Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 912-920
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5778

Election Hoax Detection on X using CNN with TF-RF and TF-IDF Weighting Features

Dila Adelia*, Widi Astuti, Kemas Muslim Lhaksmana

School of Computing, Telkom University, Bandung, Indonesia
Email: [1,*]diladelia@student.telkomuniversity.ac.id, [2]widiwdu@telkomuniversity.ac.id, [3]kemasmuslim@telkomuniversity.ac.id
Correspondence Author Email: emailpenuliskorespondensi@gmail.com
Submitted: 12/08/2024; Accepted: 15/08/2024; Published: 16/08/2024

Abstract−X social media is a microblogging platform for sharing brief thoughts and trends. It has become a focal point for expressing political views. The increased political engagement on X social media has facilitated the swift and extensive sharing of ideas. Still, it also brings the risk of spreading false information and hoaxes that can manipulate public opinion. Preventing fake news on social media is crucial because it can influence election outcomes and social stability. For example, X social media has been used during elections to spread hoaxes, such as false claims of vote tampering or misleading information about candidate qualifications. This study implements a Convolutional Neural Network (CNN) due to its advantages in recognizing complex patterns and achieving high performance in tasks like classification. The dataset used in this study consists of 2,670 tweets. The dataset is divided into three subsets: 60% for training, 20% for testing, and 20% for validation. It also uses Term Frequency Relevance Frequency (TF-RF) and Term Frequency Inverse Document Frequency (TF-IDF) weighting features to improve accuracy in detecting fake news. This study compares the TF-RF and TF-IDF weighting features using the CNN classification method on the topic of the 2024 election. The testing results indicate that both TF-RF and TF-IDF achieved similar overall performance, with TF-RF slightly excelling in recall and F1-score. At the same time, TF-IDF showed a marginally higher precision.

Keywords: X Social Media; Hoax; Convolutional Neural Network; TF-RF; TF-IDF

1. INTRODUCTION

In the digital era, X social media has emerged as a primary platform for conveying political information during elections, allowing voters to engage directly in political discussions and receive real-time information. However, the presence of social media also opens the door wide for the spread of hoaxes, which can undermine the integrity of elections. A hoax is false information deliberately created and spread as if it were true [1]. The hoaxes circulating on X social media are usually false or misleading information designed to influence public opinion during elections. Examples include false claims about certain candidates, manipulated election data, or fake news that stirs ethnic or religious tensions. Since X social media has a character limit per post, information is often presented in a brief format that can easily be misunderstood or intentionally misrepresented. Therefore, detecting hoaxes on social media, especially on the X platform, is crucial to maintaining the validity of political information and protecting the democratic process.

These hoaxes undermine the integrity of elections by creating confusion among voters, spreading distrust in the electoral process, and even potentially sparking social conflict. If left unchecked, the spread of hoaxes can disrupt the democratic process by leading voters to make decisions based on false or misleading information, ultimately affecting the election outcome. The spread of hoaxes can cause anxiety, discomfort, and even anger, especially if the information is provocative. This can incite hatred and lead to division within society [2]. Previous research has identified that the spread of hoaxes and disinformation during election periods tends to increase [3]. Therefore, measures are needed to address the spread of hoaxes in society, with an approach that can understand text within character limits.

One approach is to utilize machine learning to identify certain patterns that frequently appear in hoaxes. Machine learning can be applied through various methods, such as Support Vector Machine (SVM), Decision Tree, Self-Organizing Map, and Convolutional Neural Network (CNN) [4]. Previous research on hoax detection has been conducted by Nurhikam et al. [5] using the random forest algorithm, achieving an accuracy of 84%. Additionally, research by Fauzy et al. [6] used a Convolutional Neural Network (CNN) and successfully improved the accuracy of hoax detection in microblogging content. Therefore, this study chooses the Convolutional Neural Network (CNN) method as the main approach to detect fake news on the X social media platform.

Convolutional Neural Network (CNN) is an architecture originally designed to handle image data processing but has been successfully used in various text-processing tasks, including text classification [7]. CNN has several advantages in text classification, such as automatic feature extraction, the ability to handle big data, and more. Additionally, CNN can extract and understand complex patterns in text, which are often found in fake news content. This capability is the reason behind using the CNN method in this study.

Additionally, the author applies weighting methods using Term Frequency Relevance Frequency (TF-RF) and Term Frequency Inverse Document Frequency (TF-IDF) to enhance detection accuracy. TF-RF combines TF and RF to improve performance, focusing on the occurrence of a term within a document [8]. Conversely, TF-IDF is a weighting feature designed to represent the significance of a term within a document in a corpus.

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 912-920
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5778

TF-IDF is often used to determine how important a term is in the context of search engines, text summarization, or text classification [9]. In similar studies, using TF-RF and TF-IDF weighting features has shown significant improvements in predicting potential hoax text content [9] [10].

The difference from previous research is that [11] found that CNN only achieved an accuracy of 60% using 804 FA-KES data as the dataset. Therefore, this study aims to build upon that research to achieve more optimal results. In [12], CNN only used TF-IDF weighting features, whereas, in this study, the author uses CNN with TF-RF and TF-IDF weighting features. Previous research on fake review detection [13] compared six classification models but did not provide a deep theoretical explanation of the methods, and all models produced low accuracy. Study [10] used the TF-RF and TF-IDF as weighting features with K-Nearest Neighbor classification, achieving an accuracy of 62%. In contrast, this study uses TF-RF and TF-IDF as weighting features and CNN for classification to compare the classification results with previous studies. Study [14] used Word2Vec for word vectorization, whereas this study uses Count Vectorizer to compare the results of text representation methods with previous research.

The author chose Convolutional Neural Network (CNN) because this method has been infrequently applied to the classification of hate speech using Indonesian language [15]. CNN excels in recognizing complex patterns, achieving high performance in tasks like classification, and is very effective in handling large and complex data. The effectiveness of this research method is evaluated using a confusion matrix.

2. RESEARCH METHODOLOGY

2.1 Research Stages

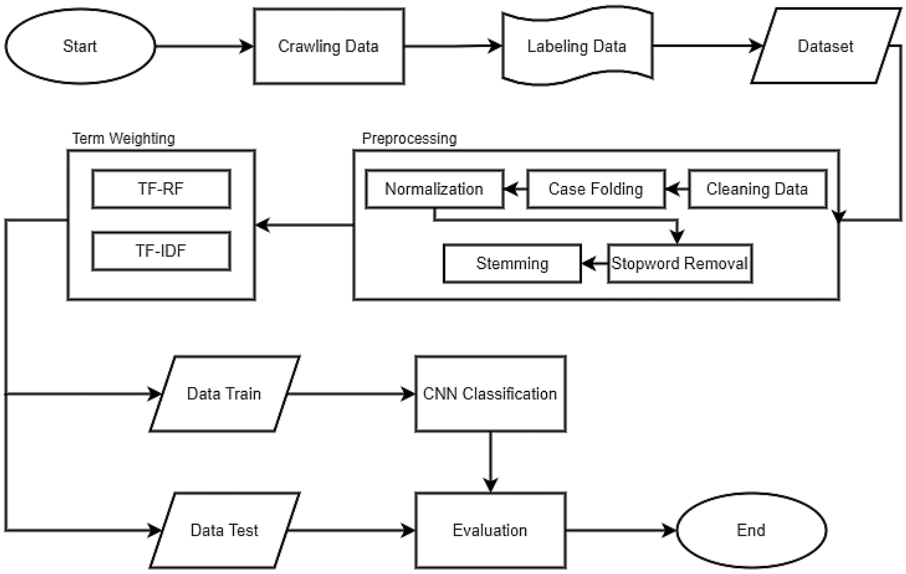

Figure 1. Flowchart of the research process

Figure 1 illustrates the process flow of the system developed in this study. The process begins with data collection using crawling techniques from X social media. Once the data is collected, the next step is labeling the data. The labeled data then undergoes preprocessing, which consists of data cleaning, case folding, normalization, stopword removal, and stemming. After that, the data is processed into vectors that can be used by the model using TF-RF and TF-IDF methods. The processed data is then divided into two parts: training data used to train the CNN model and test data used to measure model performance. The CNN model is trained with training data to perform text classification. After the model is trained, its performance is assessed using the test data to confirm its accuracy and effectiveness.

2.2 Data Crawling

Data Crawling is a phase in the data collection process. In this study, the author gathered tweets from X social media over 5 months during the election, using keywords or hashtags such as 'asal bukan anies', 'asal bukan prabowo', and 'asal bukan ganjar'. Collecting data from X social media involved utilizing the X social media API Key and implementing it with the Python programming language.

2.3 Data Labeling

The labeling process involves manually assigning labels. In this study, the labeling method used by Reiki et al. [16] is applied with two categories: index label '0' for factual news, while index label '1' for hoax news. The data labeling process is carried out by paying attention to the words in the collected tweets. If a tweet contains harsh

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 912-920
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5778

or inappropriate words, it will be labeled as 1. Conversely, if the tweet does not contain harsh words or contains words with positive meanings, it will be labeled as 0. Harsh words is defined as words that convey cursing, insulting, mocking, and similar meanings. Inappropriate sentences are those that contain provocation with bad intuition towards an issue or negative sentences that could worsen a situation. Table 1 provides examples of the labeling process.

Table 1. Data Labeling

Tweet	Label
Padahal sudah kampanye negatif dgn macam2 cara. Mulai dari seruan 4 jari...	1
#AsalBukanPrabowo Biarkan dia tetap berakhir menjadi #BangsatBangsa...	1
@OposisiCerdas Dari kmaren2 gw sih Prabowo menang no problem...	0
@aniesbaswedan Sebenarnya bukan ngotot supaya anies/ganjar menang sih...	1
@bprakosha @dennyindrayana Gue sebagai voter anies setuju akan hal ini...	0
@aewin86 @gilang_ahm31272 ini bukan soal kalah pemilu, tapi proses gibran nya itu lo bermasalah...	1
@dennyindrayana Prediksi sy, diadakan pemungutan suara ulang tetap dgn 3 paslon...	1
@LOVE_UM4M4 @prabowo DiA ngomong apa adanya Dia bukan sembarangan ngomong...	0
@RSuntya44707 @kafir_introvert Wkwkwk ngomong tuh pake bukti gitu loh bukan asal bunyi...	0
@DedynurPalakka @prabowo Terus berjuang smp pemilu ga ada kecurangan! Bukan masalah menang/kalah...	1
...	
Mundur Wir : Kita butuhnya Leader yang bisa kasih kepastian, bukan cheerleader yang cuma kasih hiburan...	1
Pak Anies itu sebetulnya kami minta sebagai Cawapres...	0

2.4 Preprocessing

This stage consists of several steps: data cleaning, case folding, normalization, stopword removal, and stemming. First, data cleaning involves eliminating noise from the data, such as emojis and punctuation. Second, case folding refers to converting all letters to lowercase. Third, normalization is the process of replacing abbreviations with their full forms. Fourth, stopword removal is the process of deleting unnecessary words. Finally, stemming is the procedure of reducing words with affixes to their base forms. Table 2 illustrates the preprocessing steps.

Table 2. Preprocessing

Preprocessing	Before	After
Data Cleaning	Yg penting asal bukan bocah haram konstitusi perusak demokrasi #AsalBukanPrabowo *The important thing is that it's not an illegitimate child of the constitution who destroys democracy #AnyoneButPrabowo*	Yg penting asal bukan bocah haram konstitusi perusak demokrasi AsalBukanPrabowo *The important thing is that it's not an illegitimate child of the constitution who destroys democracy AnyoneButPrabowo*
Case Folding	Yg penting asal bukan bocah haram konstitusi perusak demokrasi #AsalBukanPrabowo *The important thing is that it's not an illegitimate child of the constitution who destroys democracy #AnyoneButPrabowo*	yg penting asal bukan bocah haram konstitusi perusak demokrasi asalbukanprabowo *the important thing is that it's not an illegitimate child of the constitution who destroys democracy anyonebutprabowo*
Normalization	Yg penting asal bukan bocah haram konstitusi perusak demokrasi #AsalBukanPrabowo *The important thing is that it's not an illegitimate child of the constitution who destroys democracy #AnyoneButPrabowo*	yang penting asal bukan bocah haram konstitusi perusak demokrasi asalbukanprabowo *the important thing is that it's not an illegitimate child of the constitution who destroys democracy anyonebutprabowo*
Stopword Removal	Yg penting asal bukan bocah haram konstitusi perusak demokrasi #AsalBukanPrabowo *The important thing is that it's not an illegitimate child of the constitution who destroys democracy #AnyoneButPrabowo*	yang penting bukan bocah haram konstitusi perusak demokrasi asalbukanprabowo *the important thing is that it's not an illegitimate child of the constitution who destroys democracy anyonebutprabowo*
Stemming	Yg penting asal bukan bocah haram konstitusi perusak demokrasi #AsalBukanPrabowo *The important thing is that it's not an illegitimate child of the constitution who destroys democracy #AnyoneButPrabowo*	yang penting bukan bocah haram konstitusi rusak demokrasi asalbukanprabowo *the important thing is that it's not an illegitimate child of the constitution who destroys democracy anyonebutprabowo*

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 912-920
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5778

2.5 Term Weighting

a. Term Frequency – Relevance Frequency

TF-RF weighting feature is a relatively new method that has emerged as an improvement over previous methods. This method evaluates the relevance of a document by considering how often a term occurs in related categories. In relevance frequency, the weight of a term is calculated using the equation [10]:

$$tf_{td}rf = tf_{td} \times \log\left(2 + \frac{b}{\max(1, c)}\right) \tag{1}$$

Where :
$tf_{td}rf$: document weighting in the vector space model
tf_{td} : count of how many times the term t appears in the document
b : count of documents that include the term t
c : count of documents that do not include term t

b. Term Frequency – Inverse Document Frequency

TF-IDF weighting feature is a method that converts text data into numerical data to assigning weights to each word or feature [17]. TF-IDF is a method used to assess how important a word is in a document. The TF-IDF value can be determined using the following equation [18]:

$$tf_{td}idf_t = tf_{td} \times \log\left(\frac{N}{df_t}\right) \tag{2}$$

Where :
$tf_{td}idf_t$: weight of term t
tf_{td} : frequency of term t in the document
N : total number of documents
df_t : count of documents that include the term t

2.6 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is one of the components of deep learning due to its highly layered network architecture. Deep learning is a subfield of machine learning that enables computers to complete tasks as humans do [19]. CNN performs convolution operations by combining multiple processing layers, leveraging several parallel elements, and is inspired by biological neural systems [20]. The structure of CNN, illustrated in Figure 2, includes input, feature extraction, classification, and output. CNN works hierarchically, where the output of the first convolutional layer is used as input for the next convolutional layer [19].

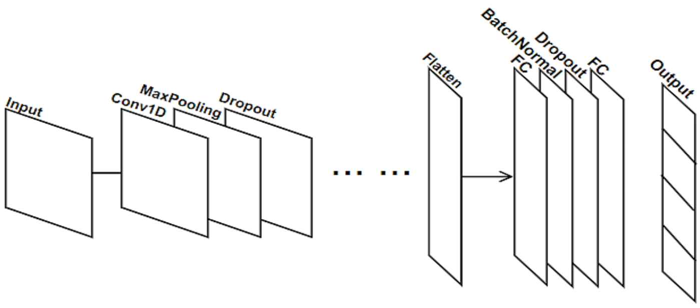

Figure 2. CNN Model Architecture

2.7 Data Splitting

The data-splitting process is carried out in two stages to ensure the model has distinct datasets for training, validation, and testing. In the first stage, the dataset is divided into two main parts: training data and testing data. 20% of the entire dataset is set aside for testing, while the remaining 80% serves as provisional training data. In the second stage, the provisional training data is further divided to obtain validation data. Specifically, 25% of the provisional training data is allocated for validation, which equates to 20% of the entire dataset. Ultimately, the training data makes up 60% of the total dataset, with 20% each allocated for validation and testing.

2.8 Evaluation

This system is utilized to evaluate the performance of the classification prediction model using a confusion matrix. The confusion matrix is presented in Table 3.

Table 3. Confusion Matrix

	Predict True	Predict False
Actual True	True Positive (TP)	False Negative (TN)
Actual False	False Positive (FP)	True Negative (TN)

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 912-920
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5778

The confusion matrix includes four categories: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

a. Accuracy
Accuracy gauges how closely predicted values align with actual values. Its function aims to assess how well the prediction matches the actual class [21]. The formula for calculating accuracy is as follows:

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

b. Precision
Precision is the accuracy level between the selected information and the answer provided by the system [21]. The formula for precision is as follows:

$$precision = \frac{TP}{TP+FP} \quad (4)$$

c. Recall
Recall measures the accuracy of the correctly identified data [21]. The formula for recall is as follows:

$$\text{recall} = \frac{TP}{FN+TP} \quad (5)$$

d. F1 Score
F1 Score is the average value of the comparison between recall and precision [21]. The formula for calculating the F1 Score is as follows:

$$\text{f1 score} = \frac{2*recall*precision}{recall+precision} \quad (6)$$

3. RESULT AND DISCUSSION

This section presents a brief overview of the dataset used in this study and provides a comprehensive analysis of the results of the tests performed. This study aims to identify fake news on social media X based on tweets. There are two scenarios in this study: the first scenario uses the Term Frequency Relevance Frequency (TF-RF) weighting feature, and the second scenario uses the Term Frequency Inverse Document Frequency (TF-IDF) weighting feature. Both scenarios employ the Convolutional Neural Network (CNN) classification method and compare the resulting performance of these two weighting features.

3.1 Dataset

This research focuses on detecting fake news from a dataset consisting of tweets related to the 2024 election. The dataset was collected from the social media platform X social media using the X social media API, with hashtags 'asal bukan anies', 'asal bukan prabowo', and 'asal bukan ganjar' as the main keywords. The data collection period lasted from December 1, 2023, to April 24, 2024, resulting in a sufficiently representative dataset for this study.

The resulting dataset consists of 2,670 tweets that have undergone manual labeling. Each tweet was labeled into two categories: the first category (labeled 1) includes tweets containing hate speech, and harsh, or inappropriate language, while the second category (labeled 0) includes tweets that do not contain hate speech and contain words with positive meanings. Of the total 2,670 labeled tweets, 1,392 tweets were classified as category 1, indicating that these tweets contain negative content. Meanwhile, the remaining 1,278 tweets were categorized as category 0, indicating that these tweets have positive content. The higher proportion of tweets with negative content compared to positive tweets indicates a negative tendency in the data collected from X social media. The data proportions used in this study can be seen in Figure 3, which provides a visualization of the distribution of tweets based on the defined categories.

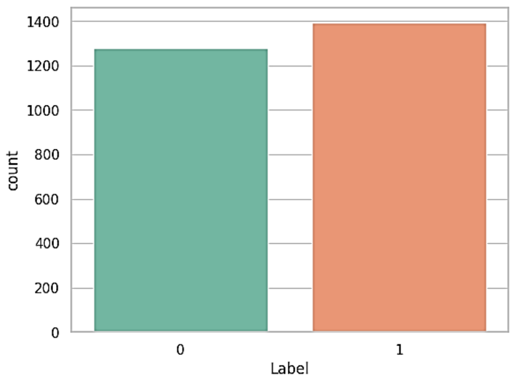

Figure 3. Data distribution

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 912-920
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5778

3.2 Training Result

CNN utilizing TF-RF and TF-IDF weighting features was tested for 10, 15, 20, 50, and 100 epochs during the testing phase. This comprehensive evaluation aimed to determine the model's performance and stability over different training durations, providing insights into the optimal number of epochs required to achieve the best predictive accuracy and generalization on the testing dataset.

a. Term Frequency – Relevance Frequency

Table 4. Comparison of accuracy and loss on TF-RF training and validation data

Epoch	Accuracy	Loss	Validation Accuracy	Validation Loss
10	0.8865	0.5306	0.9026	0.5309
15	0.8944	0.5190	0.8970	0.5004
20	0.8831	0.4686	0.9101	0.4887
50	0.8805	0.5030	0.9026	0.5112
100	0.8825	0.5195	0.8970	0.5100

In the first scenario, as seen in Table 4, the 10th epoch resulted in relatively high loss and validation loss values of 0.5306 and 0.5309, respectively. This indicates that the model is still learning to recognize patterns in the data. As the epochs progress, the loss values tend to decrease, as observed in the 20th epoch with a loss of 0.4686 and a validation loss of 0.4887, which indicates that the model is increasingly effective in understanding patterns from the training data. However, this decrease is not always consistent; for instance, in the 50th and 100th epochs, there is a slight increase in both loss and validation loss values, which could be an early indication of overfitting. Overfitting happens when the model focuses too much on the training data, making it less capable of generalizing to new, unseen data.

Additionally, the accuracy of the training data (accuracy) and validation data (validation accuracy) tends to increase in the early stages, with validation accuracy rising from 0.9026 in the 10th epoch to 0.9101 in the 20th epoch. However, after reaching a certain point, this increase in accuracy becomes insignificant and even starts to decline, as seen in the 100th epoch where the validation accuracy drops to 0.8970. This decline indicates that the model may have reached its limit in learning from the training data and is starting to show signs of overfitting. Overall, the model shows good performance in the early stages of training with increasing accuracy and decreasing loss but then experiences a decline in performance due to overfitting at higher epochs.

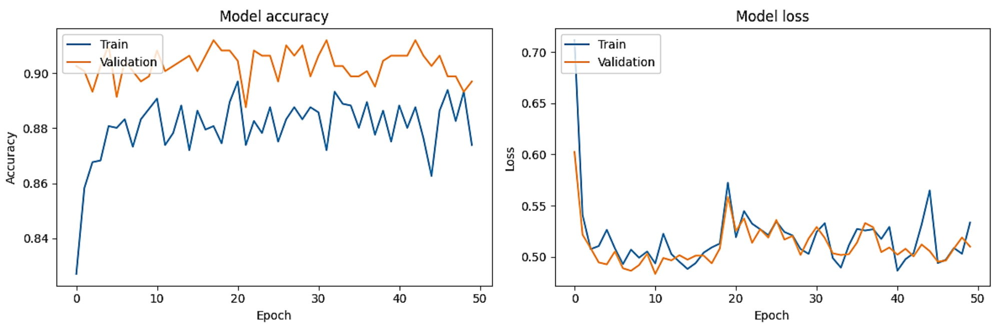

Figure 4. Accuracy and loss graphs for TF-RF training and validation data

In Figure 4, Model Accuracy, the accuracy of the training data (Train) and validation data (Validation) shows a fairly consistent increase at the beginning of the training. The training data accuracy initially rises sharply and then stabilizes around values between 0.88 and 0.90. On the other hand, the validation data accuracy shows a slightly erratic pattern but generally remains above the training data accuracy, with some peaks exceeding 0.90. This suggests that the model has good generalization capability during the initial training phases, though the erratic changes may also suggest the model is adjusting to the validation data.

In Figure 4, Model Loss, it can be seen that the loss values for both the training and validation data decrease significantly at the beginning of the training, indicating that the model is learning quickly in the early phase. However, after several epochs, the loss values tend to stabilize with minor fluctuations in both datasets, suggesting that the model has reached a stable point in the training process. It is also noticeable that the validation loss is generally below or in line with the training loss, indicating that the model is not significantly overfitting up to epoch 50.

Overall, both graphs demonstrate that the model has achieved optimal performance with high accuracy and low loss on both training and validation data. However, there are some erratic changes in accuracy and loss, particularly in the validation data, indicating variability in the model's ability to simplify to new data.

b. Term Frequency – Inverse Document Frequency

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 912-920
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5778

Table 5. Comparison of accuracy and loss on TF-IDF training and validation data

Epoch	Accuracy	Loss	Validation Accuracy	Validation Loss
10	0.9555	0.7345	0.8839	0.9212
15	0.9592	0.7051	0.8951	0.8583
20	0.9660	0.6403	0.9007	0.8922
50	0.9669	0.7858	0.9045	0.9776
100	0.9568	0.7899	0.8933	10.571

In the second scenario, as seen in Table 5, the initial epochs show a significant increase in accuracy and validation accuracy, indicating that the model is increasingly able to recognize patterns in the data. For example, accuracy rises from 0.9555 in epoch 10 to 0.9669 in epoch 50, while validation accuracy increases from 0.8839 to 0.9045. This progress shows that the model is increasingly able to make accurate predictions, both on training and validation data. However, by epoch 100, there is a slight decrease in validation accuracy to 0.8933 and an increase in validation loss to 1.0571, which could indicate overfitting. This decrease and increase highlight the importance of selecting the appropriate number of epochs to balance learning from the training data and maintaining generalization ability.

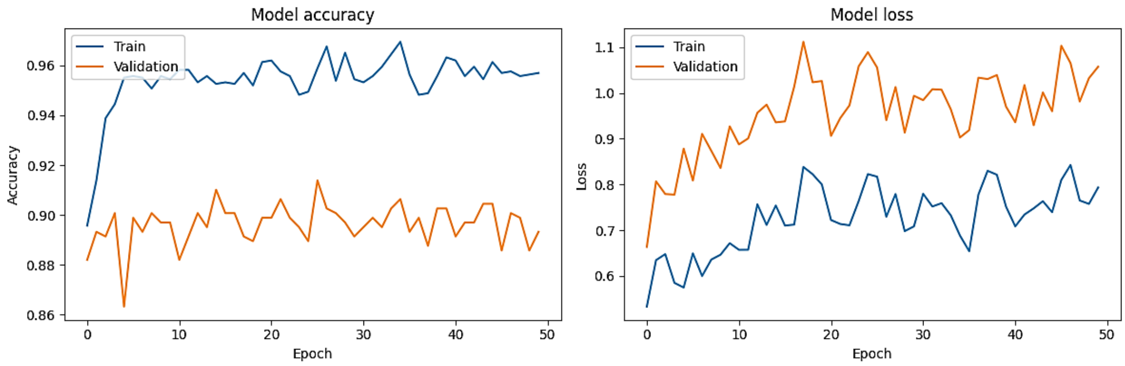

Figure 5. Accuracy and loss graphs for TF-IDF training and validation data

In Figure 5, Model Accuracy section, it is evident that the accuracy of the training data (Train) quickly increases and stabilizes around 0.96 after the first few epochs. However, the accuracy of the validation data (Validation) shows more significant fluctuations, ranging between 0.86 and 0.91. This indicates that the model performs well in predicting the data it has been trained on but does not consistently deliver the same performance on unseen data.

The Model Loss section in Figure 5 reveals a clearer pattern regarding potential overfitting. The loss on the training data (Train) decreases quite well from the beginning and stabilizes at a lower value of around 0.7. However, the loss on the validation data (Validation) shows an increasing trend, especially after the 20th epoch, with values becoming higher compared to the training loss. This increase in validation loss indicates that the model is finding it increasingly difficult to accurately predict the validation data, even though the training accuracy remains high.

Overall, the graph shows that, despite the model's good performance on the training data, there are clear indications of overfitting, particularly seen in the marked differences between accuracy and loss on the training and validation data. This underscores the importance of selecting the appropriate number of epochs and using regularization techniques or better model tuning to address overfitting.

3.3 Testing Result

Table 6. Feature Weighting Comparison

Term Weighting	TP	FP	TN	FN	Accuracy	Precision	Recall	F1 Score
TF-RF	228	33	36	237	0.8708	0.8681	0.8778	0.8729
TF-IDF	232	37	32	233	0.8708	0.8792	0.8630	0.8710

Based on Table 6, the classification results from two feature weighting methods, TF-RF and TF-IDF, show that both perform well in detecting tweets containing fake news. Each method produces different variations in the number of True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN), as well as in various other evaluation metrics such as accuracy, precision, recall, and F1 Score. These differences highlight the unique strengths and weaknesses of each approach, indicating that the choice of feature weighting method can significantly affect the overall performance of a classification model.

With the TF-RF weighting feature, the system successfully identified 228 tweets that were indeed fake news as fake news (TP). This demonstrates the system's ability to detect a large number of tweets containing

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 912-920
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5778

false information. However, 33 tweets were not fake news but were identified as fake news (FP). This indicates an error where the system mistakenly considers valid tweets as false information. Additionally, the system successfully identified 36 non-fake news tweets as non-fake news (TN), demonstrating its capability to recognize true tweets. However, there were 237 fake news tweets misidentified as non-fake news (FN). This shows a weakness in detecting all fake news, impacting recall. Overall, the TF-RF method achieved an accuracy of 0.8708, precision of 0.8681, recall of 0.8778, and an F1 Score of 0.8729.

Meanwhile, the TF-IDF method produced slightly different results. The system correctly identified 232 fake news tweets (TP), slightly higher than TF-RF. However, the number of tweets that were not fake news but identified as fake news (FP) also increased to 37. Additionally, this method showed a slightly lower ability to recognize non-fake news tweets, with only 32 correctly identified (TN). In this case, the number of fake news tweets misidentified as non-fake news (FN) was 233, slightly lower than TF-RF. Nevertheless, the accuracy remained the same at 0.8708. However, precision increased to 0.8792, although recall slightly decreased to 0.8630. The F1 Score for the TF-IDF method was 0.8710.

Based on the comparison between the TF-RF and TF-IDF methods in detecting fake news, it can be concluded that the TF-RF has better advantages in terms of recall, with a value of 0.8778. Higher recall indicates a better ability to identify all tweets that are genuinely fake news. Although TF-RF has slightly more false positives, this method ensures that more fake news is identified, minimizing the risk of missing fake news. Conversely, the TF-IDF method shows higher precision, at 0.8792, meaning it is more accurate in ensuring that tweets classified as fake news are indeed fake. However, with lower recall, this method does not capture all the existing fake news. Therefore, it is important to consider the context of use when selecting the most appropriate method for fake news detection.

4. CONCLUSION

This study aims to detect fake news on X social media content using CNN classification modeling. This study compares the weighting methods of TF-RF and TF-IDF features to minimize the spread of hoax news that harms society. Based on the testing scenarios in this study, both feature weightings achieved the same accuracy value of 0.8708. The TF-RF feature weighting delivered superior performance based on a recall comparison with TF-IDF, scoring 0.8778, indicating TF-RF's better ability to identify all tweets that are genuinely fake news. Therefore, the use of feature weighting can influence the classification accuracy of the Convolutional Neural Network method in detecting fake news. The number of epochs used significantly affects the accuracy of each feature weighting. Recommendations for future research include enhancing the preprocessing process, particularly in developing normalization dictionaries and stopword lists. The goal is to produce more standardized words, reduce the number of discarded features, and minimize ambiguous words to achieve optimal accuracy. Additionally, further evaluation of the model with a larger and more diverse dataset will be very useful for improving generalization and maintaining model performance when faced with various data variations in detecting fake news. Finally, integrating advanced natural language processing techniques, such as word embeddings or transformer-based models, could further enhance the model's ability to capture the nuances of language used in fake news.

REFERENCES

[1] C. C. Wang, “Fake News and Related Concepts: Definitions and Recent Research Development,” *Contemporary Management Research*, vol. 16, no. 3, pp. 145–174, Sep. 2020, doi: 10.7903/CMR.20677.
[2] V. Oktaviana Yamin, A. Tenriawaru, L. Ode Saidi, and G. Arviana Rahman, “Penerapan Naïve Bayes Classifier dengan Algoritma Nazief dan Adriani Untuk Deteksi Hoaks,” *Prosiding Seminar Nasional Pemanfaatan Sains dan Teknologi Informasi*, vol. 1, no. 1, pp. 335–344, 2023.
[3] S. Cinta Insani, N. Alisya Zahwa Khuzaimah, V. Zia Devita Maryadi, and T. Alya Hafizha, “Meninjau Etika Masyarakat Indonesia Dalam Bermedia Sosial Di Masa Pemilu Menggunakan Etika Media Sosial,” *Jurnal Pendidikan, Seni, Sains dan Sosial Humaniora*, vol. 1, no. 2, pp. 1–25, 2023, doi: 10.11111/nusantara.xxxxxxx.
[4] F. Rakan Tama and Y. Sibaroni, “Fake News (Hoaxes) Detection on Twitter Social Media Content through Convolutional Neural Network (CNN) Method,” *JINAV: Journal of Information and Visualization*, vol. 1, no. 1, pp. 2746–1440, 2020, doi: 10.35877/454RI.jinav2125.
[5] A. S. Nurhikam, R. Syaputra, S. Rohman, S. R. Priyambodo, and N. Agustina, “Deteksi Berita Palsu Pada Pemilu 2024 Dengan Menggunakan Algoritma Random Forest,” *Journal of Computer and Information Technology*, vol. 7, no. 1, pp. 41–50, 2023.
[6] A. R. I. Fauzy and B. S. Erwin, “Detecting Fake News on Social Media Combined with the CNN Methods,” *JURNAL RESTI (RekayasaSistemdanTeknologiInformasi)*, vol. 7, no. 2, pp. 271–277, 2023, doi: 10.29207/resti.v7i1.4889.
[7] A. D. Cahyani, A. K. Ramdani, and Y. Sibaroni, “Hoax Detection of Covid-19 News using Convolutional Neural Network and Support Vector Machine,” *Intl. Journal on ICT*, vol. 9, no. 2, pp. 177–185, 2023, doi: 10.21108/ijoict.v9i2.872.
[8] Z. Tang, W. Li, Y. Li, W. Zhao, and S. Li, “Several alternative term weighting methods for text representation and classification,” *Knowl Based Syst*, vol. 207, Nov. 2020, doi: 10.1016/j.knosys.2020.106399.

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 4, August 2024, Page 912-920
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i4.5778

[9] A. Thakkar and K. Chaudhari, “Predicting stock trend using an integrated term frequency–inverse document frequency-based feature weight matrix with neural networks,” *Applied Soft Computing Journal*, vol. 96, p. 106684, Nov. 2020, doi: 10.1016/j.asoc.2020.106684.

[10] A. N. Assidyk, E. B. Setiawan, and I. Kurniawan, “Analisis Perbandingan Pembobotan TF-IDF dan TF-RF pada Trending Topic di Twitter dengan Menggunakan Klasifikasi K-Nearest Neighbor,” *e-Proceeding of Engineering*, vol. 7, no. 2, pp. 7773–7781, 2020.

[11] J. A. Nasir, O. S. Khan, and I. Varlamis, “Fake news detection: A hybrid CNN-RNN based deep learning approach,” *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100007, Apr. 2021, doi: 10.1016/j.jjimei.2020.100007.

[12] D. A. N. Taradhita and I. K. G. D. Putra, “Hate Speech Classification in Indonesian Language Tweets by Using Convolutional Neural Network,” *Journal of ICT Research and Applications*, vol. 14, no. 3, pp. 225–239, 2021, doi: 10.5614/itbj.ict.res.appl.2021.14.3.2.

[13] A. Awalina, F. A. Bachtiar, F. Utaminingrum, and P. Korespondensi, “Perbandingan Pretrained Model Transformer pada Deteksi Ulasan Palsu,” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 9, no. 3, pp. 597–604, 2022, doi: 10.25126/jtiik.202295696.

[14] P. Song, C. Geng, and Z. Li, “Research on Text Classification Based on Convolutional Neural Network,” *International Conference on Computer Network, Electronic and Automation, ICCNEA*, pp. 229–232, Sep. 2019, doi: 10.1109/ICCNEA.2019.00052.

[15] G. B. Herwanto, A. M. Ningtyas, I. G. Mujiyatna, K. E. Nugraha, and I. N. Prayana Trisna, “Hate Speech Detection in Indonesian Twitter using Contextual Embedding Approach,” *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 15, no. 2, pp. 177–188, Apr. 2021, doi: 10.22146/ijccs.64916.

[16] M. K. A. Reiki, Y. Sibaroni, and E. B. Setiawan, “Comparison of Term Weighting Methods in Sentiment Analysis of the New State Capital of Indonesia with the SVM Method,” *International Journal on Information and Communication Technology (IJoICT)*, vol. 8, no. 2, pp. 53–65, Jan. 2022, doi: 10.21108/ijoict.v8i2.681.

[17] H. Liu, X. Chen, and X. Liu, “A Study of the Application of Weight Distributing Method Combining Sentiment Dictionary and TF-IDF for Text Sentiment Analysis,” *IEEE Access*, vol. 10, pp. 32280–32289, 2022, doi: 10.1109/ACCESS.2022.3160172.

[18] R. N. Aufa, S. S. Prasetiyowati, and Y. Sibaroni, “The Effect of Feature Weighting on Sentiment Analysis TikTok Application Using The RNN Classification,” *Building of Informatics, Technology and Science (BITS)*, vol. 5, no. 1, pp. 345–353, Jun. 2023, doi: 10.47065/bits.v5i1.3597.

[19] M. M. Taye, “Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions,” May 01, 2023, doi: 10.3390/computers12050091.

[20] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, “A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects,” *IEEE Trans Neural Netw Learn Syst*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.

[21] Yuliska, H. D. Qudsi, H. L. Juanda, U. K. Syaliman, and F. N. Nina, “Analisis sentimen pada data saran mahasiswa terhadap kinerja departemen di perguruan tinggi menggunakan Convolutional Neural Network,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 5, pp. 1067–1076, 2021, doi: 10.25126/jtiik.202184842.