleaned, the next step is to convert the text to numeric form using Bag of Words (BoW) feature extraction to calculate the frequency of word occurrence and arrange them in the form of feature vectors that can be used by machine learning algorithms [13].

BoW is popularly used in various NLP tasks such as text classification and sentiment analysis due to its simplicity and efficiency in converting unstructured text data into structured numeric data. However, BoW does not consider word order or word meaning in depth, making it sometimes less effective in capturing complex contexts.

E. Model Classification

This study compares four classification algorithms for sentiment analysis, namely:

- Support Vector Machine (SVM)
  Effective for high-dimensional data such as text and capable of handling non-linear decision boundaries which generally have high performance in text classification, such as sentiment analysis [21].
- Naïve Bayes
  Fast and efficient algorithm in training and prediction, suitable for text classification, and insensitive to small data [22].
- Decision Tree
  Easy to implement, can handle categorical and numeric data and does not require much preprocessing (normalization) [23].
- Random Forest

It is an ensembling of many decision trees, so it is more accurate and stable, can reduce overfitting in handling large and complex data [24].

The four algorithms will be implemented using the Orange Data Mining application version 3.38.1, where each algorithm is configured and connected to the Test and Score component for the evaluation process with the K-fold Cross Validation approach to ensure the reliability of the results.

F. Evaluation Metrics

To assess the performance of each algorithm, several standard evaluation metrics are used in classification, namely [6]:

- Accuracy: Proportion of correct predictions from the total data.
- Precision: The ability of the model to correctly identify the positive class.
- Recall: The ability of the model to capture all actual positive data.
- F1-Score: The average harmonic value between Precision and Recall.
- Confusion Matrix: Visualization of the classification results showing the correct and incorrect prediction results of each class.

G. Analysis and Conclusion

The evaluation results of the four algorithms will be compared to see which algorithm is the most effective in analyzing application review sentiment.

III. RESULT AND DISCUSSION

Before classification, the review data collected through PlayStore Scraper, totaling 1,905 entries, first went through a manual labeling process. Two independent annotators assessed each review and classified it into three sentiment categories: positive, negative, and neutral. Consistency between annotators was measured using Cohen's Kappa, which produced a value of 0.87 (very good category). If there was a difference, a discussion was held to reach a consensus. After the data cleaning process (removing duplicates, spam, and empty reviews), the final amount of data analyzed was 1,500 reviews., then the next process was to analyze the data using the orange data mining design as follows:
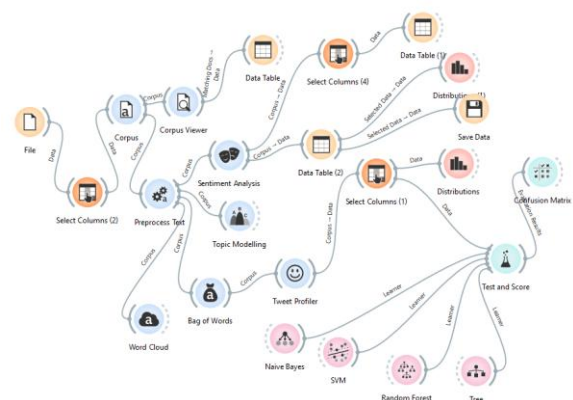

Figure 2. Design Orange Data Mining

A. *Data Preprocessing*

Data preprocessing is carried out in the Preprocess Text tool which is used at the following levels:
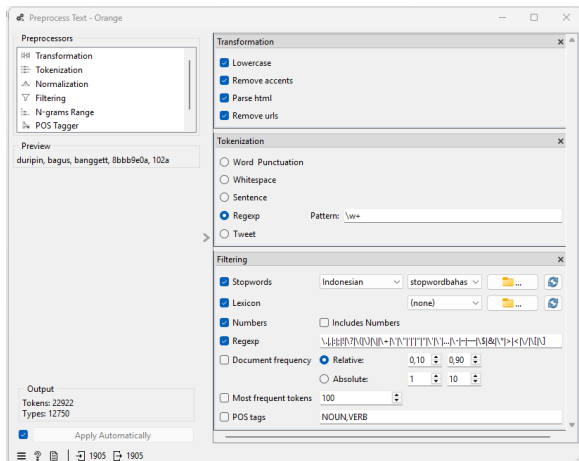

Figure 3. Preprocess Text (Orange Data Mining)

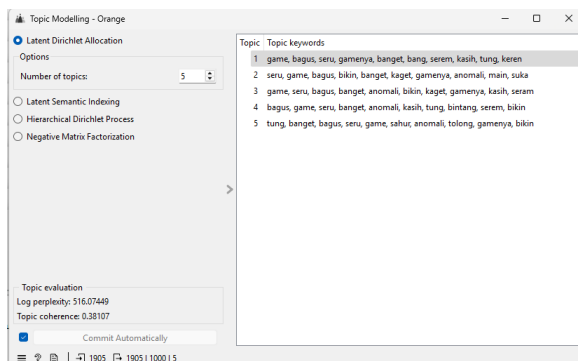After going through the Text Preprocess, the topic modeling obtained is:

Figure 4. Topic Modelling (Orange Data Mining)

In topic modeling, the topic keywords that are in the application reviews are displayed, namely: "game, bagus, seru, banget, serem, kasih, keren".

In the sentiment analysis tools, the Multilingual sentiment with language Indonesian, method is used in Indonesian as shown in the Figure 5:
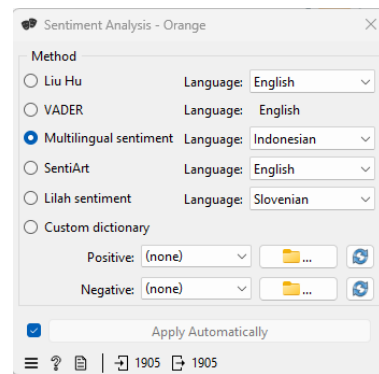

Figure 5. Sentiment Analysis (Orange Data Mining)

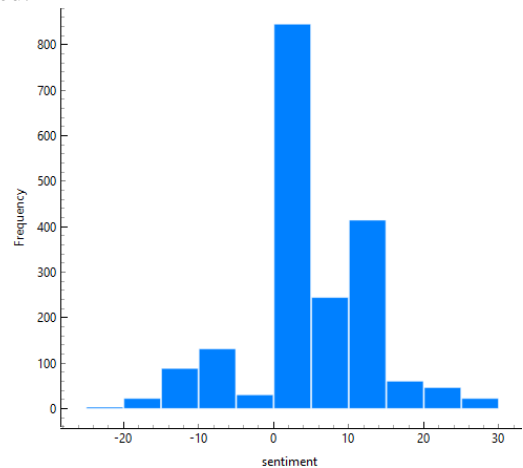From Figure 5, the following distribution graph results are obtained:

Figure 6. Graph Distributions Sentiment

In Figure 6 it can be seen that the majority of application reviews have positive sentiment with a total of 85.61% and application reviews with negative sentiment of 14.39%. This dominance of positive sentiment indicates that the majority of users appreciate the unique concept and absurd humor presented by the application, even though there are elements of fear or strangeness.

Apart from that, the tools in Preprocess Text also get Word Cloud results as in Figure 7:
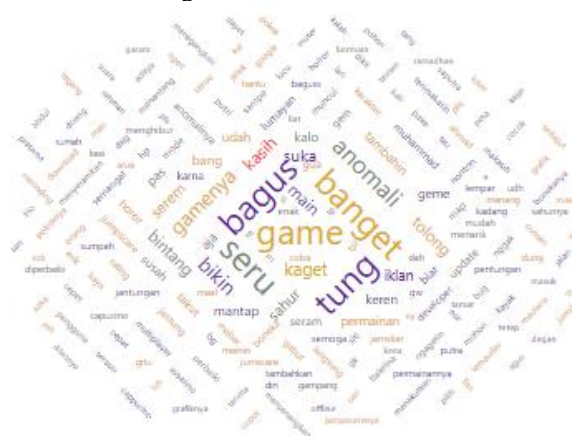

Figure 7. Word Cloud

The word cloud results in figure 7 show that the most common words in the application reviews are "game, bagus, seru, banget, anomali, kasih, kaget".

The results of word cloud and topic modeling show dominant words such as "game", "good", "fun", "banget", "scary", "anomaly", and "shocked". These words represent the main characteristics of the application, namely a combination of elements of entertainment, strangeness (anomaly), and surprise (shocked/scary). This pattern is consistent with the Brainrot phenomenon, where absurd narratives and exaggerated expressions are the characteristics of the digital culture adopted by this application. The presence of the words "anomaly" and "scary" in the context of reviews also indicates user acceptance of the absurd and chaotic elements which are indeed the main attractions of the application.

B. Feature Extraction

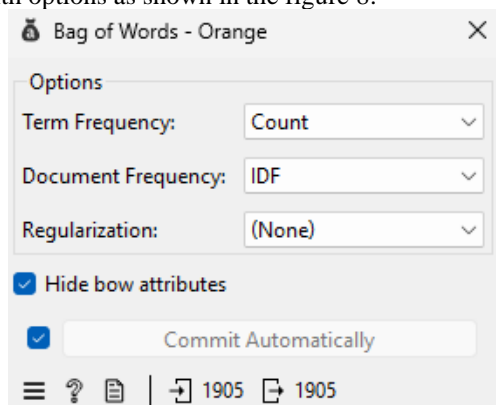At the Feature Extraction stage, the Bag of Words tool is used with options as shown in the figure 8:

Figure 8. Bag of words (Orange Data Mining)

Different from sentiment analysis in Preprocess Text in testing the classification algorithm using tweet profiler with the Ekman emotions option because it will test the emotional expression of application users when giving reviews on the "Hantu Tung Tung Tung Sahur 3D" application. Emotion analysis using Tweet Profiler (although originally developed for Twitter data) remains relevant because this model detects short text-based emotional expressions that are also common in app reviews.

The results of the tweet profiler referring to the emotions of the application users are shown in the figure 9.
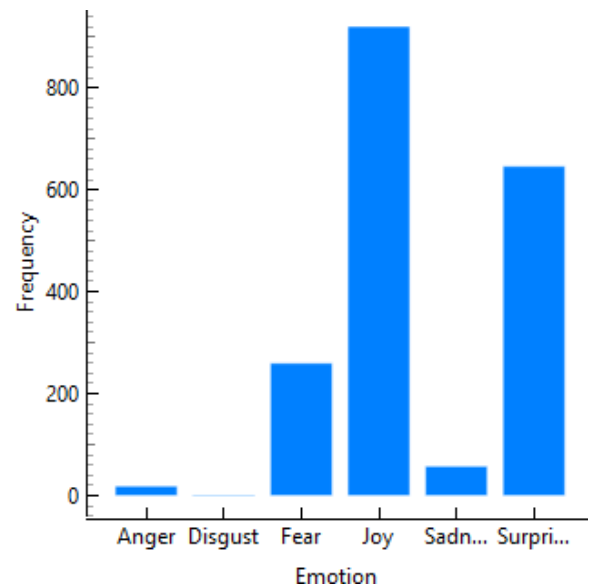

Figure 9. Garph Figure Emotion

With emotions anger 1% disgust 0.1% fear 13.65% joy 48.29% sadness 3.04% surprise 33.91%. this shows that users of the application "Hantu Tung Tung Tung Sahur 3D" when giving reviews of the application, the majority showed expressions of joy. This strengthens the argument that the app succeeds in delivering an experience that is not only entertaining, but also surprising and a little suspenseful—a hallmark of Brainrot's narrative.

C. Model Classification

At the Model Classification stage, four algorithms were compared, Support Vector Machine (SVM), Naive Bayes, Decision Tree, and Random Forest using the Test and Score tool on Orange Data Mining. The results are shown in the table.

Table I. Test and Score Result

Algorithms	AUC	F1	Precision	Recall
SVM	0,503	0,266	0,354	0,306
Random Forest	0,529	0,355	0,476	0,491
Decision Tree	0,498	0,357	0,361	0,356
Naïve Bayes	0,504	0,018	0,229	0,010

while the results of the Confusion Matrix Support Vector Machine (SVM):

		Predicted					
	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	0	0	4	3	0	12	19
Disgust	0	0	0	0	0	2	2
Fear	0	0	83	28	0	149	260
Joy	0	0	251	105	0	564	920
Sadness	0	0	12	7	0	39	58
Surprise	0	0	160	91	0	395	646
Σ	0	0	510	234	0	1161	1905

the results of the Confusion Matrix Random Forest:

Predicted

	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	0	0	0	18	1	0	19
Disgust	0	0	0	2	0	0	2
Fear	0	0	1	255	0	4	260
Joy	0	0	2	891	0	27	920
Sadness	0	0	0	57	0	1	58
Surprise	0	0	0	603	0	43	646
Σ	0	0	3	1826	1	75	1905

the results of the Confusion Matrix Decision Tree:

Predicted

	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	1	0	3	10	1	4	19
Disgust	0	0	0	2	0	0	2
Fear	7	0	43	123	7	80	260
Joy	17	0	165	447	21	270	920
Sadness	1	0	3	32	4	18	58
Surprise	5	0	120	331	7	183	646
Σ	31	0	334	945	40	555	1905

the results of the Confusion Matrix Naive Bayes:

Predicted

	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	0	18	0	1	0	0	19
Disgust	0	2	0	0	0	0	2
Fear	1	250	0	9	0	0	260
Joy	6	896	0	18	0	0	920
Sadness	1	55	0	2	0	0	58
Surprise	1	637	0	8	0	0	646
Σ	9	1858	0	38	0	0	1905

The results show that Support Vector Machine (SVM) is not very effective in this model, possibly due to the complexity of the review data or class imbalance. Random Forest shows the most stable and good performance among the four algorithms, although its accuracy is generally still moderate. Although its AUC on Decision Tree is low, the F1-score which is almost the same as Random Forest shows that Decision Tree can also be an alternative, but it needs parameter tuning. Naïve Bayes has poor performance, possibly failing because the assumption of independence between features is not met or because of the imbalance of sentiment labels. Random Forest is the best algorithm in this experiment based on F1, Precision, and Recall metrics. Support Vector Machine (SVM) and Decision Tree have moderate performance, but can still be improved. while Naïve Bayes performs very poorly in this case, not suitable for this review data directly. In the confusion matrix, Random Forest shows the highest number of correct predictions for the positive class, while SVM and Naïve Bayes tend to make many prediction errors in the minority (negative) class. This indicates that ensemble models such as Random Forest are more adaptive to variations in language style and cultural expressions that are typical of this application.

Previous studies Artanto, (2024) and Sari & Wibowo, (2019) have shown that algorithms such as SVM and Naïve Bayes can achieve high accuracy on conventional review data. However, in the context of applications with absurd narratives and digital culture such as "Hantu Tung Tung Tung Sahur 3D", the performance of both algorithms decreased drastically. This confirms that the characteristics of language and cultural expressions greatly affect the effectiveness of the model. This finding is a new contribution, while also emphasizing the importance of developing models that are more sensitive to the context of digital culture.

This study shows that Sentiment analysis in applications with absurd cultural contexts requires special attention to preprocessing and algorithm selection. Ensemble models such as Random Forest are superior in handling data with variations in language and cultural expressions. Visualizations such as word clouds and emotion analysis can enrich the understanding of user interaction patterns in viral digital culture.

IV. CONCLUSION

Based on a sentiment analysis of 1,905 user reviews for the "Hantu Tung Tung Tung Sahur 3D" app, the study found that the Bag of Words (BoW) feature extraction method is effective in converting unstructured text into numerical features suitable for machine learning algorithms. Among the four classification algorithms tested—Support Vector Machine (SVM), Naïve Bayes, Decision Tree, and Random Forest—the Random Forest algorithm achieved the highest F1 score (0.355) and recall (0.491), as well as the highest AUC value (0.529). However, it is important to note that these performance metrics remain relatively low and are only slightly better than random classification, indicating that the classification task is challenging given the nature of the data. The relatively modest performance across all models can be attributed to several factors, including class imbalance in sentiment labels, complexity, and mismatching of labels in review texts inherent in the BoW feature representation that does not capture semantic context or word order in italian brainroot anomalous cultural sentiment labeling. In particular, the Naïve Bayes algorithm performed the worst (F1 score of 0.018), likely due to its strong independence assumption that does not hold in this dataset.

Therefore, although Random Forest performed relatively better in this particular dataset, it is premature to generalize its superiority to all viral app review datasets, especially considering that this study only focused on one culturally unique app anomaly italian brainroot. Further research should explore more sophisticated feature extraction techniques such as word embeddings with Word2Vec, GloVe or TF-IDF weighting to better capture semantic nuances and reduce feature sparsity. Additionally, leveraging deep learning models such as Long Short-Term Memory (LSTM) networks can improve the ability to model sequential dependencies and contextual information in short, unstructured texts such as user reviews.

In conclusion, this study highlights the potential and challenges of applying machine learning to sentiment analysis in the context of the viral digital cultural phenomenon Italian brainroot. It underlines the need for more sophisticated

modeling approaches and careful consideration of data characteristics to improve classification accuracy.

REFERENCES

[1] N. Zarawaki, "Apa Itu Tren Anomali Brainrot? Ada Meme Tung Tung Sahur!," *IDN Times*, Apr. 25, 2025. Accessed: May 05, 2025. [Online]. Available: https://www.idntimes.com/life/inspiration/nisa-zarawaki/apa-itu-tren-anomali-brainrot?page=all

[2] Z. Hardiansyah, "Daftar Nama Anomali TikTok yang Lagi Viral, Ada Tung Tung Tung Sahur," *kompas.com*, Apr. 24, 2025. [Online]. Available: https://tekno.kompas.com/read/2025/04/24/12350027/daftar-nama-anomali-tiktok-yang-lagi-viral-ada-tung-tung-tung-sahur?page=all

[3] M. A. Budi, "Ini Asal Muasal Fenomena Meme Viral Tung Tung Sahur, dari Tradisi Ramadan ke Tren Digital," *Radar Jember*, Jember, May 07, 2025. [Online]. Available: https://radarjember.jawapos.com/nasional/795980586/ini-asal-muasal-fenomena-meme-viral-tung-tung-sahur-dari-tradisi-ramadan-ke-tren-digital

[4] C. Rosanti, F. A. Artanto, and R. E. Saputra, "Perception of user opinions towards sharia mobile banking applications in Indonesia based on comments on Google Play Store," *Serambi*, vol. 7, no. 1, pp. 97–108, 2025, doi: https://doi.org/10.36407/serambi.v7i1.1486.

[5] F. A. Artanto, "Analisis Sentimen Opini Publik terhadap Fenomena Bunuh Diri Mahasiswa di Twitter Menggunakan Algoritma Naive Bayes," *Satesi J. Sains Teknol. dan Sist. Inf.*, vol. 4, no. 1, pp. 70–76, 2024, doi: 10.54259/satesi.v4i1.2908.

[6] I. Rosyadi, H. H. Kusumawardani, F. A. Artanto, A. A. A. Hardani, and F. Nafilaturrosyidah, "Clustering K-Means Dalam Pengelompokan Penjualan Produk Pada RTO Group," *Teknomatika*, vol. 13, no. 02, pp. 55–60, 2023, [Online]. Available: http://ojs.palcomtech.ac.id/index.php/teknomatika/article/view/618/439

[7] H. H. Kusumawardani, I. Rosyadi, F. A. Artanto, F. I. Arzha, and N. A. Rachmayani, "Analisis Decision Tree dalam Pengaruh Digital Marketing terhadap Penerimaan Siswa Baru," *Remik*, vol. 6, no. April, pp. 225–231, 2022.

[8] A. Fatkhudin, M. Y. Febrianto, F. A. Artanto, M. W. N. Hadinata, and R. Fahlevi, "Algoritma Decision Tree C.45 dalam analisa kelulusan mahasiswa Program Studi Manajemen Informatika UMPP," *J. Ilm. Ilmu Komput. Fak. Ilmu Komput. Univ. Al Asyariah Mandar*, vol. 8, no. 2, pp. 83–86, 2022.

[9] C. Rosanti, F. A. Artanto, and R. E. Saputra, "Analisis Sentiment Pengguna Aplikasi Mobile Banking Pada Bank Syariah Dengan Support Vector Regression," *J. Pendidik. dan Teknol. Indones.*, vol. 4, no. 8, pp. 341–347, 2024, doi: https://doi.org/10.52436/1.jpti.460.

[10] F. A. Artanto, "Support Vector Machine Berbasis Particle Swarm Optimization Pada Analisis Sentimen Anggota KPPS," *J. FASILKOM (teknologi Inf. dan ILmu KOMputer)*, vol. 14, no. 1, pp. 75–79, 2024, doi: https://doi.org/10.37859/jf.v14i1.6795.

[11] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi," *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.

[12] A. Fatkhudin, F. A. Artanto, N. A. Safli, and D. Wibowo, "Decision Tree Berbasis SMOTE dalam Analisis Sentimen Penggunaan Artificial Intelligence untuk Skripsi," *Remik Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 8, no. April, pp. 494–505, 2024, doi: 10.33395/remik.v8i2.13531.

[13] F. A. Artanto, "Implementasi Algoritma Random Forest dan Model Bag of Words Dalam Analisis Sentimen Mengenai E-Materai," *Satesi J. Sains Teknol. dan Sist. Inf.*, vol. 4, no. 2, pp. 139–145, 2024, doi: 10.54259/satesi.v4i2.3240.

[14] D. Krisnandi, R. N. Ambarwati, A. Y. Asih, A. Ardiansyah, and H. F. Pardede, "Analisis Komentar Cyberbullying Terhadap Kata Yang Mengandung Toksisitas Dan Agresi Menggunakan Bag of Words dan TF-IDF Dengan Klasifikasi SVM," *J. Linguist. Komputasional*, vol. 6, no. 2, pp. 36–41, 2023, doi: 10.26418/jlk.v6i2.85.

[15] E. Subowo, F. Adi Artanto, I. Putri, and W. Umaedi, "BLTSM untuk analisis sentimen berbasis aspek pada aplikasi belanja online dengan cicilan," *J. Fasilkom*, vol. XII, no. Ii, pp. 132–140, 2022.

[16] M. F. I. Haq, I. Rosyadi, M. Nasir, and A. Khambali, "Sentiment Analisis Ulasan Aplikasi Livin Pada Google Play Store," *J. Surya Inform.*, vol. 14, no. 1, pp. 24–29, 2024.

[17] D. A. Putri and D. A. Muthia, "Implementasi Metode Lexicon Based dan Support Vector Machine Pada Analisis Sentimen Ulasan Pengguna ChatGPT," *IJCIT (Indonesian J. Comput. Inf. Technol.*, vol. 9, no. 2, pp. 80–86, 2024.

[18] A. Wahyuningtyas, I. S. Sitanggang, and H. Khotimah, "Deteksi Spam pada Twitter Menggunakan Algoritme Naïve Bayes," *J. Ilmu Komput. dan Agri-Informatika*, vol. 7, no. 1, pp. 31–40, 2020, doi: 10.29244/jika.7.1.31-40.

[19] S. Alfaris and Kusnawi, "Komparasi Metode KNN dan Naive Bayes Terhadap Analisis Sentimen Pengguna Aplikasi Shopee," *Indones. J. Comput. Sci.*, vol. 12, no. 5, pp. 2766–2776, 2023, doi: 10.33022/ijcs.v12i5.3304.

[20] R. Isnaeni, Sudarmin, and R. Zulkifli, "Analisis Support Vector Regression (SVR) Dengan Kernel Radial Basis Function (RBF) Untuk Memprediksi Laju Inflasi Di Indonesia," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 4, no. 1, pp. 30–38, 2022, doi: 10.35580/variansiunm13.

[21] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.

[22] R. Apriani and D. Gustian, "Analisis Sentimen dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia," *J. Rekayasa Teknol. Nusa Putra*, vol. 6, no. 1, pp. 54–62, 2019, [Online]. Available: https://rekayasa.nusaputra.ac.id/article/view/86

[23] A. Fatkhudin, A. Khambali, and F. A. Artanto, "Decision Tree Dalam Mengklasifikasi Mata Kuliah Terhadap Pemahaman Sistem Pemasaran," *J. Ilm. Ilmu Komput. Fak. ...*, vol. 7, no. 2, pp. 52–55, 2021, [Online]. Available: http://ejournal.fikom-unasman.ac.id/index.php/jikom/article/view/204

[24] I. Afdhal, R. Kurniawan, I. Iskandar, R. Salambue, E. Budianita, and F. Syafria, "Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 1, pp. 122–130, 2022, [Online]. Available: http://ojs.serambimekkah.ac.id/jnkti/article/view/4004/pdf