Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 7 No. 3 (2023) 457 - 563 ISSN Media Electronic: 2580-0760

Sentiment Analysis on Social Media with Glove Using Combination CNN and RoBERTa

Diaz Tiyasya Putra[1], Erwin Budi Setiawan[2]
[1,2]Informatics, School of Computing, Telkom University
[1]diaztiyasyaputra@student.telkomuniveristy.ac.id, [2]erwinbudisetiawan@telkomuniversity.ac.id

Abstract

Twitter is a popular social media platform that allows users to share short message's opinion and engage in real-time conversations on a wide range of topics known as tweet. However, tweets often have a complicated and unclear context, which makes it difficult to determine the actual emotion. Therefore, sentiment analysis is required to see the tendency of an opinion, whether the opinion tends to be positive, negative, or neutral. Researchers or institutions can find out how the response and emotions of an issue are happening and make good decisions. With the large user of Twitter social media in Indonesia, sentiment analysis will be carried out using deep learning Convolutional Neural Network (CNN), Term Frequency-Inverse Document Frequency (TF-IDF), Robustly Optimized BERT Pretraining Approach (RoBERTa), Synthetic Minority Over-sampling Technique (SMOTE), and Global Vector (Glove). In this research, the dataset used is trending topics with hashtags related to government policies on Twitter social media and obtained through crawling. By using 30.811 data, the result shows the highest accuracy of 95.56% using CNN with a split ratio of 90:10, baseline unigram, RoBERTa, SMOTE, and Top10 corpus tweet with an increase 10.1%.

Keywords: CNN, Glove, RoBERTa, Sentiment Analysis, Twitter

1. Introduction

The development of social media has increased rapidly. By using the internet as a communication tool, people will find and easy to use it. This is the background of how people communicate from conventional to modern and digital [1]. Social media itself is a platform that is used to communicate with someone or the general public through content in the form of text, video, or photos. Its also someone can express his opinion on the information obtained. Social media has many features for finding information and communicating and will get a lot of attention from its users. According to Wearesocial Hootsuite research in January 2019, social media users in Indonesia reached 150 million, or 56% of the total population. [2]. Social media in Indonesia is diverse, from functions to their respective characteristics. Like Twitter and Facebook, both have advantages in conveying information or opinions, Twitter with the trending topic feature and Facebook with its posting feeds. According to a statista report entitled " Twitter users in Indonesia are the Most List in the World, What Rank?", Twitter users in Indonesia reached 16.32 million in early 2022[3]. With the large enough number of users, an organization or news can use sentiment analysis to assess public opinion on an issue, whether the results tend to be good, bad, or neutral. Sentiment analysis is a process that functions to determine opinions, and also emotions that are translated through text. Usually, these sentiments are later classified into positive and negative opinions [4].

Sentiment analysis is considered one of the most important subsections of Natural Research in language processing (NLP) [5]. The sentiment analysis process can be in the form of content that contains about text reviews, forums, tweets, or blogs [6]. By using sentiment analysis, information circulating on social media can be processed into more data structured [7].

Several studies on sentiment analysis have been carried out by researchers using various methods, feature extraction and different deep-learning approaches. Like the research by Muhammad Radifan Aldiansyah, the method used to analyze sentiment on the Twitter social media platform is the Convolutional Neural Network (CNN) model obtains an accuracy rate of 88.21%. Based on the suggestions from this research to study neutral sentiments on Twitter, some tweets do not include positive or negative sentiments [8].

Firman Pradana and Handri Santoso compared the sentiment classification performance of CNN and other

Accepted: 24-01-2023 | Received in revised: 15-04-2023 | Published: 01-06-2023

deep learning techniques. Their findings showed that the CNN method had the highest accuracy at 90%. Therefore, it can be concluded that the CNN method is more effective than other deep learning methods. [7]

Furthermore, in research conducted by Muhammad Mahrus Zain et al., who implemented the RoBERTa method as word embedding for sentiment analysis. This research shows that using the RoBERTa method produces good accuracy results with an average accuracy of 95% [9].

Febiana Anistya et al., research on hate speech Indonesian Twitter uses the Glove feature expansion models. In this research, Glove was used to overcome vocabulary discrepancies, and this method produced the best accuracy with an average of 88.59%, an increase of 1.25% from the previous scenario[10].

Previous research has shown that the Convolutional Neural Network (CNN) and RoBERTa method has a superior level of accuracy compared to other learning methods. The research will contribute to analyzing public sentiment data of Twitter users using the CNN method for modelling, combining the feature extraction from TF-IDF N-gram with the RoBERTa and feature expansion from Glove. This is the first research in sentiment analysis using two ways of labelling in sentiment analysis with Glove, CNN and RoBERTA models.

2. Research Methods

The design system of the Sentiment Analysis is shown in Figure 1.

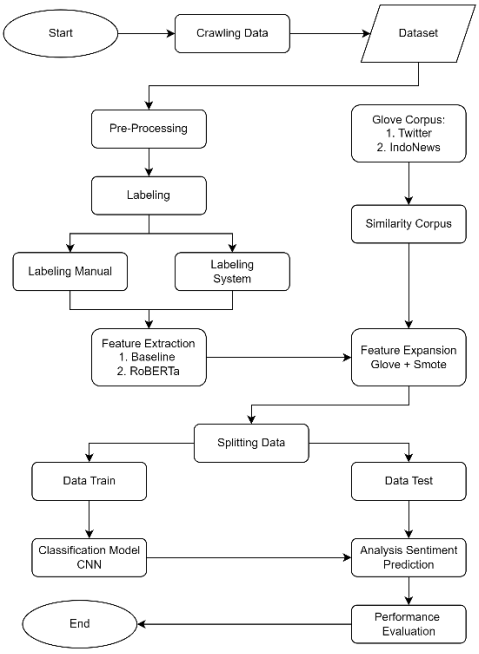

Figure 1. Sentiment Analysis System

2.1. Crawling Data

Crawling is a data collection process to obtain information from a page on a Twitter website, and the information obtained will be stored offline or locally [11]. In this research, the data consists of Indonesian-language tweets with 5 trending topics on October 2022 related to Indonesian government policies namely politics, pertamina, BPJS, pertalite, and bansos. Crawling was conducted on October 7 to 9, 2022 using snscrape. The features obtained for the dataset are tweetid, username, and tweet. Table 1 is the topic contained in this research.

Table 1. Sentiment Topics

Topics	Values
Politics	8.313
Pertamina	8.492
BPJS	10.135
Pertalite	583
Bansos	10.000

2.2. Labelling Data

The labelling process is carried out using a dataset of 30.811 originating from Twitter media. The tweet would be labeled by three respondents labeled 0 as negative tweets that can be in the form of hate speech and humiliation. One as neutral tweet that are objective words that do not contain positive or negative elements, and two as positive can be in the form of motivation or knowledge. An example of labelling is shown in Table 2.

Table 2. Example of Labelling Manual

Tweet	Respondent 1	Respondent 2	Respondent 3	Result
debus berkedok tentara	0	1	0	0
bansos ilang subsidi bbm pertalite kualitas	1	0	1	1
pecat tni polri	2	1	2	2
hormat panglima tni cawapres	0	0	0	0

In this research, the system also carried out labelling using the corpus that had been made. The corpus consists of 2 sources, and the first source was obtained by conducting a survey of 900 words contain unigram, bigram and trigram for 30 people with a range of labels ranging from -5 to 5, and the second source was obtained from Paper Fajri Koto on Github as many as 11.320 corpus with the same range of -5 to 5 [12]. The corpus will be transformed into three labels by examining past data, using the following mapping: values in the range of -5 to -1 will be assigned 0, a value of 0 will become 1, and values in the range of 1 to 5 will

DOI: https://doi.org/10.29207/resti.v7iX3.4892

be labeled as 2. The results of the conversion from manual labelling to system labelling can be observed in Table 3.

Table 3. Example of Labelling System

Tweet	Labelling Manual	Labelling System
debus berkedok tentara	0	1
bansos ilang subsidi bbm	1	1
pertalite kualitas	2	2
pecat tni polri	0	0
hormat panglima tni cawapres	0	2

Table 4 shows the distribution of the number of sentiments using labelling manual and the labelling system.

Table 4. Class Distribution

Sentiment	Labelling	
	Manual	System
Negative	10.804	7.736
Neutral	9.590	3.579
Positive	10.417	19.496

2.3. Pre-processing Data

Most recent research on sentiment analysis focuses on user-generated text. Pre-processing data are steps that function to process data in the form of a collection of texts, it needs to remove noisy information to clean, normalize, and classify speech [13]. Pre-processing in this reseach will be carried out through several stages, namely data cleaning, case folding, stopwords, tokenizing, and stemming.

The first step, data cleaning is a process that processes data by removing punctuation marks, emoticons, URLs, or characters other than text. The second step, Case folding is a process that makes all letters the same, for example, lowercase letters, by using lowercase function [14]. The third step, stopwords is a processing function to delete words that do not have a function and have a clear meaning. The fourth step, tokenizing is separating each word in a sentence. Finally, the fifth step, stemming is a process that changes words with affixes to their basic forms. The following is an example of the pre-processing carried out, which can be seen in Table 5.

Table 5. Pre-processing Data

Pre-processing	Tweet	Result
Data Cleaning	"Aksi massa yang enggk ngefek! Istana tegaskan tak ada opsi perppu Batalkan omnibus law di cipta kerja #MESKIDIDEMO RAKYAT #Jakarta #Gedungdprbengkulu #Bem_si #Bengkulu #TolakUUCiptaKerjaa https://t.co/VfWuR7xgDg https://t.co/DMzXVSaytE	Aksi massa yang enggk ngefek Istana tegaskan tak ada opsi perppu Batalkan omnibus law di cipta kerja
Case Folding	Aksi massa yang enggk yang ngefek Istana tegaskan tak ada opsi perppu Batalkan omnibus law di cipta kerja	aksi massa yang enggk ngefek istana tegaskan tak ada opsi perppu batalkan omnibus law di cipta kerja
Stopwords	aksi massa yang enggk ngefek istana tegaskan tak ada opsi perppu batalkan omnibus law di cipta kerja	aksi massa enggk ngefek istana tegaskan tak ada opsi perppu batalkan omnibus law cipta kerja
Tokenizing	aksi massa enggk ngefek istana tegaskan tak ada opsi perppu batalkan omnibus law cipta kerja	[‘aksi’, ‘massa’, ‘enggk’, ‘ngefek’, ‘istana’, ‘tegaskan’, ‘tak’, ‘ada’, ‘opsi’, ‘perppu’, ‘batalkan’, ‘omnibus’, ‘law’, ‘cipta’, ‘kerja’]
Stemming	[‘aksi’, ‘massa’, ‘enggk’, ‘ngefek’, ‘istana’, ‘tegaskan’, ‘tak’, ‘ada’, ‘opsi’, ‘perppu’, ‘batalkan’, ‘omnibus’, ‘law’, ‘cipta’, ‘kerja’]	[‘aksi’, ‘massa’, ‘enggk’, ‘efek’, ‘istana’, ‘tegas’, ‘tak’, ‘ada’, ‘perppu’, ‘batal’, ‘omnibus’, ‘law’, ‘cipta’, ‘kerja’]

2.4. Feature Extraction TF-IDF

Term Frequency – Inverse Document Frequency or TF-IDF is used in this research for feature extraction in evaluating the weight of important words in a document. This method works by counting the words that appear frequently and converting them into vector form, vectorization approaches can greatly affect the results [15], [16]. In addition to these functions, TF-IDF can create an N-grams. N-grams are n-character chunks that divide a sentence into several word parts. In this research, there are five types of N-grams. The first unigram is a token consisting of 1 word, and the second is a bigram which is a token consisting of 2 words. The third is a trigram which is a token consisting of 3 words, and the fourth uni-bigram is a token consisting of a range of unigrams and bigrams. And the last, uni-trigram, which is a token that consists of the unigram to trigram range. Table 6 shows an example of an N-grams word “aku daftar bpjs” generated by TF-IDF.

Table 6. N-grams Example

N-Gram	Result
Unigram	[aku], [daftar], [bpjs]
Bigram	[aku daftar], [daftar bpjs]
Trigram	[aku daftar bpjs]
Uni-Bigram	[aku], [daftar bpjs]
Uni-Trigram	[aku], [aku daftar bpjs]

2.5.Word Embedding RoBERTa

Word embeddings are representations of text data where each word is represented as a dense vector of

DOI: https://doi.org/10.29207/resti.v7iX3.4892

numbers. With the dataset that has been labeled, word embedding will be carried out. This method has a function to represent a word into a vector as semantic approach using the Robustly Optimized BERT Pretraining Approach (RoBERTa) method. It will be connected to the TF-IDF feature extraction. This method is an improvement from the previous method, namely Bidirectional Encoder Representations from Transformers (BERT). It extends the key hyperparameters to train with much larger learning rates and mini-batches [17].

Roberta's method consists of 4 stages. The first step, the model is trained using Dynamic Masking architecture which generates a mask pattern each time a sequence is entered into the model. The second step, the performance of tasks is improved by eliminating Next Sentence Prediction (NSP) and prompting users to consider the relationship between two sentences. The third step, optimization speed and end-task performance can be increased by training with a large number of batches. RoBERTa architecture has increased the batch size from 256 BERT sequences to 8,000 sequences to achieve this. The last step, Byte-Pair Encoding (BEP). It combines character and word representation to handle large standard vocabularies more efficiently[18]. The RoBERTa model used in this research a base called flax sentence embedding obtained from the RoBERTa large pretraining Huggingface to process Indonesian-language tweets.

2.6.Feature Expansion Glove

The expansion feature is a feature that is used to deal with distribution problems contained in data, in a word that has an ambiguous meaning [10]. One method that can be used for feature expansion is Glove. GloVe is a model that utilizes count data while capturing common meaningful linear substructures based on log-bilinear prediction [10]. With this method researchers can create a corpus of text and look for similar words, which are sorted based on the most similar rank.

The function of this method is to change the weight in feature extraction, which is 0 or ambiguous, to a new word available in the corpus. This research uses the 15 highest ranks of the searched words. Table 7 is the Glove corpus used in this research and Figure 2 is an example of the word similarity from the word "Bencana" in the resulting corpus.

Table 7. Corpus Glove IndoNews

Topics	Value
Goverments	13744
Politics	14388
Social	2890

Rank 1 Korban	**Rank 2** Cianjur	**Rank 3** Gempa	**Rank 4** Peduli	**Rank 5** Banjir
Rank 6 Evakuasi	**Rank 7** Longsor	**Rank 8** Covid	**Rank 9** Alam	**Rank 10** Intimidasi

Figure 2. Similarity Words of "Bencana"

2.7.SMOTE

Synthetic Minority Over-sampling Technique (SMOTE) is a method that is often used to deal with class imbalances in data. SMOTE was carried out on the system dataset. This method has a way of working by using classes that have a small or minority number reproduced using synthetic data from data replication [19]. Using SMOTE, the performance of classification models can be improved and reduce the risk of model bias towards the majority class.

In this research, it was used to balance the minority labels (negative and neutral) with the majority label (positive) as shown in Table 4 and the result in Table 8. By creating synthetic data, the resulting label system was balanced 19,496 positive, 19,496 neutral, and 19,496 negatives. As previously there were 19,496 positive, 7,736 negative, and 3,579 neutral labels.

Table 8. Result SMOTE

Topics	Value
Negative	19.496
Neutral	19.496
Positive	19.496

2.8.Convolutional Neural Network (CNN) Model

The modeling technique used in this research is Convolutional Neural Algorithm Network (CNN). CNN model commonly used in deep learning frameworks, this model become very popular tools in recent years, especially in the image processing community[20]. The basic structure of a convolutional neural network for image includes convolution layer, pooling layer, and fully connected layer [21]. But with the feature extraction, text can be processed to classification using this method.

The architecture of a CNN for text processing consists of 5 layers: First the Input layer have function to inputs data. Second the Convolutional layer that consist of a set filter to processes 1D text. Third the Flatten, this layer converts the 1D feature matrix into a processable matrix. Fourth the Dense layer which receives inputs from the neuron layer. Last the Output layer, which the outputs is probabilities from the hidden layer. Figure 3 shows the CNN architecture used to analyze tweet sentiment on Twitter.

DOI: https://doi.org/10.29207/resti.v7iX3.4892

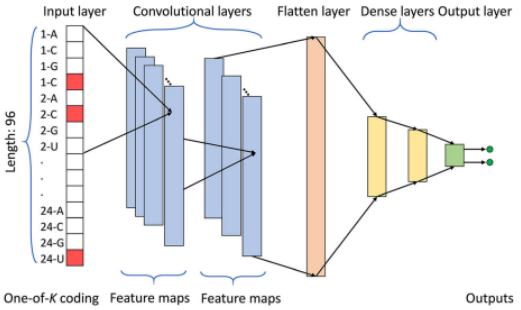

Figure 3. CNN Architecture[22]

2.9. Evaluation Matrix

The last stage in this research is evaluating the performance of the system being built. Evaluation matrix used to assess the performance or effectiveness of a particular system. This evaluation produces a value accuracy and F1-score. F1-score is used to measure the performance of a model by having the best results from a specified method [23]. In addition to the f1-score results in the form of precision and recall are also generated. Precision functions for the ratio of the positive prediction result from the total positive predictions and recall functions for the ratio of the predicted result of the overall positive predictions. Accuracy is intended to indicate the percentage of inputs that the CNN successfully predicted. The following is the calculation of the f1-score, precision, recall, and accuracy.

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

3. Results and Discussions

This research aims to obtain a model with the highest accuracy and analyze sentiment in social media with the CNN as a modeling method, TF-IDF N-gram as a baseline, RoBERTa as semantic approach, SMOTE to balance the label, and Glove as the expansion feature. There are four scenarios, and for all scenarios carried out 5 times by taking the average. Scenario 1 is to determine the data splitting ratio by testing the ratio comparison and the labelling used to serve as a baseline. Scenario 2 is to determine the best results using TF-IDF as a feature extraction between unigram, bigram, trigram, uni-bigram, or uni-trigram to find the optimal N-grams. Scenario 3 is to combine baseline from previous scenario with word embedding RoBERTa as semantic approach after TF-IDF. Combining the two extractions improves the system's performance because the more complex data, the easier the model understands the context. Scenario 4 is to create an expansion feature by using the similarity of corpus Glove from 30.811 tweets with 31.023 IndoNews dataset and SMOTE to handle label data imbalances to helps to reduce the bias in the model towards the majority class and increase the overall performance.

3.1.Result

The first scenario was carried out to find the best baseline by comparing the splitting ratio of the training data and the testing data using TF-IDF unigram. The splitting ratios used are 90:10, 80:20, 70:30, and 60:40 [24]. A ratio of 90:10 indicates that the dataset is splitting into 90% training data and 10% test data. It can be seen in Table 9 the difference in accuracy using manual labelling and labelling systems even though the ratio used is the same. Table 9 shows that the best splitting ratio is 90:10 and the labelling used is the labelling system which produces an accuracy of 86.79%. It selected as the baseline were then used for the next scenario.

Table 9. Comparison Ratio Between Manual and System

Ratio Splitting	Accuracy(%)	
	Labelling Manual	Labelling System
90:10	*69.10*	*86.79*
80:20	66.42	85.81
70:30	67.56	85.24
60:40	66.20	84.18

The second scenario uses TF-IDF as feature extraction by testing unigram, bigram, trigram, uni-bigram, and uni-trigram. It can be seen in Table 10 that the results of the performance of the N-gram with the ratio 90% train data and 10% test size. The unigram produced the highest accuracy of 86.79% followed by uni-trigram with 85.92% accuracy and trigram produces the lowest compared to the others. For the next scenario will use the unigram because it produces optimal accuracy than the others baseline.

Table 10. Feature Extraction TF-IDF

Baseline	Accuracy(%)
Unigram	86.79
Bigram	68.56
Trigram	62.20
Uni-Bigram	85.82
Uni-Trigram	85.92

Furthermore, the third scenario employs a combination of unigram TF-IDF and RoBERTa word embedding as a semantic approach method. The results show that combination leads to improvement in accuracy. As seen in Table 11, the combination of both unigram TF-IDF and RoBERTa results in an accuracy value of 87.75%, which is higher than when either method is used individually.

Table 11. TF-IDF with RoBERTa

Scenario	Accuracy(%)
Baseline + CNN	86.79
Baseline + RoBERTa + CNN	*87.75 (+1.1)*

DOI: https://doi.org/10.29207/resti.v7iX3.4892

Following the results of the previous scenario with a combination of baseline (unigram) and RoBERTa to improve the extraction features, the Glove expansion feature using the corpus similarity was added. The corpus uses a dataset of 30.811 tweets and 31.023 an IndoNews dataset of 3 topics that is government, politics and social. The first combination is a combination of baseline, RoBERTa, corpus tweet and SMOTE. The second combination is a combination of baseline, RoBERTa, corpus indonews and SMOTE. In this stage, the researcher tests the similarity of words with a similarity value, which will be replaced according to the top 1, top 5, top 10, and top 15 similarity ranks. This research also adds the SMOTE method to deal with class imbalances in the labelling system. Table 12 results from the Glove expansion feature coupled with the SMOTE method continues to increase up to 95.56%.

Table 12. Feature Expansion Glove and SMOTE

Rank	Accuracy(%)	
	Baseline + RoBERTa + Corpus Tweet + SMOTE	Baseline + RoBERTa + Corpus IndoNews + SMOTE
Top 1	94.81 (+9.23)	94.36 (+8.71)
Top 5	94.58 (+8.97)	95.02 (+9.48)
Top 10	*95.56 (+10.1)*	95.40 (+9.91)
Top 15	95.24 (+9.73)	95.17 (+9.65)

3.2.Discussion

Previous research by [8], used Convolutional Neural Network (CNN) on Twitter sentiment obtains an accuracy rate of 88.21%. The sentiment research needs to clasify the neutral sentimen, because some tweet has no negative or positive sentiments.

In this research, to deal with sentiments that are not positive and negative, 1 sentiment are added namely neutral. Testing was carried out using four scenarios. To produce maximum accuracy results, the researchers used a combination of CNN classification models to determine the best splitting ratio and TF-IDF feature extraction to determine the baseline. Added RoBERTa to improve baseline. This research also uses the Glove feature expansion by using two corpora: the Twitter corpus and the IndoNews corpus. Data imbalance results in bias, causing the classification model to have a tendency to only predict the majority class. This can be observed in Table 4, where the system data used in this research is imbalanced due to labelling. However, the research has addressed this issue by implementing an oversampling technique called SMOTE to improve the data.

Based on the four scenarios that have been tested, first scenario produces the best performance using a labeling system with a ratio of 90:10 with unigram as a baseline with an accuracy of 86.79%. The use of labelling by the system greatly influences the previous performance results, namely manual labelling. Then, the second scenario is a continuation of the previous scenario. That compares N-grams, namely unigrams, bigrams, trigrams, uni-bigrams, and uni-trigrams. The optimal accuracy for the baseline using CNN is unigram, with an accuracy of 86.79%. Both scenario 1 and scenario 2 generate optimal values by utilizing the unigram, resulting in an equivalent increase from the baseline. The third scenario is the addition of RoBERTa to optimize and improve baseline accuracy with CNN. This scenario produces an accuracy of 87.75%, the accuracy is increased by 1.1%. These results prove that combining feature extraction between TF-IDF and RoBERTa has succeeded in increasing the performance of the system being built.

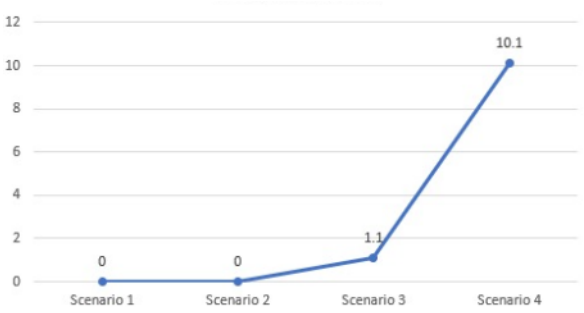

Figure 4. Scenario Accuracy

And the last scenario uses a 90:10 ratio, unigram baseline, RoBERTa, SMOTE, and Glove expansion features. The best result is using the Twitter corpus using the top 10 similarities, which is used to replace the 0 value on TF-IDF resulting in an accuracy of 95.56%, it can be seen from Figure 4 the accuracy increases quite rapidly by 10.1%

4. Conclusion

In this research, the researchers tried to analyze sentiment on social media especially Twitter. There is five topics that used in this research related to Indonesian government policies with 3 sentiment label that is positive, negative, and neutral using Convolutional Neural Network (CNN) as a modeling method, feature extraction from Term Frequency-Inverse Document Frequency (TF-IDF) with N-Gram as a baseline. Robustly Optimized BERT Pretraining Approach (RoBERTa) word embedding as semantic approach to improve the baseline, Global Vectors (Glove) as expansion feature, and Synthetic Minority Over-sampling Technique (SMOTE) to balance the labels.

The data used in this research were 30.811 data originating from tweets. The final results of this research using the model combination of TF-IDF N-gram, RoBERTa, Glove, SMOTE, and CNN resulted in an accuracy of 95.56%. Based on the result, combining a multiple method, using corpus and labelling with

DOI: https://doi.org/10.29207/resti.v7iX3.4892

system for sentiment analysis can improve the accuracy. For further research, recommended to conduct research using the optimization method such as Different Evolution to obtain the better accuracy.

References

[1] T. Siswanto, "Optimalisasi Sosial Media Sebagai Media Pemasaran Usaha Kecil Menengah," *Liquidity*, vol. 2, no. 1, pp. 80–86, 2018, doi: 10.32546/lq.v2i1.134.

[2] databoks.katadata.co.id, "Berapa Pengguna Media Sosial Indonesia?," 2019. https://databoks.katadata.co.id/datapublish/2019/02/08/berapa-pengguna-media-sosial-indonesia (accessed Apr. 22, 2022).

[3] databoks.katadata.co.id, "Pengguna Twitter Indonesia Masuk Daftar Terbanyak di Dunia, Urutan Berapa?," 2022. [Online]. Available: https://databoks.katadata.co.id/datapublish/2022/03/23/pengguna-twitter-indonesia-masuk-daftar-terbanyak-di-dunia-urutan-berapa

[4] M. Cindo and D. P. Rini, *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) Literatur Review: Metode Klasifikasi Pada Sentimen Analisis*. 2019. [Online]. Available: https://seminar-id.com/semnas-sainteks2019.html

[5] J. Tao and X. Fang, "Toward multi-label sentiment analysis: a transfer learning based approach," *J. Big Data*, vol. 7, no. 1, pp. 1–26, 2020, doi: 10.1186/s40537-019-0278-0.

[6] W. A. Prabowo and C. Wiguna, "Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM," *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 149, 2021, doi: 10.30865/mib.v5i1.2604.

[7] F. Pradana Rachman, H. Santoso, and R. Artikel, "Jurnal Teknologi dan Manajemen Informatika Perbandingan Model Deep Learning Untuk Klasifikasi Sentiment Analysis Dengan Teknik Natural Languange Processing Info Artikel ABSTRAK," vol. 7, no. 2, pp. 103–112, 2021, [Online]. Available: http://http//jurnal.unmer.ac.id/index.php/jtmi

[8] M. R. Aldiansyah and P. S. Sasongko, "Twitter Sentiment Analysis about Public Opinion on 4G Smartfren Network Services Using Convolutional Neural Network," *ICICOS 2019 - 3rd Int. Conf. Informatics Comput. Sci. Accel. Informatics Comput. Res. Smarter Soc. Era Ind. 4.0, Proc.*, pp. 3–8, 2019, doi: 10.1109/ICICoS48119.2019.8982429.

[9] M. Mahrus Zain, R. Nathamael Simbolon, H. Sulung, and Z. Anwar, "Analisis Sentimen Pendapat Masyarakat Mengenai Vaksin Covid-19 Pada Media Sosial Twitter dengan Robustly Optimized BERT Pretraining Approach," *J. Komput. Terap.*, vol. 7, no. Vol. 7 No. 2 (2021), pp. 280–289, 2021, doi: 10.35143/jkt.v7i2.4782.

[10] Febiana Anistya and Erwin Budi Setiawan, "Hate Speech Detection on Twitter in Indonesia with Feature Expansion Using GloVe," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 6, pp. 1044–1051, 2021, doi: 10.29207/resti.v5i6.3521.

[11] A. Wattimena, "Analisis Sentimen Teks Bahasa Indonesia Pada Media Sosial Menggunakan Algoritma Convolutional Neural Network (Studi Kasus : E-Commerce)," 2018, [Online]. Available: https://repository.its.ac.id/52636/1/05211440000134-Undergraduate_Theses.pdf

[12] F. Koto and G. Y. Rahmaningtyas, "Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs," *Proc. 2017 Int. Conf. Asian Lang. Process. IALP 2017*, vol. 2018-Janua, no. December, pp. 391–394, 2018, doi: 10.1109/IALP.2017.8300625.

[13] H. T. Duong and T. A. Nguyen-Thi, "A review: preprocessing techniques and data augmentation for sentiment analysis," *Comput. Soc. Networks*, vol. 8, no. 1, pp. 1–16, 2021, doi: 10.1186/s40649-020-00080-x.

[14] D. T. Hermanto, A. Setyanto, and E. T. Luthfi, "Algoritma LSTM-CNN untuk Binary Klasifikasi dengan Word2vec pada Media Online," *Creat. Inf. Technol. J.*, vol. 8, no. 1, p. 64, 2021, doi: 10.24076/citec.2021v8i1.264.

[15] N. Buslim, B. Busman, N. S. Sinatrya, and T. S. Kania, "Analisa Sentimen Menggunakan Data Twitter, Flume, Hive Pada Hadoop dan Java Untuk Deteksi Kemacetan di Jakarta," *J. Online Inform.*, vol. 3, no. 1, p. 1, 2018, doi: 10.15575/join.v3i1.141.

[16] R. K. Bania, "COVID-19 Public Tweets Sentiment Analysis using TF-IDF and Inductive Learning Models," *Infocomp*, vol. 19, no. 2, pp. 23–41, 2020.

[17] A. Barua, S. Thara, and K. P. Soman, *nalysis of Contextual and Non-contextual Word Embedding Models for Hindi NER with Web Application for Data Collection*, vol. 223, no. 4643. 2021. doi: 10.1126/science.223.4643.1350.a.

[18] A. E. Putra and W. Maharani, "Depression Levels Detection Through Twitter Tweets Using RoBERTa Method," *J. Inf. Syst. Res.*, vol. 3, no. 4, pp. 453–459, 2022, doi: 10.47065/josh.v3i4.1872.

[19] E. Sutoyo and M. A. Fadlurrahman, "Penerapan SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Television Advertisement Performance Rating Menggunakan Artificial Neural Network," *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 3, p. 379, 2020, doi: 10.26418/jp.v6i3.42896.

[20] M. Zulqarnain, R. Ghazali, Y. M. M. Hassim, and M. Rehan, "A comparative review on deep learning models for text classification," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 19, no. 1, pp. 325–335, 2020, doi: 10.11591/ijeecs.v19.i1.pp325-335.

[21] W. Meng, Y. Wei, P. Liu, Z. Zhu, and H. Yin, "Aspect Based Sentiment Analysis with Feature Enhanced Attention CNN-BiLSTM," *IEEE Access*, vol. 7, pp. 167240–167249, 2019, doi: 10.1109/ACCESS.2019.2952888.

[22] Q. Zhao *et al.*, "Prediction of plant-derived xenomiRs from plant miRNA sequences using random forest and one-dimensional convolutional neural network models," *BMC Genomics*, vol. 19, no. 1, pp. 1–13, 2018, doi: 10.1186/s12864-018-5227-3.

[23] K. L. Kohsasih *et al.*, "Analisis Perbandingan Algoritma Convolutional Neural Network Dan Algoritma Multi-Layer Perceptron Neural Dalam Klasifikasi Citra Sampah," *J. Technol. Informatics dan Comput. Syst.*, vol. 10, no. 2, pp. 22–28, 2021, [Online]. Available: http://ejournal.stmik-time.ac.id

[24] B. Anthony, C. Martani, and E. B. Setiawan, "JURNAL RESTI Five Personality on Twitter," vol. 5, no. 158, pp. 1072–1078, 2022.

DOI: https://doi.org/10.29207/resti.v7iX3.4892