JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

Bank Central Asia (BBCA) Stock Price Sentiment Analysis On Twitter Data Using Neural Convolutional Network (CNN) And Bidirectional Long Short-Term Memory (BI-LSTM)

Mansel Lorenzo Nugraha, Erwin Budi Setiawan*

Informatics, Telkom University, Bandung, Indonesia
Email: [1]mansel@student.telkomuniversity.ac.id, [2,*]erwinbudisetiawan@telkomuniversity.ac.id
Correspondence Author Email: erwinbudisetiawan@telkomuniversity.ac.id

Abstract-Stock investing has become popular among the public. Although this stock investment has significant risks, every year, investors keep increasing because the return from stocks is also quite promising. Social media also supports this stock investing, which can give information extensively and very quickly, so it can affect the stock price. The Efficient Market Hypothesis (EMH) theory defines that market information reflects stock prices. In this research, sentiment analysis uses a dataset crawled from Twitter to process the sentiment into helpful information. All the tweets related to stock prices are collected for sentiment analysis according to the appropriate sentiment type, whether it's a positive or negative sentiment. Many believe that sentiment influences stock price movements. This sentiment analysis process uses a hybrid method named Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (Bi-LSTM) with feature expansion Word2Vec. Afterwards, the hybrid method analysis will establish the final accuracy obtained. This research uses 27.930 data and shows the hybrid CNN Bi-LSTM method result is 95.74%.

Keywords: Convolutional Neural Network; Sentiment Analysis; Bidirectional Long Short-Term Memory; Twitter; Stock

1. INTRODUCTION

Stock investment has become popular among the public, and people often choose it because it can generate a high profit. Companies select stock for their financial need. Stock is an investment instrument investors favour because of its attractive return [1]. Stock is also added to by the number of investors, which keeps increasing yearly. People now realize the importance of investment, accompanied by increasingly advanced digital technology. The data obtained from the Central Securities Depository of Indonesia (KSEI) explained that, until August 2022, there were already 9.54 million stock investors. This number is very high because, in December 2021, the recorded number of investors stood at only 7.48 million. It means that stock market investors have increased by 27.38% [2]. The Efficient Market Hypothesis (EMH) theory defines that market information reflects stock price [3].

One of the most prominent social media platforms for sentiment analysis is Twitter. On Twitter, we can find specific keywords about a topic to determine whether the tweet is positive or negative [4].

Sentiment analysis is obtaining internet and social media data using text analytics. Sentiment analysis seeks to ascertain a person's perspective on social media, whether positive, negative, or neutral, with attention to the discussed topic [5]. Pradana et al.'s (2020) earlier investigation discovered a connection between the tone of a tweet and the movement of a company's stock price. A tweet from someone that contains a complaint tends to affect the stock price, as a tweet with many complaints will lead to a lower stock price, whereas positive tweets will lead to a higher stock price [6]. This sentiment is advantageous for organizations, corporations, and governments, but text mining and sentiment analysis are needed[7]. One example is the alleged Bank Central Asia's (BCA) tax objection in 2014, which led to negative perceptions of Bank Central Asia [8]. Sentiment analysis is performed on social media news and issues to determine how they influence changes in stock prices.

In this research, the researcher observes the impact of tweets on the stock price movement of Bank Central Asia (BBCA) using a hybrid method of Bidirectional Long Short-Term Memory (Bi-LSTM) and Convolutional Neural Network (CNN) with feature expansion Word2Vec. The researcher uses this hybrid method because it can enhance prediction accuracy using the method used [9]. This research uses the feature expansion Word2Vec, which converts words to vectors [10]. The advantage of using Word2Vec is that it can process small or large amounts of data [11]. The researcher expects this research to provide good accuracy results from the hybrid method and obtain sentiment information related to BBCA stock price movements to help investors make decisions.

Wang Yue and Lei Li [12] introduce this combination method of CNN and Bi-LSTM with additional Word2Vec feature expansion. The result obtained from this method is 91.48%, which is way bigger than other methods used in the same research. The accuracy result obtained from the only LSTM method is 79.83%, and with CNN, the accuracy obtained is 81.25% with Bi-LSTM, 85.69%, and with CNN-LSTM, 87.44%. Researchers conclude that using a combination method of CNN and Bi-LSTM with Word2Vec feature expansion is easier and more accurate. It also raises the accuracy of sentiment analysis for short text.

Kai Zhou and Fei Long [13] did a sentiment analysis of Chinese products using CNN as feature extraction in text and Word2Vec as feature expansion. They obtained high results after the metric evaluation with F1-Score, Precision, and Recall. Overall accuracy obtained from the hybrid method of CNN and Bi-

Submitted: **01/05/2023**; Accepted: **11/07/2023**; Published: **23/07/2023**

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

LSTM dominates positive and negative metrics over other methods used in the research. These methods are LSTM, CNN, Bi-LSTM, and hybrid CNN-LSTM.

Tam et al. [14] also researched using different combined approaches, including CNN, LSTM, Bi-LSTM, and ConvBi-LSTM, for sentiment classification with two datasets and using GloVe and Word2Vec as feature expansions. The result of this research shows that using Word2Vec feature expansion on the tweets dataset achieves higher accuracy in the ConvBi-LSTM method, which has an accuracy of 93.76%, a CNN result of 91.89%, an LSTM result of 90.94%, a Bi-LSTM result of 91.52%, and a CNN-LSTM method with more than 86% accuracy.

Gandhi et al. [15] did the sentiment analysis for a dataset named IMDB, which was obtained from Kaggle.com and uses feature expansion Word2Vec. This research indicates that the LSTM method's accuracy is 88.02% while the CNN method is 87.72%.

Dedi et al.[16] did the sentiment classification for a dataset crawl from www.Finance.detik.com. This research uses CNN and LSTM for the method and Word2Vec as feature expansion. The accuracy result of the LSTM method is 51%, hybrid LSTM-CNN is 53%, and CNN-LSTM is 62%.

Based on the five journals above, the conclusion is that using CNN and Bi-LSTM methods can provide high accuracy and be applied to big data. This research refers to the journals above and uses CNN combined with Bi-LSTM methods and Word2Vec as feature expansion. The researcher does not use the four scenarios directly explained in section 3, which are not used in the five journals above. The amount of the dataset used in this research is higher than in 3 previous research by Wang Yue and Lei Li, Kai Zhou, and Fei Long, and Dedi et al., whereas the amount of the dataset is lower than in two previous research by Gandhi et al. and Tam et al. Based on the different scenarios for this model, the researcher predicts that the overall accuracy in this hybrid method will reach more than 90%.

2. RESEARCH METHODOLOGY

2.1 System Design

This system plan for BBCA stock price sentiment analysis uses the hybrid methods CNN and Bi-LSTM and the feature expansion Word2Vec. First, the dataset is crawled and labelled 2.0 as positive, 1.0 as neutral, and 0.0 as negative. The next step is data pre-processing, feature expansion with Word2Vec, metrics evaluation, and rank Spearman correlation. Figure 1 displays the flowchart system.

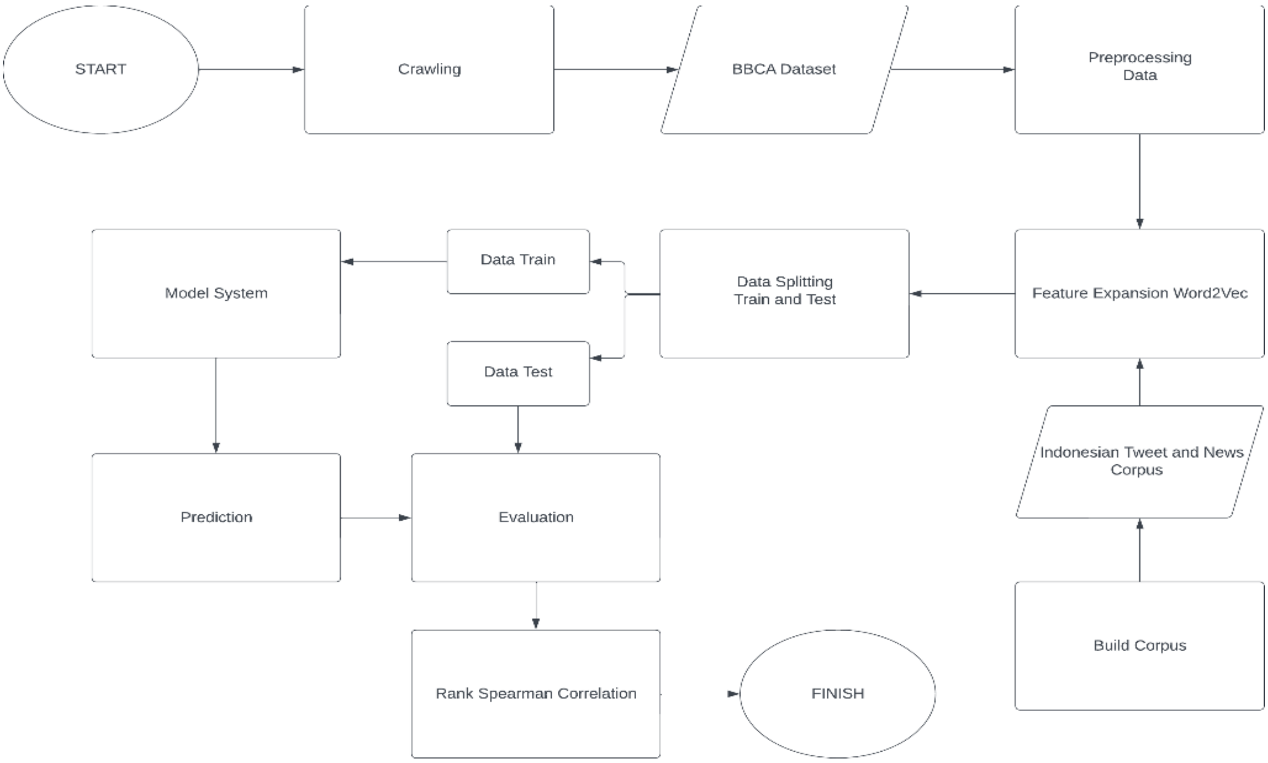

Figure 1. Flowchart System

2.2. Crawling and Data Labelling

The researcher crawled the dataset from January 2019 to February 2023. The keywords used by the researcher are BCA Error, Pelayanan BCA, Saham BBCA, BBCA, and Bank BCA. At first, the crawling got 27.930 data and

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

became 23.148 data after double tweet removal and data pre-processing. The dataset was manually labelled 2.0 as positive tweets, 1.0 as neutral, and 0.0 as negative. Five people in a research group conducted the manual labelling, and the research team divided the group to label specific data ranges. Afterwards, the group discussed each tweet containing ambiguous sentences to assign the predetermined labels. The group has also agreed to conduct independent rechecks, although not in a formal manner. Table 1 shows the labelled tweet by the group.

Table 1. Labelling Process

Data	Label
@budakfmcg Kalo kata orang sih ya kalo sampai di acc cc bca mau ajuin cc bank lain jadi gampang banget karena terkenal susahnya cc bca itu 😅	0.0
@Strategi_Bisnis Aku mau ganti kartu atm miss info banget. Kalo di mandiri kata cs twt ga perlu bawa surat ket hilang. Pas sampe bank ditolak harus ada surat ket hilang. Agak males antree di kantor polisi lama, di bank mandiri juga harus antree. Kalo di BCA ga perlu pake surat ket hilang.	2.0

2.3. Data Pre-processing

The researcher pre-processed data for the processed dataset, which was unstructured before becoming structured. Pre-processing steps in this research are case folding, data cleaning, tokenizing, stemming, and stopword removal.

2.3.1. Case Folding

The first step of data pre-processing is case folding. The researchers converted every capital letter in the dataset to lowercase. Table 2 shows the result of case folding.

Table 2. Case Folding Result

Tweet	Case Folding
@budakfmcg Kalo kata orang sih ya kalo sampai di acc cc bca mau ajuin cc bank lain jadi gampang banget karena terkenal susahnya cc bca itu 😅	@budakfmcg kalo kata orang sih ya kalo sampai di acc cc bca mau ajuin cc bank lain jadi gampang banget karena terkenal susahnya cc bca itu 😅
@Strategi_Bisnis Aku mau ganti kartu atm miss info banget. Kalo di mandiri kata cs twt ga perlu bawa surat ket hilang. Pas sampe bank ditolak harus ada surat ket hilang. Agak males antree di kantor polisi lama, di bank mandiri juga harus antree. Kalo di BCA ga perlu pake surat ket hilang.	@strategi_bisnis aku mau ganti kartu atm miss info banget. kalo di mandiri kata cs twt ga perlu bawa surat ket hilang. pas sampe bank ditolak harus ada surat ket hilang. agak males antree di kantor polisi lama, di bank mandiri juga harus antree. kalo di bca ga perlu pake surat ket hilang.

2.3.2. Data Cleaning

The second step is data cleaning. In this step, the researcher removes every URL, number, emoji, symbol, username, and punctuation. Table 3 shows the data cleaning result.

Table 3. Data Cleaning Result

Case Folding	Data Cleaning
@budakfmcg Kalo kata orang sih ya kalo sampai di acc cc bca mau ajuin cc bank lain jadi gampang banget karena terkenal susahnya cc bca itu 😅	kalo kata orang sih ya kalo sampai di acc cc bca mau ajuin cc bank lain jadi gampang banget karena terkenal susahnya cc bca itu
@strategi_bisnis aku mau ganti kartu atm miss info banget. kalo di mandiri kata cs twt ga perlu bawa surat ket hilang. pas sampe bank ditolak harus ada surat ket hilang. agak males antree di kantor polisi lama, di bank mandiri juga harus antree. kalo di bca ga perlu pake surat ket hilang.	aku mau ganti kartu atm miss info banget kalo di mandiri kata cs twt ga perlu bawa surat ket hilang pas sampe bank ditolak harus ada surat ket hilang agak males antree di kantor polisi lama di bank mandiri juga harus antree kalo di bca ga perlu pake surat ket hilang

2.3.3. Tokenization

The third step is tokenization. In this step, the researcher converts every word in sentences into tokens. Table 4 shows the tokenization result.

Table 4. Tokenization Result

Data Cleaning	Tokenization
kalo kata orang sih ya kalo sampai di acc cc bca mau ajuin cc bank lain jadi gampang banget karena terkenal susahnya cc bca itu	[kalo, kata, orang, sih, ya, kalo, sampai, di, acc, cc, bca, mau, ajuin, cc, bank, lain, jadi, gampang, banget, karena, terkenal, susahnya, cc, bca, itu]

Submitted: 01/05/2023; Accepted: 11/07/2023; Published: 23/07/2023

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

Data Cleaning	Tokenization
aku mau ganti kartu atm miss info banget kalo di mandiri kata cs twt ga perlu bawa surat ket hilang pas sampe bank ditolak harus ada surat ket hilang agak males antree di kantor polisi lama di bank mandiri juga harus antree kalo di bca ga perlu pake surat ket hilang	[aku, mau, ganti, kartu, atm, miss, info, banget, kalo, di, mandiri, kata, cs, twt, ga, perlu, bawa, surat, ket, hilang, pas, sampe, bank, ditolak, harus, ada, surat, ket, hilang, agak, males, antree, di, kantor, polisi, lama, di, bank, mandiri, juga, harus, antree, kalo, di, bca, ga, perlu, pake, surat, ket, hilang]

2.3.4. Stopword Removal

The fourth step is stopword removal. In this step, the researcher removes words that generally appear insignificant. This step focuses on the critical word. Table 5 shows the stopword removal result.

Table 5. Stopword Removal Results

Tokenization	Stopword Removal
[kalo, **kata**, orang, sih, ya, kalo, **sampai**, **di**, acc, cc, bca, mau, ajuin, cc, bank, **lain**, **jadi**, gampang, banget, **karena**, terkenal, susahnya, cc, bca, **itu**]	[kalo, orang, sih, ya, kalo, acc, cc, bca, ajuin, cc, bank, gampang, banget, terkenal, susahnya, cc, bca]
[**aku**, **mau**, ganti, kartu, atm, miss, info, banget, kalo, **di**, mandiri, **kata**, cs, twt, ga, **perlu**, bawa, surat, ket, hilang, pas, sampe, bank, ditolak, **harus**, **ada**, surat, ket, hilang, **agak**, males, antree, **di**, kantor, polisi, **lama**, **di**, bank, mandiri, **juga**, **harus**, antree, kalo, **di**, bca, ga, **perlu**, pake, surat, ket, hilang]	[ganti, kartu, atm, miss, info, banget, kalo, mandiri, cs, twt, ga, bawa, surat, ket, hilang, pas, sampe, bank, ditolak, surat, ket, hilang, males, antree, kantor, polisi, bank, mandiri, antree, kalo, bca, ga, pake, surat, ket, hilang]

2.3.5. Stemming

The final process of data pre-processing is stemming. This step converts every word to its basic form based on Indonesian spelling. The researcher uses the Sastrawi library for the Indonesian Language stemming process. Table 6 shows the stemming result.

Table 6. Stemming Result

Stopwords Removal	Stemming
[kalo, orang, sih, ya, kalo, acc, cc, bca, ajuin, cc, bank, gampang, banget, **terkenal**, **susahnya**, cc, bca]	[kalo, orang, sih, ya, kalo, acc, cc, bca, ajuin, cc, bank, gampang, banget, kenal, susah, cc, bca]
[ganti, kartu, atm, miss, info, banget, kalo, mandiri, cs, twt, ga, bawa, surat, ket, hilang, pas, sampe, bank, **ditolak**, surat, ket, hilang, males, antree, kantor, polisi, bank, mandiri, antree, kalo, bca, ga, pake, surat, ket, hilang]	[ganti, kartu, atm, miss, info, banget, kalo, mandiri, cs, twt, ga, bawa, surat, ket, hilang, pas, sampe, bank, tolak, surat, ket, hilang, males, antree, kantor, polisi, bank, mandiri, antree, kalo, bca, ga, pake, surat, ket, hilang]

2.4. Word2Vec

Word2Vec is an approach from word embedding, a technique to convert words to a numeric vector shape, making it easier for natural language processing. Word2Vec has two models: CBOW (Continuous Bag of Words Model). This first model extracts words from a sentence and predicts the values in real-time. The second model is Skip-Gram, which uses present words for context prediction[12]. Figure 2 shows the Word2Vec architecture.

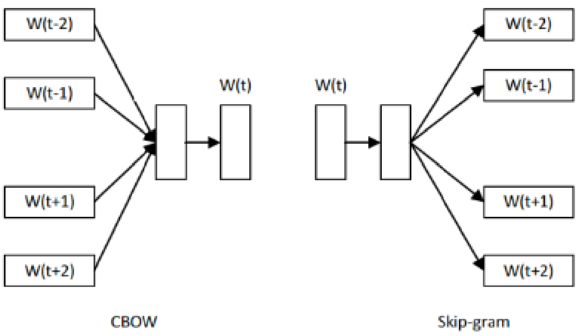

Figure 2. Word2Vec Architecture

Word2Vec, used as feature expansion in this research, sets the embedding vector size to 100, then trains with a window of 10, a min_count of 1, and 4 workers. This Word2Vec is also built by Indonesian Tweet, Indonesian News, and Indonesian Tweet + News Corpus[17]. It generates word embedding vectors for each term in the tweets and input features for each model.

Submitted: 01/05/2023; Accepted: 11/07/2023; Published: 23/07/2023

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

2.5. Split Data

After Word2Vec modelling, the researcher split the data into training and testing and shaped it into vectors. This data split is used for modelling and predicting with CNN, Bi-SLTM, and hybrid CNN Bi-LSTM. The ratio comparisons used for this research are 90:10, 20:80, and 70:30[18].

2.6. CNN Modelling

A convolutional Neural Network (CNN) is a feed-forward neural network with a convolutional structure. Classification with this method uses Convolutional Neural Networks for feature extraction [12]. The parts of Convolutional Neural Networks consist of four layers, as shown in Figure 3. The explanation of the layers is [19]:

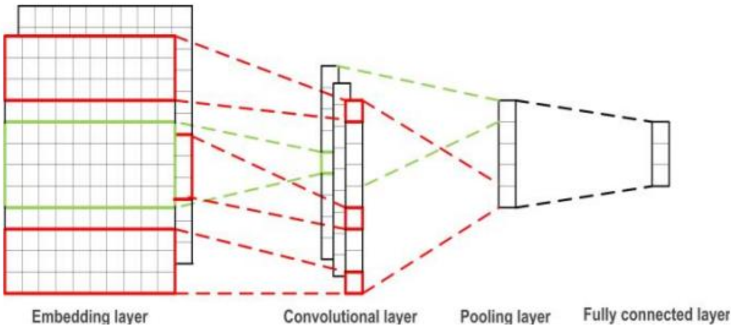

Figure 3. CNN Architecture

2.6.1. Embedding Layer

The embedding layer consists of a word vector's matrix corresponding to words in sentences arranged sequentially from top to bottom.

2.6.2. Convolutional Layer

The embedding layer consists of many feature maps with convolutional processes, which produce one-column feature maps.

2.6.3. Pooling Layer

Subsampling is another name for it, which can decrease the amount of data input. Max pooling is the most widely used pooling layer[19].

2.6.4. Fully Connected Layer

This Fully Connected Layer connects the output of this layer to all levels that came before it. This layer can be called the output layer.

2.7. Bi-LSTM Modelling

Bi-LSTM is a development of LSTM. There are two layers in this model, namely forward and backward layers. The forward layer processes and understands words from beginning to end, whereas the backward layer is the inverse of the forward layer. This model helps understand sentence patterns because every word in the document is processed sequentially, and this model also understands and takes previous words. This model is suitable for context understanding in tweets. Researchers generally use this Bi-LSTM model for long-term learning. This model also consists of gates, namely input, forget, and output [20]. Figure 4 shows the Bi-LSTM architecture.

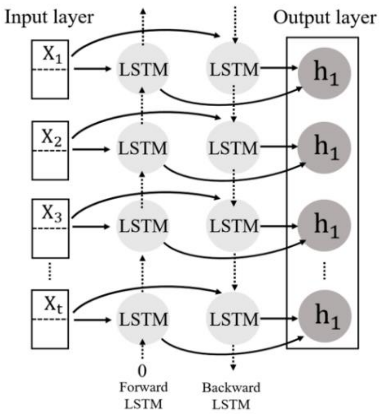

Figure 4. Bi-LSTM Architecture

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

2.8. Hybrid CNN Bi-LSTM Modelling

The researcher did this hybrid CNN Bi-LSTM method after testing only Bi-LSTM and CNN models. The Bi-LSTM and CNN layers are combined and tested. The last step compares the best model for BBCA stock price sentiment analysis.

2.9. Evaluation

This stage aims to determine the accuracy of the hybrid method used in this study, with an evaluation conducted using evaluation metrics. Commonly used evaluation metrics include F1-Score, precision, and recall. This evaluation compares the method's predictions with the actual data. The valid positive rate refers to the number of issues projected to be positive and positive. In contrast, false positive results from inaccurate prediction as a positive. For the negative class, true negative and false negative have meaningful connections[21]. Below, the researcher explains the formula and the explanation of the metrics.

2.9.1. Precision

Precision is a class's prediction value in proportion to all its actual instances. Its purpose is to observe the classification results in percentages. The following formula (1) shows the precision equation.

$$\textit{Precision}\ (\mathrm{p}):\ P = \frac{TP}{TP+FP} \tag{1}$$

2.9.2. Recall

The recall is the accuracy value of predicting a class based on the total count of actual occurrences of that class. Its purpose is to observe the proportion of finished, appropriate classification results using the method employed. The following formula (2) shows the recall equation.

$$Recall\ (r){:}\ r = \frac{TP}{TP+FN} \tag{2}$$

2.9.3. F1-Score

The evaluation calculation for F1-Score involves computing this metric by combining the precision and recall values. The following formula (3) shows the F1-Score equation.

$$F1\text{-}Score\ (f_1){:}\ f_1 = \frac{2}{\frac{1}{P}+\frac{1}{r}} = \frac{2(TP+FP)(TP+FN)}{TP} \tag{3}$$

2.9.4. Overall Accuracy

The researcher calculates the model's accuracy using overall accuracy (OA). The following formula (4) shows the overall accuracy equation.

$$Overall\ Accuracy\ (O_A){:}\ O_A = \frac{TP}{TP+FP+TN\_FN} \tag{4}$$

2.10. Rank Spearman Correlation

Rank Spearman correlation is a non-parametric method for assessing the relationship between two independent variables. Data dispersion does not influence this method as an advantage. This method uses data rank to lessen the sensitivity to outlier data[22]. The following formula (5) shows the formula of rank Spearman correlation.

$$rs = 1 - \frac{6\Sigma\, d_i^2}{n^3 - n'} \tag{5}$$

Table 7 shows Schober et al.'s classification of rank Spearman [23].

Table 7. Rank Spearman Correlation

Correlation Coefficient	Correlation Strength
0-0.1	No Correlated
0.1-0.39	Minimal Correlation
0.4-0.69	Medium Correlation
0.7-0.89	Powerful Correlation
0.9-1	Effective Correlation

3. RESULT AND DISCUSSION

3.1. Data Distribution

The crawled and pre-processed tweet data amounted to 23.148, using the Indonesian Language. The tweets contained the opinions of netizens about the stock prices of BBCA. Figure 5 shows the data distribution.

Submitted: 01/05/2023; Accepted: 11/07/2023; Published: 23/07/2023

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

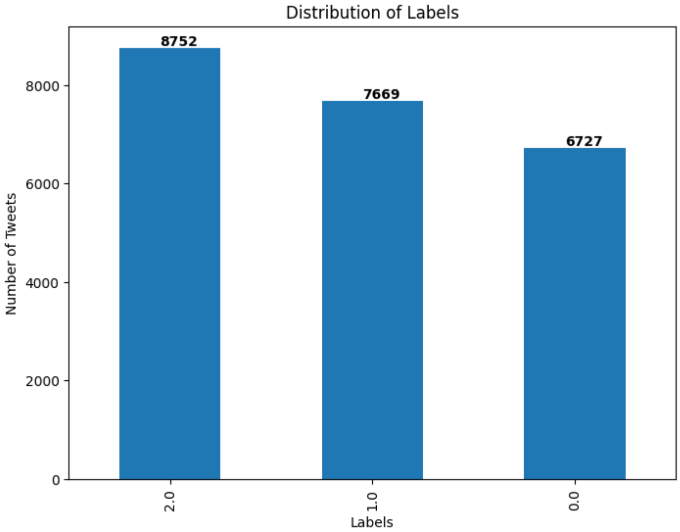

Figure 5. Data Distribution

As shown in Figure 5, the positive label (2.0) distribution is 8.752, the neutral label (1.0) is 7.669, and the negative label (0.0) is 6.727. Csv format stores all data.

3.2. Result and Testing Analysis

In this research, there are four scenarios. The first scenario tests different ratios without Word2Vec feature expansion using ratios of 90:10, 80:20, and 70:30. The researcher uses the ratio that produces the best accuracy for the following scenario. Then, scenario 2 performs oversampling testing with SMOTE using the best ratio from the previous scenario. The following scenario was tested by adding the Word2Vec expansion feature. The last scenario is to set the hyperparameter for all models with feature expansion Word2Vec.

3.2.1 Scenario 1 (Testing Without Word2Vec)

The researcher obtained the data in this experiment after pre-processing the data from the previous crawling results. The researcher got the best ratio without using Word2Vec, which was a 90:10 ratio (90% data train and 10% data test) with an overall accuracy of 95.24% for CNN, 94.34% for Bi-LSTM, and 95.41% for the hybrid CNN Bi-LSTM. Below, the tables show the complete result of all ratios.

Table 8. Scenario 1 Ratio 90:10

Model	Class	F1-Score	Precision	Recall	Overall Accuracy
	Positive	0.95	0.95	0.95	
CNN	Negative	0.97	0.98	0.96	95.24%
	Neutral	0.94	0.93	0.95	
Bi-LSTM	Positive	0.94	0.92	0.96	
	Negative	0.97	0.96	0.97	94.34%
	Neutral	0.93	0.95	0.91	
Hybrid CNN Bi-LSTM	Positive	0.95	0.94	0.96	
	Negative	0.98	0.98	0.97	95.41%
	Neutral	0.94	0.95	0.93	

The ratio of 90:10 is the best for the first scenario. The accuracy result by using all methods is higher than other ratios. Using CNN, the accuracy is 95.24%, Bi-LSTM is 93.34%, and hybrid CNN Bi-LSTM is 95.41%. The hybrid method of CNN Bi-LSTM is the highest accuracy in this ratio. The F1-Score, Precision, and Recall also showed the best result, with all the percentages higher than 90%.

Table 9. Scenario 1 Ratio 80:20

Model	Class	F1-Score	Precision	Recall	Overall Accuracy
CNN	Positive	0.92	0.93	0.90	
	Negative	0.95	0.95	0.96	92.35%
	Neutral	0.90	0.90	0.91	
Bi-LSTM	Positive	0.91	0.90	0.92	
	Negative	0.94	0.94	0.95	91.56%
	Neutral	0.90	0.91	0.88	
Hybrid CNN Bi-LSTM	Positive	0.92	0.91	0.93	

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

Model	Class	F1-Score	Precision	Recall	Overall Accuracy
	Negative	0.95	0.96	0.95	92.37%
	Neutral	0.91	0.91	0.90	

By using the ratio of 80:20, all the used methods show the best accuracy, but a little bit smaller than the ratio of 90:10. The CNN method accuracy in this scenario is 92.35%, Bi-LSTM is 91.56%, and hybrid CNN Bi-LSTM got the 92.37% accuracies. The hybrid is still the best model in this scenario.

Table 10. Scenario 1 Ratio 70:30

Model	Class	F1-Score	Precision	Recall	Overall Accuracy
	Positive	0.88	0.88	0.89	
CNN	Negative	0.93	0.93	0.92	89.10%
	Neutral	0.87	0.86	0.87	
Bi-LSTM	Positive	0.88	0.87	0.89	
	Negative	0.93	0.94	0.92	88.85%
	Neutral	0.86	0.86	0.86	
Hybrid CNN Bi-LSTM	Positive	0.89	0.88	0.90	
	Negative	0.93	0.95	0.91	89.53%
	Neutral	0.87	0.87	0.88	

The ratio of 70:30 is the last ratio of scenario 1, and the accuracy is still high but lower than the two ratios used before. In this ratio, the CNN accuracy result is 89.10%, Bi-LSTM is 88.85%, and hybrid CNN Bi-LSTM is 89.53%. The hybrid model of CNN Bi-LSTM is still the best accuracy.

3.2.2. Scenario 2 (Oversampling Testing)

In scenario 1, the researcher obtained the best result using a 90:10 ratio (90% data training and 10% data test). Next, the researcher performed oversampling using SMOTE. As shown in Table 11, the overall accuracy is slightly lower than in the previous scenario, except for the CNN method. Balancing the last condition data for testing can result in lower accuracy. The results obtained for CNN are 95.36%, Bi-LSTM 93.40%, and hybrid CNN Bi-LSTM 95.20%.

Table 11. Scenario 2 Oversampling SMOTE

Model	Class	F1-Score	Precision	Recall	Overall Accuracy
	Positive	0.95	0.93	0.97	
CNN	Negative	0.97	0.98	0.97	95.36%
	Neutral	0.94	0.96	0.92	
Bi-LSTM	Positive	0.92	0.96	0.89	
	Negative	0.96	0.95	0.97	93.40%
	Neutral	0.92	0.89	0.95	
Hybrid CNN Bi-LSTM	Positive	0.95	0.95	0.95	
	Negative	0.97	0.96	0.98	95.20%
	Neutral	0.94	0.95	0.93	

3.2.3. Scenario 3 (Using Word2Vec Testing)

In the previous scenario, the accuracy obtained decreased, but not significantly. In this scenario, the researcher conducts a test by adding Word2Vec feature expansion. Firstly, the researcher builds Word2Vec by Indonesian Tweet, Indonesian News corpus, and Indonesian Tweet + News corpus. The embedding was set to 100, representing each word as a 1D vector with 100 numbers. The Word2Vec feature expansion improved the accuracy of the first two scenarios with enhanced accuracy for two methods: CNN by 95.37% and hybrid CNN Bi-LSTM by 95.64%. Table 12 shows the complete results of scenario three.

Table 12. Scenario 3 Using Word2vec

Model	Class	F1-Score	Precision	Recall	Overall Accuracy
	Positive	0.95	0.94	0.96	
CNN	Negative	0.98	0.98	0.97	95.37%
	Neutral	0.94	0.95	0.93	
Bi-LSTM	Positive	0.93	0.95	0.91	
	Negative	0.96	0.95	0.97	93.79%
	Neutral	0.93	0.91	0.94	
Hybrid CNN Bi-LSTM	Positive	0.95	0.94	0.96	
	Negative	0.98	0.98	0.98	95.64%
	Neutral	0.94	0.96	0.93	

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

The researcher obtained the top 10 similarity words from the word inputted before. In this case, the term 'Saham' is used to test this similar word. Table 13 shows the top 10 similar words.

Table 13. Top 10 Similar Words

Top 10 Similar Words of 'Saham'	TOP 1	TOP 2	TOP 3	TOP 4	TOP 5	TOP 6	TOP 7	TOP 8	TOP 9	TOP 10
	bursa	nabung	bluechip	stocksplit	investasi	tlkm	bbri	jual	deposito	unvr

As shown in Table 13, the similarities of the word 'Saham' is obtained. The words of similarity are bursa, nabung, bluechip, stocksplit, investasi, tlkm, bbri, jual, deposito, unvr.

Table 14. Accuracy and F1-Score Result

	Indonesian Tweet Corpus		Indonesian News Corpus		Indonesian Tweet+ Indonesian News Corpus	
	Overall Accuracy	F1-Score	Overall Accuracy	F1-Score	Overall Accuracy	F1-Score
TOP 1	0.950	0.907	0.951	0.908	0.951	0.908
TOP 5	0.948	0.906	0.949	0.907	0.951	0.908
TOP 10	0.950	0.908	0.947	0.908	0.950	0.907

Here is the accuracy and F1-Score result for getting the top 1, top 5, and top 10 similarities of the 'Saham' word by feature expansion Word2vec. Above, Table 14 shows the result. The accuracy of all testing by Indonesian Tweet Corpus, Indonesian News Corpus, and Indonesian Tweet + News Corpus is relatively high and delivers the best accuracy. For all, accuracy got 0.90 higher.

3.2.4. Scenario 4 (Hyperparameter Testing)

This scenario increases the parameters from the previous scenarios. This testing succeeded in increasing the accuracy from the last three scenarios, with CNN accuracy of 95.99%, Bi-LSTM accuracy of 95.82%, and CNN Bi-LSTM accuracy of 95.74%. In this final scenario, the researcher increased the epoch by 50 from the previous 10, increased the batch size from 32 to 64, increased the filter size from 128 to 256 on the Conv1D layer, and increased the number of units on the Dense layer to 256 from the previous 128. Table 15 shows the complete result of scenario four.

Table 15. Hyperparameter

Model	Class	F1-Score	Precision	Recall	Overall Accuracy
	Positive	0.96	0.95	0.97	
CNN	Negative	0.98	0.98	0.98	95.99%
	Neutral	0.95	0.96	0.94	
Bi-LSTM	Positive	0.95	0.95	0.96	
	Negative	0.98	0.98	0.98	95.82%
	Neutral	0.95	0.95	0.94	
Hybrid CNN Bi-LSTM	Positive	0.95	0.96	0.95	
	Negative	0.98	0.98	0.97	95.74%
	Neutral	0.95	0.94	0.96	

3.3. Discussion

After conducting four scenario tests, the conclusion is each scenario test affects the model's accuracy. In the first scenario, the modelling using the original dataset without feature expansion with data ratios of (0.1) 90:10, (0.2) 80:20, and (0.3) 70:30. The researcher obtained the best accuracy at a ratio of 90:10 (0.1), the split was done, which is higher than the other ratios tested. Scenario 2 used oversampling with SMOTE, and in this scenario test, the accuracy decreased slightly but not significantly because the data was already quite balanced when using scenario 1. In scenario 3, the researcher added feature expansion using Word2Vec, and good results were obtained, with the accuracy increasing significantly compared to the previous two scenarios. The researcher got the top 10 similar words in this third scenario by building Indonesian Tweet, Indonesian News, and Indonesian Tweet+News corpus. In the last scenario, the researcher got high accuracy after testing with Word2Vec and hyperparameters. Figure 6 shows the accuracy comparison.

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 936-947
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6120

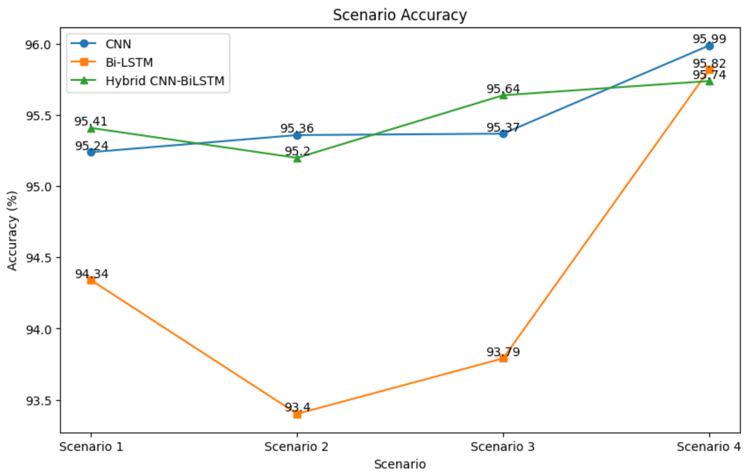

Figure 6. Scenario Comparison

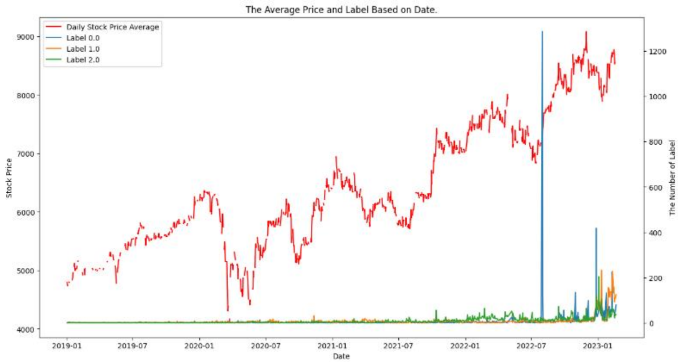

Figure 7. Visualization of Daily Stock Price Movement

As seen in Figure 7, from January 2019 to January 2021, the stock price tended to increase despite some declines. The labels also showed fluctuations but were not too significant during those months. The stock slowly began to rise in July 2021 and reached its peak in January 2023. The distribution of the three labels also moved significantly during those months of increase. However, in July 2022, the researcher observed a spike in the negative label when the condition of BBCA's stock briefly declined. Tweets were not the primary influence, as economic factors and government policies also affected the movement of stock prices. For January 2023, the three labels experienced a significant surge along with the peak of BBCA's stock price that month.

```
[ ]  text = "BBCA sedang tidak biasa"
     predicted_label = predict_label(model, tokenizer, text)
     print("Predicted Label:", predicted_label)

     1/1 [==============================] - 0s 63ms/step
     Predicted Label: 1
```

Figure 8. Labelling Result by CNN Bi-LSTM

This figure shows the predicted label using the hybrid CNN – BiLSTM method. As seen in Figure 8, the predicted result of the sentence is 1, which means that sentence has neutral sentiment. The sentence inputted before was 'BBCA sedang tidak biasa', which means BBCA is abnormal.

```
Spearman rank correlation for positive label:
 0.9875082995408913
Spearman rank correlation for negative label:
 0.988478422466455
Spearman rank correlation for neutral label:
 0.9912843598563686
```

Figure 9. Rank Spearman Correlation

The researcher assesses the relationship between the labels in the tweet data and the movement of the BBCA stock prices using the Spearman correlation rank. As shown in Figure 8, the correlation obtained in this study for the positive label is 0.987. The negative label is 0.988; for the neutral label, it is 0.991. Based on these results, according to the correlation classification by Schober et al. [23], all three labels effectively correlate with BBCA stock prices.

4. CONCLUSION

This research has made developed for sentiment analysis using the hybrid CNN and Bi-LSTM methods with Word2Vec feature expansion. In the initial scenario with a ratio of 90:10 and 80:20, both resulted in high accuracy above 90%. Different parameters and ratios have a significant impact on modelling. In scenario 2, using SMOTE oversampling to balance the data can also change accuracy and other metric performances. After adding feature expansion with Word2Vec and doing the corpus building with Indonesian Tweet Corpus, Indonesian News Corpus, and Indonesian Tweet + News Corpus, the researcher obtained the similarity word that has a connection with stock and BBCA by showing the best ten similarities of the word, the accuracy of all three models increased from the previous results. For CNN, it grew from 95.24% in the initial scenario to 95.99% in the last testing scenario with hyperparameters. For Bi-LSTM, it increased from 94.34% to 95.82%. And hybrid CNN Bi-LSTM went up from 95.41% to 95.74%. The researcher achieved the highest accuracy in scenario 4, where different parameters significantly impacted model performance. For future work, it can change more parameters such as dropout, kernel size, and regularization to prevent overfitting. The accuracy of all tested models was very high and met the researcher's expectations. For future research, researchers can use different methods, such as hybrid CNN, Bi-LSTM, and SVM, along with other feature expansion techniques, such as GloVe.

REFERENCES

[1] "PT Bursa Efek Indonesia." https://www.idx.co.id/id/produk/saham (accessed Nov. 27, 2022).
[2] A. Z. Malik, E. Utami, and S. Raharjo, "Analisis Sentiment Twitter Terhadap Capres Indonesia 2019 dengan Metode K-NN," *J. Inf. Politek. Indones. Surakarta*, vol. 5, no. 2, pp. 1–7, 2019.
[3] A. Kiky, S. Tinggi, and I. E. Wiyatamandala, "Kajian Empiris Teori Pasar Efisien (Efficient Market Hypothesis) Pada Bursa Efek Indonesia," *Maret*, vol. 6, no. 2, pp. 138–156, 2018.
[4] "Twitter Sentiment Analysis in Real-Time." https://monkeylearn.com/blog/sentiment-analysis-of-twitter/ (accessed Nov. 27, 2022).
[5] "Analisis Sentimen (Sentiment Analysis) : Definisi, Tipe dan Cara Kerjanya." https://lp2m.uma.ac.id/2022/02/21/analisis-sentimen-sentiment-analysis-definisi-tipe-dan-cara-kerjanya/ (accessed Nov. 27, 2022).
[6] M. G. Pradana, A. C. Nurcahyo, P. H. Saputro, and K. Barat, "PERUSAHAAN STIM Shanti Bhuana Sentimen Negatif merupakan opini memiliki kesan kurang baik bagi pihak yang disebut , sebaliknya sentimen positif komplain adalah bentuk ekspresi formal tentang ketidaksukaan atau ketidakpuasan terhadap beberapa aspek yang di," *Ilmiah, J. Vol, Edutic*, vol. 6, no. 2, 2020.
[7] D. Yosmita Praptiwi, "Analisis Sentimen Online Review Pengguna E-Commerce Menggunakan Metode Support Vector Machine Dan Maximum Entropy," 2018.
[8] "Kasus Pajak Bikin Saham Turun, Ini Respons Bos BCA - Saham Liputan6.com." https://www.liputan6.com/saham/read/2040307/kasus-pajak-bikin-saham-turun-ini-respons-bos-bca (accessed Nov. 30, 2022).
[9] C. Zhao, X. Huang, Y. Li, and M. Y. Iqbal, "A double-channel hybrid deep neural network based on CNN and BiLSTM for remaining useful life prediction," *Sensors (Switzerland)*, vol. 20, no. 24, pp. 1–15, 2020, doi: 10.3390/s20247109.
[10] "Cara Kerja Word2Vec. Apa itu Word2Vec? | by Afrizal Firdaus | Medium." https://medium.com/@afrizalfir/mengenal-word2vec-af4758da6b5d (accessed Nov. 30, 2022).
[11] "Advantages of word2vec | Python Natural Language Processing." https://subscription.packtpub.com/book/big-data-and-business-intelligence/9781787121423/6/ch06lvl1sec53/advantages-of-word2vec (accessed Nov. 30, 2022).
[12] W. Yue and L. Li, "Sentiment Analysis using Word2vec-CNN-BiLSTM Classification," *2020 Seventh Int. Conf. Soc. Networks Anal. Manag. Secur.*, pp. 1–5, 2020, doi: 10.1109/SNAMS52053.2020.9336549.
[13] K. Zhou and F. Long, "Sentiment analysis of text based on cnn and bi-directional LSTM model," *ICAC 2018 - 2018 24th IEEE Int. Conf. Autom. Comput. Improv. Product. through Autom. Comput.*, no. September, pp. 1–5, 2018, doi: 10.23919/IConAC.2018.8749069.
[14] S. Tam, R. Ben Said, and Ö. Tanriöver, "A ConvBiLSTM Deep Learning Model-Based Approach for Twitter Sentiment Classification," *IEEE Access*, vol. 9, pp. 41283–41293, 2021, doi: 10.1109/ACCESS.2021.3064830.
[15] U. D. Gandhi, P. Malarvizhi Kumar, G. Chandra Babu, and G. Karthick, "Sentiment Analysis on Twitter Data by Using Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM)," *Wirel. Pers. Commun.*, no. 0123456789, 2021, doi: 10.1007/s11277-021-08580-3.
[16] 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.
[17] I. S. Program, "Aspect-Based Sentiment Analysis on Twitter Using Long Short-Term Memory Method," vol. 7, pp. 1–10, 2023, doi: 10.30865/mib.v5i1.2293.
[18] G. Kamil and E. B. Setiawan, "Aspect-Level Sentiment Analysis on Social Media Using Gated Recurrent Unit (GRU)," vol. 99, no. 99, pp. 1–9, 2023, doi: 10.47065/bits.v9i9.999.
[19] Z. Wang and Z. Qu, "Research on Web text classification algorithm based on improved CNN and SVM," *Int. Conf. Commun. Technol. Proceedings, ICCT*, vol. 2017-Octob, pp. 1958–1961, 2018, doi: 10.1109/ICCT.2017.8359971.
[20] O. Setyawan and H. F. Pardede, "Sentiment analysis for event-based stock price predictions using bidirectional long short

term memory," vol. 6, no. 1, pp. 50–58, 2022, doi: 10.52362/jisicom.v6i1.772.
[21] M. A. Cifci, S. Hussain, and P. J. Canatalay, "Hybrid Deep Learning Approach for Accurate Tumor Detection in Medical Imaging Data," *Diagnostics*, vol. 13, no. 6, p. 1025, 2023, doi: 10.3390/diagnostics13061025.
[22] A. K. J and S. Abirami, "Aspect-based opinion ranking framework for product reviews using a Spearman's rank correlation coefficient method," *Inf. Sci. (Ny).*, vol. 460–461, pp. 23–41, 2018, doi: 10.1016/j.ins.2018.05.003.
[23] P. Schober and L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation," *Anesth. Analg.*, vol. 126, no. 5, pp. 1763–1768, 2018, doi: 10.1213/ANE.0000000000002864.

Submitted: **01/05/2023**; Accepted: **11/07/2023**; Published: **23/07/2023**