IMPLEMENTATION OF BILSTM-SVM ALGORITHM TO DETECT FAKE NEWS ON TEXT-BASED MEDIA

Felix Liman[1], Carsten[2*], Sufiandy Sufinata[3], Syanti Irviantina[4], Sunaryo Winardi[5]
Fakultas Informatika, Program Studi Teknik Informatika, Universitas Mikroskil, Medan
Universitas Mikroskil, Jl. Thamrin No. 112, 124, 140, Medan
[1]181110131@students.mikroskil.ac.id, [2]181111366@students.mikroskil.ac.id (*Corresponding Author),
[3]181111374@students.mikroskil.ac.id,[4]syanti@mikroskil.ac.id, [5]sunaryo.winardi@mikroskil.ac.id

Abstract. **Online media is one of the places where news can spread quickly and everyone can access it easily and freely. Not only real or valid news is spread on online media, but fake news can also be easily spread on online media, and readers sometimes do not realize that the news they read is fake. As a result, wrong opinions arise that can lead to disputes, as well as divisions between individuals or groups. This study implements the BiLSTM-SVM algorithm to detect fake news that is spread on one of the online media, namely Twitter. The steps taken are tidying up the news text (text preprocessing), converting every word from the news text into numbers in vector form (word embedding), processing the numbers, and then classifying the results of the processing with the BiLSTM-SVM model formed with TensorFlow 2.0 help, and see the performance generated by the BiLSTM-SVM algorithm. The results obtained include an accuracy rate of 86% and an F1 Score value of 87.5% in detecting news from data validation with the same news topic.**

INTRODUCTION

Online media have not only changed the way information is provided but also changed the way individuals consume it. Currently, the dissemination of information or news through online media is not limited to well-known news sites. Everyone who has access to the internet can contribute to the dissemination of information. However, a large amount of information or news that is disseminated individually or in groups is not necessarily true and follows the facts that occurred. Information or news that contains things that are uncertain or that are not following the facts that occur is called fake news [1]. The negative impact caused by the spread of fake news refers to people's emotions, causing wrong opinions, as well as causing negative propaganda that deliberately manipulates views and thoughts to form wrong behavior [2]. The wrong opinion is then led by unscrupulous individuals, causing disputes and divisions between individuals and certain groups [3]. Therefore, the problem of spreading fake news must be addressed immediately.

One way to deal with the problem of spreading fake news is to use artificial intelligence to detect manipulated facts [4]. A model formed from an algorithm can be measured its performance by measuring the accuracy generated by the model. High-performance models will produce accurate predictions [5].

Previous research using the BiLSTM-SVM algorithm was conducted by Jihen Fourarti, Mohamed Othmani, and Hela Ltifi (2022) to detect tremor which is the main symptom of Parkinson's disease. By using 2 different datasets, the BiLSTM-SVM algorithm results in an accuracy of 98.93% in the first dataset (Physionet dataset) and 97.22% in the second dataset (John Radcliffe Hospital dataset) [6]. Therefore, the proposed solution is to implement the BiLSTM-SVM algorithm to establish a system to detect fake news. With the LSTM variant that is considered superior, BiLSTM (Bidirectional LSTM) will process the data and then classify it using the SVM algorithm into 2 categories [7] namely fake news or not.

LITERATURE REVIEW
A. Fake News
Fake news is often referred to as a hoax that describes everything as false, especially in online media. The origin of word hoax is believed to have existed since the 1620s. According to John Tillotson (1630 - 1694), the word hoax comes from the word hocus from the hocus pocus mantra [8].

According to Henri Subiakto (2019), an Expert Staff to the Minister for Law at the Ministry of Communication and Information (Kominfo), the hallmark of fake news or hoaxes is that the source of the information is unclear and deceives a certain race, ethnicity or religion. In addition, information that does not contain the complete 5W+1H element and readers are asked to disseminate this information is also another characteristic of fake news [9].

B. BiLSTM-SVM

BiLSTM-SVM is a combination of 2 artificial intelligence algorithms, namely BiLSTM (Bidirectional Long Short Term Memory) and SVM (Support Vector Machine) for learning and classification. The basic algorithm of BiLSTM-SVM is the LSTM and SVM algorithms which are used side by side to take advantage of both algorithms [6], where the SVM algorithm is very efficient in the learning process but does not support the reasoning of relationships in sentences, and the LSTM algorithm which is a specialization of RNN (Recurrent Neural Network) can do reasoning and accommodate dependencies between words in sentences which can cover the disadvantage of SVM [10]. The LSTM-SVM architecture is shown in Figure 1 below.
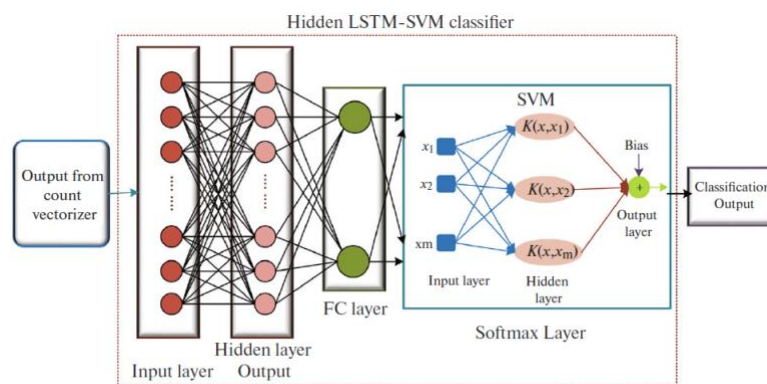

Figure 1. LSTM-SVM Architecture [16]

BiLSTM will replace LSTM in this LSTM-SVM architecture to form BiLSTM-SVM because BiLSTM is rated better than LSTM [11]. BiLSTM has the same structure as LSTM, the difference is in the addition of a backward layer [12].

Bidirectional Long Short Term Memory (BiLSTM) is a development of the LSTM architecture. BiLSTM will process data sequences in 2 directions [13] i.e. forward (from past to future) and backward (from future to past) [14] with two separate hidden layers and the process is in the opposite direction, namely the forward layer and backward layer [15]. Then the processing results from the two layers can be merged in several ways such as average, sum, multiplication, and concatenation [12].

The LSTM-SVM classifier in this architecture contains 4 layers such as input layer, hidden layer, fully connected layer, and softmax layer. Initially, the extracted data will be entered into the LSTM input layer and then forwarded to the hidden layer. In LSTM, the gated cell or gated unit is part of the hidden layer of the LSTM. The hidden layer in LSTM consists of 4 layers. Each layer will interact with one another to produce an output cell known as the cell state. After that, the output will be passed to the next hidden layer. LSTM has 1 layer which will be activated by tanh and 3 gates which will be activated by sigmoid function. After the hidden layer process is complete, the output will be forwarded to the fully connected layer. Each node of the fully connected layer learns its own set of weights on all nodes in the layer. Then the output from the fully connected layer will be forwarded to the softmax layer [16].

In general, the softmax layer is the last layer of the LSTM network. The proposed model architecture is LSTM with the difference in using the SVM algorithm for classification [16]. The basic idea behind using SVM is to implement a supervised learning algorithm that makes it easy to find the optimal hyperplane of the feature space to differentiate classes. The resulting feature vector will be used as input to the SVM classifier for classification [6].

C. Related Works

In the study by Jihen Fourarti, Mohamed Othmani, and Hela Ltifi (2022) entitled "A hybrid model based on bidirectional long-short term memory and support vector machine for rest tremor classification", the classification of tremor which is the main symptom of Parkinson's disease was carried out using the algorithm BiLSTM-SVM. The research begins by preprocessing the input in the form of normalizing all data in millimeters per second because the dataset file varies in units of tremor speed [6].

After that, the model is formed using the BiLSTM-SVM algorithm. The input given will go through 2 hidden Bidirectional LSTM layers with 256 nodes for the first BiLSTM layer followed by 128 nodes for the second BiLSTM layer, then followed by 1 hidden dropout layer with a rate of 0.2 to avoid overfitting in the dense layer and the last layer of the model is substituted by linear SVM to carry out the classification process [6].

The training and testing process is carried out by distributing 70% of training data and 30% of testing data. In the BiLSTM-SVM model that was formed, the Adam optimizer and the loss hinge function were used plus L2 Regularization. The model was trained for 30 epochs with a batch size of 64 [6].

The tremor classification test using the BiLSTM-SVM model using 2 different datasets resulted in an accuracy of 98.93% in the first dataset (Physionet dataset) and 97.22% in the second dataset (John Radcliffe Hospital dataset) [6].

METHODS

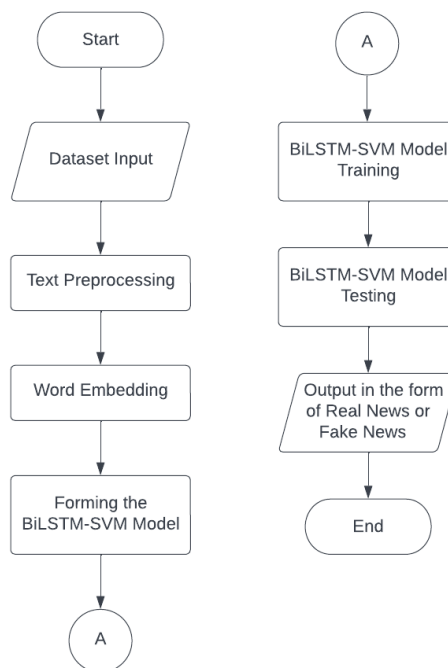Methods in this study which include training and testing of the dataset follow the flow in Figure 2 below:

Figure 2. Training and Testing Dataset Flowchart

A. Dataset Input

The process in this research begins with input the news dataset. The news dataset discusses the topic of Covid-19 and comes from GitHub uploaded by Sourya Dipta Das. The dataset in the form of training data and testing data that will be used has the totals of 6,420 data and its stored in a CSV file with the file name Constraint_Train.csv.

B. Text Preprocessing

In the developed system, there are several steps of text preprocessing that are applied, which consist of:
1. Removing URL, is the step of removing the URL element in the text [17].
2. Removing HTML tags, is the step of removing the HTML element tag in the text [18].

3. The stage of tidying sentences separated by periods without spaces.
4. Removing contractions, is the step of tidying up the abbreviated standard word in the text [19].
5. Cleaning non-ASCII characters, is the step of removing non-ASCII characters in the text.
6. Removing punctuation, is the step of eliminating elements that could reduce or misinterpret the meaning of the sentence in the text. These elements are in the form of punctuation marks in sentences [19].
7. Decompounding and spell correction, is the step of tidying up words that do not have spaces and correcting words that have typos [20].
8. Case folding, is the step of converting all capital letters in the text to lowercase [13].
9. Replacing numbers, is the stage of replacing numeric numbers with word numbers in the text.
10. Stopword removal, is the stage of removing stopword words that contain low information from the text so that readers can focus on only important words. Some examples of stopwords in English are an, been, the [21].
11. Lemmatization, is the step of remapping words that have changed word form into the basic form of the word [22].
12. Tokenization, is the step of breaking sentences one by one into a list of words [23].

C. Word Embedding

In this study, the word embedding used is Word2Vec with a corpus formed based on the existing dataset, then produces a vector representation of each word in the text corpus as output [24]. Word2Vec was chosen because it is rated better than other models such as CountVectorizer, FastText, and Glove [25]. In this study, the Word2Vec architecture used is Continuous Bag of Words (CBOW) with the word vector dimensions chosen to represent each word being 200 [26], [27]. Training on the Word2Vec model is carried out with the help of the Gensim library. Each length of the text in the news will be generalized so that it can be processed by the BiLSTM model. Each news will be padded with a size of 1000 words. 1000 words were chosen because the latest news articles from famous media have an average length of about 1000 words [28]. Then an array with a 3-dimensional size will be formed with text that has 1000 words with each word having 200 types of numbers called a weighted matrix, a matrix that has been weighted based on words.

D. Forming The BiLSTM-SVM Model

The BiLSTM-SVM model was built with the help of Keras from TensorFlow, starting from the formation of the Sequential model.

In the Sequential model that has been formed, an embedding layer is added to accommodate data from the corpus with "Embedding()", which will process the input in the form of a weighted matrix. In the next layer, a BiLSTM layer is added to perform the BiLSTM algorithm with "Bidirectional(LSTM())" which uses the merge concatenate method. The next layer is added with a dropout layer which serves to reduce the possibility of overfitting with "dropout(0.5)". Then a Dense layer is added which acts as a fully connected layer that uses the ReLU activation function which aims to make the trained and formed model able to learn non-linear dependencies. The last layer which is the layer as a classification is done with "Dense()" with 1 node using an activation function of Tanh because the activation function of Tanh can produce outputs that are in the range of -1 to 1.

The Sequential Model uses the hinge loss function which aims to represent the SVM layer [29] as a classification layer using Adam optimizer with a learning rate of 0.001 [30].

E. BiLSTM-SVM Model Training

After the model has been formed, training can be carried out to train the model in finding correlations or studying the data provided. The data from the Constraint_Train.csv file will be divided in a ratio of 0.8:0.2 where 80% of the data will be used as training data and the remaining 20% will be used as testing data. 80% of the training data that has passed the text preprocessing and word embedding process will be entered into the model that has been designed to be trained. The training process is carried out for 30 epochs with a batch size of 30 [30].

F. BiLSTM-SVM Model Testing

After the training process is complete, the model can be tested by entering the news text from the Constraint_Train.csv file as much as 20% to see the accuracy and test the performance of the model. In testing the model, the output will be obtained in the form of a classification result, namely "Real News /

Hoax" which is represented by 1 and -1 labels. After the testing process is carried out, the output of the testing data will be categorized into 4 conditions and processed using a confusion matrix to evaluate the proposed model.

RESULT AND DISCUSSION
In this study, the tests carried out are as follows:
1. Testing the data testing using 1,284 data or equivalent to 20% of the total training data and testing data provided. In testing the testing data, a threshold value test scenario will also be carried out starting from 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9. After testing the threshold value for 9 scenarios, it was found that the highest accuracy was found in the test with a threshold value of 0.3. The summary of the values of accuracy, precision, recall, and F1 Score from the threshold test can be seen in Table 1 below.

Table 1. Accuracy, Precision, Recall, and F1 Score Values from Threshold Scenario Testing

Threshold	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
0.1	88.55	89.29	88.35	88.82
0.2	88.7	89.57	88.35	88.95
0.3	88.78	89.83	88.2	89
0.4	88.7	89.93	87.89	88.9
0.5	88.55	89.9	87.59	88.73
0.6	88.47	90.01	87.29	88.63
0.7	88.39	90.12	86.98	88.52
0.8	88.47	90.39	86.83	88.58
0.9	88.47	91.04	86.08	88.49

The length of time required to perform 1 test scenario of the threshold value is about 13 seconds with an average of 10 milliseconds for each data.

2. Testing new data apart from training data and data testing, namely data validation by using 100 pieces of news with the same topic. The new data test will use the best threshold value from the threshold value test scenario that has been carried out when testing data testing. The length of time it takes to test 100 pieces of data is about 1 second with an average of 11 milliseconds for each data. From the tests carried out, the confusion matrix can be obtained as follows:

Table 2. Confusion Matrix of BiLSTM-SVM Model for Testing with New Data

		Actual Class	
		Positive	Negative
Predicted Class	**Positive**	49	6
	Negative	8	37

From Table 2, the following categories are obtained:
a. 49 data are categorized as True Positive (TP), namely real news that is detected as real news.
b. 6 data are categorized as False Positive (FP), namely fake news that is detected as real news.
c. 8 data are categorized as False Negative (FN), namely real news that is detected as fake news.
d. 37 data are categorized as True Negative (TN), namely fake news that is detected as fake news.

After obtaining the confusion matrix from the BiLSTM-SVM model with new data, it can be calculated to get the values of accuracy, precision, recall, and F1 Score.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} = \frac{49 + 37}{49 + 6 + 8 + 37} = \frac{86}{100} = 0.86 \times 100\% = 86\%$$

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{49}{49 + 6} = \frac{49}{55} = 0.8909 \times 100\% = 89.09\%$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{49}{49 + 8} = \frac{49}{57} = 0.8596 \times 100\% = 85.96\%$$

$$\text{F1 Score} = 2 \times \frac{(\text{recall} \times \text{precision})}{(\text{recall} + \text{precision})} = 2 \times \frac{(0.8596 \times 0.8909)}{(0.8596 + 0.8909)} = 0.875 \times 100\% = 87.5\%$$

The results of the calculations above will be presented in Table 3.

Table 3. Calculation Results of Accuracy, Precision, Recall, F1 Score for Testing with New Data

Metrics	Value (%)
Accuracy	86%
Precision	89.09%
Recall	85.96%
F1 Score	87.5%

The process of testing the BiLSTM-SVM model using the new data is able to produce an accuracy value of 86%, a precision value of 89.09%, a recall value of 85.96%, and an F1 Score value of 87.5%.

CONCLUSION
The conclusion obtained from this study is the BiLSTM-SVM algorithm with a threshold = 0.3 can detect fake news or not from testing data with an accuracy rate of 88.78% and an F1 Score value of 89%. The BiLSTM-SVM algorithm with a threshold = 0.3 also can detect fake news or not from data validation with the same news topic with an accuracy rate of 86% and an F1 Score value of 87.5%.

FUTURE WORKS
Suggestions that can be given and that might be helpful for further development on the same topic is that an algorithm can be developed for text preprocessing which will replace the stopword removal section that removes context from the news by using Sentiment Analysis which treats news according to the existing context.

REFERENCES
[1]   C. Juditha, "Interaksi Komunikasi Hoax di Media Sosial Serta Antisipasinya," *J. Pekommas*, vol. 3, no. 1, pp. 31–34, 2018.

[2]   D. Susilawati and W. D. Putri, "Begini Dampak Berita Hoax | Republika Online," 2017. https://www.republika.co.id/berita/trendtek/internet/17/04/11/oo7uxj359-begini-dampak-berita-hoax (accessed Nov. 02, 2021).

[3]   Jasmi, "Mengenal Ciri-ciri Berita Hoax dan Dampaknya | QuBisa," 2021. https://www.qubisa.com/article/ciri-ciri-dan-dampak-hoax#showContent (accessed Feb. 25, 2022).

[4]   B. Agung, "Saat Hoaks yang Dibuat Makin Canggih, Kecerdasan Buatan Jadi Harapan Melawannya," 2019. https://dailysocial.id/post/saat-hoaks-yang-dibuat-makin-canggih-kecerdasan-buatan-jadi-harapan-melawannya (accessed Nov. 07, 2021).

[5]   K. Barkved, "How To Know if Your Machine Learning Model Has Good Performance | Obviously AI," 2022. https://www.obviously.ai/post/machine-learning-model-performance (accessed Aug. 05, 2022).

[6]   J. Fourati, M. Othmani, and H. Ltifi, "A hybrid model based on bidirectional long-short term memory and support vector machine for rest tremor classification," *Signal, Image Video Process.*, no. March, 2022, doi: 10.1007/s11760-022-02180-9.

[7]   P. S. Reddy, D. Elizabeth Roy, P. Manoj, M. Keerthana, and P. V. Tijare, "A study on fake news detection using naïve bayes, SVM, neural networks and LSTM," *J. Adv. Res. Dyn. Control Syst.*, vol. 11, no. 6 Special Issue, pp. 942–947, 2019.

[8]   A. Y. Triartanto, "Kredibilitas Teks Hoax Di Media Siber," *J. Komun.*, vol. 6, no. 2, pp. 33–36, 2015, [Online]. Available: http://ejournal.bsi.ac.id/ejurnal/index.php/jkom/article/view/2464/1700.

[9]   M. Zulfikar, "Hoaks Masih Laku di Pasaran, Berikut 4 Ciri Hoaks Menurut Kominfo - National Geographic," 2019. https://nationalgeographic.grid.id/read/131824796/hoaks-masih-laku-di-pasaran-berikut-4-ciri-hoaks-menurut-kominfo (accessed Mar. 25, 2022).

[10]   A. Cimino and F. Dell'Orletta, "Tandem LSTM-SVM approach for sentiment analysis," *CEUR Workshop Proc.*, vol. 1749, 2016, doi: 10.4000/books.aaccademia.2003.

[11]   A. Graves, S. Fernández, and J. Schmidhuber, "Bidirectional LSTM networks for improved phoneme classification and recognition," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 3697 LNCS, pp. 799–804, 2005, doi: 10.1007/11550907_126.

[12]   E. Zvornicanin, "Differences Between Bidirectional and Unidirectional LSTM | Baeldung on Computer Science," 2022. https://www.baeldung.com/cs/bidirectional-vs-unidirectional-lstm (accessed May 21, 2022).

[13]   R. Dwi, W. Santosa, M. A. Bijaksana, and A. Romadhony, "Implementasi Algoritma Long Short-Term Memory (LSTM) untuk Mendeteksi Penggunaan Kalimat Abusive Pada Teks Bahasa Indonesia," *J. Tugas Akhir Fak. Inform.*, vol. 8, no. 1, pp. 691–702, 2021.

[14]   R. Aggarwal, "Bi-LSTM. What is a neural network? Just like our… | by Raghav Aggarwal | Medium," 2019. https://medium.com/@raghavaggarwal0089/bi-lstm-bc3d68da8bd0 (accessed Mar. 29, 2022).

[15]   Z. Cui, R. Ke, Z. Pu, and Y. Wang, "Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction," pp. 1–11, 2018, [Online]. Available: http://arxiv.org/abs/1801.02143.

[16]   N. S. Kiruthika and G. Thailambal, "Dynamic Light Weight Recommendation System for Social Networking Analysis Using a Hybrid LSTM-SVM Classifier Algorithm," *Opt. Mem. Neural Networks*, vol. 31, no. 1, pp. 59–75, 2022, doi: 10.3103/s1060992x2201009x.

[17]   U. Verma, "Text Preprocessing for NLP (Natural Language Processing),Beginners to Master | by Ujjawal Verma | Analytics Vidhya | Medium," 2020. https://medium.com/analytics-vidhya/text-preprocessing-for-nlp-natural-language-processing-beginners-to-master-fd82dfecf95 (accessed May 21, 2022).

[18]   D. Yadav, "NLP: Building Text Cleanup and PreProcessing Pipeline | by Dinesh Yadav | Towards Data Science," 2020. https://towardsdatascience.com/nlp-building-text-cleanup-and-preprocessing-pipeline-eba4095245a0 (accessed Apr. 29, 2022).

[19]   K. Rastogi, "Text Cleaning Methods in NLP - Analytics Vidhya," 2022. https://www.analyticsvidhya.com/blog/2022/01/text-cleaning-methods-in-nlp/ (accessed Apr. 27, 2022).

[20]   W. Garbe, "1000x faster Spelling Correction. Sub-millisecond compound aware… | by Wolf Garbe | Towards Data Science," 2017. https://towardsdatascience.com/symspellcompound-10ec8f467c9b (accessed Aug. 23, 2022).

[21]   A. Fardeen, "11 Techniques of Text Preprocessing Using NLTK in Python - MLK - Machine Learning Knowledge," 2021. https://machinelearningknowledge.ai/11-techniques-of-text-preprocessing-using-nltk-in-python/ (accessed May 21, 2022).

[22]   K. Ganesan, "All you need to know about text preprocessing for NLP and Machine Learning - KDnuggets," 2019. https://www.kdnuggets.com/2019/04/text-preprocessing-nlp-machine-learning.html (accessed Apr. 27, 2022).

[23]   S. Chakravarthy, "Tokenization for Natural Language Processing | by Srinivas Chakravarthy | Towards Data Science," 2020. https://towardsdatascience.com/tokenization-for-natural-language-processing-a179a891bad4 (accessed Jul. 24, 2022).

[24]   R. Fauzi Rahmadzani, "Cara Kerja Word Embedding Word2Vec | Catatan Penelitian #5," 2020. https://rifqifai.com/cara-kerja-word-embedding-word2vec/ (accessed May 05, 2022).

[25]   B. Ajose-Ismail, … O. A.-I., and undefined 2020, "Performance Analysis of Different Word Embedding Models for Text Classification," *Eprints.Federalpolyilaro.Edu.Ng*, vol. 3, no. 6, pp. 1016–1020, 2020, [Online]. Available: http://eprints.federalpolyilaro.edu.ng/1650/.

[26]   S. Tazzyman, "Neural Network models | NLP-guidance." https://moj-analytical-services.github.io/NLP-guidance/NNmodels.html (accessed May 24, 2022).

[27]   W. Widayat, "Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning," *J. Media Inform. Budidarma*, vol. 5, no. 3, p. 1018, 2021, doi: 10.30865/mib.v5i3.3111.

[28]   M. Tauberg, "How Smart is Your News Source?. Text Data Analysis of 21 Different News… | by Michael Tauberg | Towards Data Science," 2019. https://towardsdatascience.com/how-smart-is-your-news-source-1fe0c550c7d9 (accessed May 24, 2022).

[29]   J. Brownlee, "How to Choose Loss Functions When Training Deep Learning Neural Networks," 2019. https://machinelearningmastery.com/how-to-choose-loss-functions-when-training-deep-learning-neural-networks/ (accessed Jul. 03, 2022).

[30]   O. Alharbi, "A Deep Learning Approach Combining CNN and Bi-LSTM with SVM Classifier for Arabic Sentiment Analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, pp. 165–172, 2021, doi: 10.14569/IJACSA.2021.0120618.