Terakreditasi SINTA Peringkat 2
Surat Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, danTeknologi, Nomor: 158/E/KPT/2021
Masa berlaku mulai Volume 5 Nomor 2 Tahun 2021 sampai Volume 10 Nomor 1 Tahun 2026

Terbit online padalaman web jurnal: **http://jurnal.iaii.or.id**

JURNAL RESTI
(RekayasaSistemdanTeknologiInformasi)
Vol. 7 No. 2 (2023) 271 - 277 ISSN Media Elektronik: 2580-0760

Detecting Fake News on Social Media Combined with the CNN Methods

Anindika Riska Intan Fauzy[1], Erwin Budi Setiawan[2]
[1,2]Informatics, School of Computing, Telkom University
[1]intananindika@student.telkomuniveristy.ac.id, [2]erwinbudisetiawan@telkomuniversity.ac.id

Abstract

Social media platforms are created to facilitate human social life as technology develops. Twitter is one of the most popular and frequently used social media for exchanging information. This social media platform disseminates real-time and complete information. Unfortunately, there are not a few tweets that contain false information or are often referred to as hoaxes. Those hoaxes that existed on Twitter are very troubling for society. Fake news or hoaxes can cause misunderstandings in receiving information. Therefore, this research aimed at developing a system that can detect hoaxes on Twitter to anticipate their spread, which can be detrimental to related parties. The system being developed uses a deep learning approach with the Convolutional Neural Network (CNN), Term Frequency-Inverse Document Frequency (TF-IDF), Bidirectional Encoder Representations from Transformers (BERT), and Global Vectors (GloVe). The results of this study display the fake news detected by the system using the CNN method with baseline, BERT, and GloVe. The data have been adjusted to the keywords related to fake news and spread on online media, such as Hoax or Not from Detik.com, CekFakta from Kompas.com, etc. The results show the highest accuracy of 98.57% using CNN with a split ratio of 90:10, baseline unigram-bigram, BERT, and Top10 corpus tweet+IndoNews with an increase of 4.7%.

Keywords: *hoax, CNN; baseline, BERT; GloVe*

1. Introduction

Social media are essential in people's lives in current technological developments. Social media are an information medium that spreads news faster [1]. However, not all information scattered in them is valid and factual [2]. News can come from all over the world through online social media. Information contained in a reported statement can range from harmless to serious causes of various reactions. Restricting misinformation can prevent and protect users from receiving misleading information in the form of fake news [3]. In addition, social media can be used as means of spreading fake news or hoaxes [4]. Much of the information being spread is fake and can trigger various reactions from the users [5]. Twitter is one of the social media with a high-speed rate of information dissemination [6]. According to statista.com's portal data and statistics, in 2021, the number of Twitter users in Indonesia was around 16.32 million [7]. The use of social media has two interrelated impacts, positive and negative, on the users. The positive side of social media is being able to make friends widely and not being limited in terms of time and space. Therefore, we can receive more and more information [8]. On the other hand, the negative side of freer use of social media is that there are no filters on social media to screen the information whether the information is credible or not [9]. Hoaxes are fake news, information, or facts engineered for a specific purpose, ranging from jokes to real intentions [10]. The uncertainty about information being factual or not leads to misunderstandings in receiving information. The spread of fake news through social media, such as Twitter, has a direct negative impact on users and parties harmed by this fake news. Fake news can also be used as a weapon to bring down and harm certain parties based on issues spread on social media that are irrelevant and intentional [7].

Several studies have carried out hoax detection systems using various methods, from feature expansion to deep learning approaches. In this research, contributing to preventing the spread of hoaxes, a fake news detection system was developed, which functions to identify whether the news is a hoax or not. Research conducted previously to define fake news shows the conclusion that from this study, the use of the Convolutional Neural Network (CNN) method has a better accuracy level than the other methods in papers [11] and [12]. So the CNN method was chosen as a classification model combined with the other models.

DiterimaRedaksi: 22-01-2023 | SelesaiRevisi: 13-02-2023 | Diterbitkan Online: 15-03-2023

Then one of them is the research conducted by Sharma et al. to detect fake news using Bidirectional Encoder Representations from Transformers (BERT) produces the best accuracy compared to other methods of 98% [13]. Another research conducted by Anistya et al. used Global Vectors (GloVe) as feature expansion. The GloVe is used to reduce vocabulary mismatches in hate speech on Twitter. The result shows that the corpus built using GloVe and baseline (TF-IDF N-gram) produced the highest accuracy, 88.59%, 1.25% higher than the predetermined baseline [14].

Based on previous research, the result used the CNN, TF-IDF N-gram, BERT, and GloVe methods produced better accuracy than the other method. The aim of using this approach was to obtain an optimal accuracy by combining several methods that produce good accuracy based on previous studies. The combinations to be used are CNN as a classification model, TF-IDF N-gram as a baseline or feature extraction, BERT as an improvisation from a baseline, and Glove as a feature expansion.

2. Research Methods

The system design of the hoax detection is shown in Figure 1.

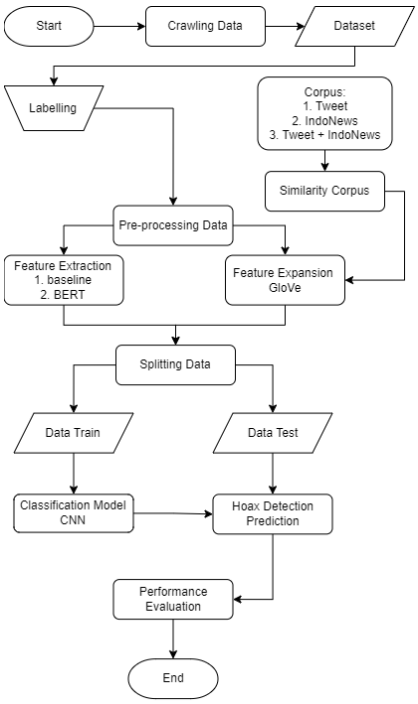

Figure 1. Hoax Detection System

2.1. Crawling Data

The data for this research was collected from Indonesian-version of Twitter, using the tools available in the Python programming language, namely snscrape [15]. The said data was taken and gathered from a collection of tweets. The topics used were the tragedy of the Kanjuruhan and the murder case of Brigadir J, which, for the moment, are still currently being discussed. Both of these topics are very prone to hoaxes or fake news. Before crawling data on Twitter, the researchers had made sure that the fake news being diplayed on online media and spread via social media was truly hoax. The online media used were "Hoax or Not" from Detik.com, "Cek Fakta" from Kompas.com, "Cek Fakta" from Liputan 6.com, "Cek Fakta" from Suara.com, and Turnbackhoax.id. The keywords used for crawling the data had been confirmed that they were not valid or were hoaxes for each topic, as shown in Tables 1 and 2.

Table 1. Keywords Hoax for Kanjuruhan Topic

No	Keyword
1.	Pemukulan jadi penyebab tragedi
2.	FIFA bekukan sepak bola Indonesia
3.	Komunis uji coba gas beracun
4.	Pemain Prancis sindir penggunaan gas air mata
5.	Kesaksian penjual dawet

Table 2. Keywords Hoax for Sambo Topic

No	Keyword
1.	Putri Candrawathi pingsan eksepsi ditolak hakim
2.	Wasiat Ferdy Sambo sebelum meninggal
3.	Bharada Eliezer divonis bebas
4.	Pintu rahasia di Rumah Sambo
5.	Putri Candrawathi divonis mati
6.	Sel mewah Ferdy Sambo
7.	Jenazah Ferdy Sambo dikirim ke Magelang
8.	Ferdy Sambo akan dieksekusi mati
9.	Ferdy Sambo nyaris tewas
10.	Ferdy Sambo sujud ke Jokowi mintaampun
11.	Ferdy Sambo melarikan diri dari Mako Brimob
12.	Putri Candrawathi bunuh diri di rumahnya
13.	Kamaruddin Simanjuntak disekap di Bunker
14.	Polisi sita puluhan tengkorak dari ruang rahasia
15.	Anggota DPR RI disuap Ferdy Sambo
16.	Kuat Ma'ruf dibawa kerumah
17.	Dua anak Ferdy Sambo dijemput paksa
18.	Putri Candrawathi minta ampun
19.	Jendral Andika Perkasa panggil tukang di rumah Ferdy Sambo
20.	Arwah Brigadir J beri kesaksian
21.	Kapolri temukan mayat perempuan tanpa busana
22.	Ferdy Sambo satu sel dengan Napoleon Bonaparte
23.	Dua organ Brigadir J dijual Ferdy Sambo
24.	Sel tahanan Ferdy Sambo kosong
25.	Ferdy Sambo divonis bebas

Following the hoax keywords above, the topics used in the crawling process can be seen in Table 3, with a total of 25.325 tweets, including hoaxes and non-hoaxes.

Table 3. List of Topics

Topic	Data Amount
Kanjuruhan	2.699
Sambo	22.626

2.2. Labelling

The labeling process was done manually and analyzed by the researchers referring to hoax features. The description of each of the nine hoax features is shown in Table 4.

DOI: https://doi.org/10.29207/resti.v7i1.4889

Table 4. Description of Each Hoax Feature [16]

Feature	Description
Username	Whether the username is set with the real name or pseudonym, contains numbers / symbols, contains elements of hatred or not.
Display Name	Another identifier used by Twitter users.
Following > Followers	The following number is more than the number of followers.
Verified Account	Whether the account has been verified or not.
Retweet	The number of retweets from the tweet.
Bio	One component of complete Twitter profile information. It is used to give others a brief about who you are, list your interests, or promote your business.
Profile Image	The profile picture explains who the owner of the account is through the photo.
Location	The location which is shown in the tweets uploaded.

Labeling was carried out based on the data set before entering the following process. For this system, the hoax detector gave label 1 for tweets that contain hoaxes and 0 for tweets that contain non-hoaxes. Labeling was performed using hoax indicators such as influencing one's views, cornering related parties, causing disputes, and containing provocative sentences or hate speech [17]. This labeling stage can be seen in Table 5.

Table 5. Class Distribution

Topic	Label	Data Amount
Kanjuruhan	Hoax	1.420
	Non-Hoax	1.279
Sambo	Hoax	11.038
	Non-Hoax	11.588

The balance of the existing data can affect the results. Therefore, to achieve optimal results, the data used must be balanced. The distribution of the data between hoaxes and non-hoaxes classes is balanced. This class balance can be seen in Figure 2.

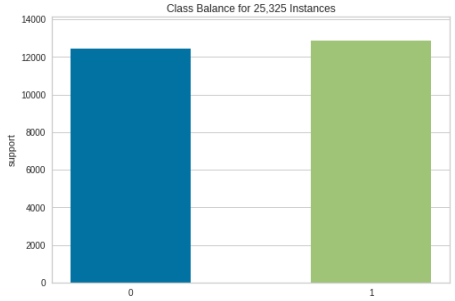

Figure 2. Class Balance

2.3. Data Pre-processing

The data pre-processing is the first stage to process the raw data before the next. The purpose of this stage is to simplify data processing in the classification stage [17]. The pre-processing stage that was being carried out includes data cleaning, case folding, stopwords, stemming, and then tokenizing. Data cleaning is removing or eliminating emoticons, punctuation marks, characters (@, #, $, %, ^, &, *, etc.), URL, and number. Case folding is a process of changing all letters into uppercase or lowercase. Stopwords removes or filters words on tweets that are not needed. Stemming is a process of eliminating affixes, either prefixes or suffixes, so the resulting word is primary. Tokenizing is the process of cutting the existing sentences into words [14].

2.4. Bidirectional Encoder Representation from Transformers (BERT)

A collection of tweets that had been processed in the data pre-processing stage was then going through the process of feature extraction in Nature Language Processing (NLP) [18]. One of the methods of feature extraction is Bidirectional Encoder Representation from Transformers (BERT). In the existing researches, there are two types of BERT models that have been investigated for context-specific tasks: (1) BERT Base is smaller in size, computationally affordable, and not applicable to complex text mining operations and (2) BERT Large is extensive in size, computationally expensive, can be used for complex text mining operations, and crunches large text data to deliver the best result [19].

The model used in this research was the Indonesian BERT model, namely IndoBERT. Previously, the trained tokenizer and the model used was indobenchmark/indobert-base-p1. It means the BERT model used was the BERT Base. The feature extraction representated in each tweet in the data set can be seen in Figure 3.

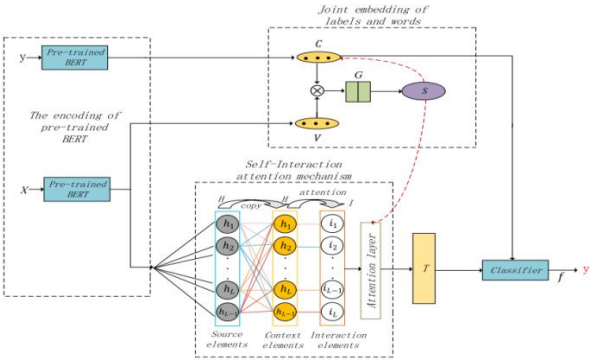

Figure 3. Feature Extraction Represented by BERT [20]

The dataset using BERT for its feature extraction was then processed into two stages: (1) becoming a vector of the sentence; (2) being pre-trained by the BERT tokenizer and then becoming a token. The result was to give a label to the token that can be used for model classification. Both of these processes were later included in the next step [18].

DOI: https://doi.org/10.29207/resti.v7i1.4889

2.5. N-gram

The use of the N-gram model as feature extraction for separating any desired words has been widely applied in various activities, such as predicting the correct spelling of finite words. In the case of spelling correction, N-gram is where a group of words have a certain length, then displayed as N-word. The tweet representation using the N-gram feature included Unigram, Bigram, and Trigram [14].

2.6. Term Frequency-Inverse Document Frequency (TF-IDF)

Term Frequency-Inverse Document Frequency (TF-IDF) is one of the feature extraction methods. Feature extraction is a function parameter used to represent each word in vector. TF-IDF is one of the weighting methods that combines together Term Frequency (TF) and Inverse Document Frequency (IDF) to do the weighting on the position of each word in any document [21]. Term Frequency (TF) shows the number of words that appears in a document, while Inverse Document Frequency (IDF) is the reverse process of Term Frequency (TF) which shows how often the word appears in the document. If a word appears more often, the weighted value will be smaller [14]. The TF-IDF Equation 1 was used in this research.

$$W_{ij} = tf_{ij} \times IDF_j, with\ IDF_j = (\log \frac{N}{df}) \quad (1)$$

With Wij is the weight of the document to the j-word, $tfij$ is the number of words searched for in a document, $IDFj$ is the Inverse Document Frequency, N is the entire documents, and the last df is the whole documents that contain the searched words.

2.7. Global Vectors (GloVe)

One method of feature expansion is Global Vectors (GloVe). The GloVe is an unsupervised learning algorithm for obtaining a vector of word representations [14]. This method came from a Stanford University research. The construct of the GloVe model utilizes the main benefit of count data while simultaneously capturing the meaningful linear substructures existed like Word2Vec. But with the same corpus, vocab, window size and training time GloVe consistently outperforms Word2Vec [22]. Algorithm 1 shows the steps of the feature expansion used.

Algorithm 1. GloVe

```
Input:textVector, corpusGloVe
Output:list similarity text
foreach val € GboW_vector do
    temp <- val.copy()
    for j <- 0 to val.length()do
     T <- cv.vocabulary_.get(similar[j])
if(val[j] = 0) and (similar[j] ≠
Null) and (t ≠ Null):
if(temp[t] = 1):
        val[j] <- 1
    end if
    end if
    end for
end for
```

The corpus built using GloVe is tweets, news, and combinations of tweets and news. The news used as data here is a compilation of news articles from Indonesian media outlets. They included CNN Indonesia, Tempo, Koran Sindo, Kompas and Republika. The result of the corpus that was built using vocab from GloVe can be seen in Table 6.

Table 6. GloVeCorpus

Corpus	Value Vocabulary
Tweet	20.733
IndoNews	79.346
Tweet+IndoNews	96.358

The word similarity was obtained with the corpus that was successfully built using the GloVe method [23]. Table 7 shows the top ten similar words to "Nyawa" in the similarity corpus built from the tweets. The value of each word determined the similarity of each keyword.

Table 7. Similarity Words of "Nyawa"

Rank	Similar Word	Value
1	Bayar	0.8986
2	Balas	0.8647
3	Layang	0.7990
4	Regang	0.7880
5	Cabut	0.7847
6	Hutang	0.7718
7	Rampas	0.7546
8	Hilang	0.7531
9	Lenyap	0.7424
10	Rugi	0.7013

The top ten similar words in Table 7 is an example of similar words used to continue the feature expansion process (GloVe) on the representation vector obtained from the feature extraction process (TF-IDF N-gram). The process to complete these vectors will be carried out by replacing vectors with a value of 0 with similar words in the GloVe corpus list [24].

2.8. Convolutional Neural Network (CNN)

The modeling technique used in this research was the Convolutional Neural algorithm Network (CNN). The implementation of the algorithm used the hardware library with language Python programming. The following algorithm shows the steps of the modeling based on the prior study. CNN consists of these layers: (1) Input layer is where you can input data into the models. The number of neurons or nerves in this layer corresponds with the total features in the existing data. (2) Convolutional layers consist of a set of filters and kernels. In this research, 1D convolutional layer is used because the data are one-dimensional texts as tweets. (3) Flatten layer is in between the convolutional and dense layers. This layer transforms a one-dimensional matrix of features into a vector that can be included in a fully connected neural network classifier. (4) Dense

DOI: https://doi.org/10.29207/resti.v7i1.4889

layer is merely an ordinary layer of neurons in a neural network. Each neuron receives input from all neurons in the previous layer so it can be tightly connected. (5) Output layer is the output of the hidden layer later entered into an output variable with the resulting value probability that will be displayed for the results of this CNN method [19]. The architecture of the one-dimensional CNN is shown in Figure 4.

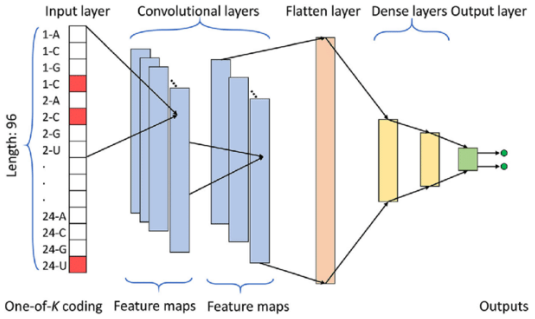

Figure 4. Architecture of 1D-CNN [25]

2.9. Performance Evaluation

The last stage was the evaluation stage. At this stage, the results of the modeling were issued previously using the confusion matrix method. The result was precision, recall, f1-score, and accuracy of the tweet data set that had been carried out at the classification stage with the CNN method. Precision is the ratio of accurate positive predictions compared to all positive predicted results; recall is the ratio of accurate positive predictions compared to all accurate positive data, while f1-score is the ratio of the average flat weighted precision and recall. Accuracy is the ratio of correct predictions (positive and negative) for the entire data [26]. Results testing was performed by testing the data based on the model built using training data. Here are the equation 2, 3, 4 nd 5 of precision, recall, f1-score, and accuracy for confusion matrix.

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

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

$$F1-Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (5)$$

TP is True Positive, FP is False Positive, TN is True Negative, and FN is False Negative. True Positive is when a tweet is predicted as a positive hoax and the result is a true hoax. False Positive is when a tweet is predicted as a positive hoax and the result is a false or non-hoax. True Negative is when a tweet is predicted as a negative hoax and the result is a true non-hoax; and False Negative is when a tweet is predicted as a negative hoax, and the result is a true hoax [27].

3. Results and Discussions

In this research, we applied CNN as a classification model, TF-IDF N-gram as a baseline, BERT, and GloVe as the expansion feature. The hoax detection system was built with four scenarios. The first scenario determined the resulting splitting ratio that was subsequently used in the next scenario. The second scenario compared baselines and found the best one for use in the next scenario. Next, the third scenario improved the baseline by combining it with another model. Last, the fourth scenario determined the highest accuracy of the system built by using corpus from tweets, IndoNews, and tweet+IndoNews. The accuracy results were the average five times iterations for each scenario.

3.1. Result

This study aims to obtain a model with the highest accuracy. The first scenario was done to determine the best splitting ratio in a test using CNN as a model and TF-IDF as the baseline. The splitting data ratios were 90:10, 80:20, 70:30, and 60:40 [18]. The results of the first scenario can be seen in Table 8. The table shows a splitting ratio of 90:10, which means 90% of data trained and 10% of data tested produced the highest accuracy. By using the same baseline, TF-IDF might produce different accuracy. As we can see, by using parameters minimum and maximum features could produce higher accuracy. The highest accuracy for the first scenario is 94.15%. With this, the splitting ratio of 90:10 and selected parameters as the baseline were then used for the next scenario.

Table 8. Results Comparison of the First Scenario

Ratio Splitting	Accuracy (%)	
	Using Min & Max Features	Without Min & Max Features
90:10	**94.15**	93.43
80:20	94	92.94
70:30	93.45	92.75
60:40	93.49	92.49

Testing was carried out to produce better results by comparing the baseline results. The combined baselines used are unigram, bigram, trigram, unigram bigram, unigram trigram and unigram bigram trigram. The results comparison among baselines can be seen in Table 9. The table shows that the combination of unigram and bigram produced the highest accuracy among the other baselines. The accuracy for the second scenario reached 94.55%. This baseline combination was also used for the next scenario.

Table 9. Results Comparison of the Second Scenario

Baseline	Accuracy (%)
Unigram	94.15
Bigram	91.58
Trigram	83.74
Unigram, Bigram	**94.55**
Unigram, Trigram	94.54
Unigram, Bigram, Trigram	91.76

DOI: https://doi.org/10.29207/resti.v7i1.4889

In the third scenario, not only the baseline but also the data was used for extraction. The extracted results from BERT were combined with the results from the previous baseline. The results of this combination can be seen in Table 10. This combination between baseline (unigram, bigram) and BERT produced a higher accuracy than if it only used the baseline. The accuracy for the third scenario is 95.56%.

Table 10. Results Comparison of the Third Scenario

Scenario	Accuracy (%)
CNN + baseline	94.55 (+0.42)
CNN + baseline + BERT	**95.56 (+1.49)**

Following the results of the previous scenario with a combination of baseline (unigram, bigram) and BERT to improve the extraction features, the GloVe expansion feature was added. Using the corpus that had been built with GloVe by looking for similar words and then connecting them with the existing expansion features, the results can be seen in Table 11. The first combination is a combination of baseline and corpus tweets. Then, the second combination is a combination between baseline and corpus IndoNews. And the last combination is a combination of baseline and corpus tweet+IndoNews, and this combination produced the highest accuracy similar to the top ten similar words in the corpus of tweet+IndoNews. The highest accuracy for the last scenario reached 98.57%.

Table 11. Results Comparison of the Fourth Scenario

Rank	Accuracy (%)		
	Baseline + Tweet	Baseline + IndoNews	Baseline + Tweet IndoNews
Top 1	98.00 (+4.09)	98.22 (+4.32)	98.45 (+4.57)
Top 5	98.31 (+4.42)	98.41 (+4.52)	98.48 (+4.59)
Top 10	98.35 (+4.46)	98.44 (+4.56)	**98.57 (+4.69)**
Top 15	98.28 (+4.39)	97.87 (+3.95)	96.98 (+3.01)

3.2. Discussion

Because the data set used in this study was balanced between hoaxes and non-hoaxes labels, the accuracy results are in line with the researchers' expectations. To produce maximum accuracy results, the researchers used a combination of CNN classification models, producing the best splitting ratio, the best baseline, added BERT to improve the baseline and using GloVe with various corpuses. Based on this result, a confusion matrix for one of the best scenarios can be seen in Figure 5, with axis x as Predicted Label and axis y as True Label.

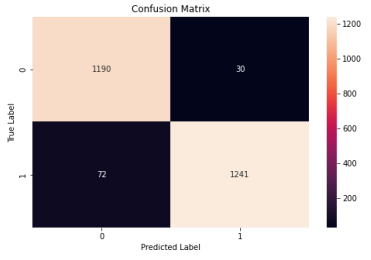

Figure 5. Confusion Matrix

The number of True Negative (TN) is 1190. Next, the number of False Positive (FP) is 30. Then the number of False Negative (FN) is 72. Lastly, the number of True Positive (TP) is 1241. Based on the result in the confusion matrix calculation above, the predictions where hoax label was given to TP and non-hoax label was given to TN are correct. According to these calculations, the accuracy obtained is optimal because the TP and TN values are directly proportional to its results.

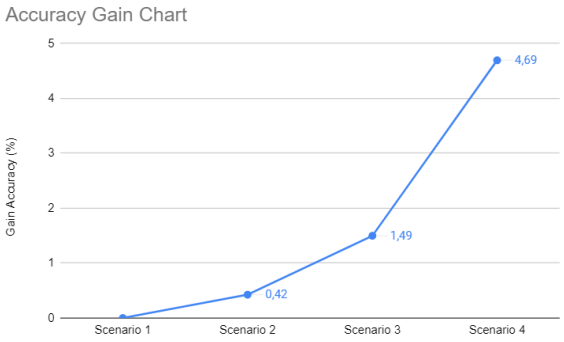

Figure 6. Accuracy Gain Chart

Based on the research done by combining the CNN method Figure 6 is a graph of the increase in accuracy in each scenario that has been carried out. The value taken is the most optimal accuracy value in each scenario. In the first scenario, a combination of baseline (unigram) by comparing the best data splitting ratio using CNN, which has the highest accuracy value of 94.15%. Then, the second scenario is a continuation of the previous scenario, namely using the best ratio, namely 90:10. The combination of the best ratios will use to compare the best baseline using CNN, which has the highest accuracy value of 94.55%. As has been done before, the third scenario is a continuation of the second scenario, namely a combination of the best ratio and the best baseline. But this time, BERT will add to optimize the baseline by using CNN. In this scenario, the accuracy obtained is 95.56%. The last scenario is the best combination of each scenario that is carried out using a 90:10 ratio, unigram-bigram baseline, BERT, and then adding GloVe. In the fourth scenario, the best use of the GloVe corpus is the tweet+IndoNews with a ranking of similar words, namely the top 10, which means the 10 most similar words. The accuracy obtained in this scenario is 98.57%, with an increase of 4.7% from the first scenario.

4. Conclusion

In this study, the researchers tried to predict fake news or hoaxes on social media, especially Twitter, using Convolutional Neural Network (CNN) as a modeling method, Term Frequency-Inverse Document Frequency (TF-IDF) with N-gram as baseline, Bidirectional Encoder Representations from Transformers (BERT) to improve the baseline, and Global Vectors (GloVe) as

DOI: https://doi.org/10.29207/resti.v7i1.4889

the expansion feature. The data the researchers collected consisted of 25.325 tweets from crawling results using snscrape tools. The data was balanced between hoaxes and non-hoaxes labels. The result of this research shows that the model combination of TF-IDF N-gram, BERT, GloVe, and CNN managed to achieve an accuracy of 98.57% with an increase in accuracy of 4.7%. Based on the test results, it can be concluded that combining multiple methods can improve the accuracy. In addition, the more corpuses used, the more optimal the accuracy will be. It can be recommended for further researches alongside with combining this system with the other methods such as Genetic Algorithm to obtain a better accuracy.

Reference

[1] 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)," *Indones. J. Comput.*, vol. 5, no. 3, pp. 23–36, 2020, doi: 10.34818/INDOJC.2020.5.3.506.

[2] R. K. Kaliyar, K. Fitwe, P. Rajarajeswari, and A. Goswami, "Classification of Hoax/Non-Hoax News Articles on Social Media using an Effective Deep Neural Network," *Proc. - 5th Int. Conf. Comput. Methodol. Commun. ICCMC 2021*, no. Iccmc, pp. 935–941, 2021, doi: 10.1109/ICCMC51019.2021.9418282.

[3] L. Nashif, "Detecting and Identifying Fake News on Twitter," no. c, pp. 37–40, 2021.

[4] B. Irena and Erwin Budi Setiawan, "Fake News (Hoax) Identification on Social Media Twitter using Decision Tree C4.5 Method," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 4, pp. 711–716, 2020, doi: 10.29207/resti.v4i4.2125.

[5] A. R. Jamaludin and E. B. Setiawan, "Deteksi Berita Hoax Di Media Sosial Twitter Dengan Ekspansi Fitur Menggunakan Glove," *eProceedings …*, vol. 9, no. 3, pp. 1847–1854, 2022, [Online]. Available: https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/17986%0Ahttps://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/17986/17615

[6] Y. Servatius, D. Langobelen, and E. B. Setiawan, "Ekspansi Fitur menggunakan GloVe pada Sistem Pendeteksi Hoax di Twitter," pp. 1–11, 2021.

[7] P. K. Laksana Utama, "Identifikasi Hoax pada Media Sosial dengan Pendekatan Machine Learning," *Widya Duta J. Ilm. Ilmu Agama dan Ilmu Sos. Budaya*, vol. 13, no. 1, p. 69, 2018, doi: 10.25078/wd.v13i1.436.

[8] I. Zukhrufillah, "Gejala Media Sosial Twitter Sebagai Media Sosial Alternatif," *Al-I'lam J. Komun. dan Penyiaran Islam*, vol. 1, no. 2, p. 102, 2018, doi: 10.31764/jail.v1i2.235.

[9] C. Boididou, S. Papadopoulos, M. Zampoglou, L. Apostolidis, O. Papadopoulou, and Y. Kompatsiaris, "Detection and visualization of misleading content on Twitter," *Int. J. Multimed. Inf. Retr.*, vol. 7, no. 1, pp. 71–86, 2018, doi: 10.1007/s13735-017-0143-x.

[10] C. S. Sriyano and E. B. Setiawan, "Pendeteksian Berita Hoax Menggunakan Naive Bayes Multinomial Pada Twitter dengan Fitur Pembobotan TF-IDF," *e-Proceeding Eng. Vol.8, No.2*, vol. 8, no. 2, pp. 3396–3405, 2021.

[11] M. K. Balwant, "Bidirectional LSTM Based on POS tags and CNN Architecture for Fake News Detection," *2019 10th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2019*, pp. 1–6, 2019, doi: 10.1109/ICCCNT45670.2019.8944460.

[12] A. A. Kurniawan and M. Mustikasari, "Implementasi Deep Learning Menggunakan Metode CNN dan LSTM untuk Menentukan Berita Palsu dalam Bahasa Indonesia," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, p. 544, 2021, doi: 10.32493/informatika.v5i4.6760.

[13] S. Sharma, M. Saraswat, and A. K. Dubey, "Fake News Detection Using Deep Learning," *Commun. Comput. Inf. Sci.*, vol. 1459 CCIS, pp. 249–259, 2021, doi: 10.1007/978-3-030-91305-2_19.

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

[15] F. S. Darusman, A. A. Arifiyanti, and S. F. A. Wati, "Sentiment Analysis Pedulilindungi Tweet Using Support Vector Machine Method," *Appl. Technol. Comput. Sci. J.*, vol. 4, no. 2, pp. 113–118, 2022, doi: 10.33086/atcsj.v4i2.2836.

[16] C. W. Kencana, E. B. Setiawan, and I. Kurniawan, "Hoax Detection on Twitter using Feed-forward and Back-propagation Neural Networks Method," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 3, pp. 648–654, 2017.

[17] W. Pamungkas and S. Suryani, "Deteksi Hoax Untuk Berita Hoax Covid 19 Indonesia Menggunakan CNN," vol. 8, no. 5, pp. 10264–10276, 2021.

[18] B. Anthony, C. Martani, and E. B. Setiawan, "Naïve Bayes-Support Vector Machine Combined BERT to Classified Big Five Personality on Twitter," vol. 5, no. 158, pp. 1072–1078, 2022.

[19] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," *Multimed. Tools Appl.*, vol. 80, no. 8, pp. 11765–11788, 2021, doi: 10.1007/s11042-020-10183-2.

[20] Y. Dong, P. Liu, Z. Zhu, Q. Wang, and Q. Zhang, "A Fusion Model-Based Label Embedding and Self-Interaction Attention for Text Classification," *IEEE Access*, vol. 8, pp. 30548–30559, 2020, doi: 10.1109/ACCESS.2019.2954985.

[21] 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 Eng.*, vol. 7, no. 2, pp. 7773–7781, 2020.

[22] P. M. Brennan, J. J. M. Loan, N. Watson, P. M. Bhatt, and P. A. Bodkin, "GloVe: Global Vectors for Word Representation," *Br. J. Neurosurg.*, vol. 31, no. 6, pp. 682–687, 2017, doi: 10.1080/02688697.2017.1354122.

[23] A. M. Zakaria and E. B. Setiawan, "Aspect-Based Analysis of Telkomsel User Sentiment on Twitter Using the Random Forest Classification Method and Glove Feature Expansion," *J. Teknol. dan Sist. …*, no. September, 2022, [Online]. Available: https://jtsiskom.undip.ac.id/article/view/14558

[24] E. B. Setiawan, D. H. Widyantoro, and K. Surendro, "Feature expansion using word embedding for tweet topic classification," *Proceeding 2016 10th Int. Conf. Telecommun. Syst. Serv. Appl. TSSA 2016 Spec. Issue Radar Technol.*, no. 2011, 2017, doi: 10.1109/TSSA.2016.7871085.

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

[26] R. C. Riana and Y. Sibaroni, "Hoax Detector of Covid 19 Indonesia in twitter using Rocchio Classification Method," vol. 8, no. 5, pp. 10427–10439, 2021.

[27] I. M. Mubaroq and E. B. Setiawan, "The Effect of Information Gain Feature Selection for Hoax Identification in Twitter Using Classification Method Support Vector Machine," *Indones. J. …*, vol. 5, no. September, pp. 107–118, 2020, doi: 10.21108/indojc.2020.5.2.499.

DOI: https://doi.org/10.29207/resti.v7i1.4889