

# Classification of Public Sentiment on Fuel Price Increases Using CNN

Anak Agung Istri Arinta Maharani <sup>1)\*</sup>, Sri Suryani Prasetyowati <sup>2)</sup>, Yuliant Sibaroni <sup>3</sup>

<sup>1,2,3)</sup> Fakultas Informatika, Universitas Telkom, Bandung, Indonesia

<sup>1)</sup>[arintamaharani@student.telkomuniversity.ac.id](mailto:arintamaharani@student.telkomuniversity.ac.id), <sup>2)</sup>[srisuryani@telkomuniversity.ac.id](mailto:srisuryani@telkomuniversity.ac.id),

<sup>3)</sup>[yuliant@telkomuniversity.ac.id](mailto:yuliant@telkomuniversity.ac.id)

**Submitted** : Jun 27, 2023 | **Accepted** : Jun 30, 2023 | **Published** : July 1, 2023

**Abstract:** The government's policy of changing fuel prices is carried out every year. The public gave responses to this policy categorized as positive, negative, or neutral sentiments. The community's response was conveyed through tweets on the Twitter application. Based on the public's response to the policy, sentiment classification can be done using data mining classification techniques. Even though there has been a systematic approach, challenges still exist in classifying sentiments with a high degree of accuracy on the public's response to the fuel price change policy. Tweets on Twitter as a source of sentiment data require complex processing and a proper understanding of context. Therefore, it is necessary to develop more sophisticated and accurate classification methods to understand better people's responses in the context of fuel price policies. Some research has been carried out on classification techniques using deep learning and machine learning methods. In general, deep learning methods get better results, and this research will be approached using the CNN method. The system stages start from crawling data, labeling, and preprocessing, which consists of cleaning, case folding, tokenization, normalization, removing stopwords and stemming, classification using CNN, and evaluation using 10-Cross Validation. The dataset used is 17,270. The results show that the developed classification system is relatively high, with the highest accuracy of 87%, 93% recall, 93% precision, and 90% F1 score. An in-depth analysis of the classification results and an understanding of sentiment toward rising fuel prices can also provide valuable insights.

**Keywords:** Sentiment Classification, Fuel up, CNN, SMOTE, K-Fold Cross Validation

## INTRODUCTION

In 2022, the government will again set a policy of increasing fuel prices as in previous years. This decision was prompted by the ongoing war in Europe between Russia and Ukraine, which causes an increase in oil prices (Wardani et al., 2022). Peralite's price was initially IDR 7,650 per liter, increasing to IDR 10,000 per liter. Meanwhile, the price of diesel, which was originally IDR 5,150 per liter, increased to IDR 6,800 per liter, and the price of Pertamina, which was originally IDR 12,500 per liter, became IDR 14,500 per liter (Sihombing, 2022). The government's move to increase fuel prices was based on the amount of subsidies which had reached Rp 502 trillion and were deemed not on target. Government subsidies should be used to help the underprivileged, but most of the benefits are enjoyed by those who can afford it, such as those who have private cars (Rodani, 2022).

With regard to these policies, the public submitted responses and categorized them as positive, negative, and opinions neutral. Some people do not agree with this policy because it is considered burdensome for those in the lower classes. Even though the government provides Direct Cash Assistance (BLT), this assistance is only valid for the short term or temporarily (dpr.go.id, 2022) However, some people agree with this policy, because it can increase state revenues.

Many public responses regarding the increase in fuel prices were conveyed via social media Twitter. Twitter is a medium for users to write reviews, including related reviews of rise in fuel prices. With these various responses, the classification of public sentiment regarding the increase in fuel prices through Twitter can help the government make policies according to community needs. This sentiment classification aims to provide objective insights into the public comments regarding the increase in fuel prices.

\*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Research related to sentiment analysis in classification was carried out by (Kumar & Garg, 2020 regarding aspect-based sentiment analysis, using various models such as CNN, Naïve Bayes, SVM, and ANN. The results obtained from using the CNN model are better than other methods with an accuracy of 72.17%. Whereas in the study (Listyarini & Anggoro, 2021), the accuracy achieved was higher, amounting to 90% with an epoch of 100. Meanwhile in the study (Yuliska et al., 2021), the accuracy obtained was relatively high at 98% for the two sentiment classes and 91% for the third sentiment class.

Using K-Fold Cross Validation and CNN in research (Irawan & Rochmah, 2022) obtained 99.61% and 96.53% accuracy when using Naïve Bayes. Whereas without using K-Fold, the accuracy obtained was 98.66% with CNN and 94.66% with Naïve Bayes. In research (Alhakiem & Setiawan, 2022), Hanif and Erwin used SMOTE to get an F1-Score of 95.89% for the signal aspect and 93.02% for the service aspect. These results are better than those without SMOTE, producing an F1-Score of 93.37% for the signal aspect and 84.64% for the service aspect. Based on the studies above, CNN is a suitable method for sentiment classification. However, in this study, there is no classification of sentiment regarding the increase in fuel prices using CNN by adding validation in the form of K-Fold Cross Validation and SMOTE for unbalanced data sets. This research will provide valuable insights to the government and society in dealing with the issue of rising fuel prices and facilitate better decision-making and a deeper understanding of the impact of the policy.

## LITERATURE REVIEW

### Related Work

The utilization of Convolutional Neural Network (CNN) has been explored in several prior studies (Kumar & Garg, 2020), (Listyarini & Anggoro, 2021), (Yuliska et al., 2021). Specifically, research (Kumar & Garg, 2020) focuses on aspect-based sentiment analysis and investigates the effectiveness of CNN, Naïve Bayes, Support Vector Machines (SVM), and Artificial Neural Network (ANN) methods in order to identify the optimal model. The results of this study show an accuracy value of 72% using CNN. This accuracy value is more significant than using the Naïve Bayes, SVM, and ANN models, each of which obtains an accuracy of 67%, 61%, and 66%. Research (Listyarini & Anggoro, 2021) uses a dataset from tweets on Twitter regarding public concerns about the implementation of regional elections in 2020 due to the Covid-19 pandemic, obtaining accurate results using CNN, namely 90% with epoch 100. Epoch 100 has higher accuracy than epoch 50, with results accuracy of 82% and 86% at epoch 75. The research was also conducted (Yuliska et al., 2021) using a dataset in the form of student suggestions on the performance of departments in tertiary institutions. This study uses two sentiment classes and three sentiment classes. The results of this study indicate that using two classes obtains higher accuracy results, namely 97%, compared to 3 sentiment classes, namely 91%.

The research was also carried out (Irawan & Rochmah, 2022) using K-Fold Cross Validation with 99.61% accuracy when using CNN and 96.53% when using Naïve Bayes; this shows higher numbers than before using K-Fold Cross Validation which only got 98.66% accuracy using CNN and 94.66% use naïve Bayes. In addition to using K-Fold Cross Validation, using SMOTE can also increase accuracy by balancing unbalanced datasets as was done in research (Alhakiem & Setiawan, 2022) using SMOTE to get F1-Score results of 95.89% for signal aspects and 93.02% for service aspects. This is better than before using SMOTE, which brought an F1-Score of 93.37% for the signal aspect and 84.64% for the service aspect.

### CNN

The Convolution Neural Network (CNN) model is used in this system. The model will undergo a training and testing process for as many k iterations because the K-Fold Cross Validation method is applied to get a comprehensive model performance for the entire dataset. K-Fold Cross Validation is used as a validation method to test and evaluate the performance of models that have been built (Lee & Sibaroni, 2023). Figure 1 shows an illustration of the K-Fold Cross Validation (Lee & Sibaroni, 2023).

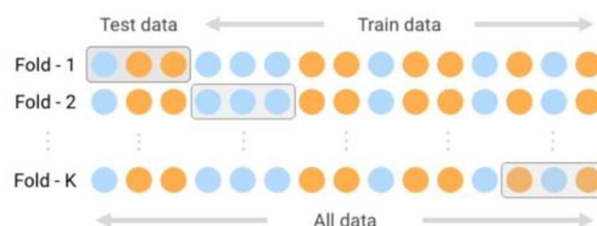


Fig. 1 Illustration of K-Fold Cross Validation

Convolution Neural Network (CNN) results from artificial neural networks and deep learning methods. The CNN structure consists of one or more convolutional layers, often followed by a subsampling layer, and then

\*name of corresponding author



connected by one or more fully connected layers as in ordinary neural networks (Fonda et al., 2020). Although CNN is generally used in computer vision for image classification, it also performs well in text classification (Jacovi et al., 2018). Figure 2 shows an illustration of a CNN consisting of an embedding layer, convolutional layer, pooling layer, and fully connected layer (Ramadhan & Setiawan, 2023).

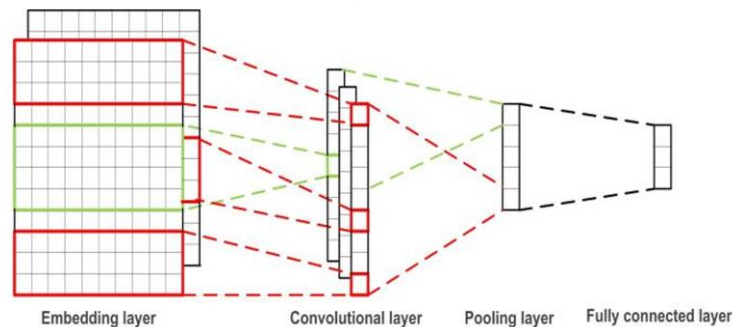


Fig. 2 Illustration of CNN

The following is the definition of the CNN section above:

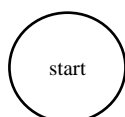
- **Embedding layer:** the embedding layer is a matrix consisting of word vectors arranged sequentially based on the words in the sentence. If the sentence includes  $m$  words and each word vector has  $n$  dimensions, then the embedded layer will be a matrix with size  $m \times n$ .
- **Convolutional layer:** the embedding layer can generate multiple feature maps via convolution operations, where the size of the convolution window is  $k \times n$ . Here,  $k$  represents the number of words longitudinally in the text, while  $n$  represents the dimensions of the word vector. Using the convolution window, the embedding layer will generate a number of feature maps with one column as a result.
- **Pooling layer:** pooling layer is also known as a sub-sampling layer, which serves to reduce the size of the input data. Various pooling methods can be used in a Convolutional Neural Network (CNN), but the most common is max-pooling.
- **Fully connected layer :** Typically, the last layer in a neural network is connected to one or more fully connected layers, and the output of the fully connected layer is the network's final output.

### SMOTE

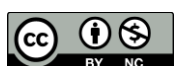
SMOTE is a technique used to balance the distribution of sample data in the minority class by selecting and synthesizing the sample data until the number is balanced with the number of samples in the majority class (Siringoringo, 2018). The application of the SMOTE method can result in overfitting. Overfitting can occur because the data in the minority class is replicated so that the same training data may appear more than once (Nikmatul Kasanah et al., 2019).

### METHOD

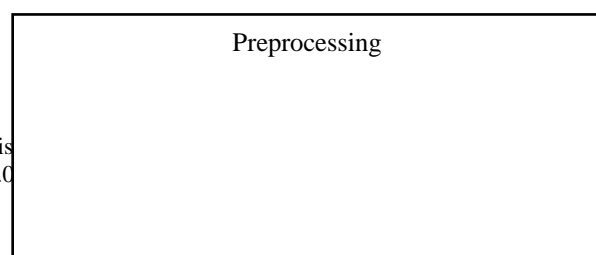
This study follows the steps shown in the flowchart illustrated in Figure 3. The flowchart describes the design of a system for conducting aspect-based sentiment analysis using the convolutional neural network (CNN) method. This study uses a case study of the increased fuel prices obtained from Twitter.



\*name of corresponding author



This is an Creative Commons License This  
Commons Attribution-NonCommercial 4.0



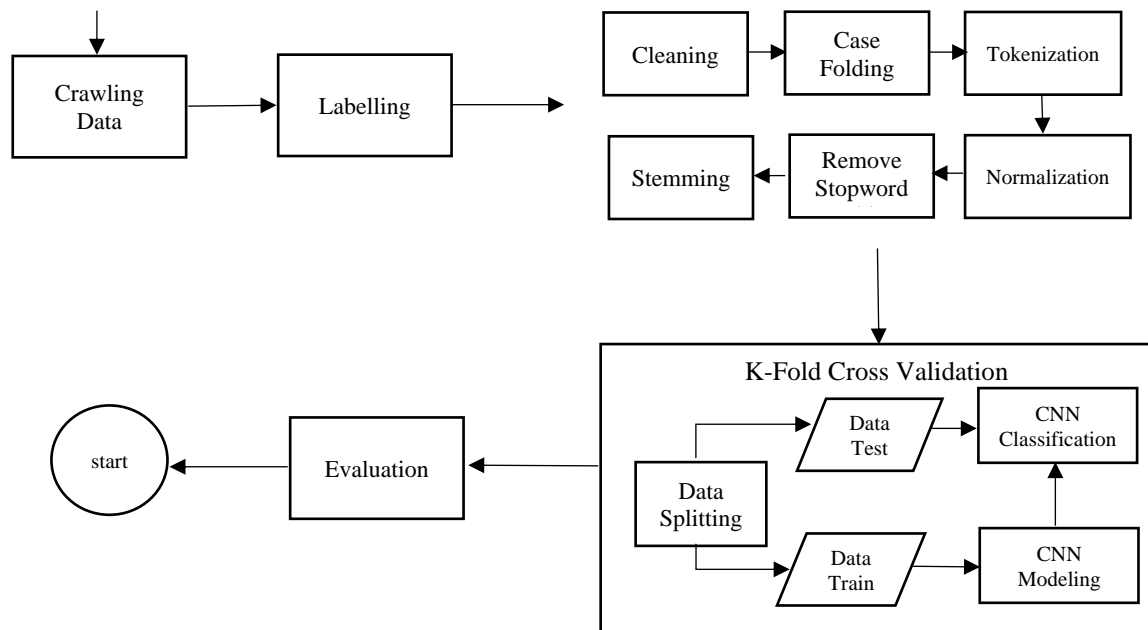


Fig. 3 System design flowchart using CNN

**Crawling Data**

The data is collected through crawling or retrieving from the Twitter social media application. In this process, the author uses the Twitter API Key, which stands for Application Programming Interface. The author retrieves tweet data by using the keyword "BBM naik." The dataset obtained is 17.270.

**Labeling**

The labeling process is done manually for each tweet contained in the dataset collected based on the crawling results. Sentiment is labeled with the number 1 if it has a positive sentiment, a label -1 if it has a negative sentiment, and a label 0 if it is considered neutral. Table 1 is an example of the results of labeling, while table 2 is the amount of data used

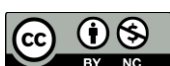
Table 1. Example of Labeling Results

Tweet	Label
Tidak ada yang berbalik dari ucapan dan keputusan Presiden @jokowi. 16 Agustus 2022 bicara subsidi energi Rp 502 T, dan tetap disediakan sd akhir tahun. Tgl 3 September menaikkan harga BBM utk mengurangi risiko subsidi bengkak sd Rp 700 T. Ini yg dialihkan, al utk bansos.	0
Lha piye toh,, harga minyak mentah bakal turun lagi, kok mas @jokowi masih ngotot naikkin harga BBM ? Pembantu2 situ itu kebanyakan ABS, doyan beri data palsu & manipulatif, masak trend ginian aja ndak bisa diperkirakan ketika naikkan harga BBM ?	-1
Logika waras saja, Kalau kenaikan Harga BBM itu membuat Rakyat Sejahtera, kenapa banyak yang demo menolaknya? Liat tuh anggota DPR yang sudah pada sejahtera, engga ada tuh yang Demo.	1

Table 2. Data Used

Label	Total Data
Negative	6662

\*name of corresponding author



Neutral	7960
Positive	2648

### Preprocessing

The data preprocessing stage involves several additional steps to change the sentences in the dataset into standard sentences. The author gets the dataset from tweets on Twitter, so there may be a lot of non-standard words. This stage aims to prevent data duplication, overcome data loss (missing values), and correct errors in the dataset (Asroni et al., 2018). At this stage, the data will undergo a process of cleaning or cleaning data so that the data can be processed and undergo a data mining process. The processes carried out at this stage are cleaning, case folding, tokenization, normalization, removing stopwords, and stemming. The following is the process carried out at the preprocessing stage.

#### A. Cleaning

This cleaning process aims to clean the data. At this stage, removing punctuation marks, deleting numbers, and changing sentences that use capital letters to lowercase letters are carried out. In addition, cleaning is also carried out if there are excessive spaces in the text.

#### B. Case Folding

The case folding process aims to make uniform text by changing uppercase letters to lowercase letters. This stage is applied to input tweets that have been cleaned.

#### C. Tokenization

Tokenization is converting sentences in tweets into a group of tokens or fragments of words which are then stored in an array. At this stage, the tweets have gone through the case folding step.

#### D. Normalization

The purpose of normalization in the pre-processing stage is to transform data into a uniform or standard form. Normalization removes irrelevant variations or differences in the data, facilitating subsequent processing and analysis. In this normalization using the tqdm library. tqdm is used to monitor the text normalization process.

#### E. Remove Stopwords

At this stage, the goal is to pick and choose the words that matter. In general, words that appear repeatedly or have a high frequency are considered unimportant words. By removing these unimportant words, essential words become the main focus. Input in this process are tweets that have gone through the tokenization stage. The stopwords dictionary used is obtained from the Kaggle site.

#### F. Stemming

Stemming in the pre-processing stage aims to reduce words to their base or root forms. Stemming removes word endings or prefixes to transform words with the same root into a uniform format. The library used for stemming is the " Sastrawi " library.

### SMOTE

In Table 2, it can be seen that the data needs to be balanced between classes. This can be overcome by using the Synthetic Minority Over-sampling Technique (SMOTE).

### CNN Modeling

After getting the data that has gone through the preprocessing process and has been cleaned, the data is used to build a sentiment analysis model using the Convolutional Neural Network (CNN) method.

### Evaluation

In this study, an evaluation will be carried out using an evaluation matrix after building the CNN model to measure results and determine the model's effectiveness. The evaluation matrix also plays an essential role in assessing the performance and quality of a model that has been built (Munawar et al., 2020). The evaluation matrix used consists of the following:

#### A. Confusion Matrix

The confusion matrix describes the amount of test data correctly and incorrectly classified by the model. This matrix explains the model's performance in classifying data (Normawati & Prayogi, 2021), (Nawangsih et al., 2021). An illustration of the confusion matrix can be seen in Table 3 (Novitasari & Dwifabri Purbolaksono, 2021).

\*name of corresponding author



Table 3. Illustration of Confusion Matrix

Actual Class	Positive Predictions	Negative Predictions
Actual Positive	True Positive	True Negative
Actual Negative	False Positive	False Negative

## B. Accuracy

Accuracy is performed to measure the ratio between the number of correct predictions and the total number of predictions made by the model.

$$accuracy = \frac{TP+TN}{(TP+FP+FN+TN)} \quad (1)$$

## C. Recall

The recall measures the model's ability to identify all positive events correctly.

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

## D. Precision

Precision is performed to measure the model's ability to identify only positive events correctly.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

## E. F1-Score

The F1-score is the harmonic average of precision and recall and balances the two metrics.

$$F1 \text{ score} = \frac{2 \times (\text{recall} \times \text{precision})}{(\text{recall} + \text{precision})} \quad (4)$$

## RESULT

In this study, crawling data from the Twitter application was carried out, and 17.270 data were obtained. After the data is collected, it is labeled to determine class sentiment by using labels -1 as negative, 0 as neutral, and 1 as positive. Data that has been labeled as not necessarily clean for processing. It is necessary to clean it through the preprocessing stage to make the data more optimal when processed. Clean data will be entered into the CNN model, and testing and training will be carried out for k iterations using the k-fold cross validation method. The data will be evaluated by calculating the evaluation matrix and experiments using two scenarios consisting of:

- i. Scenario I : looking for the best epoch value.
- ii. Scenario II : effect of SMOTE

In scenario I, data training uses epochs 5, 10, and 20. The best results are shown in Table 4 which gets 87% accuracy, 86% recall, 87% precision, and 87% F1-Score

Table 4. Best Epoch Results

Epoch	Accuracy	Recall	Precision	F1- Score
5	86%	86%	86%	86%
10	87%	86%	87%	86%
20	86%	86%	86%	86%

In scenario II, the SMOTE technique is added to balance minority data. The results can be seen in Table 5 that using SMOTE can improve recall, precision, and F1-Score.

Table 5. Implementing SMOTE

Metode	Accuracy	Recall	Precision	F1- Score
CNN + SMOTE	87%	93%	93%	90%
CNN	87%	86%	87%	86%

## DISCUSSIONS

Based on the two scenarios that have been done, each scenario has different results. In scenario I, three epoch numbers are taken, namely epoch 5, 10, and 20; With these figures, it can be seen that epoch 10 has better results than epochs 5 and 20, namely 87% accuracy, 86% recall, 87% precision, and 86% F1-Score. In scenario II, the SMOTE technique was applied to balance minority data, and the results obtained were in line with expectations,

\*name of corresponding author



increased recall, precision, and F1-Score, namely 87% accuracy, 93% recall, 93% precision, and 90% f1-score. Figure 4 shows a graph of the improvement when using SMOTE.

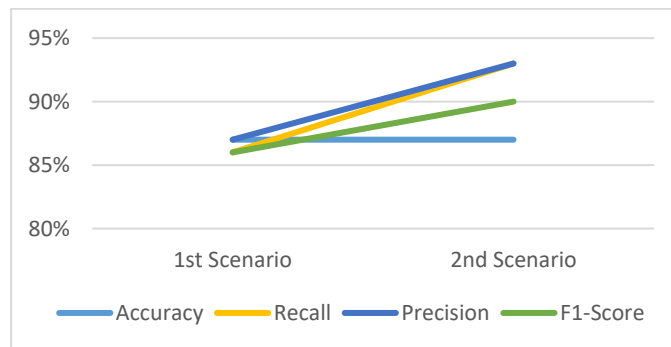


Fig. 4 Evaluation Chart

### CONCLUSION

In this study, sentiment analysis was carried out on rising fuel prices using the Convolution Neural Network (CNN) method. The dataset used results from tweets on the Twitter application with the keyword "BBM naik." The datasets are 17,270, with 6,662 negative labels, 7,960 neutral labels, and 2,648 positive labels. This study uses K-Fold Cross Validation with a total of  $k = 10$  and uses epoch 10, with the highest accuracy value of 87%. In this study, applying the SMOTE technique to balance unbalanced data effectively improves the performance of the CNN model by obtaining 87% accuracy, 93% recall, 93% precision, and 90% f1-score. It can be concluded that the K-fold cross validation with the CNN model is effective using the SMOTE technique.

### REFERENCES

- Alhakiem, H. R., & Setiawan, E. B. (2022). Aspect-Basled Sentiment Analysis on Twitter Using Logistic Regression with FastText Feature Expansion. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 6(5), 840–846. <https://doi.org/10.29207/resti.v6i5.4429>
- Asroni, A., Fitri, H., & Prasetyo, E. (2018). Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik). *Semesta Teknika*, 21(1). <https://doi.org/10.18196/st.211211>
- dpr.go.id. (2022, September 19). *Nur Azizah: Kenaikan Harga BBM Memberatkan Rakyat*. Dpr.Go.Id.
- Fonda, H., Irawan, Y., Febriani, A., Informatika, S., & Pekanbaru, H. T. (2020). KLASIFIKASI BATIK RIAU DENGAN MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORKS (CNN) 1 2 3 Email : 1 2 3. In *JIK* (Vol. 9, Issue 1). <http://jik.htp.ac.id>
- Irawan, F. A., & Rochmah, D. A. (2022). Penerapan Algoritma CNN Untuk Mengetahui Sentimen Masyarakat Terhadap Kebijakan Vaksin Covid-19. *JURNAL INFORMATIKA*, 9(2). <http://ejournal.bsi.ac.id/ejournal/index.php/ji>
- Jacovi, A., Shalom, O. S., & Goldberg, Y. (2018). *Understanding Convolutional Neural Networks for Text Classification*.
- Kumar, R., & Garg, S. (2020). Aspect-Based Sentiment Analysis Using Deep Learning Convolutional Neural Network. In *Advances in Intelligent Systems and Computing* (Vol. 933, pp. 43–52). Springer Verlag. [https://doi.org/10.1007/978-981-13-7166-0\\_5](https://doi.org/10.1007/978-981-13-7166-0_5)
- Lee, H. M., & Sibaroni, Y. (2023). Comparison of IndoBERTweet and Support Vector Machine on Sentiment Analysis of Racing Circuit Construction in Indonesia. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(1), 99–106. <https://doi.org/10.30865/mib.v7i1.5380>
- Listyarini, S. N., & Anggoro, D. A. (2021). Analisis Sentimen Pilkada di Tengah Pandemi Covid-19 Menggunakan Convolution Neural Network (CNN). *Jurnal Pendidikan Dan Teknologi Indonesia*, 1(7), 261–268. <https://doi.org/10.52436/1.jpti.60>
- Munawar, Z., Putri, N. I., & Musadad, D. Z. (2020). MENINGKATKAN REKOMENDASI MENGGUNAKAN ALGORITMA PERBEDAAN TOPIK. *Jurnal Sistem Informasi, J-SIKA*, 1(2).
- Nawangsih, I., Melani, I., & Fauziah, S. (2021). PREDIKSI PENGANGKATAN KARYAWAN DENGAN METODE ALGORITMA C5.0 (STUDI KASUS PT. MATARAM CAKRA BUANA AGUNG). *Jurnal Pelita Teknologi*, 16(2), 24–33.

\*name of corresponding author



- Nikmatul Kasanah, A., Muladi, & Pujiyanto, U. (2019). Penerapan Teknik SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Objektivitas Berita Online Menggunakan Algoritma KNN KNN. *Jurnal RESTI*, 3(2), 196–201.
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. In *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 5, Issue 2).
- Novitasari, F., & Dwifabri Purbolaksono, M. (2021). Sentiment Analysis Aspect Level on Beauty Product Reviews Using Chi-Square and Naïve Bayes. *JOURNAL OF DATA SCIENCE AND ITS APPLICATIONS*, 4(1), 18–030. <https://doi.org/10.34818/JDSA.2021.4.72>
- Ramadhan, A. I., & Setiawan, E. B. (2023). Aspect-based Sentiment Analysis on Social Media Using Convolutional Neural Network (CNN) Method. *Building of Informatics, Technology and Science (BITS)*, 4(4). <https://doi.org/10.47065/bits.v4i4.3103>
- Rodani, A. (2022, September 12). *Menyikapi Kenaikan Harga BBM secara Bijak*. Djkn.Kemenkeu.Go.Id.
- Sihombing, J. (2022, September 15). *Kenaikan Harga BBM : Jahat atau Sepakat..???* Djkn.Kemenkeu.Go.Id.
- Siringoringo, R. (2018). *KLASIFIKASI DATA TIDAK SEIMBANG MENGGUNAKAN ALGORITMA SMOTE DAN k-NEAREST NEIGHBOR* (Vol. 3, Issue 1).
- Wardani, W., Suriana, Arfah, S., Zulaili, & Lubis, P. (2022). Dampak kenaikan Bahan Bakar Minyak (BBM) Terhadap Inflasi dan Implikasinya Terhadap Makroekonomi Indonesia. *AFoSJ-LAS*, 2(3), 63–70. <https://j-las.lemkomindo.org/index.php/AFoSJ-LAS/index>
- Yuliska, Hidayatul Qudsi, D., Hakim Lubis, J., Umam Syaliman, K., & Fadilah Najwa, N. (2021). ANALISIS SENTIMEN PADA DATA SARAN MAHASISWA TERHADAP KINERJA DEPARTEMEN DI PERGURUAN TINGGI MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 8(5), 1067–1076. <https://doi.org/10.25126/jtiik.202184842>

\*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.