Journal of Information Systems and Informatics
Vol. 6, No. 4, December 2024 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v6i4.929 Published By DRPM-UBD

Analysis of Labeling and Class-Balancing Effects on Clash of Champions Sentiment Using LSTM and BERT

Audi Ilham Atmaja[1], Maimunah[2], Pristi Sukmasetya[3]

[1,2,3]Informatics Engineering Study Program, Muhammadiyah University of Magelang, Magelang City, Central Java, Indonesia
Email: [1]audiatmaja@gmail.com, [2]maimunah@unimma.ac.id, [3]pristi.sukmasetya@ummgl.ac.id

Abstract

Advances in digital technology have changed the way people interact and access information, including in education. One educational event that has caught the public's attention is Clash of Champions by Ruangguru, designed to increase young people's interest in learning through an interactively presented competition. The purpose of this study is to use posts on X social media to examine public opinion on the event. Using TweetHarvest, 1,891 tweets were gathered and preprocessed (cleaning, case folding, normalization, tokenization, stopword removal, stemming, and English translation). A total of 12 experimental scenarios were created by combining VADER and TextBlob labeling strategies with class balancing techniques (undersampling and SMOTE), and the LSTM and BERT models were evaluated for each scenario. The best results were achieved by combining VADER, SMOTE, and BERT, yielding an accuracy of 97.73%, with precision, recall, and F1-scores of 98%, 98%, and 96% (positive), 99% (neutral), and 98% (negative), respectively. These findings highlight the efficacy of transformer-based models like BERT in addressing class imbalance and improving sentiment classification. The integration of SMOTE effectively mitigated class imbalance, providing consistent and accurate performance across all sentiment categories.

Keywords: Clash of Champions, Sentiment Analysis, VADER, SMOTE, BERT

1. INTRODUCTION

In the era of the digital revolution, information and communication technology is developing rapidly, changing the way people interact and access information through fast and simple media, and enabling interaction between technology products [1]. Social media, as one of the main outcomes of digital progress, has become not only a means of communication but also a platform for individuals to express their views openly and in real-time. These platforms have become an important part of society's communication, changing the way experiences are shared, reviews are given, and public opinion is formed, even influencing behavior in various aspects of life [2].

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

The widespread use of the internet also affects the field of education and academic achievement. Rapid internet access makes it easier for young people to obtain information and learning materials, supports independent learning, and enables interaction with a wider learning community [3]. Recent studies have shown that the high activity of internet users, including the use of smartphones in academic and educational activities, has a positive correlation to the improvement of academic skills and achievement, and has a significant impact on improving access and quality of education in the digital era [4].

One platform that optimally utilizes this technology is Ruangguru, an edutech platform that focuses on developing digital-based education in Indonesia. Since its establishment in 2014, Ruangguru has been a pioneer in providing interactive and high-quality online learning content, including through engaging and easy-to-understand animated videos, to improve the quality of education [5]. The platform has gained much attention for its commitment to improving access to education using technology.

As a strategy to reach more users, Ruangguru launched an Edutainment Show called Clash of Champions (CoC), which is packed with interesting visuals and marketing strategies, thus becoming a new buzz in teen social media which was previously dominated by non-educative content. The show featured outstanding teenagers from domestic and foreign universities competing in logic, memorization, and calculation skills, and created a new hegemony for teen shows in Indonesia [6]. Clash of Champions also became a widely discussed topic on social media, encouraging users to give their opinions, responses and sentiments towards the show.

The use of X as a promotional platform gives audiences direct access to express opinions about Clash of Champions, which can be analyzed to determine the perceptions, appeal, and learning interests that are expected to grow through this event. Sentiment analysis is an important tool to understand public opinion towards the show [7]. Positive, negative and neutral sentiments expressed through tweets can give an idea of how the show is being received by the public.

Numerous research has shown how deep learning-based models and natural language processing may be used for social media sentiment analysis. Bidirectional Encoder Representations from Transformers (BERT) was used in a deep learning study to determine if tweets about the Clash of Champions event were neutral, bad, or favorable by Ruangguru, with the results showing the dominance of positive sentiments above 85% which reflects the enthusiasm of the community for the event [8]. On the other hand, BERT was also utilized in research related to sentiment analysis of 2020 Jakarta flood tweets to support emergency response and improve situational awareness, yielding a 79% testing accuracy [9].

Journal of Information Systems and Informatics
Vol. 6, No. 4, December 2024 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v6i4.929 Published By DRPM-UBD

Furthermore, studies comparing the sentiment analysis of COVID-19-related tweets in Nepal using the Long Short Term Memory (LSTM), Naïve Bayes, and Support Vector Machine (SVM) approaches where the LSTM model recorded the highest accuracy of 79% [10].

Previous research shows that Sentiment study of COVID-19-related social media data using TextBlob, Valence Aware Dictionary and Sentiment Reasoner (VADER) yields inconsistent findings. On a collection of Arabic-translated tweets concerning vaccinations, the performance of TextBlob and VADER only reached 75% and 70% accuracy, affected by limitations in detecting negation as well as the polarity of comparative and superlative adjectives [11]. Conversely, the PeduliLindungi app's public sentiment analysis showed that VADER was more effective than TextBlob in capturing opinions on social media, providing useful information for the improvement of the app [12].

Several studies have shown the importance of handling data imbalance in sentiment analysis on social media. The use of random majority under-sampling (RMU) method on Twitter dataset resulted in competitive accuracy and reduced processing time by 50%, although the accuracy was slightly lower than the method without resampling [13]. On the other hand, overcoming data imbalance with the application of the Synthetic Minority Over-sampling Technique (SMOTE) in the MBKM "Merdeka Belajar Kampus Merdeka" program response through social media successfully improved the quality of model training, with SVM achieving the highest accuracy of 91% [14].

Previous sentiment analysis methods, such as TextBlob and VADER, often struggle with certain linguistic challenges, including handling negations, detecting the nuances of comparative and superlative adjectives, and interpreting informal language common on social media. Furthermore, social media datasets, which are typically large-scale and often imbalanced in class distribution, pose additional challenges for achieving robust and reliable sentiment classification. Addressing these limitations requires advanced modeling techniques, such as transformer-based models and effective data balancing strategies, to ensure consistent performance across all sentiment categories.

Based on previous research, this study aims to compare and derive several methods, including TextBlob and VADER labeling methods, class balancing techniques using SMOTE and random undersampling, and modeling with LSTM and BERT on sentiment analysis performance in analyzing public opinion about Clash of Champions. By addressing the limitations of traditional sentiment analysis methods and tackling data imbalance, this study seeks to determine the most effective approach for sentiment classification in large-scale social media datasets, particularly in the context of the education and entertainment industry in Indonesia.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

2. METHODS

A number of steps are taken in this research approach stage which include several steps or processes to obtain the results and conclusions of the research conducted. Figure 1 presents the overall research flow:

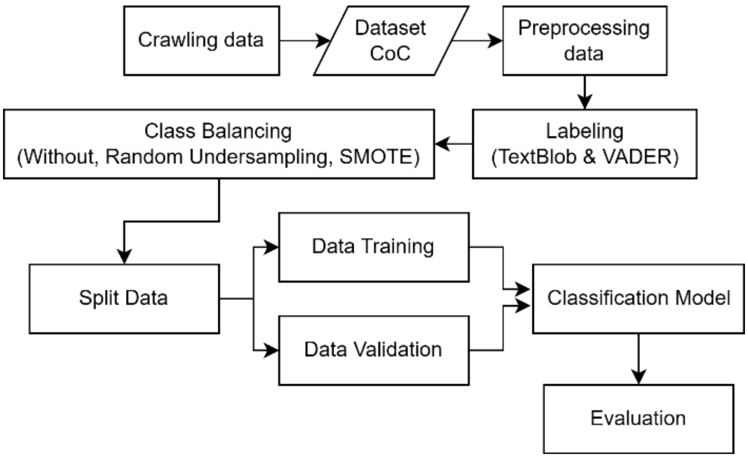


Figure 1. Research flow

2.1. Data Crawling

The initial phase of this study approach is data crawling. The data crawling process is obtained from platform X by using the Twitter Application Programming Interface (API) of Tweet Harvest which allows direct and automatic access to a large number of relevant tweets/comments [15]. The data crawling process involved extracting tweets/comments containing the keyword "clash of champions" and several other keyword variations related to the clash of champions' event. The data crawling period used is from June 29, 2024 to August 18, 2024, where the range of determining this period is taken from the first episode of the clash of champions event program starts until the final episode stage. This step is taken to collect representative and effective data so that the dynamics of public opinion will be obtained after going through many series of episodes that have been aired.

2.2. Data Preprocessing

In the next stage, data preprocessing, various steps are taken to clean and prepare the text data so that it is ready for further analysis. The first step is to eliminate unnecessary characters, such as mentions, hashtags, RTs, URLs, HTML, numbers, emojis, symbols, punctuation marks, new lines, and excessive spaces [16], with the aim of removing noise so that the text is more focused on relevant content. Next, the

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

process of case folding involves changing all of the text to lowercase for consistency, avoiding meaningless differences in writing [17]. Non-standard words are then normalized to standard form using a specialized dictionary of slang language [18]. After that, a tokenization process, which serves to separate sentences into word units or tokens, is applied for word-based analysis. Common words that don't have much significance are called stopwords, are removed to make the keywords more prominent [19], then stemming is performed to return the terms to their most basic form so that variations in word form do not affect the analysis results [20]. Finally, the text was translated into English as needed for further analysis [21]. These steps aim to optimize the quality of the text data so that it is more easily understood by the learning algorithm.

2.3. Labeling

The labeling stage is performed to classify text data based on sentiment using two main methods, namely TextBlob and VADER lexicon. First, the TextBlob method is applied to identify the sentiment polarity of each text. TextBlob provides two main analysis metrics, namely polarity and subjectivity. Positive numbers indicate positive emotion, negative values indicate negative sentiment, and values near 0 are regarded as neutral. Polarity scores range from -1 to 1 [22]. The second method used is the VADER lexicon which calculates sentiment scores based on four categories: compound, neutral, negative, and positive. For every paragraph, this compound score which combines positive, negative, and neutral scores is normalized between -1 and 1, with values above 0 denoting positive sentiment, values below 0 denoting negative emotion, and values below 0 denoting neutral sentiment [23].

2.4. Class Balancing

The next stage is class balancing to overcome the imbalance of sentiment data that can affect model performance. The two methods used are without balancing and resampling with random undersampling and SMOTE techniques. In the no-balancing method, data is used according to the results of the labeling process from the TextBlob and VADER methods without changing the original distribution of data between sentiment classes. This aims to assess the model's performance on the data's initial distribution. Data from the majority class is used in the random undersampling technique is randomly removed to balance the number of classes with the minority class [24].

SMOTE randomly selects data from a minority class and searches for the k closest data within the same class to generate new samples [25]. Since SMOTE is often used to numerical data, vectorization is necessary to produce a numerical representation of each word as a feature for usage in text datasets [26]. Compared to the Bag of Words (BoW) method, SMOTE feature extraction utilizing the Term Frequency-Inverse Document Frequency (TF-IDF) technique yields higher

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

classification accuracy [27]. The TF-IDF method extracts the weight of each term in the text data, thus improving the learning model's performance [28]. TF-IDF calculation consists of two components, namely TF and IDF, each of which is calculated separately before being combined as shown in Equation 1 to 3.

$$TF(t,d)= \frac{N_t}{N(T,d)} \quad (1)$$

TF is the frequency of occurrence of a word in a document, where t is the term whose frequency is calculated, d is the document containing the term, nt is the number of occurrences of term t in document d, and N (T, d) is the total number of terms in document d.

$$IDF=log\frac{D}{n_d} \quad (2)$$

IDF measures how important a term is in the corpus, where D is the total number of documents in the corpus, and n is the number of documents containing the term t.

$$TF\text{-}IDF=TF* IDF \quad (3)$$

The TF-IDF value is calculated as the product of TF and IDF values.

2.5. Split Data

Data sharing is done in a ratio of 80:20, where 80% is used to train the model and 20% for validation. This division prevents overfitting and ensures the model can generalize well. Test data is important to assess the model's ability to classify new data [29]. This consistent split is applied to each model scenario to ensure accurate evaluation and balanced comparison.

2.6. Classification Models

This research implements two sentiment classification models with and without resampling techniques. The first model, LSTM, is chosen for its effectiveness in handling sequential data, especially in natural language processing tasks. LSTM is a type of Recurrent Neural Network (RNN) that replaces the traditional RNN nodes with LSTM cells, allowing it to store information through three gates: input, forget, and output gates. This ability enables LSTM to learn long-term dependencies in the data, which is crucial for understanding the context of sentiment in longer sentences or documents [30].

The second model, BERT, is selected for its state-of-the-art performance in natural language understanding tasks, particularly in sentiment analysis. BERT is a

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

transformer-based architecture developed by Google, designed to better understand the context of words in a sentence by considering both the preceding and succeeding words in its bidirectional training. This gives BERT an edge in capturing complex relationships between words and sentiments. BERT was trained on large-scale corpora, including Wikipedia, making it particularly powerful in understanding the nuanced meanings and context in text. Given its superior performance in various NLP benchmarks, BERT was chosen to compare against LSTM as a modern deep learning alternative for sentiment analysis. The two versions of BERT, BERTBASE and BERTLARGE, differ in their layers and parameter size, with BERTBASE (12 layers and 110 million parameters) being chosen for this study due to its computational efficiency while still achieving excellent results on NLP tasks [31].

2.7. Evaluation

In the evaluation stage, this research uses confusion matrix to measure the accuracy of the classifier. This method allows determining the accuracy, specificity, and sensitivity of the formed classes. Confusion Matrix is a matrix of size $N \times N$, where N is the number of classes in the classification data, with the rows and columns of the matrix representing the actual number of classes [32]. The results will produce important metrics such as Precision, Recall, Accuracy, and F1-Score [33] which are obtained using Equations (4), (5), (6), and (7).

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (5)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (6)$$

$$F1\text{-}Score = \frac{2 \times Precision \times Recall}{Precision+Recall} \times 100\% \quad (7)$$

The evaluation metrics of the classification models include TP (True Positive) which indicates a correct classification, TN (True Negative) which indicates a wrong classification, FP (False Positive) which indicates a correct classification but is actually wrong, and FN (False Negative) which indicates a wrong classification even though it should be correct. After calculating the evaluation metrics for each model, the results are compared to assess their respective performance and effectiveness in sentiment classification. This process aims to identify the most optimal method in achieving the best performance in sentiment analysis.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

3. RESULTS AND DISCUSSION

3.1 Data Crawling

The data crawling process from X produces 1891 comment data that has been filtered and saved with the .csv extension so that clean data is obtained and ready to use. Table 1 is an example of 5 crawling data results that have been filtered and cleaned.

Table 1. Data Crawling

username	full_text
herewithm333	seneng banget dengan adanya COC ini orang tua jadi termotivasi buat lebih perhatian sama akademik anak²nya thank u tim @ruangguru #ClashofChampions https://t.co/Emr0mAjTCX
sadgurl1257	jujur kalo diliat dari game pertama kebanyakan matematika nya, dan peserta di coc yang unggul di math ya axel sama sandy, jujur kasian sama tim yang lain yang bukan basic di matematika.. kebetulan mereka juga dari berbagai jurusan kan
sskrtyss	GILAAAAA Keren bangetttttt COC big appreciation to tim ruangguru karena sudah memproduksi Clash of Champions. Jujur yah as a watcher aku termotivasi sesuai sm tujuannya kalian juga membentuk tim COC ini. Dan seperti apa yg ditayangkan diakhir the season 2 akan dinantikan
ikramchannel9840	ngak pintar2 amat. jadi mapres karena rajin kuliah ajah. standar sih otaknya😂😂. maaf bukan sombong tapi emang gitu😂🙏🙏. Soaly mudah bangat tpi menyelesainya lama bangat🙏🙏
MRianHardianto	@cacingsauruz Ruangguru Clash of Champions Episode 5 https://t.co/FRVmLGVgay

3.2. Data Preprocessing

The data preprocessing process is supported by several libraries to assist in data cleaning. Some of the libraries used include RegEx for the filtering process using Regular Expression, pandas for data manipulation and analysis, Natural Language Toolkit (NLTK) for tokenization and stopwords removal, Sastrawi for the Indonesian stemming process, and googletrans to translate text from Indonesian to English.

The next stage is to perform the cleaning process, which is cleaning the data in the full_text column from the crawling results to ensure the consistency of the same sentence. Table 2 shows an example of data that has been cleaned in this stage.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

Table 2. Cleaning

Before Cleaning	After Cleaning
seneng banget dengan adanya COC ini orang tua jadi termotivasi buat lebih perhatian sama akademik anak^{2}nya thank u tim @ruangguru #ClashofChampions https://t.co/Emr0mAjTCX	seneng banget dengan adanya COC ini orang tua jadi termotivasi buat lebih perhatian sama akademik anak anaknya thank u tim
jujur kalo diliat dari game pertama kebanyakan matematika nya, dan peserta di coc yang unggul di math ya axel sama sandy, jujur kasian sama tim yang lain yang bukan basic di matematika.. kebetulan mereka juga dari berbagai jurusan kan	jujur kalo diliat dari game pertama kebanyakan matematika nya dan peserta di coc yang unggul di math ya axel sama sandy jujur kasian sama tim yang lain yang bukan basic di matematika kebetulan mereka juga dari berbagai jurusan kan
GILAAAAA Keren bangetttttt COC big appreciation to tim ruangguru karena sudah memproduksi Clash of Champions. Jujur yah as a watcher aku termotivasi sesuai sm tujuannya kalian juga membentuk tim COC ini. Dan seperti apa yg ditayangkan diakhir the season 2 akan dinantikan	GILAAAAA Keren bangetttttt COC big appreciation to tim ruangguru karena sudah memproduksi Clash of Champions Jujur yah as a watcher aku termotivasi sesuai sm tujuannya kalian juga membentuk tim COC ini Dan seperti apa yg ditayangkan diakhir the season akan dinantikan
ngak pintar2 amat. jadi mapres karena rajin kuliah ajah. standar sih otaknya😂😂. maaf bukan sombong tapi emang gitu😂🙏🙏. Soaly mudah bangat tpi menyelesainya lama bangat🙏🙏	ngak pintar amat jadi mapres karena rajin kuliah ajah standar sih otaknya maaf bukan sombong tapi emang gitu Soaly mudah bangat tpi menyelesainya lama bangat
@cacingsauruz Ruangguru Clash of Champions Episode 5 https://t.co/FRVmLGVgay	Ruangguru Clash of Champions Episode

The next stage is case folding, which is the process of converting all letters into lowercase letters. Table 3 shows an example of data results after case folding.

Table 3. Case Folding

Before Case folding	After Case folding
seneng banget dengan adanya COC ini orang tua jadi termotivasi buat lebih perhatian sama akademik anak anaknya thank u tim	seneng banget dengan adanya coc ini orang tua jadi termotivasi buat lebih perhatian sama akademik anak anaknya thank u tim
jujur kalo diliat dari game pertama kebanyakan matematika nya dan peserta di coc yang unggul di math ya axel sama sandy jujur kasian sama tim yang lain yang bukan basic di matematika kebetulan mereka juga dari berbagai jurusan kan	jujur kalo diliat dari game pertama kebanyakan matematika nya dan peserta di coc yang unggul di math ya axel sama sandy jujur kasian sama tim yang lain yang bukan basic di matematika kebetulan mereka juga dari

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Before Case folding	**After Case folding**
	berbagai jurusan kan
GILAAAAA Keren bangetttttt COC big appreciation to tim ruangguru karena sudah memproduksi Clash of Champions Jujur yah as a watcher aku termotivasi sesuai sm tujuannya kalian juga membentuk tim COC ini Dan seperti apa yg ditayangkan diakhir the season akan dinantikan	gilaaaaa keren bangetttttt coc big appreciation to tim ruangguru karena sudah memproduksi clash of champions jujur yah as a watcher aku termotivasi sesuai sm tujuannya kalian juga membentuk tim coc ini dan seperti apa yg ditayangkan diakhir the season akan dinantikan
ngak pintar amat jadi mapres karena rajin kuliah ajah standar sih otaknya maaf bukan sombong tapi emang gitu Soaly mudah bangat tpi menyelesainya lama bangat	ngak pintar amat jadi mapres karena rajin kuliah ajah standar sih otaknya maaf bukan sombong tapi emang gitu soaly mudah bangat tpi menyelesainya lama bangat
Ruangguru Clash of Champions Episode	ruangguru clash of champions episode

The next stage is normalization using a dictionary from Kaggle which contains 15,184 standard and non-standard word pairs. However, after reading and analyzing the dataset, it was found that many words were not listed in the dictionary. Therefore, 253 customized words were added to the dataset, bringing the total in the dictionary to 15,437 lines. Table 4 shows an example of data results after normalization.

Table 4. Normalization Result

Before Normalization	**After Normalization**
seneng banget dengan adanya coc ini orang tua jadi termotivasi buat lebih perhatian sama akademik anak anaknya thank u tim	senang banget dengan adanya coc ini orang tua jadi termotivasi buat lebih perhatian sama akademik anak anaknya thank lu tim
jujur kalo diliat dari game pertama kebanyakan matematika nya dan peserta di coc yang unggul di math ya axel sama sandy jujur kasian sama tim yang lain yang bukan basic di matematika kebetulan mereka juga dari berbagai jurusan kan	jujur kalau dilihat dari game pertama kebanyakan matematika ya dan peserta di coc yang unggul di matematika ya axel sama sandy jujur kasihan sama tim yang lain yang bukan basic di matematika kebetulan mereka juga dari berbagai jurusan kan
gilaaaaa keren bangetttttt coc big appreciation to tim ruangguru karena sudah memproduksi clash of champions jujur yah as a watcher aku termotivasi sesuai sm tujuannya kalian juga membentuk tim coc ini dan seperti apa yg ditayangkan diakhir the season akan dinantikan	gila keren banget coc big appreciation tapi tim ruangguru karena sudah memproduksi clash of champions jujur ya as a watcher aku termotivasi sesuai sama tujuannya kalian juga membentuk tim coc ini dan seperti apa yang ditayangkan diakhir the season akan dinantikan

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Before Normalization	**After Normalization**
ngak pintar amat jadi mapres karena rajin kuliah ajah standar sih otaknya maaf bukan sombong tapi emang gitu soaly mudah bangat tpi menyelesainya lama bangat	tidak pintar sekali jadi mapres karena rajin kuliah saja standar sih otaknya maaf bukan sombong tapi memang begitu soalnya mudah banget tapi menyelesainya lama banget
ruangguru clash of champions episode	ruangguru clash of champions episode

The next step is tokenization. Table 5 shows an example of data that has gone through the tokenization process.

Table 5. Tokenization Result

Before Tokenization	**After Tokenization**
senang banget dengan adanya coc ini orang tua jadi termotivasi buat lebih perhatian sama akademik anak anaknya thank lu tim	['senang', 'banget', 'dengan', 'adanya', 'coc', 'ini', 'orang', 'tua', 'jadi', 'termotivasi', 'buat', 'lebih', 'perhatian', 'sama', 'akademik', 'anak', 'anaknya', 'thank', 'lu', 'tim']
jujur kalau dilihat dari game pertama kebanyakan matematika ya dan peserta di coc yang unggul di matematika ya axel sama sandy jujur kasihan sama tim yang lain yang bukan basic di matematika kebetulan mereka juga dari berbagai jurusan kan	['jujur', 'kalau', 'dilihat', 'dari', 'game', 'pertama', 'kebanyakan', 'matematika', 'ya', 'dan', 'peserta', 'di', 'coc', 'yang', 'unggul', 'di', 'matematika', 'ya', 'axel', 'sama', 'sandy', 'jujur', 'kasihan', 'sama', 'tim', 'yang', 'lain', 'yang', 'bukan', 'basic', 'di', 'matematika', 'kebetulan', 'mereka', 'juga', 'dari', 'berbagai', 'jurusan', 'kan']
gila keren banget coc big appreciation tapi tim ruangguru karena sudah memproduksi clash of champions jujur ya as a watcher aku termotivasi sesuai sama tujuannya kalian juga membentuk tim coc ini dan seperti apa yang ditayangkan diakhir the season akan dinantikan	['gila', 'keren', 'banget', 'coc', 'big', 'appreciation', 'tapi', 'tim', 'ruangguru', 'karena', 'sudah', 'memproduksi', 'clash', 'of', 'champions', 'jujur', 'ya', 'as', 'a', 'watcher', 'aku', 'termotivasi', 'sesuai', 'sama', 'tujuannya', 'kalian', 'juga', 'membentuk', 'tim', 'coc', 'ini', 'dan', 'seperti', 'apa', 'yang', 'ditayangkan', 'diakhir', 'the', 'season', 'akan', 'dinantikan']
tidak pintar sekali jadi mapres karena rajin kuliah saja standar sih otaknya maaf bukan sombong tapi memang begitu soalnya mudah banget tapi menyelesainya lama banget	['tidak', 'pintar', 'sekali', 'jadi', 'mapres', 'karena', 'rajin', 'kuliah', 'saja', 'standar', 'sih', 'otaknya', 'maaf', 'bukan', 'sombong', 'tapi', 'memang', 'begitu', 'soalnya', 'mudah', 'banget', 'tapi', 'menyelesainya', 'lama', 'banget']
ruangguru clash of champions episode	['ruangguru', 'clash', 'of', 'champions', 'episode']

The next step is to perform word removal with Stopword Removal. Table 6 shows an example of data results after stopwords.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Table 6. Stopwords Result

Before Stopwords	After Stopwords
['senang', 'banget', 'dengan', 'adanya', 'coc', 'ini', 'orang', 'tua', 'jadi', 'termotivasi', 'buat', 'lebih', 'perhatian', 'sama', 'akademik', 'anak', 'anaknya', 'thank', 'lu', 'tim']	['senang', 'banget', 'coc', 'orang', 'tua', 'termotivasi', 'perhatian', 'akademik', 'anak', 'anaknya', 'thank', 'lu', 'tim']
['jujur', 'kalau', 'dilihat', 'dari', 'game', 'pertama', 'kebanyakan', 'matematika', 'ya', 'dan', 'peserta', 'di', 'coc', 'yang', 'unggul', 'di', 'matematika', 'ya', 'axel', 'sama', 'sandy', 'jujur', 'kasihan', 'sama', 'tim', 'yang', 'lain', 'yang', 'bukan', 'basic', 'di', 'matematika', 'kebetulan', 'mereka', 'juga', 'dari', 'berbagai', 'jurusan', 'kan']	['jujur', 'game', 'kebanyakan', 'matematika', 'ya', 'peserta', 'coc', 'unggul', 'matematika', 'ya', 'axel', 'sandy', 'jujur', 'kasihan', 'tim', 'basic', 'matematika', 'jurusan']
['gila', 'keren', 'banget', 'coc', 'big', 'appreciation', 'tapi', 'tim', 'ruangguru', 'karena', 'sudah', 'memproduksi', 'clash', 'of', 'champions', 'jujur', 'ya', 'as', 'a', 'watcher', 'aku', 'termotivasi', 'sesuai', 'sama', 'tujuannya', 'kalian', 'juga', 'membentuk', 'tim', 'coc', 'ini', 'dan', 'seperti', 'apa', 'yang', 'ditayangkan', 'diakhir', 'the', 'season', 'akan', 'dinantikan']	['gila', 'keren', 'banget', 'coc', 'big', 'appreciation', 'tim', 'ruangguru', 'memproduksi', 'clash', 'of', 'champions', 'jujur', 'ya', 'as', 'a', 'watcher', 'termotivasi', 'sesuai', 'tujuannya', 'membentuk', 'tim', 'coc', 'ditayangkan', 'diakhir', 'the', 'season', 'dinantikan']
['tidak', 'pintar', 'sekali', 'jadi', 'mapres', 'karena', 'rajin', 'kuliah', 'saja', 'standar', 'sih', 'otaknya', 'maaf', 'bukan', 'sombong', 'tapi', 'memang', 'begitu', 'soalnya', 'mudah', 'banget', 'tapi', 'menyelesainya', 'lama', 'banget']	['pintar', 'mapres', 'rajin', 'kuliah', 'standar', 'sih', 'otaknya', 'maaf', 'sombong', 'mudah', 'banget', 'menyelesainya', 'banget']
['ruangguru', 'clash', 'of', 'champions', 'episode']	['ruangguru', 'clash', 'of', 'champions', 'episode']

The next step is stemming. Table 7 shows an example of data results after stemming.

Table 7. Stemming Result

Before Stemming	After Stemming
['senang', 'banget', 'coc', 'orang', 'tua', 'termotivasi', 'perhatian', 'akademik', 'anak', 'anaknya', 'thank', 'lu', 'tim']	senang banget coc orang tua motivasi perhati akademik anak anak thank lu tim
['jujur', 'game', 'kebanyakan', 'matematika', 'ya', 'peserta', 'coc', 'unggul', 'matematika', 'ya', 'axel', 'sandy', 'jujur', 'kasihan', 'tim', 'basic', 'matematika', 'jurusan']	jujur game banyak matematika ya serta coc unggul matematika ya axel sandy jujur kasihan tim basic matematika jurus
['gila', 'keren', 'banget', 'coc', 'big', 'appreciation', 'tim', 'ruangguru', 'memproduksi', 'clash', 'of', 'champions', 'jujur', 'ya', 'as', 'a', 'watcher', 'termotivasi', 'sesuai', 'tujuannya', 'membentuk', 'tim', 'coc', 'ditayangkan', 'diakhir', 'the', 'season',	gila keren banget coc big appreciation tim ruangguru produksi clash of champions jujur ya as a watcher motivasi sesuai tuju bentuk tim coc tayang akhir the

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Before Stemming	**After Stemming**
'dinantikan']	season nanti
['pintar', 'mapres', 'rajin', 'kuliah', 'standar', 'sih', 'otaknya', 'maaf', 'sombong', 'mudah', 'banget', 'menyelesainya', 'banget']	pintar mapres rajin kuliah standar sih otak maaf sombong mudah banget selesa banget
['ruangguru', 'clash', 'of', 'champions', 'episode']	ruangguru clash of champions episode

The last step is to translate the stemming results from Indonesian to English. Table 8 shows an example of data results after translation.

Table 8. Translate Result

Before Translate	**After Translate**
senang banget coc orang tua motivasi perhati akademik anak anak thank lu tim	I'm really happy that parents are motivated to pay attention to their children's academics thank you team
jujur game banyak matematika ya serta coc unggul matematika ya axel sandy jujur kasihan tim basic matematika jurus	Honestly the game has a lot of mathematics yes and COC is superior in mathematics Axel Sandy honestly I feel sorry for the basic math skills team
gila keren banget coc big appreciation tim ruangguru produksi clash of champions jujur ya as a watcher motivasi sesuai tuju bentuk tim coc tayang akhir the season nanti	crazy really cool COC big appreciation Ruangguru team production of Clash of Champions honestly as a watcher motivation according to the aim of forming the COC team to show at the end of the season
pintar mapres rajin kuliah standar sih otak maaf sombong mudah banget selesa banget	smart mapres diligent in standard studies sorry brain arrogant very easy very comfortable
ruangguru clash of champions episode	Ruangguru clash of champions episode

Preprocessing begins with cleaning the full_text column to remove unwanted elements. In the case folding process, all letters are converted to lowercase. Nonstandard words are normalized using a custom dictionary. Tokenization separates words in each sentence, followed by stopwords removal to eliminate meaningless words. Stemming reduces words to their base form, and finally, the text is translated from Indonesian to English using googletrans for further analysis.

3.3. Labeling

The labeling process uses 2 automatic labeling methods, namely TextBlob and VADER, which produce 3 sentiment classes, namely positive, neutral, and negative. Table 9 shows an example of data labeling results with TextBlob.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

Table 9. TextBlob Labeling Result

Comments	Polarity	Labels
I'm really happy that parents are motivated to pay attention to their children's academics thank you team	0.8	Positive
look at the child's report card	0.0	Neutral
Admin's suggestion is that the format is like a master chef event with lots of comments and game composition comments and it's not fun to understand how to teach the game	-0.3167	Negative

An example of labeling results with VADER can be seen in Table 10 below.

Table 10. VADER Labeling Result

Comments	Compound	Labels
I'm really happy that parents are motivated to pay attention to their children's academics thank you team	0.8439	Positive
look at the child's report card	0.0	Neutral
Admin's suggestion is that the format is like a master chef event with lots of comments and game composition comments and it's not fun to understand how to teach the game	-0.0521	Negative

3.4. Class Balancing

To analyze the effect of class balancing on the amount of data in each sentiment category, the distribution of labeling results using TextBlob and VADER is shown in Table 11, both before and after the resampling process.

Table 11. Class Balancing Distribution

Labeling Method	Sentiment Class	Without Balancing	Undersampling	SMOTE
TextBlob	Positive	1030	290	1012
TextBlob	Neutral	569	290	1012
TextBlob	Negative	290	290	1012
VADER	Positive	1498	182	1458
VADER	Neutral	209	182	1458
VADER	Negative	182	182	1458

The table above displays the distribution of data before and after the application of the class balancing technique. In the initial condition without balancing, there is an imbalance in the amount of data between sentiment classes, especially in the majority class. To overcome this, two balancing techniques were applied: random undersampling and SMOTE. Random undersampling equalizes the amount of data in each class based on the minority class, while SMOTE increases the minority class data to be equal to the majority class. With these two techniques, it is expected that the classification model can provide more balanced results on

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

each sentiment class.

In addition, to support resampling techniques such as SMOTE, a numerical representation of the text data is required. One of the techniques used to generate numerical representation is TF-IDF. For example, the following is a TF-IDF calculation for three documents containing responses to the "Clash of Champions" event:

Data 1 : clash of champions is really exciting
Data 2 : click clash of champions scroll
Data 3 : damn watch clash of champions insecure

The following is the calculation of TF, IDF, and TF-IDF values for some words in the document:

Table 12. TF, IDF, and TF-IDF Calculation Results

Word	TF1	TF2	TF3	IDF	TF-IDF1	TF-IDF2	TF-IDF3
clash	0.17	0.2	0.17	0	0	0	0
of	0.17	0.2	0.17	0	0	0	0
champions	0.17	0.2	0.17	0	0	0	0
is	0.17	0	0	0.48	0.08	0	0
really	0.17	0	0	0.48	0.08	0	0
exciting	0.17	0	0	0.48	0.08	0	0
click	0	0.2	0	0.48	0	0.1	0
scroll	0	0.2	0	0.48	0	0.1	0
damn	0	0	0.17	0.48	0	0	0.08
watch	0	0	0.17	0.48	0	0	0.08
insecure	0	0	0.17	0.48	0	0	0.08

In Table 12 for example, for the word "clash", the TF value for Data 1 is 0.17, indicating the frequency of the word in the document. The IDF value for "clash" is 0 because this word is present in all documents, which means it does not provide any distinguishing information between documents. The TF-IDF result is calculated by multiplying TF and IDF, resulting in a value of 0 for Data 1.

3.5. Split Data

From the split data results with a division ratio of 80% training data and 20% validation data, the distribution results are obtained in Table 13.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Table 13. Split Data Distribution Table

Labeling Method	Total Data	Training (80%)	Validation (20%)
TextBlob (Without Balancing)	1889	1511	378
TextBlob (Undersampling)	870	696	174
TextBlob (SMOTE)	3036	2428	608
VADER (Without Balancing)	1889	1511	378
VADER (Undersampling)	546	436	110
VADER (SMOTE)	4374	3499	875

3.6. Classification Model

This research uses LSTM and BERT, the results of both methods can be seen below. First, using the LSTM model by performing hyperparameter tuning as shown in table 14.

Table 14. Hyperparameters of the LSTM Model

Hyperparameter	Value
Model Architecture	LSTM
Embedding Dimension	128
LSTM Units	128
Dropout Layer	SpatialDropout1D
Loss Function	Sparse Categorical Crossentropy
Optimizer	Adam
Epochs	5

The LSTM model uses an embedding dimension of 128 and 128 LSTM units to capture text information effectively. SpatialDropout1D is applied to reduce overfitting. The loss function used is Sparse Categorical Crossentropy, chosen for its ability to automatically adjust the learning rate, thus speeding up the training process. The model was trained for 5 epochs to achieve convergence.

Table 15. LSTM Model Scenario Results

Scenario	Precision	Recall	F1-score	Accuracy
TextBlob + LSTM	84%	76%	79%	83.33%
TextBlob + Undersampling + LSTM	75%	75%	75%	75.29%
TextBlob + SMOTE + LSTM	93%	92%	92%	92.11%
VADER + LSTM	77%	70%	73%	87.83%
VADER + Undersampling + LSTM	66%	64%	65%	65.45%
VADER + SMOTE + LSTM	96%	96%	96%	96.23%

Table 15 shows the classification results using the LSTM model with a combination of labeling methods and class balancing techniques. The highest result is obtained in the combination of VADER + SMOTE + LSTM with an

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Accuracy value of 96.23%, while the lowest result is in the combination of VADER + Undersampling + LSTM with an Accuracy of 65.45%. The second is using the BERT model with hyperparameter tuning as shown in Table 16.

Table 16. Hyperparameters of the BERT Model

Hyperparameter	Value
Model Architecture	BERT (bert-base-uncased)
Maximum Sequence Length	256
Optimizer	BertAdam
Learning Rate	3e-5
Weight Decay	0.01
Batch Size	16
Epochs	5

The BERT model uses bert-base-uncased with a maximum sequence length of 256 tokens. Optimizer BertAdam and learning rate 3e-5 were chosen for training stability, with weight decay 0.01 to avoid overfitting. A batch size of 16 and number of epochs of 5 were chosen for training efficiency.

Table 17. BERT Model Scenario Results

Scenario	Precision	Recall	F1-score	Accuracy
TextBlob + BERT	90%	82%	85%	87.71%
TextBlob + Undersampling + BERT	88%	86%	87%	86.77%
TextBlob + SMOTE + BERT	94%	94%	94%	94.24%
VADER + BERT	85%	69%	74%	88.49%
VADER + Undersampling + BERT	90%	90%	90%	92.22%
VADER + SMOTE +BERT	98%	98%	98%	97.73%

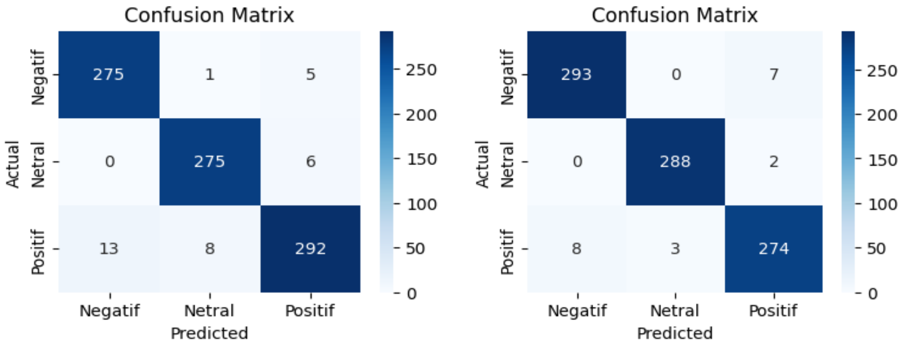
Table 17 above shows the classification results using the BERT model for various scenarios of combinations of labeling methods and class balancing techniques. The highest result is obtained in the combination of VADER + SMOTE + BERT with an Accuracy value of 97.73%, while the lowest result is in the combination of TextBlob + Undersampling + BERT with an Accuracy of 86.77%.

3.7. Evaluation

The LSTM model (Figure 2a) shows fairly good classification results in each class, with a total of 275 correct predictions for the Negative class, 275 for the Neutral class, and 292 for the Positive class. In the Negative class, there was a small amount of misclassification, with 6 incorrect predictions (5 classified as Positive and 1 as Neutral). The Neutral class also had the same number of errors, with 6 incorrect predictions where all of them were classified to the Positive class.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Meanwhile, the Positive class showed the highest number of correct predictions, 292, but there were still 21 errors (13 classified as Negative and 8 as Neutral). Although the LSTM is able to capture the data pattern well, it still has limitations on the Positive class, which requires improvement to optimize prediction accuracy across all classes. The following are the best confusion matrix results from each scenario for both models.



(a) (b)

Figure 2. Model confusion matrix results: (a) LSTM, (b) BERT

The BERT model (Figure 2b), provides more consistent and accurate results in each class. In the Negative class, the BERT model made 293 correct predictions, with 7 mispredictions, all of which were classified into the Positive class. The Neutral class also showed high performance with 288 correct predictions and only 2 errors, both of which were classified to the Positive class. In the Positive class, the BERT model recorded 274 correct predictions, with 11 errors of which 8 were classified as Negative and 3 as Neutral. From these results, it can be seen that the BERT model has the advantage of handling unbalanced data distributions, as well as providing more accurate classification results in each class than the LSTM. This shows that transformer-based approaches, such as BERT, are very effective in better distinguishing sentiment variations, making it a highly recommended choice for complex sentiment analysis.

The results shown in Table 18 illustrate that the overall model performance improves when the data distribution in each class is more balanced, especially in the BERT model with the VADER + SMOTE + BERT scenario, which achieved the highest accuracy of 98%. In this best scenario of the BERT model, each class has very high and consistent Precision, Recall, and F1-score values. The Neutral class, for example, achieved Precision, Recall, and F1-score of 99% each with 290 samples as support. For the Positive and Negative classes, BERT also performed well with F1-score of 96% and 98%, respectively, demonstrating BERT's ability to handle data with a more balanced class distribution. The

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

following are the best classification report results from each scenario for both models.

Table 18. Classification report results of the best scenario of each model

Scenario	Class	Precision	Recall	F1-score	Support	Accuracy
	Positive	96%	93%	95%	313	
VADER + SMOTE	Neutral	97%	98%	97%	281	96%
+ LSTM	Negative	95%	98%	97%	281	
	Positive	97%	96%	96%	285	
VADER + SMOTE	Neutral	99%	99%	99%	290	98%
+ BERT	Negative	97%	98%	98%	300	

On the other hand, in the VADER + SMOTE + LSTM scenario, the LSTM model also performed quite well with 96% accuracy. In this scenario, the Neutral and Negative classes had an F1-score of 97%, while the Positive class had an F1-score of 95%. This improvement indicates that data balancing through the SMOTE method plays an important role in reducing bias towards the majority class and improving the model's ability to recognize patterns in each class, including the minority class.

3.8. Discussion

This study aimed to evaluate the impact of labelling methods, class balancing techniques approaches on the performance of sentiment analysis models for the Clash of Champions show. The results indicate that the BERT model, particularly in the VADER + SMOTE scenario, achieved the highest accuracy of 97.73%, outperforming other models across all metrics. This confirms the effectiveness of transformer-based models in capturing complex linguistic and contextual nuances, making them suitable for sentiment analysis tasks involving varied and ambiguous expressions.

Among the different scenarios, models utilizing SMOTE for class balancing consistently outperformed those without balancing, demonstrating the importance of addressing class imbalances in sentiment analysis. Interestingly, while the LSTM model demonstrated robust performance with balanced data, its confusion matrix revealed challenges in differentiating Neutral and Positive sentiments. This pattern suggests that LSTM may struggle with contextually ambiguous phrases, as it relies heavily on sequential patterns rather than bidirectional contextual understanding. In comparison, BERT showed fewer misclassifications across all sentiment classes, underscoring its ability to generalize well across diverse scenarios.

Analysis of the confusion matrix provides valuable insights into potential areas

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

for improvement. Misclassifications in the Neutral class, particularly in models without SMOTE, highlight the impact of class imbalances on prediction reliability. These findings suggest that future research could focus on exploring alternative balancing techniques, such as ADASYN, to further refine class distributions, as well as integrating hybrid approaches that combine the strengths of different modelling techniques to enhance overall robustness.

The findings have significant implications for real-world applications. In scenarios where high accuracy is critical, such as sentiment analysis for social media campaigns, adopting advanced models like BERT alongside effective preprocessing techniques can lead to superior results. Future research could explore the use of domain-specific pretraining for transformer models to better align with the sentiment nuances of specific contexts, testing the models on larger and more diverse datasets to examine their generalizability, and investigating the integration of ensemble techniques to combine the strengths of multiple modelling approaches.

4. CONCLUSION

Based on the sentiment analysis of the Clash of Champions event by Ruangguru through social media X, the results show that the sentiment analysis of 1,891 tweets that initially had an unbalanced class distribution achieved the best approach with the BERT model, using the VADER labeling method and SMOTE technique for class balancing. Preprocessing steps included cleaning, case folding, normalization, tokenization, stopword removal, stemming, and translation to English to ensure text compatibility with labeling methods using TextBlob and VADER, which are designed for English. The BERT model produced the highest accuracy of 97.73% with a consistent F1-score across all classes (positive, neutral, and negative), demonstrating a good ability to handle sentiment variation after the data distribution was balanced using SMOTE. The use of the SMOTE technique proved to provide better results compared to undersampling and no data balancing, where the combination of SMOTE with VADER resulted in the highest performance across all models, both LSTM and BERT. In addition, the analysis also shows that VADER is more effective than TextBlob in classifying sentiment on this dataset, with higher accuracy and F1-score. The distribution of the analysis results indicates that the majority of sentiments are positive, showing the dominance of public enthusiasm for the event, while negative sentiments contain criticisms and suggestions that are useful for the development of the event. Compared to the LSTM model, BERT proved to be more effective in identifying complex sentiment variations on social media. This research provides insight for Ruangguru in understanding public perceptions more accurately, so that it can help decision-making for the improvement of similar events in the future. For future research, further optimization of the preprocessing stage, refinement of the classification model,

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

and experimentation with other hyperparameter tuning techniques are expected to improve the quality of sentiment analysis in the future.

REFERENCES

[1] S. Adha, "Faktor Revolusi Perilaku Konsumen Era Digital : Sebuah Tinjauan Literatur," *Jipis*, vol. 31, no. 2, pp. 134–148, 2022, doi: 10.33592/jipis.v31i2.3286.
[2] S. A. Zaxrie, N. Rina, S. Thoibah, K. Putri, and M. Sosial, "Peran Media Sosial Sebagai Saluran Komunikasi Digital Dalam Kehumasan The Role of Social Media As A Digital," vol. 3, no. 1, pp. 26–39, 2024.
[3] A. P. Anggraini and F. U. Najicha, "Pengembangan Wawasan Nusantara Sebagai Muatan Pendidikan Kewarganegaraan Generasi Muda Melalui Pemanfaatan Internet," *J. Pendidik. Ilmu Pengetah. Sos.*, vol. 14, no. 1, pp. 174–180, 2022, doi: 10.37304/jpips.v14i1.4747.
[4] F. A. R. Manurung, A. Padriansyah, E. R. A. Sitorus, M. Hasanah, and K. Saleh, "The Effect of Smartphone Use on Academic Achievement and Social Skills of Elementary School Students: Case Study at UPTD SDN 10 Bangun Sari Village," *J. IPTEK Bagi Masy.*, vol. 4, no. 1, 2024.
[5] T. M. Assalamah, "Strategi Viral Marketing Melalui Konten Edutainment Clash of Champions by Ruangguru Viral Marketing Strategy Through Edutainment Content Clash of Champions by Ruangguru," vol. 7, no. 3, doi: 10.1080/019722409032.....154.
[6] K. A. Sari and G. Reftantia, "Hegemoni ' Clash of Champions Ruang Guru ' di Tengah Maraknya Tayangan Non Edukatif," pp. 130–145, 2024.
[7] M. Hudha, E. Supriyati, and T. Listyorini, "Analisis Sentimen Pengguna Youtube Terhadap Tayangan #Matanajwamenantiterawan Dengan Metode Naïve Bayes Classifier," *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 1, pp. 1–6, 2022, doi: 10.33387/jiko.v5i1.3376.
[8] S. Helmiyah and A. Verdian, "Analisis Sentimen Terhadap Minat Belajar pada Tayangan Acara CoC by Ruangguru Berdasarkan Tweets Menggunakan Metode NLP dan Model BERT: Analisis Sentimen Terhadap Minat Belajar pada Tayangan Acara CoC by Ruangguru Berdasarkan Tweets Menggunakan Metode NLP," *J. Pendidik. Rosalia*, vol. 7, no. 2, pp. 138–149, 2024.
[9] W. Maharani, "Sentiment Analysis during Jakarta Flood for Emergency Responses and Situational Awareness in Disaster Management using BERT," *2020 8th Int. Conf. Inf. Commun. Technol. ICoICT 2020*, 2020, doi: 10.1109/ICoICT49345.2020.9166407.
[10] M. Tripathi, "Sentiment Analysis of Nepali COVID19 Tweets Using NB, SVM AND LSTM," *J. Artif. Intell. Capsul. Networks*, vol. 3, no. 3, pp. 151–168, 2021, doi: 10.36548/jaicn.2021.3.001.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[11] B. M. Alenzi, M. B. Khan, M. H. A. Hasanat, A. K. J. Saudagar, M. Alkhathami, and A. Altameem, "Automatic Annotation Performance of TextBlob and VADER on Covid Vaccination Dataset," *Intell. Autom. Soft Comput.*, vol. 34, no. 2, pp. 1311–1331, 2022, doi: 10.32604/iasc.2022.025861.

[12] F. Illia, M. P. Eugenia, and S. A. Rutba, "Sentiment Analysis on PeduliLindungi Application Using TextBlob and VADER Library," *Proc. Int. Conf. Data Sci. Off. Stat.*, vol. 2021, no. 1, pp. 278–288, 2022, doi: 10.34123/icdsos.v2021i1.236.

[13] S. N. Almuayqil, M. Humayun, N. Z. Jhanjhi, M. F. Almufareh, and N. A. Khan, "Enhancing Sentiment Analysis via Random Majority Under-Sampling with Reduced Time Complexity for Classifying Tweet Reviews," *Electron.*, vol. 11, no. 21, 2022, doi: 10.3390/electronics11213624.

[14] E. R. N. Mustaqim, U. Pagalay, and ..., "Prediksi Tingkat Kepercayaan Masyarakat Terhadap Pilpres 2024 Menggunakan Tf-Idf Dan Bow Menggunakan Metode Svm," *Mandalika*, pp. 515–530, 2024.

[15] N. Mardiah, L. Marlina, Z. Sitorus, and M. Iqbal, "Analysis Of Indonesian People ' s Sentiment Towards 2024 Presidential Candidates On Social Media Using Naïve Bayes Classifier and Support Vector Machine," vol. 6, no. 2, pp. 950–960, 2024, doi: 10.47065/bits.v6i2.5766.

[16] E. Priansyah and T. Sutabri, "IJM: Indonesian Journal of Multidisciplinary Analisis Sentimen Berbasis Naïve Bayes Pada Media Sosial Twitter Terhadap Hasil Pemilu Indonesia 2024," *IJM Indones. J. Multidiscip.*, vol. 2, pp. 128–138, 2024.

[17] A. R. Isnain, H. Sulistiani, B. M. Hurohman, A. Nurkholis, and S. Styawati, "Analisis Perbandingan Algoritma LSTM dan Naive Bayes untuk Analisis Sentimen," *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 2, p. 299, 2022, doi: 10.26418/jp.v8i2.54704.

[18] Muhammad Fernanda Naufal Fathoni, Eva Yulia Puspaningrum, and Andreas Nugroho Sihananto, "Perbandingan Performa Labeling Lexicon InSet dan VADER pada Analisa Sentimen Rohingya di Aplikasi X dengan SVM," *Modem J. Inform. dan Sains Teknol.*, vol. 2, no. 3, pp. 62–76, 2024, doi: 10.62951/modem.v2i3.112.

[19] D. B. A. N. Puspita Sari, "Analisis Perbandingan Sentimen Pengguna Twitter Terhadap Layanan Salah Satu Provider Internet Di Indonesia Menggunakan Metode Klasifikasi," *Tematik*, vol. 10, no. 2, pp. 246–251, 2022.

[20] Diana Dwi Rahayu, Muhammad Fatchan, and Alfonsus Ligouri, "Analisis Sentimen Twitter Terpilihnya Prabowo - Gibran Menggunakan Metode Neural Network," *Tematik*, vol. 11, no. 1, pp. 85–91, 2024, doi: 10.38204/tematik.v11i1.1943.

[21] S. Z. Rozaan, M. R. Andrianto, R. S. Purnama, N. I. Ramadhan, and W. Putra, "Analisis Sentimen terhadap Kenaikan UKT di Indonesia pasca Terpilihnya Capres 02 menggunakan VADER," vol. 4, pp. 88–92, 2024.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[22] M. Raees and S. Fazilat, “Lexicon-Based Sentiment Analysis on Text Polarities with Evaluation of Classification Models,” pp. 1–18, 2024.
[23] V. Vinardo and I. Wasito, “Two-Stage Sentiment Analysis on Indonesian Online News Using Lexicon-Based,” *Sinkron*, vol. 8, no. 4, pp. 2109–2119, 2023, doi: 10.33395/sinkron.v8i4.12769.
[24] N. Husin, “Komparasi Algoritma Random Forest, Naïve Bayes, dan Bert Untuk Multi-Class Classification Pada Artikel Cable News Network (CNN),” *J. Esensi Infokom J. Esensi Sist. Inf. dan Sist. Komput.*, vol. 7, no. 1, pp. 75–84, 2023, doi: 10.55886/infokom.v7i1.608.
[25] W. I. Sabilla and C. Bella Vista, “Implementasi SMOTE dan Under Sampling pada Imbalanced Dataset untuk Prediksi Kebangkrutan Perusahaan,” *J. Komput. Terap.*, vol. 7, no. 2, pp. 329–339, 2021, doi: 10.35143/jkt.v7i2.5027.
[26] L. D. Cahya, A. Luthfiarta, J. I. T. Krisna, S. Winarno, and A. Nugraha, “Improving Multi-label Classification Performance on Imbalanced Datasets Through SMOTE Technique and Data Augmentation Using IndoBERT Model,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 3, pp. 290–298, 2024, doi: 10.25077/teknosi.v9i3.2023.290-298.
[27] V. Rupapara, F. Rustam, H. F. Shahzad, A. Mehmood, I. Ashraf, and G. S. Choi, “Impact of SMOTE on Imbalanced Text Features for Toxic Comments Classification Using RVVC Model,” *IEEE Access*, vol. 9, pp. 78621–78634, 2021, doi: 10.1109/ACCESS.2021.3083638.
[28] M. Mujahid *et al.*, “Sentiment analysis and topic modeling on tweets about online education during covid-19,” *Appl. Sci.*, vol. 11, no. 18, 2021, doi: 10.3390/app11188438.
[29] A. Alim Murtopo, M. Aditdya, P. Septiana Ananda, and G. Gunawan, “Penerapan Computer Vision Untuk Mendeteksi Kelengkapan Atribut Siswa Menggunakan Metode Cnn,” *PROSISKO J. Pengemb. Ris. dan Obs. Sist. Komput.*, vol. 11, no. 2, pp. 247–258, 2024, doi: 10.30656/prosisko.v11i2.8752.
[30] D. Melati, Herfina, and Mulyati, “Penerapan Metode Long Short-Term Memory (LSTM) dalam Analisis Sentimen terhadap Pelaksanaan Pilkada di Masa Pandemi COVID-19,” *J. Inform. dan Komput.*, vol. 22, no. 1, pp. 22–28, 2024, doi: 10.35508/jicon.v12i1.9899.
[31] M. P. Geetha and D. Karthika Renuka, “Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model,” *Int. J. Intell. Networks*, vol. 2, no. June, pp. 64–69, 2021, doi: 10.1016/j.ijin.2021.06.005.
[32] N. Puspitasari, A. Septiarini, and A. R. Aliudin, “Metode K-Nearest Neighbor Dan Fitur Warna Untuk Klasifikasi Daun Sirih Berdasarkan Citra Digital,” *PROSISKO J. Pengemb. Ris. dan Obs. Sist. Komput.*, vol. 10, no. 2, pp. 165–172, 2023, doi: 10.30656/prosisko.v10i2.6924.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[33] A. Syahril *et al.*, "Perbandingan Metode Decision Tree Dan K-Nearest Neighbor Terhadap Ulasan Pengguna Aplikasi Mypertamina Menggunakan Confusion Matrix," *J. Inf. Syst. Res.*, vol. 5, no. 4, pp. 1085–1094, 2024, doi: 10.47065/josh.v5i4.5639.