**Building of Informatics, Technology and Science (BITS)**
Volume 4, No 2, September 2022 Page: 441−447
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1907

# Detection of Radicalism Speech on Indonesian Tweet Using Convolutional Neural Network

**Faiza Aulia Rahma Putra[1,*], Yuliant Sibaroni[2]**

School of Computing, Telkom University, Bandung, Indonesia
Email: [1,*]faizaaulia@student.telkomuniversity.ac.id, [2]yuliant@telkomuniversity.ac.id
Email Penulis Korespondensi: faizaaulia@ student.telkomuniversity.ac.id
Submitted:**21/07/2022**; Accepted:**18/08/2022**; Published**: 30/09/2022**

**Abstract**−The ease of disseminating information today is inseparable from the rapid development of information technology. Unfortunately, radical groups also use this condition to spread propaganda and recruit members through social media such as Facebook and Twitter. Therefore, detecting radicalism-related content on social media is essential for the prevention of online radicalization, given the ease with which information can be spread that can affect social media users. Several studies to classify radicalism speech have been carried out using machine learning algorithms such as K-Nearest Neighbor (KNN) and Support Vector Machine (SVM). However, only a few used the Indonesian language and even utilized a small dataset. This study proposed to detect radicalism speech in Indonesian tweets using Convolutional Neural Network (CNN) and Word2Vec as feature extraction. The dataset is a collection of Indonesian-language tweets obtained through tweet crawling. CNN modeling was conducted using several scenarios with the number of filter parameter values = 100 and 300, and kernel size parameter value = 3, 5, 7, 9. From the training process using those scenarios, the most optimal model is obtained with parameter filters = 300 and kernel size = 7, producing the best accuracy of 87.87% and average accuracy of 86.93%. Based on the best model obtained, an evaluation was carried out on the test data, which resulted in an accuracy of 87.15%.

**Keywords**: Radicalism Detection; Convolutional Neural Network; Word2Vec; Twitter; Text Classification

## 1. INTRODUCTION

The development of information technology as it now makes the dissemination of information impactful [1]. Nevertheless, it is also become an advantage for radical groups to spread their propaganda and recruit members through social networks such as Facebook and Twitter [2]–[4]. According to some experts, radicalism is an ideology of changing the social and political system using violence. The point of this radicalism is the attitude of a person or group of people who use violence to bring out the desired change, and generally, this group wants drastic change to happen immediately [5]. Generally, radicalism exists in social interaction spaces physically, such as the house of worship or other public area. However, in recent years, this has begun to move to the digital world via the internet, where some terrorist groups also use social media to spread their propaganda [4]. For example, extremist groups such as Al-Qaeda and ISIS have popularized their digital magazine on social media and spread their propaganda worldwide, which served as inspiration for the 2010 Boston Marathon bombers. In 2018, Twitter declared that 1.2 million user accounts had been suspended for posting terrorist propaganda [2]. Moreover, ISIS Indonesia has recruited young Indonesian from 2014 to 2019 through Twitter and Facebook [6]. Therefore, in this study, we propose a text classification model to detect whether social media content contains a radicalism speech or not.

Text classification is a supervised learning task aiming to group text documents into predetermined classes. This model learns from a labeled text document and predicts the new documents that do not yet have labels. Text classification can be single-labeled, which only groups by one class, and multi-label classification, which groups based on more than one class [7]. This study will use binary classification, which divides into two classes: **positive** class indicates that the text contains radicalism speech and **negative** class otherwise. Before the text document is classified, several steps must be carried out, such as pre-processing, tokenizing, and dimension reduction.

Convolutional Neural Networks (CNN) resemble conventional neural networks since they also consist of neurons that optimize themselves through learning [8]. In this type of network, the output of each neuron layer serves as the input of the next neuron layer. Multi-layer convolution is used to transform the nonlinear results of each layer up to the output layer. Generally, the convolutional neural network model used in text analysis consists of four layers: the embedding layer, the convolutional layer, the pooling layer, and the fully connected layer. Unlike conventional image analysis models, the CNN model's input layer for text analysis is the vector of words [9]. CNN has achieved impressive results in the text classification tasks even though this model was initially developed for the computer vision field [10].

Numerous studies on detecting textual content on social networks have been conducted previously. Research [11] by Subhan et al. (2017) used the K-Nearest Neighbor (K-NN) method to detect radicalism content on the extremist website, terror, and all content related to SARA (Ethnic, Races, and Religion issues). The collected data is labeled manually regarding the category defined by the Ministry of Communication and Informatics of the Republic of Indonesia. The best accuracy result was obtained by 66.37% at k=7. This result could be improved by combinate the classification algorithm such as Decision Tree, Support Vector Machine (SVM), and others. Research [6] by Miranda et al. (2020) used the SVM method to identify radical content on Twitter, with the highest accuracy level was 83.3%. Moreover, only 100 data have been utilized. This model categorizes tweets in Indonesian based on the number of

**Building of Informatics, Technology and Science (BITS)**
Volume 4, No 2, September 2022 Page: 441−447
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1907

occurrences of two keywords (ISIS and/or Syria) in the document. Another term associated with the radicalism content and the amount of data could be applied to enhance the model's classification performance.

Long Short Term Memory (LSTM) and CNN are used in research [12] by Ahmad et al. (2019) to detect radicalism content in social networks. The dataset was gathered from Twitter and numerous English and Arabic web forums containing extremist-related keywords, namely, ISIS, suicide, bomb, etc. This research applies the word embedding method to provide the model's input. Other research [13] by Taradhita and Putra (2021) uses CNN and Term Frequency-Inverse Document Frequency (TF-IDF) as the feature extraction to detect hate speech content on Indonesian tweets. Even though the trained model can detect hate speech content effectively, other feature extraction techniques such as Word2Vec could be implemented to enhance the model's accuracy. Word2Vec can produce words with semantic similarities to surrounding words [14], as opposed to TF-IDF which indicates the relevance of keywords to particular documents [15]. Research [16] by Kim (2014) classifies text on some benchmark datasets using CNN and Word2Vec. Even though the model only employs a simple 1-layer of convolution, it can perform well.

This study aims to pursue the best possible CNN parameters for detecting radicalism speech in Indonesian tweets and evaluate their performance. Apart from research [13], which employs TF-IDF as the feature extraction, this study will use Word2Vec as the feature extraction. To accomplish this, we will experiment with different Word2Vec models and several scenarios of CNN parameter values, such as the number of filters and kernel size (window size).

# 2. RESEARCH METHODOLOGY

## 2.1 Design System

The system to be constructed is radicalism speech detection on Indonesian tweets using Word2Vec and CNN classification method. The system development process starts with collecting the dataset from Twitter which will then be carried out in the preprocessing stage. After the preprocessing stage is done, the dataset will be split into a train set and a test set. The train set will be used in Word2Vec feature extraction and CNN modeling, while the test set will be used to evaluate the CNN model. Figure 1 below shows the diagram of the research processes.

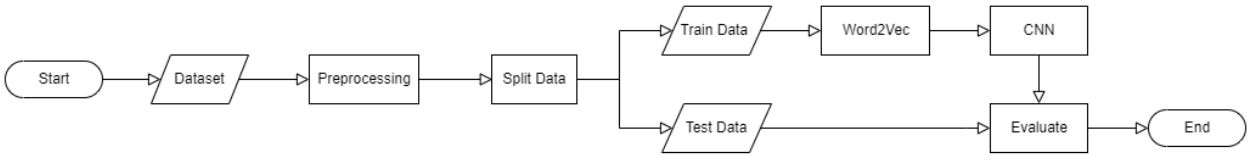


**Figure 1.** Research Processes

### 2.1.1 Dataset

In the research mentioned above, the dataset contains terms associated with radicalism in general, such as ISIS, Syria, bomb, and suicide. This study uses the dataset of Indonesian tweets with keywords related to radicalism issues specifically in Indonesia, such as **HTI**, **Khilafah**, **Syariah**, **Papua Merdeka**, etc. This data was collected through the crawling process using the Twint open-source library with Python programming language. The dataset is limited to tweets posted between January 2016 and Mei 2017; no user background checks are performed during the crawling process. The collected data will be manually labeled with label 1 for positive class, indicating the tweets containing radicalism-related utterances, and label 0 for negative class, indicating otherwise. Three annotators are used during the labeling, and the commonly occurred label will be used as the final label. This labeling is based on several criteria of the radical website by the Indonesian National Counterterrorism Agency (BNPT), one of which is the desire to make rapid change using religiously motivated violence [17].

### 2.1.2 Preprocessing

Since the collected data is unstructured, several preprocessing steps are required before being used for Word2vec and CNN modeling. The following steps are performed during the preprocessing:

a. Text Cleansing
   The first step is removing unnecessary characters or text such as URL, punctuation, HTML tag, a space character, numeral, and emojis. Next, the data will undergo the case-folding process where all of the text is converted to lower case. The purpose of the case-folding is to normalize the data in the same form case.
b. Duplicate Handling
   During the exploratory data analysis, many tweets were identical but contained different URLs. Therefore, duplicate data will be removed once there is no URL in the tweets.
c. Normalization & Stemming
   The stemming process is transforming the words into their root forms. In this stage, the normalization of slang words is also carried out to become a standard language. This step is done using the Sastrawi python library.
d. Stopwords Removal

**Building of Informatics, Technology and Science (BITS)**
Volume 4, No 2, September 2022 Page: 441−447
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1907

The process of removing words that have no meaning; these words frequently occur in documents but do not contribute to the model's learning process. This step uses the Indonesian stopword dictionary from the Natural Language Toolkit (NLTK) library.

e. Tokenization
Tokenization is splitting each document into the smallest unit widely recognized as a token. These tokens are used to be the input of the word embedding process.

### 2.1.3 Word2Vec

Mikolov et al. (2013) introduced a word embedding method called Word2Vec. Unlike the non-distributed (high-dimensional, sparse) representation of words typical of bag-of-words (BoW) or one-hot encoding, Word2Vec is a distributed (low-dimensional, dense) representation of words [18]. Word2Vec comes with two types of shallow neural network architectures, as shown in Figure 2: Skip-gram and Continuous Bag of Words (CBoW). It utilized distributed (low-dimensional, dense) representation of words to identify together words with similar meanings. This study uses the CBoW Word2Vec model, which predicts the center word $w_0$ based on a representation of the context words $w_{-c}, \dots, w_{-1}, w_1, w_c$ as the input. Furthermore, the projection layer will sum up the embeddings of the context word to produce the output [19], [20]. The output of the Word2Vec model is a vocabulary where each word is associated with a vector.

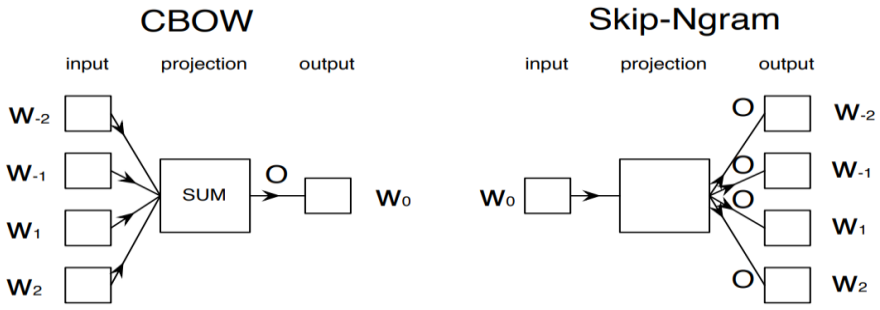


**Figure 2.** Word2Vec Models [19]

### 2.1.4 CNN (Convolutional Neural Network)

As depicted in Figure 3 below, the CNN model for text classification consists of 4 main layers. The embedding layer is the first layer, which accepts a word vector representation as input. This vector comes from the Word2Vec output, with the sentence matrix's size of $s \times d$, where $s$ is the number of words in a sentence and $d$ is word vector dimensionality. In this architecture, the length of the sentences will be padded to maintain the same length. The second layer is the convolutional layer. The idea of this CNN is to filter the information from the input and produce the feature map.

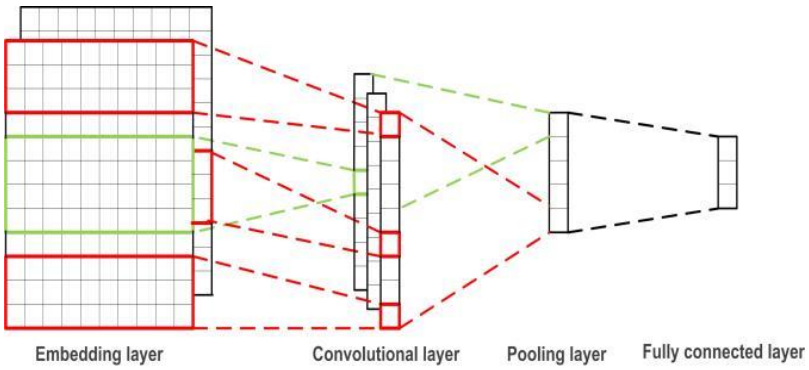


**Figure 3.** Typical CNN Model Structure [9]

Let $\mathbf{A} \in \mathbb{R}^{s \times d}$ is the sentence matrix, and $\mathbf{A}_{i:j}$ represents the sub-matrix of $\mathbf{A}$ from row $i$ to $j$. The convolution operation performed involves the *filter* $\mathbf{w} \in \mathbb{R}^{h \times d}$ that applied to a window of $h$ words to generate a new feature [21]. This operation will be performed repeatedly in order to generate new feature $c_i$ by applying the filter on each sub-matrix of $\mathbf{A}$ with the following equation (1).

$$c_i = f(\mathbf{w} \cdot \mathbf{A}_{i:\, i+h-1} + b) \quad (1)$$

Where $f$ is an activation function such as ReLU is used in this study, $i = s - h + 1$, and $b \in \mathbb{R}$ is a bias term. The generated *feature map* will be shown in equation (2):

$$\mathbf{c} = [c1, c2, \dots, c_{s-h+1}] \quad (2)$$

**Building of Informatics, Technology and Science (BITS)**
Volume 4, No 2, September 2022 Page: 441−447
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1907

with $\mathbf{c} \in \mathbb{R}^{s-h+1}$. Dropout is one regularization method that randomly settings values in the weight vector to 0. Using dropout could impede the hidden layer's co-adaptation and minimize model overfitting [21]. The next layer is the pooling layer. The feature map's dimensionality in this layer will be reduced by aggregating the information. One of the pooling types is max pooling, which obtained the required feature of a sentence by selecting the maximum value [12].

The next layer is the flatten layer. The flatten layer transforms the pooled feature map into a single vector and then feeds it to the densely connected layer. And the last layer is the dense layer, commonly known as the fully connected layer. This layer specifies which features are most associated with a specific class by giving each class's final probabilities based on the pooling layer result. This study used the sigmoid activation function to predict the probabilities of positive and negative classes between 0 and 1 [13].

Overall, the CNN model in this has the following configurations: embedding layer with a size of 300, convolutional layer with ReLU activation function, dropout layer on 0.5 rates, one max pooling layer, flatten layer, and two dense layers which use ReLU activation function with 128 units and the sigmoid activation function to classify the feature class.

#### 2.1.5 Evaluation

The CNN model will be evaluated using the confusion matrix and the classification report. The confusion matrix shows performance measurement values such as accuracy, precision, and recall through the following equation (3) - (5) below.

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

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

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

Where TP is True Positive that indicates a radicalism speech predicted as radicalism speech by the model, FP is False Positive where non-radicalism speech is detected as radicalism speech by the model. TN is True Negative where the model detects non-radicalism speech as non-radicalism, and FN is False Negative where the model detects non-radicalism speech as radicalism speech.

## 3. RESULT AND DISCUSSION

### 3.1 Dataset

The total amount of collected data is 5493 rows, which total of negative class data is 4474, while the positive class is only 1019. The data sample and its class are presented in Table 1 below.

**Table 1.** Dataset Sample

| Tweet | Class |
|---|---|
| INDONESIA sudah bukan lagi negara Hukum yang Demokrasi, tapi Negara Islam. | Positive |
| Indonesia sekarang negara islam.. #islam #isis | Positive |
| Sedih nama islam di indonesia sudah jelek. | Negative |
| Kr Indonesia tidak hanya Islam #IndonesiaGakButuhHTI | Negative |
| #IndonesiaGaButuhHTI Indonesia damai bersama masyarakat Islam tradisional... | Negative |

The collected dataset remains in its unstructured form, therefore preprocessing steps are carried out. After preprocessing is done, the structure of the dataset is inevitably affected. The total data is reduced to 3968, of which 3102 are positive, and 866 are negative. The results of each preprocessing step are displayed in Table 2 as follows.

**Table 2.** Preprocessing Steps and Its Result

| Preprocessing Step | Input | Output |
|---|---|---|
| Text Cleansing | @ShintaAleyda @jokowi masihkah ini berlaku ??? paham komunis yg jelas Anti Pancasila & Anti NKRI didiamkan saja ??? https://t.co/Y3Uqedwa1N | masihkah ini berlaku paham komunis yg jelas anti pancasila anti nkri didiamkan saja |
| Normalization & Steming | masihkah ini berlaku paham komunis yg jelas anti pancasila anti nkri didiamkan saja | masih ini laku paham komunis yang jelas anti pancasila anti nkri diam saja |
| Stopword Removal | masih ini laku paham komunis yang jelas anti pancasila anti nkri diam saja | laku paham komunis anti pancasila anti nkri diam |
| Tokenization | laku paham komunis anti pancasila anti nkri diam | laku, paham, komunis, anti, pancasila, anti, nkri, diam |

**Building of Informatics, Technology and Science (BITS)**
Volume 4, No 2, September 2022 Page: 441−447
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1907

### 3.2 Word2Vec

Before jumping into the CNN modeling experiment, we try to compare two different models of Word2Vec. The first model is a pre-trained Word2Vec model with an Indonesian Wikipedia corpus, and the second model is retrained Word2Vec model using the corpus from the radicalism dataset. Table 3 shows the difference between those models for finding the most similar words to a keyword, such as "HTI." It shows that the second model was better, with similar words more relevant to a keyword than the first model, which the "HTI" is one of the organizations frequently associated with the issue of radicalism in Indonesia.

**Table 3.** Word2Vec Model Comparison

| Pre-trained Word2Vec | | Re-trained Word2Vec | |
|---|---|---|---|
| Similar Word | Similarity | Similar Word | Similarity |
| jamsostek | 54.16% | ormas | 60.9% |
| pengusahaan | 53.51% | kamibersamahti | 59.22% |
| hph | 53.30% | khilafah | 57.68% |
| perum | 52.68% | hizbut | 55.95% |
| migas | 52.65% | tahrir | 55.93% |
| agraria | 51.41% | ummat | 53.26% |
| bpr | 51.41% | dibubarin | 52.06% |
| pkp | 50.40% | ancam | 49.68% |
| perbenihan | 50.14% | islam | 49.55 % |
| berdikari | 49.90% | nkri | 48.60% |

We also defined a simple neural network with one hidden layer to check how those models perform. As a result, the second model performed better with the best train accuracy of 92.63% and validation accuracy of 88.41%, whereas the first model with 88.41% of train accuracy and 86.40% of validation accuracy. With this result, we decide to use the second model in the CNN as the embedding matrix.

### 3.3 CNN Model

This study has built the CNN model for radicalism speech classification using one convolution layer with a ReLU activation function. Dropout was also applied to the model with a value of 0.5. The training process of the CNN model has been carried out using several scenarios on parameter values: filters = 100, 300, and kernel size (window size) = 3, 5, 7, 9. All of those scenarios are being trained using cross-validation with k=5. The training data is 80% of the entire dataset, which consists of 2481 negative classes and 693 positive classes. The result of the training process for each model is shown in Table 4 below.

**Table 4.** Accuracy of Each CNN Model

| Model | Filters | Kernel Size | Average Accuracy |
|---|---|---|---|
| 1 | 100 | 3 | 86.77% |
| 2 | 100 | 5 | 86.58% |
| 3 | 100 | 7 | 86.70% |
| 4 | 100 | 9 | 86.14% |
| 5 | 300 | 3 | 86.83% |
| 6 | 300 | 5 | 86.70% |
| **7** | **300** | **7** | **86.93%** |
| 8 | 300 | 9 | 86.17% |

Table 4 indicates that the CNN model number 7 with parameter filter=300 and kernel size=7 yields the best result based on the average accuracy. Table 5 below shows the accuracy of model number 7 for each training fold, where the best accuracy is obtained at the value of k=4. Based on these results, CNN model number 7 will be used to detect radicalism speech in the Indonesian tweets.

**Table 5.** Model 7 Accuracy

| Fold | Accuracy |
|---|---|
| 1 | 86.46% |
| 2 | 85.83% |
| 3 | 87.40% |
| **4** | **87.87%** |
| 5 | 87.07 % |
| Average | 86.93% |

**Building of Informatics, Technology and Science (BITS)**
Volume 4, No 2, September 2022 Page: 441−447
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1907

### 3.3 Evaluation

The evaluation process is carried out using a test set of data which is 20% of the entire dataset, with 621 negative and 173 positive classes. The CNN model number 7 performs well with an accuracy of 87.15%, with the detail of the confusion matrix as follows: TP=83, FP=12, TN=609, FN=90. Based on that confusion matrix, the value of the model's precision=87.37% and recall=47.98%. The recall has a reasonably low value for the overall model; it can be an outcome of the unbalanced data between the positive and negative classes, so the model needs more balance and varied data to be trained, especially for the positive class.

## 4. CONCLUSION

The following are the conclusion that summarizes the findings of the conducted research. The CNN model is able to detect radicalism speech in Indonesian tweets with an accuracy of 87.15%. The accuracy tends to rise with the increasing filter parameter value since more information is extracted to form a new feature. However, increasing the number of filters will also increase the model's complexity and the running time because more convolution operations are performed. Although the CNN model has performed exceptionally well, it can be improved by adding more varied data to the dataset so the model can train more effectively, especially on the positive data classes, and it would be preferable if there were an equal amount of positive and negative classes.

## REFERENCES

[1] A. Rekik, S. Jamoussi, and A. ben Hamadou, "Violent Vocabulary Extraction Methodology: Application to the Radicalism Detection on Social Media," in Computational Collective Intelligence, 2019, pp. 97–109. doi: https://doi.org/10.1007/978-3-030-28374-2_9.

[2] M. Nouh, J. R. C. Nurse, and M. Goldsmith, "Understanding the Radical Mind: Identifying Signals to Detect Extremist Content on Twitter," in 2019 IEEE International Conference on Intelligence and Security Informatics (ISI), 2019, pp. 98–103. doi: 10.1109/ISI.2019.8823548.

[3] S. Aldera, A. Emam, M. Al-Qurishi, M. Alrubaian, and A. Alothaim, "Online Extremism Detection in Textual Content: A Systematic Literature Review," IEEE Access, vol. 9, pp. 42384–42396, 2021, doi: 10.1109/ACCESS.2021.3064178.

[4] M. Fernandez, M. Asif, and H. Alani, "Understanding the Roots of Radicalisation on Twitter," in WebSci 2018 - Proceedings of the 10th ACM Conference on Web Science, May 2018, pp. 1–10. doi: 10.1145/3201064.3201082.

[5] I. Nur, A. H. Nawawie, H. Fajarwati, and H. Chusna, "Embracing Radicalism and Extremism in Indonesia with the Beauty of Islam," Asian Research Journal of Arts & Social Sciences, pp. 1–18, Jan. 2020, doi: 10.9734/arjass/2020/v10i230141.

[6] E. Miranda, M. Aryuni, Y. Fernando, and T. M. Kibtiah, "A Study of Radicalism Contents Detection in Twitter: Insights From Support Vector Machine Technique," in 2020 International Conference on Information Management and Technology (ICIMTech), 2020, pp. 549–554. doi: 10.1109/ICIMTech50083.2020.9211229.

[7] V. K. Vijayan, K. R. Bindu, and L. Parameswaran, "A comprehensive study of text classification algorithms," in 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2017, pp. 1109–1113. doi: 10.1109/ICACCI.2017.8125990.

[8] J. Teuwen and N. Moriakov, "Convolutional neural networks," in Handbook of Medical Image Computing and Computer Assisted Intervention, Elsevier, 2019, pp. 481–501. doi: 10.1016/B978-0-12-816176-0.00025-9.

[9] A. Jacovi, O. S. Shalom, and Y. Goldberg, "Understanding Convolutional Neural Networks for Text Classification," Sep. 2018, [Online]. Available: http://arxiv.org/abs/1809.08037

[10] Z. Wang and Z. Qu, "Research on Web text classification algorithm based on improved CNN and SVM," in 2017 IEEE 17th International Conference on Communication Technology (ICCT), 2017, pp. 1958–1961. doi: 10.1109/ICCT.2017.8359971.

[11] Muh. Subhan, A. Sudarsono, and A. Barakbah, "Preprocessing of Radicalism Dataset to Predict Radical Content in Indonesia," in 2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), 2017, pp. 270–275. doi: 10.1109/KCIC.2017.8228598.

[12] S. Ahmad, M. Z. Asghar, F. M. Alotaibi, and I. Awan, "Detection and classification of social media-based extremist affiliations using sentiment analysis techniques," Human-centric Computing and Information Sciences, vol. 9, no. 1, p. 24, 2019, doi: 10.1186/s13673-019-0185-6.

[13] 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, Feb. 2021, doi: 10.5614/itbj.ict.res.appl.2021.14.3.2.

[14] F. Enríquez, J. A. Troyano, and T. López-Solaz, "An approach to the use of word embeddings in an opinion classification task," Expert Systems with Applications, vol. 66, pp. 1–6, 2016, doi: https://doi.org/10.1016/j.eswa.2016.09.005.

[15] S. Qaiser and R. Ali, "Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents," International Journal of Computer Applications, vol. 181, Jul. 2018, doi: 10.5120/ijca2018917395.

[16] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Oct. 2014, pp. 1746–1751. doi: 10.3115/v1/D14-1181.

[17] A. Aghnia and I. Sandy, "Kriteria Situs Islam Radikal Versi BNPT," CNN Indonesia, Apr. 01, 2015. https://www.cnnindonesia.com/teknologi/20150401093434-185-43429/kriteria-situs-islam-radikal-versi-bnpt (accessed Jan. 18, 2022).

[18] T. Adewumi, F. Liwicki, and M. Liwicki, "Word2Vec: Optimal hyperparameters and their impact on natural language processing downstream tasks," Open Computer Science, vol. 12, no. 1, pp. 134–141, Jan. 2022, doi: 10.1515/comp-2022-0236.

**Building of Informatics, Technology and Science (BITS)**
Volume 4, No 2, September 2022 Page: 441−447
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1907

[19] W. Ling, C. Dyer, A. W. Black, and I. Trancoso, "Two/Too Simple Adaptations of Word2Vec for Syntax Problems," in Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, May 2015, pp. 1299–1304. doi: 10.3115/v1/N15-1142.

[20] B. Jang, I. Kim, and J. W. Kim, "Word2vec convolutional neural networks for classification of news articles and tweets," PLoS ONE, vol. 14, no. 8, Aug. 2019, doi: 10.1371/journal.pone.0220976.

[21] Y. Zhang and B. Wallace, "A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification," in Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Nov. 2017, pp. 253–263. [Online]. Available: https://aclanthology.org/I17-1026