Hoax Detection Through the Method of Convolutional Neural

Akmaezhan A. Nogaibayeva[1]
Suleyman Demirel University

Gulzhaina K.Kassymova[2]
Abal Kazakh National Pedagogical University

Sulis Triyono[3]
Universitas Negeri Yogyakarta

Binar Winantaka[4]
Universitas Negeri Yogyakarta

Corespondence : Akmarzhan A.Nogaibayeva (zhaina.kassym@gmail.com)

Abstract

The term "hoax" or "fake news" refers to the deliberate propagation of false material on social media in order to confuse and mislead readers in order to achieve an economic or political purpose. Furthermore, the increasing diversity and number of participants in the sphere of news authoring and transmission have resulted in the creation of news pieces that must be acknowledged regardless of whether they are trustworthy or not. Furthermore, hoaxes might undermine Indonesian society's social and political features. In 2016, Central Connecticut University published The World's Most Literate Nations, in which Indonesia ranked 60th out of 61 countries, indicating that Indonesian media literacy still needs to improve in terms of critically evaluating information and distinguishing between fake news and valid news. Based on this description, the research will develop the Synonym-Based Data Augmentation for Hoax Detection method and Easy Data Augmentation (EDA) approach. This study yielded an accuracy of 8.81, showing that it is effective at spotting fake news.

Keywords: CNN; Data Enhancement; Detection; EDA Approach; Hoax

Introduction

The term "hoax" refers to the deliberate dissemination of misleading material on social media in order to confuse and mislead readers in order to achieve an economic or political objective (Nayoga et al., 2021; Santoso et al., 2020). Furthermore, the increasingly diversified and numerous participants in the sphere of news production and transmission have resulted in the development of news pieces, whether credible or not (Nayoga et al., 2021; Santoso et al., 2020; Utami, 2019b). The National Intelligence Agency (BIN) estimates that roughly 60% of news in Indonesia is false (Ghinadya et al., 2020; Herrero-Diz et al., 2022; Mufid et al., 2019; Spera et al., 2019).

Furthermore, the deception has put Indonesian society's social and political components in jeopardy. People with mental illnesses are being persecuted (Christina Gregory, 2017), and Indonesian politicians' reputations are being harmed during the general election (Gusfa et al., 2020; Utami, 2019a). In 2016, Central Connecticut University published a research titled World's Most Literate Nations, in which Indonesia rated 60th out of 61 countries (Fuadi et al., 2020). According to this rating, Indonesian media literacy still needs to develop in terms of critically evaluating information and discriminating between fake and legitimate news (Jormand et al., 2021; Khidhir, 2019).

Several studies use deep learning to detect hoax, such as research by Nayoga et al. (2021), which discusses Hoax Analyzer for Indonesian News Using Deep Learning Models, using the Deep Neural Network (DNN) method, Long Short-Term Memory (LSTM), Bidirectional LSTM (BI-LSTM), Gated Recurrent Unit (GRU), Bidirectional GRU (BI-GRU), and 1-Dimensional Convolutional Neural

Network (1D Other research has been undertaken by (Apriliyanto et al., 2020; Shao et al., 2022; Singh et al., 2021; Youssef et al., 2022) on fake identification in Indonesian Language using Long Short-Term Memory Model, and the identification of fake news in Indonesian has been effective using the LSTM method.

This research also developed the synonym-based data augmentation approach to detect hoax in headlines using the Convolutional Neural Network (CNN) method (Albashish, 2022; Anton et al., 2021; Defriani et al., 2022; Putra et al., 2021). Previous research investigated the Bidirectional LSTM approach with foreign language news and obtained an F1- score of 0.2423 (Ghinadya et al., 2020). However, because the data in this study was based on foreign language news, the accuracy percentages achieved were relatively low, with a macro F1 score of only 0.24. Based on earlier research, this study employed Indonesian News data with the CNN approach to improve the augmentation process. This approach is commonly utilised in augmentation (Nayoga et al., 2021). During the detection phase, text was employed as the primary input for Natural Language Processing (NLP), which requires the same corpus of language as the information given. As a result, the non-English text corpus, including the Indonesian language text corpus, contains less vocabulary than the English text corpus, resulting in the availability of only a few non-English hoax detectors. As a result, the CNN approach was utilised in this work to detect hoax in Indonesian, and data augmentation was used to improve the CNN accuracy in detecting hoax in Indonesian.

Methods

The dataset used in this study consists of hoax news and factual news obtained from local news publishers such as detikNews, Kumparan, Kompas, Liputan6, Sindonews, Republika, Tempo, Tribunnews, Wowkeren, Kapanlagi, data Mendeley, Okezone, and TurnbackHoax.id. The data used totaled 1000 in Indonesian, with 700 real news and 300 fake news. The data was then preprocessed, which included case folding, normalisation, tokenization, filtering, and steamming. Following the creation of a clean dataset as a result of the preprocessing procedure, the data is divided into training and testing data. The Easy Data Augmentation (EDA) technique is then used to do data augmentation. This level was completed in order for the machine to learn and recognise the origin of various types of words. Following the data augmentation procedure, the CNN approach is used to detect fake news. The characteristics produced from word tokenization were incorporated into the convolutional layer at this stage, and the results were aggregated into representative figures. Following that, the confusion matrix is used to analyse and measure system performance. As a result, the study's output, namely the categorization of test data in the form of documents labelled as hoax and documents classified as true, is obtained. Figure 1 depicts the planned system.

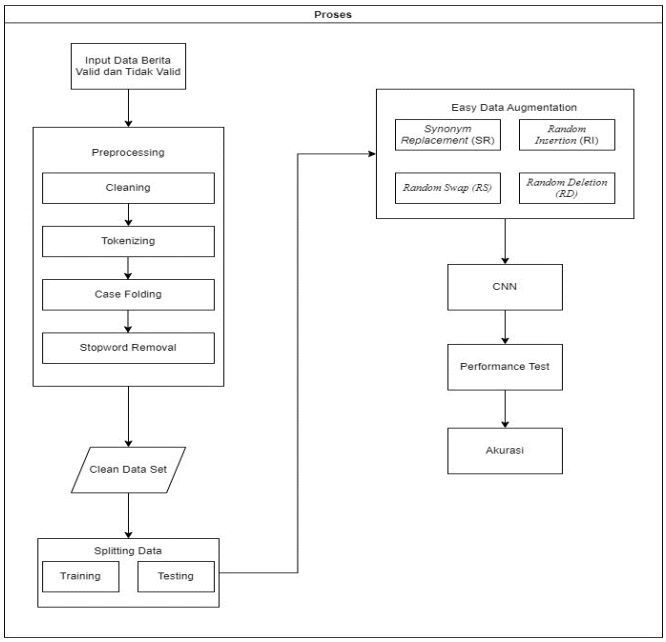

Figure 1: EDA-based system design

Augmentation

The EDA technique was utilised in this investigation, which includes four operations: random deletion (RD), random insertion (RI), random exchange (RS), and synonymous replacement (SR). The original sentence became more than four extra sentences from the prior dataset in this study. Table 1 is an example of an EDA-generated sentence.

Table 1. Examples of EDA Result Sentences

Sentence	
Not Using EDA	Prospective students of the Faculty ofMedicine have taken an exam in the morning at 09.00 WIB
Synonym Replacement (SR)	Prospective students of the Faculty ofMedicine have taken a test in the morning at 09.00 WIB
Random Deletion (RD)	Prospective students of the Faculty ofMedicine have taken the test at 09.00 WIB
Random Insertion (RI)	Prospective students of the Faculty of Medicine have taken the entrance exam in the morning at 09.00 WIB
Random Exchange(RS)	Prospective students of the Faculty ofMedicine have taken an exam in the morning at 09.00 WIB

Table 2 shows parameter suggestions for using EDA tools. The parameter Alpha () denotes "the percentage of words in a sentence that are changed by each augmentation." The amount of additional sentences generated by EDA per original sentence is denoted by Naug (Wei, J., & Zou, 2019).

Table 2. Recommended Parameters

N_{train}	α	N_{aug}
500	0.05	16
2000	0.05	8
5000	0.10	4
More	0.10	4

Method of Convolutional Neural Network (CNN)

The CNN approach is implemented by starting with pre-processed input news. Convolutional layers will include features generated from word tokenization. It is the layer where the input layer is abstracted into a feature map made up of kernels/filters (which might be several).

Convolutional findings are aggregated into representative numbers and supplied into a fully linked layer. This layer is totally linked since each neuron is connected to the neurons in the previous layer. Each activation of the previous layer is transformed into one-dimensional data before being coupled to all neurons in a completely connected layer. The resulting weights for each feature in the text can then be used to make a classification decision.

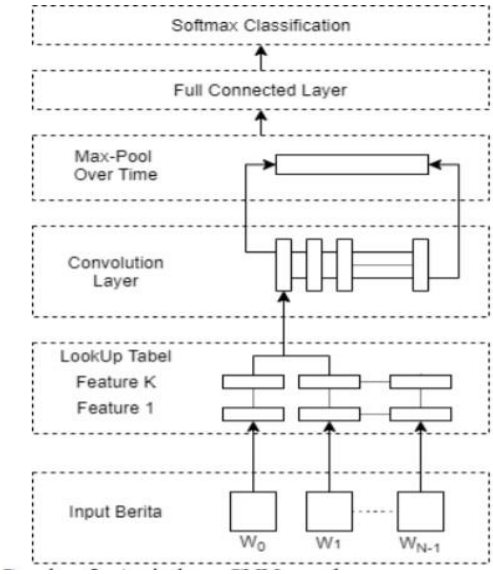

Figure 2: CNN Text Processing Architecture

Finally, the resulting weight is fed into the fully linked layer. Its CNN network component goes through its own backpropagation technique to determine the best weight (Kurniawan et al., 2021). Each neuron will be assigned a value, in this example "fact." "breaking news" or "fake news" Finally, neurons assign a value to each label and arrive at a judgement or classification result.

A CNN 1D was constructed in this study and employed in this experiment to match the dimensions of the text input data. CNNs extracted characteristics from data vectors concurrently rather than sequentially, resulting in significantly faster data processing. The 1D-CNN model's hyperparameters are as follows: 128 filter sizes with 5 default kernel sizes and steps. The 1D-CNN model architecture employs the ReLU activation function. GlobalMaxPooling 1D is the pooling layer used to sample feature maps before further processing by the classifier. Figure 3 depicts the architectural representation of the 1D-CNN model used in this investigation (Nayoga et al., 2021).

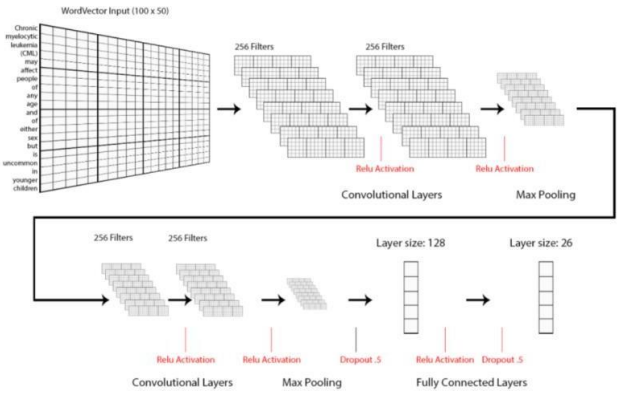

Figure 3. The illustration of 1D-CNN architecture model

Results and Discussion

The Convolutional Neural Network (CNN) approach was employed, along with data augmentation. The flowchart of the algorithm CNN with data augmentation implementation may be observed. Figure 4.

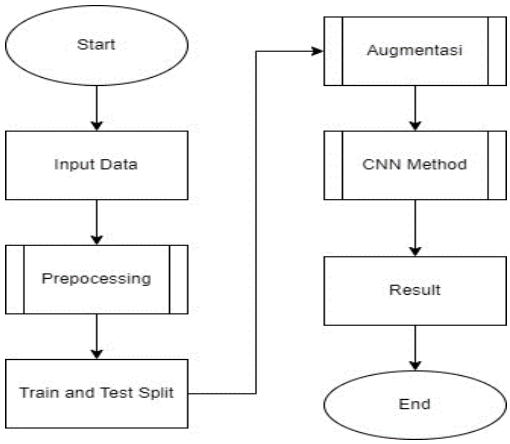

Figure 4. *CNN flowchart* with data augmentation

Figure 4 depicts the procedures for detecting hoax news using CNN with data augmentation. The first stage is the input of the dataset. This study's dataset is separated into two categories: positive and negative. In this investigation, data was collected using web scraping techniques and the Python module newspaper3k.

The following stage is data preparation. The third stage is to divide the data, namely data training and data testing, using Split Validation tools. The distribution of training data and testing data in this study was 70% training data and 30% testing data. Table 3 shows the results of partitioning the train and test data.

Table 3. Results Distribution of Data Train and Test

Stem	Label
kadrun look at hot and cold images	1
good afternoon friends facebook information rezeky where yesterday took part in the smartphone advertising event jne June special order smartphone vivo v pro pricing rp hp vivo pro, interested fb friends, join the inbox admin whatsapp	1
see govern the country	1
The president's joke wins the corona victims' burial plot in Italy, despite state-of-the-art health-care facilities failing to control the corona.	1
Please look for foolish police officers as soon as possible.	1
China police torture uighur women by strangling their necks with their feet, according to literacy, and thousands of language police take turns staring at savages.	1
Jokowi is prepared to resign.	1
Hermansyah, a former prostitute, and Irina, the wife of an IT professional	1
nezar patria pki cadres palace night meeting curfew alfian tandjung accused	0
Take the Palson semi-container truck.	0
mr immunisation contact unwell vaccine effects	0
The persecution of West Java clerics hr Bogor clerics ust sulaiman sliced Bogor madmen so many times Muslims at queue meetings are constantly on guard.	0
rambutan food hazards drinking sweet tea due to airway blockage	0
Be careful, gaeeees, as a result of playing the popular snake game right now.	1
slap news dot blogspot dot com article merdekaind appear article title report pdip sukmawati hold strong people article write	1

Table 4. Color Information on Train and Test

	Data Trains
	Test Data

The fourth stage of this investigation involves data augmentation, for which the researchers used the Easy Data Augmentation (EDA) technique. This was done in order for the machine to learn and recognise the origins of various types of words. The Easy Data Augmentation (EDA) technique has numerous stages, the first of which is Synonym Replacement (SR). It is a method for generating new data by randomly selecting words from phrases and replacing them with synonyms of the selected words. The second stage of the EDA technique is Random Insertion (RI), which selects words in phrases at random, replaces them with synonyms, and inserts the word in a different position in the text. The third stage of the EDA approach is the Random Swab (RS), which is a random exchange of samples without the use of synonyms. The EDA technique concludes with Random Deletion (RD), which is the random deletion of words in the text with a probability defined by the parameters.
The aggregate of all approaches is performed on the results of the augmentation above to obtain the ultimate outcome of the technique utilised. Following that, a similarity assessment is performed using a vulnerable Cosine similarity value ranging from -1 to 1, with the two texts being said to be similar when the angle between the vectors is close to zero and the cosine value is close to 1. Thus, each piece of news or data with a similarity score of 1 indicates that there is no difference in the terms before and after the augmentation.

The CNN approach is used to detect in the sixth step of this investigation. At this point, the Convolutional Neural Network (CNN) algorithm will be used to distinguish between hoax and legitimate news. The researcher includes Adaptive Moment Estimation (Adam) in the training model to optimise the training data, which was done to improve the accuracy of the method utilised. The researcher used epochs = 5 epochs to run the fit model. An apoch indicates how frequently the network will examine the complete data set, whereas batch_size = 64 indicates the amount of training instances in one forward or backward pass. The more RAM is requested, the greater the batch_size value. Figure 5 depicts the CNN accuracy algorithm's outcomes.

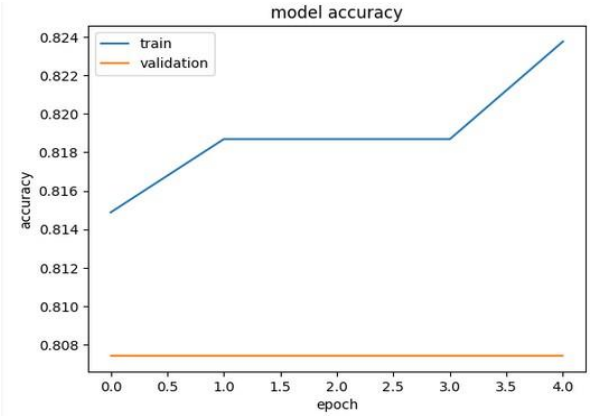

Figure 5: CNN algorithm accuracy result

Meanwhile, the loss function is utilised to display the loss model data in graphical form. The loss model is the value of the training dataset's loss calculation function that is then displayed in graphical form so that it is known how much the loss value lowers or can be interpreted function used to see the performance of the CNN model. Figure 6 depicts the CNN algorithm function's loss outcomes.

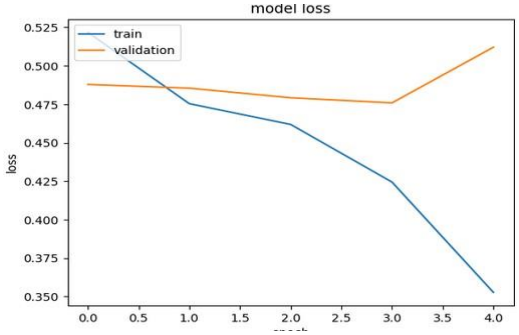

Figure 6. Loss result of CNN algorithm functions

The sixth stage is to carry out an evaluation that aims to measure the work of the model. Evaluationmethod applied is confusion matrix. The results of the confusion matrix can be seen in Table 5.

Parameter	Calculation	Results (%)
accuracy	$\frac{(TP+TN)}{(TP+FP+FN+TN)} x\,100 = \frac{(21+1029)}{(1029+17+201+21)} x\,100\%$	82.81
precision	$\frac{TP}{(TP+FP)} x\,100\% = \frac{1029}{(1029+201)} x\,100\%$	83.66
recall	$\frac{TP}{(TP+FN)} x\,100\% = \frac{1029}{(1029+17)} x\,100\%$	98.37
F1-Score	$2\frac{Presisi*Recall}{Presisi+Recall} * 100\% = 2\frac{83.66\%*98.37\%}{83.66\%+93.37\%} * 100\%$	90.42

Table 5. Calculation of Confusion Matrix

Figure 7 depicts the confusion matrix results in tabular form.

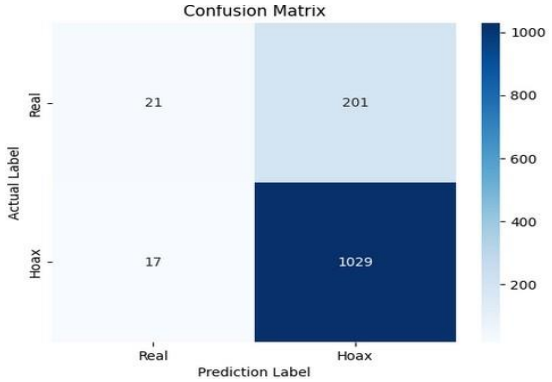

Figure 7. Results of confusion matrix in tabular form

Table 6 displays the confusion matrix results in percentage form.

Table 6. Results

Yield	Result
Accuracy	82.81%
Precision	83.66%
Recall	98.37%
F1-Score	90.42%

According to the findings of earlier study, Easy Data Augmentation for Fake News Detection employs the Convolutional Neural Network Method for Indonesian News to achieve pretty excellent accuracy results.

Conclusion

According to the study's findings, implementing the Convolutional Neural Network (CNN) algorithm with data augmentation can raise awareness about circulating material and aid in the avoidance of hoaxes. This conclusion is based on the testing findings, which showed that it had an accuracy of 82.81% in detecting bogus news. As for future research directions, it is required to evaluate the accuracy with the outcomes of this study using different algorithms or other augmentation strategies.

References

Albashish, D. (2022). Ensemble of adapted convolutional neural networks (CNN) methods for classifying colon histopathological images. *PeerJ Computer Science*, *8*. https://doi.org/10.7717/peerj-cs.1031

Anton, A., Nissa, N. F., Janiati, A., Cahya, N., & Astuti,

P. (2021). Application of Deep Learning Using Convolutional Neural Network (CNN) Method For Women's Skin Classification. *Scientific Journal of Informatics*, *8*(1). https://doi.org/10.15294/sji.v8i1.26888

Apriliyanto, A., & Kusumaningrum, R. (2020). Hoax Detection in Indonesia Language Using Long Short-Term Memory Model. *Sinergi*, *24*(3), 189.https://doi.org/10.22441/sinergi.2020.3.003

Borisov, V., Leemann, T., Sessler, K., Haug, J.,Pawelczyk, M., & Kasneci, G. (2022). Deep Neural Networks and Tabular Data: A Survey. *IEEETransactions on Neural Networks and LearningSystems*. https://doi.org/10.1109/TNNLS.2022.3229161

Buhrmester, V., Münch, D., & Arens, M. (2021). Analysisof Explainers of Black Box Deep Neural Networks for Computer Vision: A Survey. *Machine Learning and Knowledge Extraction*,*3*(4).https://doi.org/10.3390/make3040048

Christina Gregory. (2017). *Internet Addiction Disorder - Signs, Symptoms, and Treatments*. Vertical Health LLC.

Defriani, M., & Jaelani, I. J. (2022). Recognition of Regional Traditional House in Indonesia Using Convolutional Neural Network (CNN) Method. *Journal of Computer Networks, Architecture and High Performance Computing*, *4*(2). https://doi.org/10.47709/cnahpc.v4i2.1562

Ghinadya, & Suyanto, S. (2020). Synonyms-Based Augmentation to Improve Fake News Detection using Bidirectional LSTM. *2020 8th International Conference on Information and Communication Technology, ICoICT 2020*.

https://doi.org/10.1109/ICoICT49345.2020.9166230

Gusfa, H., & Kadjuand, F. E. D. (2020). Political Agonism for Indonesian Cyberpolitic: Critical Cyberculture to Political Campaign of 2019 Indonesian Presidential Election in Twitter. *Nyimak: Journal ofCommunication*, *4*(2).

https://doi.org/10.31000/nyimak.v4i2.2685

Hao, Z., Li, M., Yang, W., & Li, X. (2022). Evaluation of UAV spraying quality based on 1D-CNN model and wireless multi-sensors system. *Information Processing in Agriculture*.https://doi.org/10.1016/j.inpa.2022.07.004

Herrero-Diz, P., & Pérez-Escolar, M. (2022). Analysis of Hoaxes about COVID-19 Debunked by Maldita and Colombiacheck: Effects of the Infodemic on the Behavior of Society. *Palabra Clave*, *25*(1). https://doi.org/10.5294/pacla.2022.25.1.7

Jormand, H., Bashirian, S., Barati, M., Khazaei, S., Jenabi, E., & Zareian, S. (2021). A Qualitative Study On People's Experiences Of Covid-19 Media Literacy. *Media Literacy And Academic Research*, *4*(1). Retrieved from

https://www.ceeol.com/search/article- detail?id=946040

Khidhir, S. (2019). Indonesia is too slow! *Theaseanpost*. Kurniawan, A. A., & Mustikasari, M. (2021). Implementasi Deep Learning Menggunakan Metode CNN dan LSTM untuk Menentukan Berita Palsu dalam Bahasa Indonesia. *Jurnal Informatika Universitas Pamulang*, *5*(4),544. https://doi.org/10.32493/informatika.v5i4.6760

Montavon, G., Samek, W., & Müller, K. R. (2018). Methods for interpreting and understanding deep neural networks. In *Digital Signal Processing: A Review Journal* (Vol. 73).https://doi.org/10.1016/j.dsp.2017.10.011

Mufid, F. L., & Hariandja, T. R. (2019). Efektivitas Pasal 28 Ayat (1) UU ITE tentang Penyebaran Berita Bohong (Hoax). *Jurnal Rechtens*, *8*(2). https://doi.org/10.36835/rechtens.v8i2.533

Nayoga, B. P., Adipradana, R., Suryadi, R., & Suhartono,

D. (2021). Hoax Analyzer for Indonesian News Using Deep Learning Models. *Procedia Computer Science*, *179*, 704–712.https://doi.org/10.1016/j.procs.2021.01.059

Putra, O. V., Musthafa, A., Nur, M., & Rido, M. (2021). Classification of Calligraphy Writing Types Using Convolutional Neural Network Method (CNN). *Procedia of Engineering*

and Life Science, *2*. https://doi.org/10.21070/pels.v2i0.1136 Salahuddin, Z., Woodruff, H. C., Chatterjee, A., &

Lambin, P. (2022). Transparency of deep neural networks for medical image analysis: A review of interpretability methods. In *Computers in Biology and Medicine* (Vol. 140) https://doi.org/10.1016/j.compbiomed.2021.1051 11

Santoso, H. A., Rachmawanto, E. H., Nugraha, A.,Nugroho, A. A., Setiadi, D. R. I. M., & Basuki, R. S.(2020). Hoax classification and sentiment analysisof Indonesian news using Naïve Bayesoptimization. *Telkomnika (Telecommunication Computing Electronics and Control)*, *18*(2). https://doi.org/10.12928/TELKOMNIKA.V18I2.14744

Semiawan, C. R. (2017). *Metode Penelitian Kuantitatif*. Grasindo

Shao, X., & Kim, C. S. (2022). Unsupervised Domain Adaptive 1D-CNN for Fault Diagnosis of Bearing.*Sensors*, *22*(11). https://doi.org/10.3390/s22114156

Singh, K., Scholar, R., Mahajan, A., & Mansotra, V. (2021). 1D-CNN based Model for Classification and Analysis of Network Attacks. *International Journal of Advanced Computer Science and Applications*, *12*(11). https://doi.org/10.14569/IJACSA.2021.0121169

Spera, R., & Peña-Guzmán, D. M. (2019). The Anatomy of a Philosophical Hoax: The Politics of Delegitimation in Contemporary Philosophy. *Metaphilosophy*, *50*(1–2). https://doi.org/10.1111/meta.12343

Sugiyono. (2019). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (1st ed.). Penerbit Alfabeta.

Tang, W., Long, G., Liu, L., Zhou, T., Jiang, J., & Blumenstein, M. (2020). Rethinking 1D-CNN for Time Series Classification: A Stronger Baseline. *ArXiv*. Retrieved from https://www.arxiv- vanity.com/papers/2002.10061/

Utami, P. (2019). Hoax in modern politics: The meaning of hoax in Indonesian politics and democracy. *Jurnal Ilmu Sosial Dan Ilmu Politik*, *22*(2). https://doi.org/10.22146/jsp.34614

Utami, P. (2019b). Hoax in Modern Politics. *Jurnal Ilmu Sosial Dan Ilmu Politik*, *22*(2). https://doi.org/10.22146/jsp.34614

Wei, J. & Zou, K. (2019). EDA:Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 6382–6388. https://doi.org/10.48550/arXiv.1901.11196

Wu, Y. Y., Hu, Y. S., Wang, J., Zang, Y. F., & Zhang, Y. (2022). Toward Precise Localization of Abnormal Brain Activity: 1D CNN on Single Voxel fMRITime-Series. *Frontiers in Computational Neuroscience*, *16*.

https://doi.org/10.3389/fncom.2022.822237 Youssef, A. M., Pradhan, B., Dikshit, A., Al-Katheri, M.

M., Matar, S. S., & Mahdi, A. M. (2022). Landslide susceptibility mapping using CNN-1D and 2D deep learning algorithms: comparison of their performance at Asir Region, KSA. *Bulletin of Engineering Geology and the Environment*, *81*(4). https://doi.org/10.1007/s10064-022-02657-4

.