**Scientific Journal of Informatics**
Vol. 11, No. 2, May 2024

p-ISSN 2407-7658 https://journal.unnes.ac.id/journals/sji/index e-ISSN 2460-0040

# Modified Convolutional Neural Network for Sentiment Classification: A Case Study on The Indonesian Electoral Commission

**Slamet Riyadi[1*], Naufal Gita Mahardika[2], Cahya Damarjati[3], Suzaimah Ramli[4]**
[1,2,3]Department of Information Technology, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Indonesia
[4]Department of Computer Science, Universiti Pertahanan Nasional Malaysia, Malaysia

**Abstract.**
**Purpose:** This study aims to analyze public sentiment towards the Indonesian Electoral Commission (KPU) performance and evaluate a modified Convolutional Neural Network (CNN) model effectiveness in sentiment analysis.
**Methods:** This research employs several methods to achieve its objectives. First, data collection was conducted using web crawling techniques to gather public opinions on the performance of the Indonesian Electoral Commission for the 2024 elections, with a specific focus on platform X. A total of 5,782 data points were collected and underwent preprocessing before sentiment analysis was performed. This study uses the CNN method due to its exceptional ability to recognize patterns and features in text data through its convolutional layers. CNN is highly effective in sentiment analysis tasks because of its ability to capture local context and spatial features from text data, which is crucial for understanding the nuances of sentiment in comments. The modified CNN model was then trained and evaluated using a labeled dataset, where each comment was classified into positive, negative, or neutral sentiment categories. Modifying the CNN model involved adjusting its architecture and parameters, as well as adding layers such as batch normalization and dropout to optimize its performance. The effectiveness of the modified CNN model was assessed based on metrics such as classification accuracy, precision, recall, and F1 score. Through this methodological approach, the research aims to gain insights into public sentiment towards the KPU performance in the 2024 elections and to evaluate the effectiveness of the modified CNN model in sentiment analysis.
**Result:** The research revealed several significant findings. Firstly, most comments expressed concerns regarding performance aspects of KPU's, including transparency, fairness, and integrity. Neutral sentiment dominated the discourse, with approximately 23.66% of comments conveying dissatisfaction or skepticism towards KPU's handling of the elections. Additionally, sentiments expressed on social media platform X mirrored those found across other platforms, indicating a consistent perception of KPU performance among users. Furthermore, the evaluation of the modified CNN model demonstrated a substantial improvement in accuracy, achieving an impressive 93% accuracy rate compared to the pre-modification model's accuracy of 77%. These results suggest that the modifications made to the CNN model effectively enhanced its performance in sentiment analysis tasks related to KPU performance during the 2024 elections. These findings contribute to a deeper understanding of public sentiment toward KPU performance and underscore the importance of leveraging advanced technology, such as modified CNN models, for sentiment analysis.
**Novelty:** This study contributes novelty in several ways. Firstly, it provides insights into public sentiment towards the performance of the KPU during the 2024 General Elections, which is crucial for understanding the perception of democracy in Indonesia. Second, the study employs a mixed-methods approach, combining web crawling techniques for data collection and a modified CNN model for sentiment analysis, which offers a comprehensive and advanced methodology for analyzing sentiments on social media platforms. Thirdly, the evaluation of the modified CNN model demonstrates a significant improvement in accuracy, indicating the approach's efficacy in analyzing sentiments related to KPU performance. This study offers valuable contributions to academic research and practical applications in sentiment analysis, particularly in democratic processes and institutional performance evaluation.

**Keywords**: Sentiment analysis, KPU, CNN, Deep learning, General election
**Received** May 2024 / **Revised** May 2024 / **Accepted** June 2024



## INTRODUCTION

Public sentiment analysis is the process of identifying and categorizing opinions expressed in text, primarily to determine the public's reaction to a specific topic, whether positive, negative, or neutral. In the context of the 2024 General Election in the Republic of Indonesia, public sentiment analysis holds significant importance as it helps understand the perceptions and opinions of the public regarding the process and

*Corresponding author.
Email addresses: riyadi@umy.ac.id (Riyadi)
DOI: 10.15294/sji.v11i2.4929

performance of related institutions. This General Election marks a crucial moment for the nation's democracy. As an institution responsible for ensuring a fair, transparent, and democratic electoral process, the General Election Commission (KPU) plays a vital role. Together with institutions like the Constitutional Court, Judicial Commission, and Supreme Court, the KPU bears substantial responsibility for upholding Indonesia's democratic integrity, as mandated by Article 22E, paragraph 5, which stipulates that elections are organized by a national, permanent, and independent General Election Commission [1]. Consequently, the performance of the KPU garners widespread attention from both the public and political observers, as it is the institution accountable for overseeing elections in Indonesia. Therefore, continuous improvement and optimization of KPU performance are crucial to ensuring the democratic, fair, and smooth conduct of elections [1].

In the digital era, numerous platforms serve as avenues for individuals to express their opinions. In this context, platform X emerges as the primary social media platform for people to voice their views, particularly regarding their perceptions of KPU performance. During the 2024 elections, there were widespread rumors of potential fraud and ballot manipulation scandals in the presidential and vice-presidential elections of the Republic of Indonesia. These rumors sparked concerns and debates among the populace regarding the KPU integrity in fulfilling its duties. In such circumstances, sentiment analysis is needed to understand the general public's perceptions of KPU performance and to assess the effectiveness of modified CNN applications, as well as to ascertain whether the public perceives KPU performance as predominantly negative, positive, or neutral [2].

This study conducts sentiment analysis on the performance of the KPU in the 2024 election based on comments on social networking platform X, utilizing a modified CNN method. CNN is employed for sentiment analysis due to its capability to automatically extract features from complex inputs [3]. This is facilitated by its structure consisting of convolutional layers, allowing the model to identify important patterns in data, including in text contexts such as comments on social networking platforms. In this way, CNN can effectively capture nuances of sentiment that may be present in lengthy and varied sentences. Moreover, its effectiveness in classifying information within sentences is also evident due to its ability to learn increasingly abstract representations of data through deep layers. This enables CNN to differentiate between various types of sentiment, such as positive, negative, or neutral, even in complex contexts.

However, the application of CNN in sentiment analysis has limitations, including CNN inability to capture complex contextual meanings in sentences. CNN tends to rely on simpler or more direct patterns in text, leading to inaccuracies in understanding sentence nuances and context [4]. Additionally, CNN may be less effective in addressing class imbalance issues in sentiment data, where positive, negative, and neutral classes are unbalanced.

With the occurrence of the 2024 general elections organized by the Indonesian Electoral Commission, many people have provided opinions or responses regarding the performance of the KPU during the 2024 general elections. Some previous studies have adopted the CNN method for sentiment analysis, with variations in approach and accuracy. These studies have included the use of additional convolution layers to improve accuracy, as done by Listyarini and Anggoro (2021) [2], while others have implemented the LSTM method for sentiment analysis, as done by Irawati Setiawan et al. (2020) [5]. The range of accuracy achieved by these studies varies, ranging from 55.7% to 82%.

Based on related studies, much research on societal sentiment analysis utilize various methods and algorithms for testing and obtaining various levels of accuracy. This serves as supporting literature or references in the application and modification of the CNN method used in this research.

Therefore, this study aims to enhance CNN performance in sentiment analysis by modifying the CNN model through the addition of layers and balancing data before testing, thus improving accuracy in analyzing public sentiment towards KPU performance during the 2024 General Elections, the central focus of this research.

## METHODS

In this stage, the research process is explained from data collection to transforming it into meaningful information, with a design in Figure 1. This research consists of 5 stages: data collection, data preprocessing, data labeling, data balancing, data testing, and evaluation.

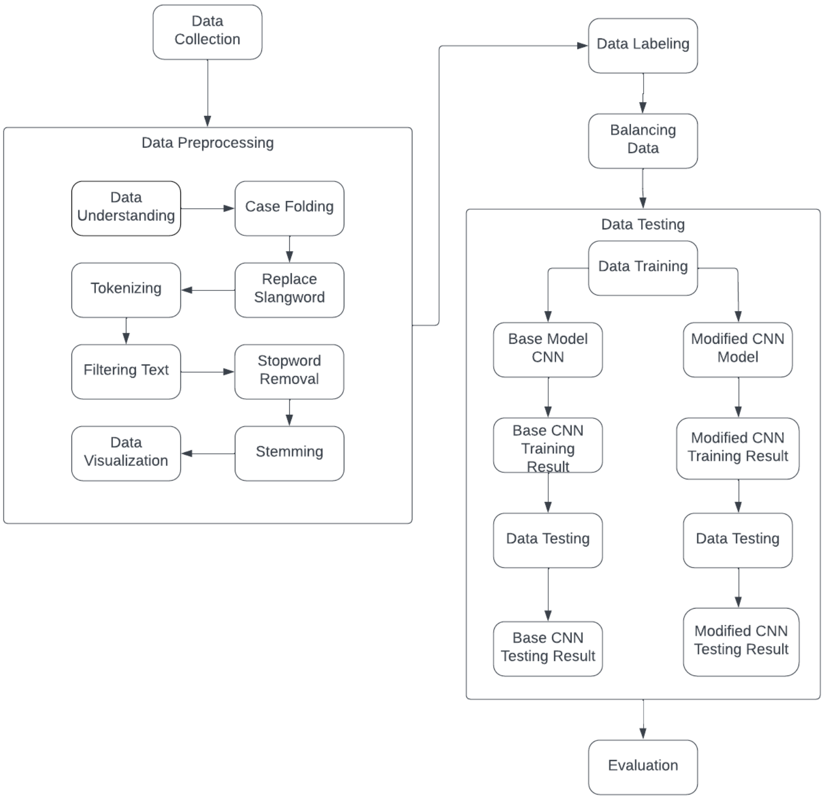



Figure 1. Research steps

**Data collection**

In Figure 1, it is explained that the research flow begins with Data Collection. In this stage, data is collected by crawling tweets from social media X using a tool called Tweet Harvest. In this study, primary data from X was utilized because X was the most suitable for gathering data, particularly regarding public opinions. This is attributed to the 'tweet' feature on X, which facilitates individuals in expressing opinions and engaging in discussions about various issues [6]. This study retrieves data such as posting date, user ID, tweet text, number of quotes, number of replies, number of retweets, number of favorites, language, user ID string, conversation ID string, username, tweet link, and obtains 5782 rows of data, ranging from December 3, 2023, to February 21, 2024. An example dataset can be seen in Table 1:

Table 1. Dataset example

| Created_at | Id_str | Full_text | Quote_count | Reply_count | Retweet_count | Favorite_count | lang | Conversation_id_string | username | Tweet_url |
|---|---|---|---|---|---|---|---|---|---|---|
| Wed Feb 21 23:57:45 +0000 2024 | 17xxxx | Tweet example | 12 | 23 | 20 | 6 | in | 17xxxxx | @usn | https://twitter.com/xxxx/xxx |

**Preprocessing data**

Data preprocessing is cleaning and transforming raw data into a format that is easier to process using Natural Language Processing (NLP) algorithms. This stage involves data understanding, case folding, replacing slang words, tokenizing, text filtering, stopword removal, stemming, and data visualization [7].

**Data understanding**

Data understanding is an in-depth process to understand the characteristics, structure, and content of the data being used. In this stage, each row and column of the entire data are examined to monitor for any defects or deficiencies. The data will be removed if any defective or inadequate data is detected [8].

**Case folding**

Case folding converts all sentences within the dataset into lowercase letters to make text processing more consistent and efficient, enabling accurate string matching without considering differences in uppercase and lowercase letters [9]. An example of the result of case-folding can be seen in Table 2 below:

Table 2. An example of the result of applying case folding

| Before | After |
|---|---|
| Biar Dikira kerja ya Min? | biar dikira kerja ya min? |

**Replace slang word**

Replace slang words is the process of changing slang or colloquial language in a sentence using a slang word dictionary for the Indonesian language [10]. This process aims to clean the collected data from slang and informal language, making it easier to analyze statistically. An example of the result of replacing slang words can be seen in Table 3 below:

Table 3. An example of the result of applying to replace slang word

| Before | After |
|---|---|
| Yg curang itu 02, yg ga bener itu sistem sirekap, yg ga bener kerjanya itu KPU, yg ga kerja itu Bawaslu. Ga ada yg nyalahin petugas KPPS. Biarlah petugas KPPS itu bahagia bisa nyicil Pajero dr gajinya 😆 | yang curang itu 02, yang enggak benar itu sistem sirekap, yang enggak benar kerjanya itu kpu, yang enggak kerja itu bawaslu. enggak ada yang menyalahkan petugas kpps. biarlah petugas kpps itu bahagia bisa mencicil pajero dari gajinya 😆 |

**Tokenizing**

Tokenizing is the process of breaking sentences into smaller units or tokens [7]. This process utilizes the nltk module to assist in breaking Indonesian sentences into individual words or tokens. An example of tokenizing results can be seen in Table 4 below:

Table 4. An example of the result of applying to tokenize

| Before | After |
|---|---|
| Copot segera sdh berapakali bikin kesalahan ..skrg ternyata lebih jahat lagi cara kerja KPU ....sy sendiri sdh muak lihat mukanya dan Rakyat yg cinta Jurdil disrluruh nusantara sdn muak juga lihat mukanya.... | ['copot', 'segera', 'sudah', 'berapakali', 'buat', 'kesalahan', 'skrg', 'ternyata', 'lebih', 'jahat', 'lagi', 'cara', 'kerja', 'kpu', 'sendiri', 'sudah', 'muak', 'lihat', 'mukanya', 'dan', 'rakyat', 'yang', 'cinta', 'jujur', 'adil', 'diseluruh', 'nusantara', 'muak', 'juga', 'lihat', 'mukanya'] |

**Filtering text**

Filtering text is removing irrelevant words in text data, such as conjunctions, symbols, prepositions, or absorbed [11]. An example of filtering text results can be seen in Table 5 below:

Table 5. An example of the result of applying filtering text

| Before | After |
|---|---|
| Guwa Kpps. Tersinggung berat dengan bacotan anda wahai Gaji ga seberapa kerja serius 24 jam non stop. Giliran sirekap yg ngaco kalian salahin Kpps. Gile !!! | guwa kpps tersinggung berat dengan bacotan anda wahai id gaji enggak seberapa kerja serius jam non stop giliran sirekap yang mengaco kalian salahkan kpps gila |

**Stopword removal**

Stopword Removal is the process of eliminating less important or non-descriptive words, allowing the analysis focus to be directed towards key words containing sentiment meanings. This process utilizes functions from the Natural Language Toolkit (NLTK) library [8]. An example of Stopword Removal results can be seen in Table 6 below:

Table 6. An example of the result of applying stopword removal

| Before | After |
|---|---|
| ['copot', 'segera', 'sudah', 'berapakali', 'buat', 'kesalahan', 'skrg', 'ternyata', 'lebih', 'jahat', 'lagi', 'cara', 'kerja', 'kpu', 'sendiri', 'sudah', 'muak', 'lihat', 'mukanya', 'dan', 'rakyat', 'yang', 'cinta', 'jujur', 'adil', 'diseluruh', 'nusantara', 'muak', 'juga', 'lihat', 'mukanya'] | [copot, berapakali, kesalahan, jahat, kerja, muak, lihat, mukanya, cinta, jujur, adil, nusantara, muak, lihat, mukanya] |

**Stemming**

Stemming is a process that is useful for extracting the base words of a text by removing its affixes, whether they are prefixes, suffixes, or infixes, to reduce the variety in the representation of a word [12]. An example of stemming results can be seen in Table 7 below:

Table 7. An example of the result of applying stemming

| Before | After |
|---|---|
| Sok kerja paling berat, taunya kerja paling santai 🤣🤣🤣🤣🤣 | sok kerja berat tau kerja santai |

**Data visualization**

Data visualization is the process of displaying data after preprocessing, data visualization can help to understand the data and gain insight from the data [13], such as commonly occurring negative and positive sentiment words in the form of a word cloud, and frequently occurring negative and positive sentiment words in the form of a graph. In Figure 2, negative sentiments frequently appearing in the dataset are depicted, with the condition that larger words indicate more frequent occurrences in the dataset. In Figure 3, positive sentiments that frequently appear in the dataset are shown under the same condition as Figure 2. Then Figure 4 displays the top 10 most frequently occurring negative sentiment words, with the highest-ranked word being "salah" appearing over 700 times, and the lowest-ranked word "bangsat" appearing less than 100 times. Figure 5 illustrates the top 10 most frequently occurring positive sentiment words, with the word "keras" appearing over 250 times and the lowest-ranked word "cepat" appearing less than 100 times.


Figure 2. Negative sentiment words in the form of a word cloud

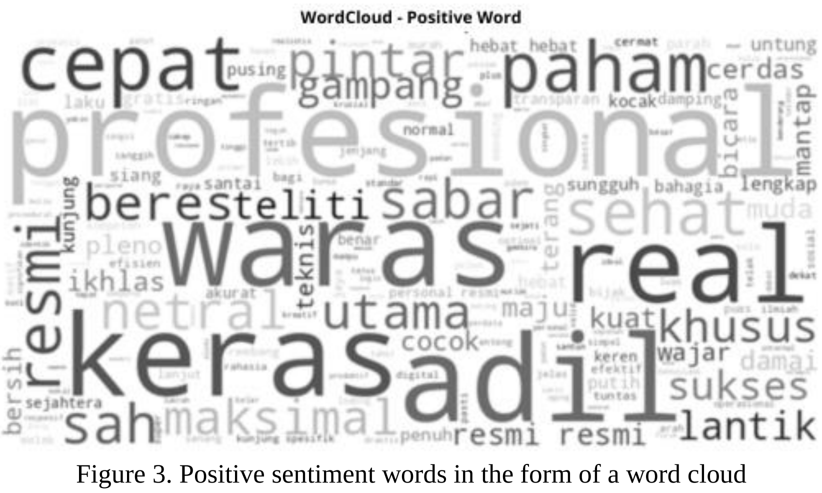



Figure 3. Positive sentiment words in the form of a word cloud

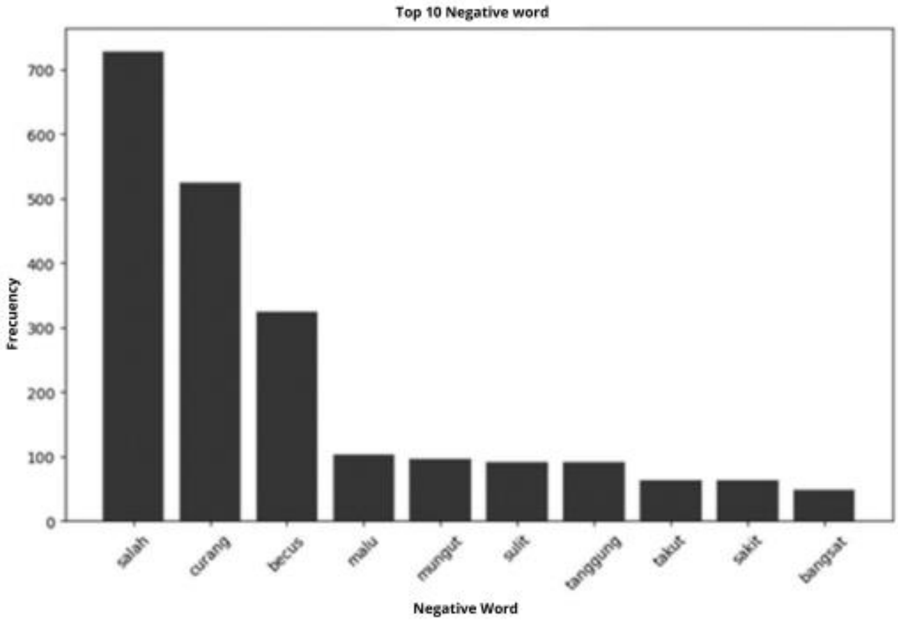



Figure 4. Top 10 most frequently occurring negative sentiments

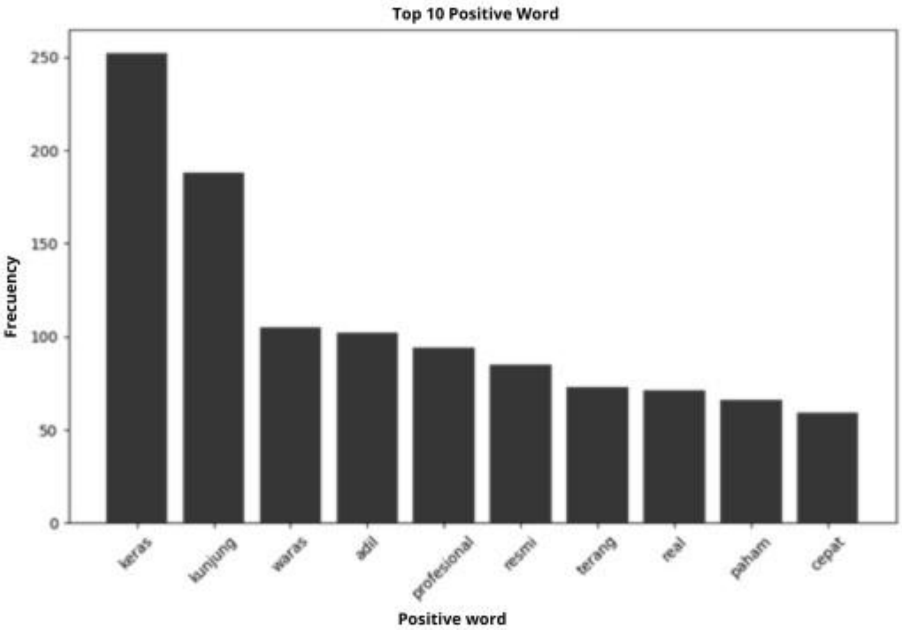



Figure 5. Top 10 most frequently occurring positive sentiments

**Labeling data**

Labeling data is assigning labels to each data row by dividing the data into three sentiment classes: positive, negative, and neutral [14].

This process utilizes an Indonesian lexicon tool by adding positive and negative dictionaries to assign a score to each word [15]. The scores are then summed up for each tweet row, with the condition that the score is > 0 for positive sentiment. For negative sentiment, the score is < 0; for neutral sentiment, the score is = 0. Labeling the data yielded 3382 neutral sentiment data, 1368 negative sentiment data, and 1032 positive sentiment data, totaling 5782 labeled data. The labeling results in the form of a pie chart can be seen in Figure 4 below.

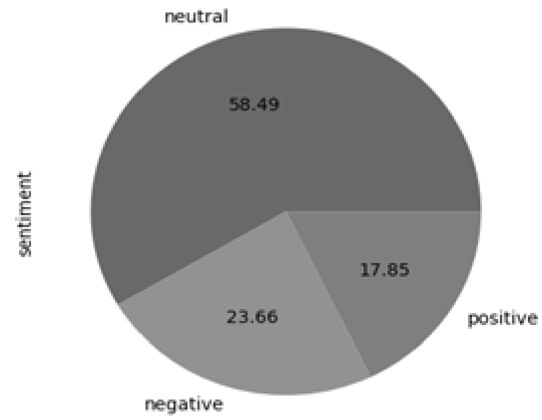



Figure 6. Display of Pie Chart for labeled data

Figure 6 shows the results of labeling the entire dataset after preprocessing, indicating that neutral sentiment accounts for 58.49%, negative sentiment for 23.66%, and positive sentiment for 17.85%.

**Balancing data**
Balancing data is a technique used when the available data for each output class is imbalanced. In this study, with 3382 data for neutral sentiment, 1368 data for negative sentiment, and 1032 data for positive sentiment, it indicates that the generated data is imbalanced.

In this research, the oversampling method was employed to balance the data because this method can preserve information from the minority class and enhance the model's ability to predict that class without sacrificing the majority class [16]. This method works by increasing the number of samples in the minority class (in this case, negative and positive sentiment) to align with the majority class (neutral sentiment). By duplicating existing samples or synthetically generating new ones, the dataset achieves a more balanced distribution. Consequently, the model can learn effectively from all classes without leaning towards the majority class. This approach improves the model's ability to capture subtleties from the minority classes, thus enhancing overall classification performance without bias towards the majority class [17].

**Data testing**
Data testing is used to measure how well the classifier correctly classifies [18]. In this stage, data testing is conducted using the base CNN model and the modified CNN model. This stage is carried out to evaluate the performance of the trained models on the dataset available.

**Modeling**
The first step in data testing is modeling. Modeling is the process of creating, training, and evaluating a CNN model. In this study, the base CNN model is used and then modified to have several core layers forming the model for optimizing the predictive performance and accuracy on the validation set for the CNN models [19]. The architecture of the CNN model formed before and after modification can be seen in Figure 8 using the summary () function.

Table 8. Model CNN layer before and after modification

| CNN Model Layers Before Modification | | | CNN Model Layers After Modification | | |
|---|---|---|---|---|---|
| *Layer (type)* | *Output Shape* | *Param#* | *Layer (type)* | *Output Shape* | *Param#* |
| *Embedding (Embedding)* | *(None, 43, 32)* | *3200* | *Embedding_2 (Embedding)* | *(None, 100, 100)* | *864900* |
| *conv1d_3 (Conv1D)* | *(None, 39, 64)* | *10304* | *conv1d_2 (Conv1D)* | *(None, 96, 128)* | *64128* |
| *max_pooling1d_1 (MaxPooling1D)* | *(None, 19, 64)* | *0* | *global_max_pooling1d_1 (GlobalMaxPooling1D)* | *(None, 19, 64)* | *0* |
| *flatten_1 (Flatten)* | *(None, 1216)* | *0* | *dense_4 (Dense)* | *(None, 64)* | *8526* |
| *dense_8 (Dense)* | *(None, 64)* | *77888* | *dropout_1 (Dropout)* | *(None, 64)* | *0* |
| *dense_9 (Dense)* | *(None, 3)* | *195* | *batch_normalization (Batch Normalization)* | *(None, 64)* | *256* |

| | | | |
|---|---|---|---|
| *Total params: 120387* | *dense_5 (Dense)* | *(None, 32)* | *2080* |
| *Trainable params: 120387* | *dropout_2 (Dropout)* | *(None, 32)* | *0* |
| *Non trainable params: 0* | *dense_6 (Dense)* | *(None, 16)* | *528* |
| | *Dense_7 (Dense)* | *(None, 3)* | *51* |
| | *Total params: 940199* | | |
| | *Trainable params: 940071* | | |
| | *Non trainable params: 128* | | |
| | *Shaded lines indicates additional layer* | | |

**Layer**

Layer in Convolutional Neural Network (CNN) are processing units that receive input from the previous layer. Generally, CNN consists of 3 layers: the input layer, hidden layer, and output layer [20]. In Figure 7, there are four layers depicted: the input layer, convolutional layer, pooling layer, and fully connected layer. The image explains the sequence of layers used in this study, with the following explanations:

- Input Layer: The input layer is the first layer in the neural network that receives raw data as input. In this study, the input layer receives text data representations to be processed by the network to learn relevant features that can be understood by the network [21].
- Convolutional Layer: This layer functions to extract important features from the input using convolution operations. This layer helps the network understand patterns and features in the input data, such as edges, textures, or more complex patterns [22].
- Pooling Layer: This layer is used to reduce the spatial dimensions of the feature representations generated by the convolutional layers. This helps reduce overfitting and excessive computations by taking the average or maximum values from specific areas in the feature representations [23].
- Fully Connected Layer: This layer is responsible for connecting every neuron in the previous layer to every neuron in the next layer. This layer is typically located at the end of the network and is used to generate classification outputs based on the features learned by the previous convolutional layers [24].

The general function of these four layers is as follows: the input layer receives raw data, the convolutional layer is used to extract important features, the pooling layer reduces the dimensionality of feature representations, and the fully connected layer generates classification outputs based on the learned features [25].

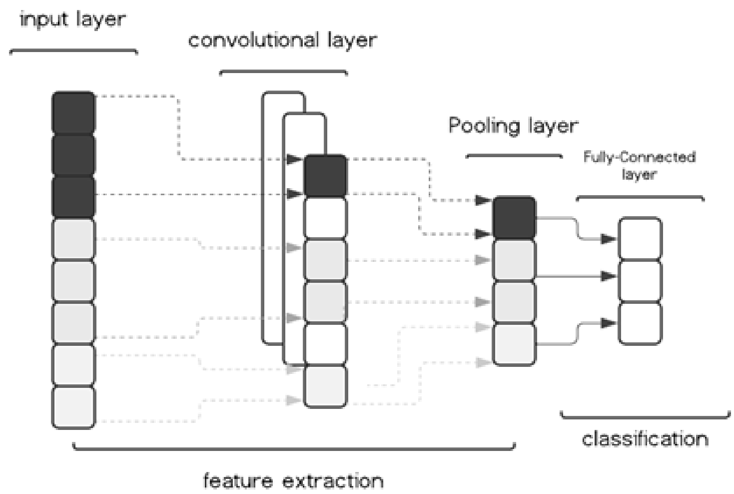



Figure 7. Architecture layer CNN

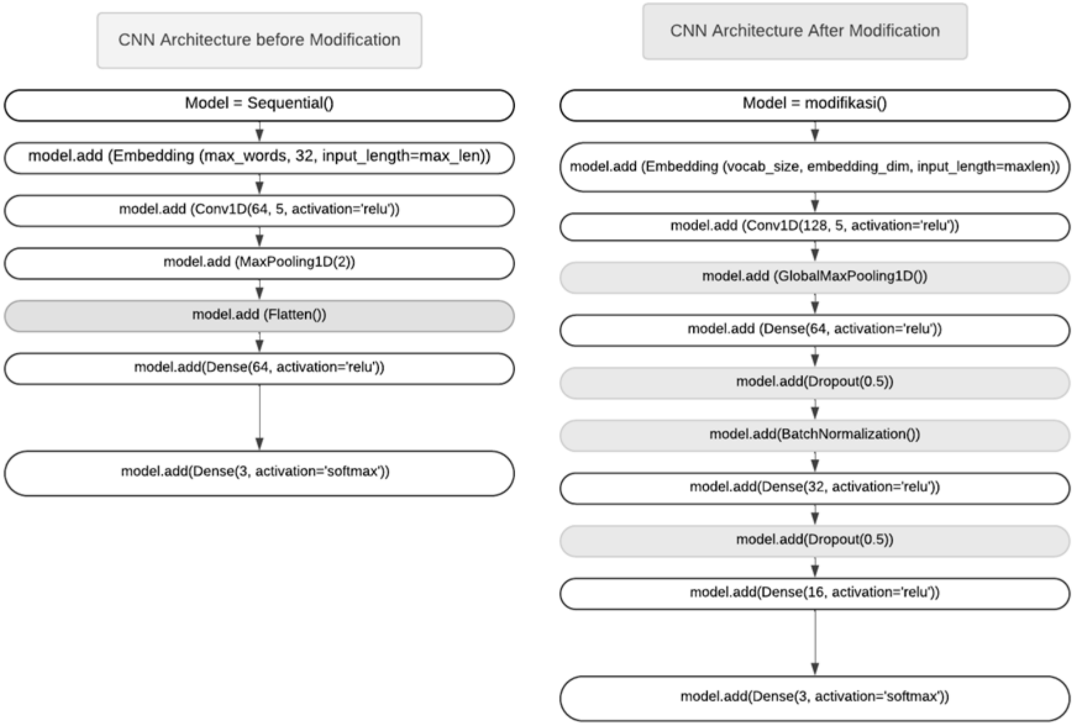



Figure 8. Architecture CNN before and after Modification

From the above figure 8, there is an explanation regarding the layers in the CNN model used in this study:

**a) Embedding layer**

The first layer is the Embedding layer. This layer converts text into vector representations that can be stored and utilized to recognize and extract important text-related patterns related to positive, negative, or neutral sentiments [26].

**b) Conv1D layer**

Conv1D is a one-dimensional layer that processes text and extracts important features from the sequence of words in the text [27].

**c) GlobalMaxPooling1D layer**

GlobalMaxPooling1D is a pooling method that collects the maximum value from each block of its input and provides translation invariance to the network. This is achieved by extracting the maximum value in the map [19].

**d) Flatten layer**

Flatten converts feature maps from the pooling layer into a vector form for input for the fully connected layer with a single hidden layer [28].

**e) Dense layer**

Dense is a fully connected layer process aimed at transforming the data dimensions to be classified linearly or in parallel. This process uses relu activation to enhance non-linear properties without affecting the receptive field of the Convolutional Layer [29].

**f) Dropout layer**

Dropout is a regularization technique in CNN that helps prevent overfitting in the model. This technique works by randomly deactivating some units in the hidden layer during the training process, thus enabling it to generalize the data being trained more effectively [30].

**g) Batch Normalization layer**

Batch Normalization is a technique that aims to improve stability and accelerate training in CNN by normalizing the input of each layer into a more stable range [31].

### Evaluation metrics

The parameters used in this study include the confusion matrix, which contains accuracy, precision, recall, and f-measure.

**a) Accuracy**

Accuracy is a metric that measures the number of correct predictions that match the original labels out of the total predictions made by the model [32]. Accuracy is calculated as:

$$Accuracy = \frac{True\ Positive + True\ Negative + True\ Neutral}{True\ Positive + False\ Positive + True\ Negative + False\ Negative + True\ Neutral + False\ Neutral} \quad (1)$$

**b) Precision**

Precision is the ratio of the number of sentiments predicted correctly to the total number of sentiments classified [33]. Precision is calculated as:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

**c) Recall**

Recall is the ratio of the number of true positive sentiments classified correctly to the total number of positive samples [34]. Recall is calculated as:

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (3)$$

**d) F-measure**

F-measure is an evaluation metric calculating the harmonic mean of recall and precision [34]. The calculation of the F-measure is:

$$F - measure = \frac{2\ x\ Precision\ x\ Recall}{Precision + Recall} \quad (4)$$

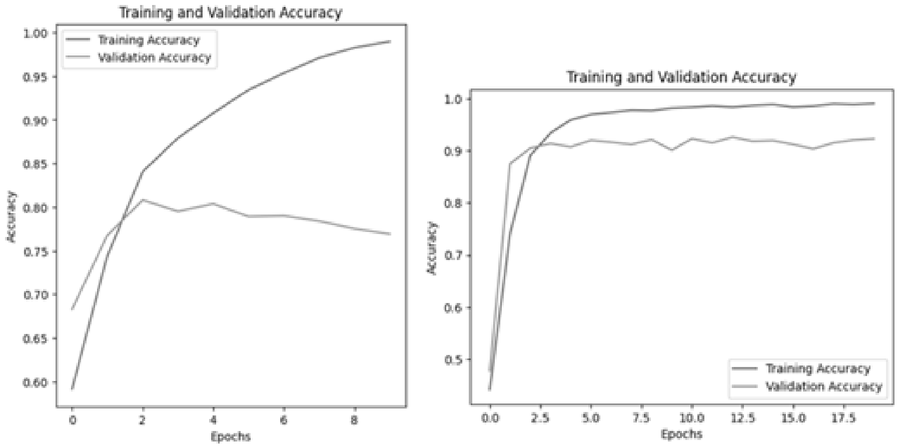
## RESULTS AND DISCUSSIONS

In this stage, the results of testing the CNN model before and after modification are discussed to determine the difference in accuracy in sentiment analysis regarding the performance of the Indonesian Electoral Commission in the 2024 general elections. The testing process of the CNN model begins with training and testing the data using the models, as seen in table 8.

Based on the reference, a comparison was made between the testing results of the CNN model before modification and the CNN model after modification. The CNN model before modification utilized embedding, convolution, pooling, flattening, and several dense layers, then compiled using the 'Adam' optimizer, 'sparse_categorical_crossentropy' loss function, and 'accuracy' metric for testing.

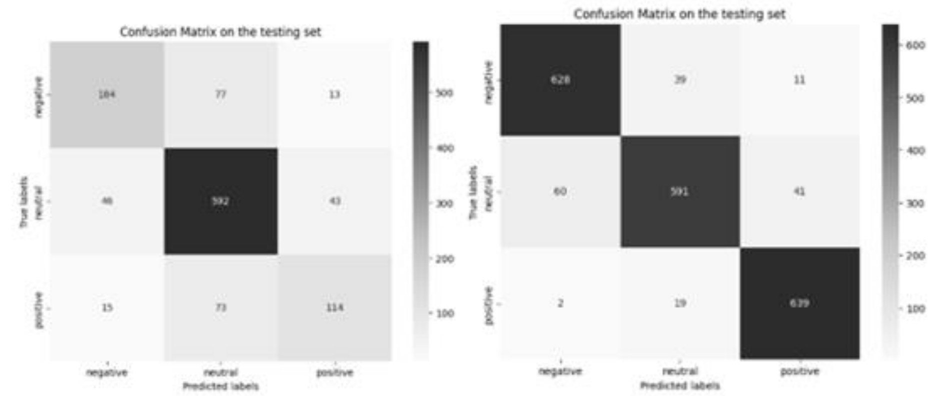
After modification, the CNN model started with data balancing using oversampling. Then, the flattened layer was removed to reduce the number of parameters to be learned in the model. Then, larger filters were added to the convolution layers, and the pooling layer was changed to GlobalMaxPooling1d to reduce data dimensions. A batch Normalization layer was added, along with a dropout layer, to reduce overfitting. The output layer with the softmax activation function was also used to generate class predictions for three sentiment categories [35].

The main difference between the testing model before and after modification is using a normalization layer. With the batch normalization layer, the model can learn more stable representations of the input data, resulting in improved performance and convergence speed during training. Moreover, the additional dropout layer enables the model to handle better overfitting [35]. The testing results of both models are compared as follows:




**Before Modification** **After Modification**

Figure 9. Training and validation results of the CNN model before and after modification




**Confusion Matrix Before Modification** **Confusion Matrix After Modification**

Figure 10. Confusion matrix results of CNN model before and after modification

Table 9. Result of testing CNN model before and after modification

| Results of Testing CNN Model Before Modification | Precision | Recall | F1-score | Support |
|---|---|---|---|---|
| Negative | 0.75 | 0.67 | 0.71 | 274 |
| Neutral | 0.80 | 0.87 | 0.83 | 681 |
| Positive | 0.67 | 0.56 | 0.61 | 202 |
| Accuracy | | | 0.77 | 1157 |
| **Results of Testing CNN Model After Modification** | **Precision** | **Recall** | **F1-score** | **Support** |
| Negative | 0.94 | 0.90 | 0.92 | 678 |
| Neutral | 0.88 | 0.91 | 0.90 | 692 |
| Positive | 0.97 | 0.97 | 0.96 | 660 |
| Accuracy | | | 0.93 | 2030 |

In Figure 9, the curve results show the difference in validation results of the CNN model before and after modification. In the model before modification, the generated curve experienced overfitting. Then, after modification, showed an increase in accuracy from epoch 1. This was due to good weight initialization and sufficient data quality, enabling the model to start learning well at the beginning of the training process and continue to show good and stable performance in subsequent epochs, whereas before modification, the model only achieved an accuracy of 77%. This improvement occurred due to the addition of layers in the

CNN, especially in the use of dropout and batch normalization, which helped stabilize the model during testing and minimize overfitting.

Then based on the Figure 10, the confusion matrix shows the CNN model before and after modification. In the confusion matrix before modification, the model estimated 184 sentiments as negative, 592 sentiments as neutral, and 114 sentiments as positive. Meanwhile, after modification, the model estimated 628 sentiments as negative, 591 sentiments as neutral, and 639 sentiments as positive.

Based on the test results in Table 9, accuracy was obtained for negative, positive, and neutral sentiments on the original and modified CNN models. The original CNN model achieved an accuracy of 77%, which increased to 93% after modification, with precision matrix details for negative sentiment at 75%, increasing to 94% after modification. Additionally, precision increased from 80% to 88% for neutral sentiment, and from 67% to 97% for positive sentiment. This indicates an increase in the number of sentiments predicted correctly compared to the total number of classified sentiments.

Regarding recall, the original model achieved 67% for negative sentiment, which increased to 90% after modification. For neutral sentiment, recall increased from 87% to 91%, and for positive sentiment, recall increased from 56% to 97%. This shows the comparison between the number of positive sentiments correctly classified as positive against the total number of positive samples increased.

Furthermore, the F1-score results were obtained, with the original model achieving 71% for negative sentiment, which increased to 92% after modification. For neutral sentiment, the F1-score increased from 83% to 90%, and for positive sentiment, it increased from 61% to 96%. This indicates an increase in the average obtained from recall and precision.

However, there are limitations or challenges in developing this model, namely the difficulty of continuously increasing the model's accuracy to reach a higher level, considering that the modified CNN model has already achieved a high accuracy rate of 93%. Therefore, further research and exploration are needed to address these challenges and are expected to provide significant contributions and important insights into public sentiment analysis, as well as additional information about modifications to CNN for sentiment analysis.

## CONCLUSION

This study aims to analyze public sentiment towards the performance of the Indonesian Electoral Commission and evaluate the effectiveness of a modified CNN model in sentiment analysis. The study focuses on sentiment analysis related to the KPU performance in the 2024 general elections using the CNN method, with a dataset comprising 5782 rows of data. The study compares the test results of the model between the base CNN model and the modified CNN model.

After the comparison, it is concluded that the modified CNN model achieved a higher accuracy result of 93%, while before modification, the model only achieved an accuracy of 77%. This improvement occurred because additional layers were added to the CNN, particularly in the use of dropout and batch normalization, which helped stabilize the model during testing and minimize overfitting.

These conclusions indicate that modifications to the CNN architecture have a positive impact on improving the model's performance in sentiment analysis of the KPU performance in the 2024 elections and are expected to assist in evaluating the KPU performance in conducting elections.

Thus, it can be concluded that the modified CNN model successfully achieved a high accuracy rate of 93%, and efficiently extracted important information while generating rich and abstract feature representations. The main result of this study demonstrates the effectiveness of CNN modifications, namely, a 16% increase in accuracy compared to the pre-modified CNN model. This is expected to provide significant contributions and important insights in public sentiment analysis, as well as additional information regarding CNN modifications for sentiment analysis.

## REFERENCES


[1] Fuad Amirullah, Syariful Alam, and M.Imam Sulistyo S, "Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini Twitter Menggunakan Naïve Bayes," *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 2, no. 3, pp. 69–76, Aug. 2023, doi: 10.55123/storage.v2i3.2293.

[2] S. N. Listyarini and D. A. Anggoro, "Analisis Sentimen Pilkada di Tengah Pandemi Covid-19 Menggunakan Convolution Neural Network (CNN)," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 1, no. 7, pp. 261–268, Jul. 2021, doi: 10.52436/1.jpti.60.

[3] F. M. Shiri, T. Perumal, N. Mustapha, and R. Mohamed, "A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU," *ArXiv*, 2023, doi: 10.48550/arXiv.2305.17473.

[4] P. Bhuvaneshwari, A. N. Rao, Y. H. Robinson, and M. N. Thippeswamy, "Sentiment analysis for user reviews using Bi-LSTM self-attention based CNN model," *Multimed Tools Appl*, vol. 81, no. 9, pp. 12405–12419, Apr. 2022, doi: 10.1007/s11042-022-12410-4.

[5] E. Irawati Setiawan *et al.*, "Analisis Pendapat Masyarakat terhadap Berita Kesehatan Indonesia menggunakan Pemodelan Kalimat berbasis LSTM," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 9, no. 1, pp. 8–17, Feb. 2020, doi: 10.22146/jnteti.v9i1.115.

[6] M. Li and Y. Shi, "Sentiment analysis and prediction model based on Chinese government affairs microblogs," *Heliyon*, vol. 9, no. 8, p. e19091, Aug. 2023, doi: 10.1016/j.heliyon.2023.e19091.

[7] T. Ahammad, "Identifying hidden patterns of fake COVID-19 news: An in-depth sentiment analysis and topic modeling approach," *Natural Language Processing Journal*, vol. 6, p. 100053, Mar. 2024, doi: 10.1016/j.nlp.2024.100053.

[8] P. L. Parameswari and Prihandoko, "PENGGUNAAN CONVOLUTIONAL NEURAL NETWORK UNTUK ANALISIS SENTIMEN OPINI LINGKUNGAN HIDUP KOTA DEPOK DI TWITTER," *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 27, no. 1, pp. 29–42, Apr. 2022, doi: 10.35760/tr.2022.v27i1.4671.

[9] K. Fithriasari, S. Z. Jannah, and Z. Reyhana, "Deep Learning for Social Media Sentiment Analysis," *MATEMATIKA*, pp. 99–111, Aug. 2020, doi: 10.11113/matematika.v36.n2.1226.

[10] YosefOwenM, "Implementation-Of-Indobert-On-UserReviewsOfThePlnMobile-Application-BasedOnIndonesianLanguageLexicon."

[11] "Preprocessing Text Data for Machine Learning."

[12] O. Olabanjo *et al.*, "From Twitter to Aso-Rock: A sentiment analysis framework for understanding Nigeria 2023 presidential election," *Heliyon*, vol. 9, no. 5, p. e16085, May 2023, doi: 10.1016/j.heliyon.2023.e16085.

[13] D. Destiawan, E. Eko Zulkaryanto, M. F. Hafidz, Y. Hisbullah, and M. S. Hasibuan, "Data Visualization of Research Publications of IIB Darmajaya Lecturers," *Proceeding of 7th ICITB*, pp. 51–55, 2021.

[14] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," *SISTEMASI*, vol. 10, no. 1, p. 222, Jan. 2021, doi: 10.32520/stmsi.v10i1.1179.

[15] O. Alsemaree, A. S. Alam, S. S. Gill, and S. Uhlig, "An analysis of customer perception using lexicon-based sentiment analysis of Arabic Texts framework," *Heliyon*, vol. 10, no. 11, p. e30320, Jun. 2024, doi: 10.1016/j.heliyon.2024.e30320.

[16] J. Prasetya, "Penerapan Klasifikasi Naive Bayes dengan Algoritma Random Oversampling dan Random Undersampling pada Data Tidak Seimbang Cervical Cancer Risk Factors," *Leibniz: Jurnal Matematika*, vol. 2, no. 2, pp. 11–22, Jul. 2022, doi: 10.59632/leibniz.v2i2.173.

[17] A. Q. Md, S. Kulkarni, C. J. Joshua, T. Vaichole, S. Mohan, and C. Iwendi, "Enhanced Preprocessing Approach Using Ensemble Machine Learning Algorithms for Detecting Liver Disease," *Biomedicines*, vol. 11, no. 2, p. 581, Feb. 2023, doi: 10.3390/biomedicines11020581.

[18] Baiq Nurul Azmi, Arief Hermawan, and Donny Avianto, “Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma Decision Tree untuk Klasifikasi Penderita Penyakit Liver,” *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 4, no. 4, pp. 281–290, Feb. 2023, doi: 10.35746/jtim.v4i4.298.

[19] M. M. A. Monshi, J. Poon, V. Chung, and F. M. Monshi, “CovidXrayNet: Optimizing data augmentation and CNN hyperparameters for improved COVID-19 detection from CXR,” *Comput Biol Med*, vol. 133, p. 104375, Jun. 2021, doi: 10.1016/j.compbiomed.2021.104375.

[20] M. Urgo, W. Terkaj, and G. Simonetti, “Monitoring manufacturing systems using AI: A method based on a digital factory twin to train CNNs on synthetic data,” *CIRP J Manuf Sci Technol*, vol. 50, pp. 249–268, Jun. 2024, doi: 10.1016/j.cirpj.2024.03.005.

[21] S. F. Handono, F. T. Anggraeny, and B. Rahmat, “IMPLEMENTASI CONVOLUTIONAL NEURAL NETWORK (CNN) UNTUK DETEKSI RETINOPATI DIABETIK,” 2020.

[22] A. S. R. Riyadi, I. . P. Wardhani, D. S. Widayati, and K. Kunci, “Classification of dog and cat images using Convolutional Neural Network (CNN),” *Universitas Gunadarma Jl. Margonda Raya*, vol. 5, no. 1, 2021.

[23] Y. Bili, “Design of a Pineapple Ripeness Detection Tool Using a Microcontroller Using the Convolutional Neural Network (CNN) Method.”

[24] O. V. Putra, A. Musthafa, M. Nur, and M. Rido, “Classification of Calligraphy Writing Types Using Convolutional Neural Network Method (CNN),” *Procedia of Engineering and Life Science*, vol. 2, Nov. 2021, doi: 10.21070/pels.v2i0.1136.

[25] I. Wulandari, H. Yasin, and T. Widiharih, “KLASIFIKASI CITRA DIGITAL BUMBU DAN REMPAH DENGAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN)”, [Online]. Available: https://ejournal3.undip.ac.id/index.php/gaussian/

[26] S. Imron, E. I. Setiawan, and J. Santoso, “Deteksi Aspek Review E-Commerce Menggunakan IndoBERT Embedding dan CNN,” *Journal of Intelligent System and Computation*, vol. 5, no. 1, pp. 10–16, Apr. 2023, doi: 10.52985/insyst.v5i1.267.

[27] M. Data, “Perbedaan Antara Conv1D, Conv2d, dan Conv3d Pada CNN.”

[28] R. Fajri and L. Atika, “Implementation of Machine Learning Using CNN for Template Image Classification,” in *Bina Darma Conference on Computer Science*,

[29] A. Zancanaro, G. Cisotto, J. R. Paulo, G. Pires, and U. J. Nunes, “CNN-based Approaches For Cross-Subject Classification in Motor Imagery: From the State-of-The-Art to DynamicNet,” in *2021 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, IEEE, Oct. 2021, pp. 1–7. doi: 10.1109/CIBCB49929.2021.9562821.

[30] A. Firmansyah, A. F. Itsnan, A. Apip, R. T. Mulliya, and P. Rosyani, “Student Attendance System Using Face Recognition with CNN Algorithm.”

[31] S. Lange, K. Helfrich, and Q. Ye, “Batch Normalization Preconditioning for Neural Network Training.”

[32] F. A. Irawan and D. A. Rochmah, “Application of the CNN Algorithm to Find Out Public Sentiment Regarding the Covid-19 Vaccine Policy,” *JURNAL INFORMATIKA*, vol. 9, no. 2, 2022.

[33] R. Yacouby and D. Axman, “Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models,” in *Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2020, pp. 79–91. doi: 10.18653/v1/2020.eval4nlp-1.9.

[34] Y. Yuliska, D. H. Qudsi, J. H. Lubis, K. U. Syaliman, and N. F. Najwa, “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, p. 1067, Oct. 2021, doi: 10.25126/jtiik.2021854842.

[35] C. Garbin, X. Zhu, and O. Marques, “Dropout vs. batch normalization: an empirical study of their impact to deep learning,” *Multimed Tools Appl*, vol. 79, no. 19–20, pp. 12777–12815, May 2020, doi: 10.1007/s11042-019-08453-9.