Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

El Niño Sentiment Analysis Using Recurrent Neural Network and Convolutional Neural Network Use GloVe

Denada Putrisia*, Kemas Muslim Lhaksmana

School of Computing, Informatics, Telkom University, Bandung, Indonesia
Email: [1,*]denadaputrisia@student.telkomuniversity.ac.id, [2]kemasmuslim@telkomuniversity.ac.id
Correspondence Author Email: denadaputrisia@student.telkomuniversity.ac.id
Submitted: **04/06/2024**; Accepted: **29/06/2024**; Published**: 30/06/2024**

Abstract−Sentiment analysis regarding the El Niño climate change is a crucial aspect in understanding public perception and response. This enables deeper identification and understanding of the sentiments evident in online conversations. Sentiment analysis through deep learning approaches using Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) methods is conducted. RNN is a type of artificial neural network designed to process sequential data such as text or time. On the other hand, CNN utilizes convolutional layers to scan text with filters to capture local features like phrases and keywords determining sentiment. Leveraging GloVe representation technique enables the representation of words in numerical vector form capturing semantic relationships among words, facilitating sentiment analysis related to El Niño on social media. The aim of this study is to evaluate the performance of RNN and CNN methods in classifying El Niño-related sentiment with and without GloVe word representation, and to develop a model that can provide accurate and reliable sentiment analysis results. The contribution of this research indicates that the accuracy of sentiment analysis has been improved and can be a significant reference for further research in the field of text analysis and natural language processing (NLP). This study also emphasizes the crucial role of word representation techniques like GloVe in enhancing the performance of deep learning models. The results of the study indicate that the RNN and CNN methods with the utilization of GloVe provide better sentiment classification related to the El Niño issue in social media data, showing that the use of RNN and CNN models with GloVe features perform better compared to not using GloVe features. The use of the RNN algorithm with 80:20 split ratio testing produced an accuracy score of 94.90%, recall of 94.90%, precision of 94.94%, and F1-Score of 94.85%. Meanwhile, the use of the CNN algorithm with 90:10 split ratio testing produced an accuracy score of 94.61%, recall of 93.61%, precision of 94.69%, and F1-Score of 94.58%. This results in the conclusion that sentiment analysis using RNN modeling with GloVe features has better performance than CNN modeling, with an average accuracy rate of 94.90%.

Keywords: Sentiment Analysis; El Niño; RNN; CNN; Glove

1. INTRODUCTION

El Niño is a natural phenomenon that has significant impacts on weather and climate across various parts of the world. Extreme weather caused by El Niño leads to droughts and water shortages, affecting staple food production. This extreme climate phenomenon once struck Indonesia, resulting in a 14-month drought (March 1997 to April 1998), causing many areas to experience water deficits [1].

Social media sentiment analysis can provide valuable insights into how the public understands climate change issues related to El Niño. Numerous previous studies have used sentiment analysis to understand public reactions to specific issues of social media. However, the climate change issue related to El Niño may not have been specifically researched, making this study potentially valuable. In recent years, advances in deep learning have produced highly accurate models for text analysis. This study compares various deep learning techniques for sentiment analysis using Twitter data. Deep learning (DL) approaches are becoming increasingly popular among academics in this field because they help solve various challenges simultaneously. Two types of neural networks are specifically used: Recurrent Neural Networks (RNN), which have been successful in natural language processing (NLP) applications, and Convolutional Neural Networks (CNN), which excel in image processing. This study evaluates and differentiates between the ensemble and combination of CNNs with a specific type of RNN known as Long Short-Term Memory (LSTM) [2].

Numerous previous studies on sentiment analysis have used various algorithms, including deep learning and machine learning. For example, Peng Cen, Kexin Zhang, and Desheng Zheng conducted a study on sentiment analysis using deep learning algorithms such as RNN, CNN, and LSTM. Their results showed that CNN and neural network models achieved good classification effects when applied to film sentiment analysis. CNN reported an accuracy of 88.22%, while RNN and LSTM achieved accuracies of 68.64% and 85.32%, respectively [3]. Another study by Abdul Mohaimin Rahat, Abdul Kahir, and Abu Kaisar Mohammad Masum compared Naïve Bayes and SVM algorithms for sentiment analysis. In their tests, SVM achieved an accuracy of 83%, whereas Naïve Bayes achieved 77% [4].

Hasan M and Setiawan E conducted a study [5]. Sentiment analysis on Twitter was used to monitor the stock price movements of Bank Central Asia (BBCA) to find its correlation. The highest accuracy achieved was 76.29%, produced by the hybrid RNN-CNN method with the highest corpus similarity of tweets. This study contributes by improving upon previous research by proposing sentiment analysis using a hybrid deep learning model of RNN and CNN with Global Vectors (GloVe) feature expansion. The dataset used was sourced from Twitter concerning Bank Central Asia (BBCA). Another study by Uly N, Hendry H, and Iriani A proposed a hybrid CNN-RNN model for diagnosing COVID-19 from X-Ray images, leveraging the combination of CNN and RNN. The methods used were CNN (ResNet50) and RNN (LSTM). The dataset consisted of 1000 images, with 500 images used as the COVID-19

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

dataset and 500 as the normal dataset. The ResNet152V2-LSTM model with GRU achieved an accuracy of 93.37% [6]. Research conducted by Safrizal Ahmad and his colleagues analyzed sentiment on "Tools & Home Improvement" product reviews on the Amazon platform using Deep Learning models, namely Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). The dataset utilized in this study comprises "Tools & Home Improvement" data, which consists of information about tools and household equipment available on Amazon. Previous research that combined CNN and LSTM architectures achieved better results by approximately 2.7%-8.5% compared to conventional models. Additionally, another study used the LSTM model to analyze Amazon product review datasets and achieved an accuracy of 93.66%, surpassing other models [7].

The research conducted by Peng Cen, Kexin Zhang, and Desheng Zheng, as well as Hasan M and Setiawan E, focuses on the application of deep learning algorithms such as RNN, CNN, and LSTM in sentiment analysis across various domains, including film sentiment analysis and stock price prediction. However, this study specifically concentrates on sentiment analysis related to El Niño and its impact on social media. Additionally, the research by Uly N, Hendry H, and Iriani A introduces a hybrid CNN-RNN model for diagnosing COVID-19 from X-Ray images. However, this study is not directly related to sentiment analysis on the issue of El Niño climate change. Hasan M and Setiawan E's study utilizes GloVe to enhance the accuracy of sentiment analysis on Bank Central Asia (BBCA) stock prices from Twitter data. Nevertheless, the focus of that research is not on sentiment analysis related to El Niño climate change. Previous studies do not specifically mention the use of GloVe in sentiment analysis, whereas this study highlights the advantages of utilizing GloVe to improve the performance of deep learning models. Previous research primarily focuses on sentiment analysis for purposes such as predicting stock price movements or diagnosing diseases, while this research aims to develop a deep learning model that can identify public sentiment and opinions related to the El Niño climate change issue on social media. This study seeks to provide insights into how the public perceives this issue and how these perceptions evolve over time, as well as to raise public awareness about climate change and the impact of El Niño. In addition, the results can offer valuable insights for policymakers, environmental scientists, and nongovernmental organizations sensitive to environmental issues.

Based on the results of previous studies, this research contributes by improving upon previous work by proposing sentiment analysis using a hybrid deep learning model of RNN and CNN with the expansion of Global Vectors (GloVe) [5]. This study focuses on the correlation results and influence of Twitter user sentiment on the El Niño Climate Change Issue. The accuracy levels of RNN and CNN are higher than those of other methods. To achieve high accuracy, the RNN (Recurrent Neural Networks) and CNN (Convolutional Neural Network) models were chosen with the expansion of Global Vectors to produce more accurate sentiment analysis. CNN is used for visual data analysis, while RNN can understand the context and sequence of data. Therefore, the combination of both can help analyze the emotional impact of Twitter users on opinions related to the El Niño Climate Change Issue.

2. RESEARCH METHODOLOGY

2.1 Research Stages

From the results of the study, a system was developed with the ability to classify sentiment based on public opinion surveys related to El Niño. The dataset used was sourced from Twitter and compared using RNN and CNN algorithms with the use of GloVe to determine their accuracy. Below is the operational flowchart of the system, as shown in Figure 1.

Based on Figure 1, the first step in planning this sentiment analysis is data collection through Twitter crawling. Once the data crawling process is completed, the acquired dataset is prepared for further implementation. Next, data labeling is conducted to determine the number of data labeled as positive, negative, and neutral. After data labeling, the data preprocessing process is performed, involving data cleaning, case folding, stemming, stopword removal, normalization, and tokenization. The next step is feature extraction. Subsequently, feature expansion is carried out. Then, the data is split. The data is divided into two parts, namely test data and train data. The next step is to train the RNN and CNN algorithm models using the train and test data to produce a sentiment model. From this sentiment model, a confusion matrix is generated, from which accuracy, recall, precision, and F1-Score values are obtained.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

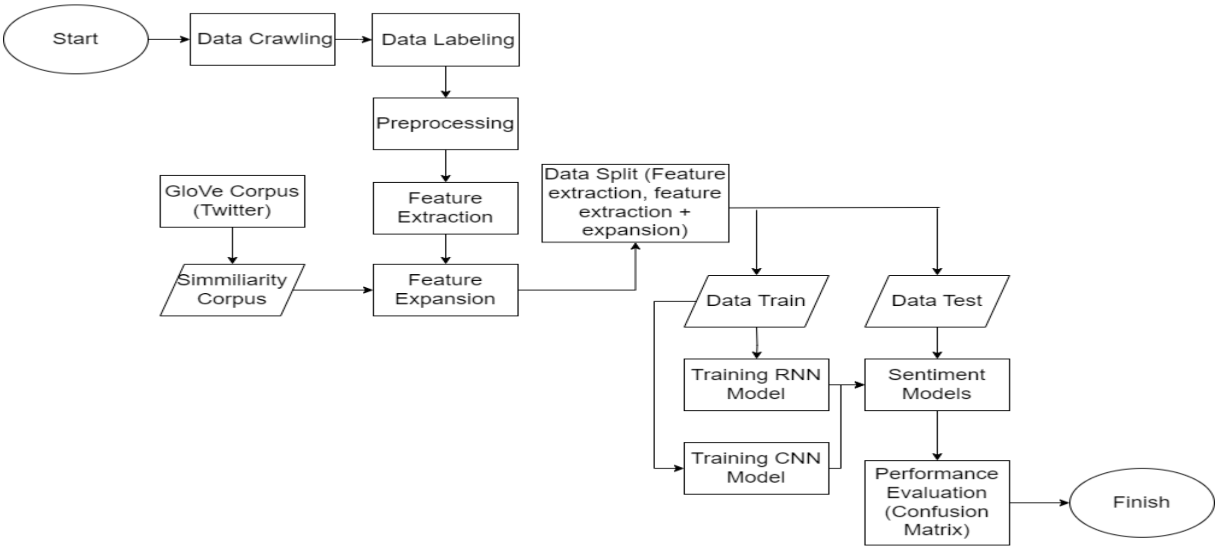

Figure 1. Flowchart of the Developed System

2.2 Data Crawling

Data collection (Data Crawling) is the process of retrieving information or data from the social media platform Twitter. The data obtained from the crawling process will be used as the dataset for this study. This process uses the Application Programing Interface (API) from the Twitter domain to collect tweet data based on selected keywords. The data extracted from this dataset comprise tweets about public opinions and views on the issue of El Niño climate change. A total of 6016 data points were used as the dataset. The keywords used to search for tweet data are listed in Table 1.

Table 1. Crawling Dataset

Keyword	Total	Keyword	Total
cuaca elnino	596	Siklus El Niño	235
ElNino	204	#ElNiño	273
dampak elnino	283	El Niño dan kebakaran hutan	523
fenomena El Niño	251	El Niño dan kekeringan	144
efek ElNino	394	elnino	656
sebab elnino	367	cuaca ElNino	709
imbas El Nino	446	iklim elnino	186
karena elnino	368	pengaruh elnino	85
Perubahan iklim El Niño	296		
Total data			**6016**

In Table 1, there are several keywords such as "cuaca elnino", "fenomena El Niño", "imbas El Nino", "iklim elnino", and "pengaruh elnino". These keywords are determined by seeking their coverage and by searching for highly specific keywords to increase the number of available tweets for that keyword.

2.3 Data Labeling

Data labeling is a part of sentiment analysis that assigns labels or categorizes text from social media into specific sentiment categories (e.g., positive, negative, neutral related to the issue of El Niño climate change). Below, the categories and total labeled data are presented in Table 2.

Table 2. Total Data Labels

Label	Total	Percentage (%)
Positive	528	8,77
Negative	228	3,78
Neutral	5260	87,4
Total Data	**6016**	**100**

In the above Table 2, it can be concluded that A positive label is assigned if the data contains positive statements such as "El Niño increases rainfall in usually dry areas" or "El Niño reduces air pollution," which have meanings that refer to positive sentiments like favorable reviews of the El Niño phenomenon. Conversely, a negative label is given to tweet data containing negative statements such as "El Niño damages marine ecosystems with warmer water temperatures" or "El Niño increases the risk of forest fires in Southeast Asia." A neutral label is applied to data that

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

does not have a measurable sentiment, either positive or negative. This is usually seen in tweets containing questions, such as "How often does El Niño occur?" or "Climate change is a global challenge," and so on.

2.4 Data Preprocessing

The data obtained from Twitter contains a lot of information that is difficult for the system to process. Therefore, the data must be preprocessed to be used effectively [8]. Data preprocessing involves several steps, such as data cleaning, converting text to lowercase (case folding), removing stopwords, stemming, normalization, and tokenization.

a. Data Cleaning

Data cleaning is a crucial step in data preprocessing that identifies, corrects, and removes inaccuracies and incompleteness in the dataset. Special characters and terms are also removed at this stage [9]. This process is illustrated in Table 3.

Table 3. Data Cleaning

Before Data Cleaning	After Data Cleaning
Repost @Kementerian Pertanian SobaTani, pemerintah terus berupaya dalam meningkatkan produksi padi secara nasional. Salah satu langkah cepat dalam menghadapi El Nino. https://t.co/RM4bLDzsTZ #NotaKesepahaman #Pompanisasi #ElNino #KemeterianPertanian https://t.co/3mjwK6TS8p [10], [11]. (Repost @Ministry of Agriculture SobaTani, the government continues to strive to increase rice production nationally. One of the quick measures taken to face El Niño is as follows. https://t.co/RM4bLDzsTZ #MemorandumOfUnderstanding#Pumping #ElNino#MinistryOfAgriculture https://t.co/3mjwK6TS8p [10], [11].)	Repost Pertanian SobaTani pemerintah terus berupaya dalam meningkatkan produksi padi secara nasional Salah satu langkah cepat dalam menghadapi El Nino [10], [11]. (Repost Agriculture SobaTani, the government continues to strive to increase national rice production. One of the quick measures taken to face El Niño [10], [11].)

Based on Table 3, data is sorted by removing white space, symbols, and specific characters in order to reduce noise and make the text easier to read.

b. Case Folding

Case folding is a preprocessing step that standardizes all characters in the text to lowercase. This process converts all character formats in the dataset to lowercase [12]. For example, the word "Repost" is changed to "repost", and "Elnino" is changed to "elnino". This process is shown in Table 4.

Table 4. Case Folding

Before Case Folding	After Case Folding
Repost Pertanian SobaTani pemerintah terus berupaya dalam meningkatkan produksi padi secara nasional Salah satu langkah cepat dalam menghadapi El Nino [10], [11]. (Repost from Agriculture SobaTani, the government continues to strive to increase rice production nationally. One quick step in facing El Nino [10], [11].)	repost pertanian sobatani pemerintah terus berupaya dalam meningkatkan produksi padi secara nasional salah satu langkah cepat dalam menghadapi el nino [10], [11]. (repost from agriculture sobatani, the government continues to strive to increase rice production nationally. one of the quick steps in facing el nino [10], [11].)

Based on Table 4, By converting all characters to a single case, whether lowercase or uppercase, case folding ensures consistency in text representation and eliminates potential discrepancies that may arise due to variations in case sensitivity. This makes it easier for algorithms to process and analyze text data accurately, regardless of the original casing used in the input text.

c. Stemming

Stemming is a technique used in information retrieval to address lexical mismatch issues in which the query term does not match the document term [13]. For example, the words "pertanian", "berupaya", and "meningkatkan" can be reduced to their root forms "tani", “upaya”, and “tingkat” through stemming. This process is shown in Table 5.

Table 5. Stemming

Before Stemming	After Stemming
repost pertanian sobatani pemerintah terus berupaya dalam meningkatkan produksi padi secara nasional salah satu langkah cepat dalam menghadapi el nino [10], [11].	['repost', 'tani', 'sobatani', 'perintah', 'upaya', 'tingkat', 'produksi', 'padi', 'nasional', 'salah', 'langkah', 'cepat', 'hadap', 'el', 'nino'] [10], [11].

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

(Repost Agriculture SobaTani, the government continues to strive to increase rice production nationally. one of the quick steps in facing el nino [10], [11].)	(['repost', 'agriculture', 'sobatani', 'government', 'effort', 'increase', 'production', 'rice', 'national', 'one', 'quick', 'step', 'face', 'el', 'nino'] [10], [11].)

Based on Table 5, This process helps in standardizing words with similar meanings, which may vary due to different suffixes, prefixes, or inflections. For example, the words "pertanian", "berupaya", and "meningkatkan" can be reduced to their root forms "tani", “upaya”, and “tingkat” through stemming.

d. Stopword Removal

Stopword removal is a process in natural language processing aimed at removing frequently occurring words that contribute little to the understanding of the text [14]. Examples of stopwords include conjunctions and common words. This process is shown in Table 6.

Table 6. Stopword Removal

Before Stop Removal	After Stop Removal
['repost', 'tani', 'sobatani', 'perintah', 'upaya', 'tingkat', 'produksi', 'padi', 'nasional', 'salah', 'langkah', 'cepat', 'hadap', 'el', 'nino'] [10], [11].	repost, pertanian, sobatani, pemerintah, berupaya, meningkatkan, produksi, padi, nasional, salah, langkah, cepat, menghadapi, el, nino [10], [11].
(['repost', 'agriculture', 'sobatani', 'government', 'efforts', 'increase', 'production', 'rice', 'national', 'one', 'quick', 'steps', 'face', 'el', 'nino'] [10], [11].)	(repost, agriculture, sobatani, government, efforts, increase, production, rice, national, one, quick, steps, face, el, nino [10], [11].)

Based on Table 6, the stopword removal stage was used to remove from the data common words like "ini," "adalah," "dan," "untuk," and others that don't have much meaning.

e. Normalization

In this phase, the process of normalizing language is conducted by addressing each word that contains non-blank characters, such as noise, and transforming them into blank and standard words. Inaccurate and noisy words are those containing regional dialects that do not conform to the Kamus Besar Bahasa Indonesia (KBBI) and colloquial language used in social media. For example, a term that frequently appears on social media is "tdk," but over time, it becomes less common and is replaced by "tidak." [15]. This process is illustrated in Table 7.

Table 7. Normalization

Before Normalization	After Normalization
repost, pertanian, sobatani, pemerintah, berupaya, meningkatkan, produksi, padi, nasional, salah, langkah, cepat, menghadapi, el, nino [10], [11].	repost, pertanian, sobatani, pemerintah, terus, berupaya, dalam, meningkatkan, produksi, padi, secara, nasional, salah, satu, langkah, cepat, dalam, menghadapi, el, nino [10], [11].
(repost, agriculture, sobatani, government, strives, to increase, rice, production, nationally, one, quick, step, in facing, el nino [10], [11].)	(repost, agriculture, sobatani, government, continues, to strive, in, increasing, rice, production, nationally, one, quick, step, in, facing, el nino [10], [11].)

Based on Table 7, this process ensures that words that may not be recognized or understood by language processing algorithms are converted into recognizable and commonly used words. In the example provided, normalization replaces OOV words like "repost," "pertanian," "sobatani," and "pemerintah" with their corresponding common words "repost," "agriculture," "government," and "continues," respectively.

f. Tokenization

Tokenization is the process of turning a word in a table, paragraph, or passage into a token, potongan word, or termmed word that is unique to itself is called tokenization [16]. The results of this process are shown in Table 8.

Table 8. Tokenization

Before Tokenization	After Tokenization
repost, pertanian, sobatani, pemerintah, terus, berupaya, dalam, meningkatkan, produksi, padi, secara, nasional, salah, satu, langkah, cepat, dalam, menghadapi, el, nino [10], [11].	repost, pertanian, sobatani, pemerintah, terus, berupaya, dalam, meningkatkan, produksi, padi, secara, nasional, salah, satu, langkah, cepat, dalam, menghadapi, el, nino [10], [11].
(repost, agriculture, Sobatani, government, continuously, strives, in, increasing, rice, production, nationally, one, quick, step, in, facing, el nino [10], [11].)	(repost, agriculture, sobatani, government, continues, to strive, in, increasing, rice, production, nationally, one, quick, step, in, facing, el nino [10], [11].)

Based on Table 8, This process divides the text into smaller units, making it easier to analyze and process. For instance, the phrase "repost, pertanian, sobatani, pemerintah" is tokenized into separate words: "repost," "agriculture," "Sobatani," and "government”.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

2.5 Feature Extraction

Feature extraction is a step that scans the features contained in text documents. This stage is a crucial factor in document processing because it greatly determines the success rate of a text mining process [17]. Term Frequency-Inverse Document Frequency (TF-IDF) has become a widely used feature extraction method that can evaluate the significance of vocabulary within a document. Term Frequency (TF) is the ratio of the number of times a word appears in a document to the total number of words in that document. Conversely, Inverse Document Frequency (IDF) is a value calculated or weighted for a word. Words such as "is", "and", and "or" are examples of frequently occurring words that do not have significant meaning [5]. When the IDF value is high, the word becomes less important. Below is the formula for calculating the IDF value.

$$TF - IDF(t, d) = tf_{t,d} \text{ x} \log\left(\frac{N}{df_t}\right) \tag{1}$$

In equation 1 above, it can be explained that $tf_{t,d}$ is the total count of word t in document d. N represents the total number of documents in the dataset. df_t is the number of documents containing the word t.

2.6 Feature Expansion

Feature expansion is a method of creating or adding new functionalities to an existing dataset. One form of feature expansion model is Global Vectors (GloVe). GloVe is a word embedding technique that produces vector representations of words based on the distribution of words in a text corpus, considering the relationships between words in a vector space. This method combines ideas from two common approaches to word embedding: co-occurrence matrix-based methods and neural network-based learning. The expansion feature used in this study is called Global Vector (GloVe). GloVe has a working method for creating a matrix that can be mentioned a few times every time a new word appears. GloVe will produce an output list with similarity terms that will later be expanded with respect to vector models [18].

2.7 Data Splitting

Data splitting is a process in machine learning that involves dividing the dataset into different subsets for training and testing purposes. The goal of data splitting is to avoid overfitting and to validate the model's performance objectively. Part of the training data will be used to train the model, while the test data will be used to evaluate the model’s final performance after training.

2.8 Recurrent Neural Network (RNN)

Recurrent Neural Networks (RNNs) come in various forms. One form is global, which uses a standard multilayer perceptron (MLP) with additional loops. Thus, this stage can leverage the nonlinear mapping performance of MLP [17]. In the RNN architecture, the network only utilizes a single layer during training, which is the tanh layer [19]. Pooling layers cause the loss of detailed local information because they only capture the most important features in the pool and ignore other features. This study uses an RNN model with two simple RNN layers: the first layer includes return sequences, dropout, and recurrent dropout, whereas the second layer does not have return sequences. These layers are connected to several supporting layers, including two dense layers and a batch normalization layer.

2.9 Convolutional Neural Network (CNN)

One type of neural network commonly used with image data is the Convolutional Neural Network (CNN). CNNs are capable of detecting and identifying objects in specific images. CNNs are divided into two main parts: the fully connected layer and the feature extraction layer. The two components of the feature extraction layer are the pooling layer and the convolutional layer. Convolutional Layer: The convolutional layer uses the principles of stride and sliding window to reduce the complexity of the human body's structure [19]. The CNN model layers used include SpatialDropout1D, followed by two Conv1D layers with 60 and 15 filters, two MaxPooling1D layers with a pool size of 2, two dropout layers, and one flattening layer with three unit layers and dense layers (as this study uses three different emotion labels) [20].

2.10 Performance Evaluation (Confusion Matrix)

The Confusion Matrix is a method for evaluating the performance of a classification model by comparing the model’s predicted results with actual test data. The Confusion Matrix assesses four parameters: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). This performance evaluation results in the F1 score and accuracy [21].

3. RESULT AND DISCUSSION

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

The section provides an overview of the experimental results obtained from testing different models in sentiment analysis research, specifically focusing on Twitter data related to El Niño. The primary focus of the study lies in comparing the effectiveness of Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) algorithms in handling diverse data with various data split ratios.

3.1 Testing Results

Model testing is a critical stage in sentiment analysis research, especially in the context of diverse Twitter data related to El Niño. In this study, the modeling approach using RNN and CNN algorithms was the main focus to test their effectiveness in handling different data with varying data split propositions. The evaluation of this study was conducted after applying RNN and CNN modeling techniques to compare three different data split propositions: 90:10, 80:20, and 70:30. Data splitting generates two subsets of the dataset: one subset is used to train the model (training set), and the other is used to evaluate the trained model's performance (testing set). This study determines how well a model can generalize to new data. The data splitting ratio is crucial for balancing training efficiency and model generalization. The evaluation's findings include the F1 score, recall, accuracy, and precision determined by calculating the weighted average section. The study's results include evaluations of accuracy, recall, precision, F1-score, and support, aiming to achieve the best results for the classification model. The following are the testing results of RNN and CNN modeling using Glove features and without Glove features, as shown in Table 9.

Table 9. Testing Results

Feature	Model	Data Split Ratio	Accuracy	Recall	Precision	F1-Score
Using GloVe	CNN	90:10	**94.61%**	94.61%	94.69%	94.58%
		80:20	92.84%	92.84%	92.81%	92.80%
		70:30	93.71%	93.71%	93.75%	93.64%
	RNN	90:10	93.98%	93.98%	94.18%	93.90%
		80:20	**94.90%**	94.90%	94.94%	94.85%
		70:30	93.56%	93.56%	93.67%	93.48%
Without GloVe	CNN	90:10	94.49%	94.49%	94.48%	94.48%
		80:20	**94.58%**	94.58%	94.78%	94.54%
		70:30	93.98%	93.98%	94.29%	93.90%
	RNN	90:10	**92.78%**	92.78%	93.27%	92.65%
		80:20	92.17%	92.17%	92.72%	92.00%
		70:30	91.51%	91.51%	91.99%	91.31%

Based on the data split testing results using GloVe features in Table 9, it can be concluded that for CNN modeling with a 90:10 ratio, the best results were achieved with an accuracy score of 94.61%, recall 94.61%, precision 94.69%, and F1-Score 94.58%. In contrast, RNN modeling with an 80:20 ratio showed the best testing results with an accuracy score of 94.90%, recall 94.90%, precision 94.94%, and F1-Score 94.85%. It can be concluded that the data split testing results using RNN modeling with GloVe features yielded the best results with the highest accuracy score of 94.90%. The testing results of RNN and CNN modeling without GloVe features showed lower evaluation values such as accuracy, precision, recall, and F1-Score. From the data split testing results without GloVe features, it can be concluded that there is a decrease in accuracy, recall, precision, and F1-Score when testing the two models without GloVe features. The best results obtained from this testing were for CNN modeling with an 80:20 ratio, achieving an accuracy score of 94.58%, recall 94.58%, precision 94.78%, and F1-Score 94.54%. Meanwhile, the RNN model testing with a 90:10 ratio only reached an accuracy score of 92.78%, recall 92.78, precision 93.27%, and F1-Score 92.65%. From the two model tests, RNN and CNN using GloVe features or without GloVe features, it can be concluded that the GloVe feature significantly helps in obtaining the best results from these two models. This is evidenced in Table 9, where GloVe can improve accuracy as the model can utilize more accurate and meaningful word representations.

This research also revealed the importance of using GloVe features to enhance model performance. The testing results indicated that models using GloVe features tended to perform better than those without these features. This is because GloVe enables the model to understand the meaning of words more deeply, thereby enhancing its ability to classify sentiments more accurately. However, there are several challenges that need to be addressed in this research. For example, although the RNN model showed better overall performance, there were some cases where the CNN model still managed to compete, especially in the 90:10 data split ratio. This suggests that other factors should be considered in model development, such as network architecture and data preprocessing techniques. The following are two comparisons of the highest and lowest accuracies when using GloVe features. In Figure 2 dan 3 below depict the accuracy plots of the best-performing RNN model, which utilized an 80:20 ratio, and the model with the lowest accuracy using GloVe features, namely the CNN model with an 80:20 ratio.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

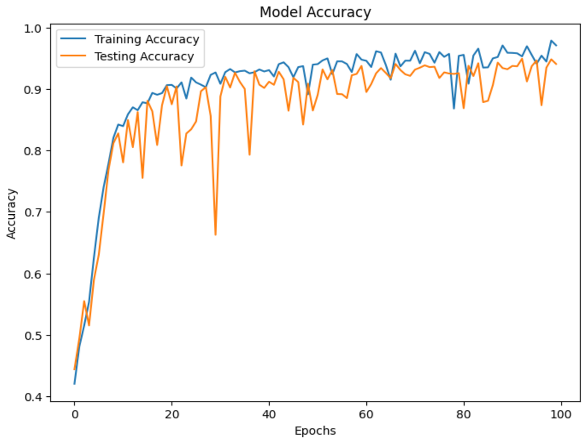

Figure 2. RNN Accuracy (80:20)

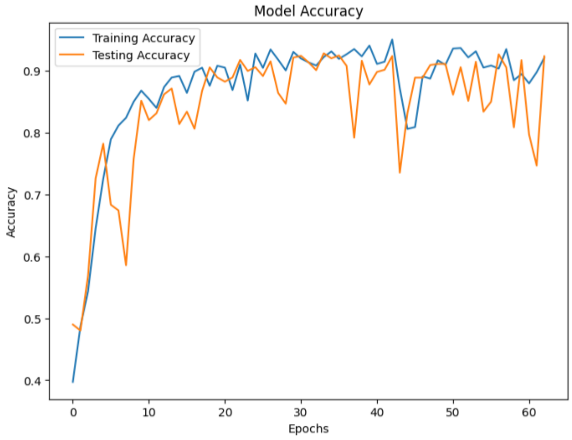

Figure 3. CNN Accuracy (80:20)

Based on Figures 2 and 3, each model underwent 300 epochs of training, and each training session experienced a significant improvement. Running 300 epochs resulted in fluctuations in the accuracy graph. From the results of the 300 epochs, it was found that the RNN modeling achieved the highest accuracy with a ratio of 80:20, whereas the lowest accuracy was observed in the CNN modeling with the same ratio. Understanding these two graphs indicates that each model experiences gradual improvement as the training session progresses, showcasing the effectiveness of the model training process in enhancing their performance. However, fluctuations in the accuracy graph suggest variations in model performance during the training process, which are influenced by factors such as dataset size, model complexity, and hyperparameter settings. It is noteworthy that RNN modeling demonstrated the highest accuracy in these experiments, implying its suitability for the particular task or dataset employed in this research. However, it is essential to recognize that model performance is influenced not only by the type of model architecture but also by various other factors, including data processing methodologies, hyperparameter optimization techniques, and the training strategies employed.

Furthermore, a comparative analysis of the modeling accuracies of RNN and CNN at the same ratio provides valuable insights into the relative strengths of each architecture. Despite RNN modeling exhibiting superior accuracy, it does not diminish the value or effectiveness of CNN architectures. Each architecture possesses unique advantages and drawbacks, necessitating careful consideration of the dataset characteristics and research objectives when selecting the appropriate architecture. In conclusion, the findings of this experiment contribute to a deeper understanding of the relative performance of RNN and CNN models in classification tasks. Further investigation into the factors influencing model performance, coupled with exploration of potential enhancement strategies, constitutes promising avenues for future research endeavors.

3.2 Sentiment Result

The following analysis delves into the sentiment results: in the form of a pie chart, this is explained in Figure 4.

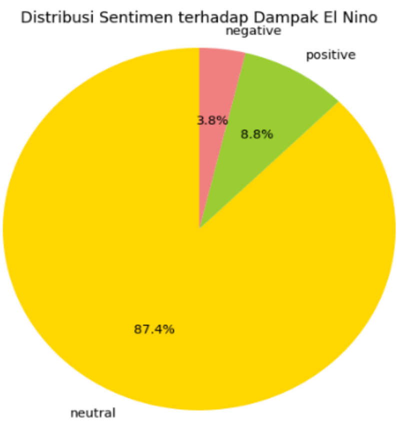

Figure 4. Sentiment Analysis Results

Based on Figure 4, the diagram above is a pie chart illustrating the distribution of sentiment toward the impact of El Nino. This diagram divides sentiment into three categories: positive, negative, and neutral. Based on the data presented in this diagram, most respondents have a neutral sentiment toward the impact of El Nino, with a percentage reaching 87.4%. This neutral sentiment indicates that most respondents do not have strong opinions or may not feel significantly affected by the El Nino phenomenon. Meanwhile, positive sentiment toward the impact of El Nino only

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

accounted for 8.8% of the total respondents. This positive sentiment may reflect the view that El Nino brings some benefits or specific opportunities although this group is a minority compared with neutral sentiment. Conversely, negative sentiment toward the impact of El Nino was recorded at 3.8%. This negative sentiment likely comes from respondents who have experienced or anticipate the adverse effects of the El Nino phenomenon, such as droughts, floods, or other climate disturbances that can affect daily life and the economy.

This distribution of sentiment provides important insights into public perceptions of the impact of El Nino. With the majority being neutral, it can be concluded that while most people may not feel directly significant impacts, there is a small group experiencing both positive and negative impacts. This information can serve as a basis for researchers and policymakers to formulate more effective communication strategies and interventions in addressing the impact of El Nino, as well as to increase awareness and understanding of this climate phenomenon among the public. Furthermore, further analysis is needed to understand the factors influencing these perceptions and sentiments, including demographic, geographic, and socioeconomic variables.

3.3 Testing Results Analysis

The analysis of the testing results was conducted using RNN and CNN modeling with and without GloVe features. Comparing these two experiments, it can be concluded that modeling with GloVe features is superior to modeling without GloVe features. The testing used three ratios: 90:10, 80:20, and 70:30, which showed significant differences in the results for the two models. The highest performance testing result was with an 80:20 split ratio for RNN modeling using GloVe features. Next, the highest result for CNN modeling was with a 90:10 split ratio among the three tested ratios. Testing was conducted across three data partitioning ratios: 90:10, 80:20, and 70:30. The outcomes exhibit noteworthy disparities in the performance of both models contingent on the data split ratios employed. Apex performance was attained at an 80:20 ratio for RNN modeling with GloVe features. This underscores the notion that a harmonized distribution of training and testing data fosters enhanced performance in the RNN model. Conversely, the optimal outcome for CNN modeling materializes at a split ratio of 90:10, suggesting that CNN modeling may exhibit heightened sensitivity to diverse data distributions. Table 10 displays the confusion matrix for the RNN modeling with the highest accuracy rate, whereas Table 11 presents the confusion matrix results for the CNN modeling with the highest accuracy rate among the three tested ratios. Here are the explanations for Tables 10 and 11. The section also includes comparative analyses, illustrated with figures and confusion matrices (Tables 10 and 11), providing insights into the relative strengths of RNN and CNN architectures. Despite RNN models exhibiting superior accuracy, CNN architectures remain valuable, each possessing unique advantages. Table 10 below the results of the confusion matrix from RNN modeling with an 80:20 ratio.

Table 10. RNN Confusion Matrix (80:20)

		Prediction		
		Positive	Neutral	Negative
Actual	Positive	1017	35	0
	Neutral	79	932	41
	Negative	0	6	1046

Based on Table 10, this confusion matrix, the following conclusions can be drawn:

a. **True Positive (TP):** The number of instances correctly predicted as positive. In this case, there are 1017 true positives.
b. **True Neutral (TN):** The number of instances correctly predicted as neutral. In this table, there are 932 true neutrals.
c. **True Negative (TN):** The number of instances correctly predicted as negative. Here, there are 1046 true negatives.
d. **False Positive (FP):** The number of instances incorrectly predicted as positive. The matrix shows 79 false positives (0 for actual negative and 79 for actual neutral).
e. **False Neutral (FN):** The number of instances incorrectly predicted as neutral. There are 41 false neutrals (35 for actual positive and 6 for actual negative).
f. **False Negative (FN):** The number of instances incorrectly predicted as negative. According to the table, there are 41 false negatives.

Table 11 provides an explanation of the confusion matrix from CNN modeling with a 90:10 ratio.

Table 11. CNN Confusion Matrix (90:10)

		Prediction		
		Positive	Neutral	Negative
Actual	Positive	511	14	1
	Neutral	49	465	12
	Negative	0	9	517

Based on Table 11, this confusion matrix, the following conclusions can be drawn:

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

a. **True Positive (TP):** The number of instances correctly predicted as positive. In this case, there are 511 true positives.
b. **True Neutral (TN):** The number of instances correctly predicted as neutral. In this table, there are 465 true neutrals.
c. **True Negative (TN):** The number of instances correctly predicted as negative. Here, there are 517 true negatives.
d. **False Positive (FP):** The number of instances incorrectly predicted as positive. The matrix shows 14 false positives (1 for actual negative and 13 for actual neutral).
e. **False Neutral (FN):** The number of instances incorrectly predicted as neutral. There were 61 false neutrals (49 foractual positive and 12 for actual negative).
f. **False Negative (FN):** The number of instances incorrectly predicted as negative. According to the table, there are 9 false negatives.

4. CONCLUSION

In this study, we used RNN and CNN algorithms to perform sentiment analysis on Twitter data related to El Niño by comparing three different data split ratios for each query. The researchers also compared model testing results using GloVe and without Glove, where RNN and CNN modeling with GloVe features resulted in better performance than without GloVe features. The RNN algorithm using an 80:20 split ratio achieved an accuracy score of 94.90%, recall 94.90%, precision 94.94%, and F1-Score 94.85%. Meanwhile, the CNN algorithm using a 90:10 split ratio achieved an accuracy score of 94.61%, recall of 93.61%, precision of 94.69%, and F1-Score of 94.58%. This demonstrates that sentiment analysis using RNN modeling with GloVe features produces a better model for handling sentiment analysis in this study than CNN modeling, with an average accuracy rate of 94.90%. In facing the challenge of sentiment analysis on diverse Twitter data related to El Niño, the use of RNN and CNN algorithms has proven to provide an effective solution. By comparing various data splitting ratios and considering the use of GloVe features, this study asserts the superiority of RNN modeling over CNN in achieving higher accuracy levels. These findings indicate that RNN, particularly with the use of GloVe features, can deliver better performance in sentiment analysis tasks. To optimize this research, future studies may further explore additional techniques to enhance sentiment analysis accuracy and address the complexity of Twitter data analysis in environmental contexts.

REFERENCES

[1] H. Amrianah, S. Al, and G. Barru, “IMPACTS OF THE EL NIÑO AND LA NIÑA EXTREME CLIMATE ON SOCIO-ECONOMIC AND POLITICS IN INDONESIA.”
[2] D. Goularas and S. Kamis, “Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data,” in 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), 2019, pp. 12–17. doi: 10.1109/Deep-ML.2019.00011.
[3] P. Cen, K. Zhang, and D. Zheng, “Sentiment analysis using deep learning approach,” J. Artif. Intell, vol. 2, no. 1, pp. 17–27, 2020.
[4] A. M. Rahat, A. Kahir, and A. K. M. Masum, “Comparison of Naive Bayes and SVM Algorithm based on sentiment analysis using review dataset,” in 2019 8th International Conference System Modeling and Advancement in Research Trends (SMART), 2019, pp. 266–270.
[5] M. K. Hasan and E. B. Setiawan, “Sentiment Analysis of Twitter Data on Bank Central Asia Stocks (BBCA) Using RNN and CNN Model with GloVe Feature Expansion,” in 2023 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), 2023, pp. 195–200.
[6] N. Uly, H. Hendry, and A. Iriani, “CNN-RNN Hybrid Model for Diagnosis of COVID-19 on X-Ray Imagery,” Digital Zone: Jurnal Teknologi Informasi dan Komunikasi, vol. 14, no. 1, pp. 57–67, 2023.
[7] S. Ahmad, A. M. Ridwan, and G. D. Setyawan, “Analisis Sentimen Product Tools & Home Menggunakan Metode Cnn Dan Lstm,” Teknokom, vol. 6, no. 2, pp. 133–140, 2023.
[8] M. D. D. Sreya and E. B. Setiawan, “Penggunaan Metode GloVe untuk Ekspansi Fitur pada Analisis Sentimen Twitter dengan Na\"\ive Bayes dan Support Vector Machine,” eProceedings of Engineering, vol. 9, no. 3, 2022.
[9] R. Hendrawan, S. Al Faraby, and others, “Multilabel classification of hate speech and abusive words on Indonesian Twitter social media,” in 2020 International Conference on Data Science and Its Applications (ICoDSA), 2020, pp. 1–7.
[10] Kementerian Pertanian Republik Indonesia, “SobaTani, pemerintah terus berupaya dalam meningkatkan produksi padi secara nasional,” https://web.facebook.com/kementanRI.
[11] bsip.pascapanen, “Pompanisasi dari Sungai-Sungai pada Lahan Sawah Kering. Pascapanen,” https://pascapanen.bsip.pertanian.go.id/.
[12] S. Lende, Y. N. Tetik, and M. P. Ayu, “Analisis Sentimen Siswa Terhadap Pelajaran Informatika di SMPK St. Yohanes Kalembu Lona Dengan Metode Naive Bayes Classifer,” Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI), vol. 6, no. 3, pp. 217–225, 2023.
[13] S. Alam and N. Yao, “The impact of preprocessing steps on the accuracy of machine learning algorithms in sentiment analysis,” Comput Math Organ Theory, vol. 25, pp. 319–335, 2019.
[14] A. W. Pradana and M. Hayaty, “The effect of stemming and removal of stopwords on the accuracy of sentiment analysis on indonesian-language texts,” Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control, pp. 375–380, 2019.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 354-364
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5284

[15] J. A. Septian, T. M. Fachrudin, and A. Nugroho, “Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor,” INSYST: Journal of Intelligent System and Computation, vol. 1, no. 1, pp. 43–49, 2019.
[16] D. Darwis, N. Siskawati, and Z. Abidin, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional,” Jurnal Tekno Kompak, vol. 15, no. 1, pp. 131–145, 2021.
[17] A. R. Jamaludin and E. B. Setiawan, “Deteksi berita hoax di media sosial twitter dengan ekspansi fitur menggunakan glove,” eProceedings of Engineering, vol. 9, no. 3, 2022.
[18] A. R. Fitriansyah and Y. Sibaroni, “Analisis Sentimen Terhadap Pembangunan Kereta Cepat Jakarta-Bandung pada Media Sosial Twitter Menggunakan Metode SVM dan GloVe Word Embedding,” eProceedings of Engineering, vol. 10, no. 2, 2023.
[19] A. S. L. Lie, D. H. Setiabudi, and I. Sugiarto, “Analisis Sentimen Dari Keywords Yang Dimasukan Pengguna Di Twitter Indonesia Untuk Penunjang Pembelajaran Strategi Komunikasi Di Program Studi Ilmu Komunikasi Universitas Kristen Petra Dengan Metode Cnn-Bidirectional Lstm,” Jurnal Infra, vol. 10, no. 2, pp. 294–299, 2022.
[20] H. K. Farid, E. B. Setiawan, and I. Kurniawan, “Implementation Information Gain Feature Selection for Hoax News Detection on Twitter using Convolutional Neural Network (CNN),” Indonesia Journal on Computing (Indo-JC), vol. 5, no. 3, pp. 23–36, 2020.
[21] D. T. Putra, E. B. Setiawan, and others, “Sentiment Analysis on Social Media with Glove Using Combination CNN and RoBERTa,” Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 7, no. 3, pp. 457–563, 2023.