ASEAN Engineering Journal

Full Paper

Detection Model for Fake News on COVID-19 in Indonesia

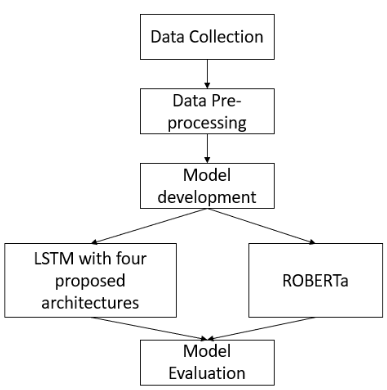
Article history
Received
16 December 2022
Received in revised form
17 May 2023
Accepted
27 May 2023
Published online
30 November 2023

Achmad Pratama Rifai[a], Yun Prihantina Mulyani[a*], Rian Febrianto[a], Hilya Mudrika Arini[a], Titis Wijayanto[a], Nurul Lathifah[b], Xiao Liu[c], Jianxin Li[c], Hui Yin[c], Yutao Wu[c], Rami Mohawesh[c]

[a]Department of Mechanical and Industrial Engineering, Universitas Gadjah Mada, Indonesia
[b]Department of Industrial Engineering, Universitas Indonesia, Indonesia
[c]School of Information Technology, Deakin University, Australia

*Corresponding author
yun.prihantina@ugm.ac.id

Graphical abstract



Abstract

Today, fake information has become a significant problem, exacerbated by the acceleration of access to information. The spread of fake information has a dangerous impact, especially regarding global health issues, for example COVID-19. People can access various resources to obtain information, including online sites and social media. One of the methods to control the spread of false information is detecting hoaxes. Many methods have been developed to identify hoaxes; most previous studies have focused on developing hoax detection methods using data from a single source in English. The present study is carried out to detect fake news in Indonesian language using multiple data sources, including traditional and social media in the context of COVID-19. The study uses Long Short-Term Memory (LSTM) and the Robustly Optimised Bidirectional Encoder Representations from Transformers Pre-Training Approach (RoBERTa). The LSTM approach is used to develop four different architectures that varied based on: (1) the use of text-only versus the use of both title and text; (2) the number of LSTM and dense layers; and (3) the activation function. The LSTM model with text-only data, a single LSTM layer and two dense layers, outperformed other LSTM architectures, achieving the highest accuracy of 92.17%. The LSTM models require a considerably short training time of 23–27 minutes for 3,847 articles and has a detection time of 3.8–4.1 ms per article. The RoBERTa classifiers outperformed all LSTM models with an accuracy of over 97% and a significantly better training time, with a margin of more than 50% compared to LSTM classifiers, although it had a slightly longer test time. Both LSTM and RoBERTa models outperformed the Naïve Bayes and SVM benchmark methods in terms of accuracy, precision, and recall. Therefore, this study shows that both LSTM and RoBERTa methods are reliable and can be reasonably implemented for real-time fake news detection.

Keywords: Fake news detection, COVID-19 misinformation, fake news in Indonesian, machine learning, LSTM, RoBERTa

1.0 INTRODUCTION

Information cannot be separated from every aspect of human life. Today, people can quickly access information from various different sources, including newspapers, television, radio, news portals and websites, and social media platforms such as Twitter, Facebook, Instagram, and others. Nevertheless, some of this information is false or part of a hoax. Several studies have focused on anticipating the emergence of false or hoax information on the internet, for instance in Arabia [1], Indonesia [2,3], India [21], and other countries. It has become a particular concern for researchers to prevent readers from being easily deceived by false information.

False or hoax information can be found circulating in several media, including news portal websites and social media. Therefore, researchers have focused on the detection of hoaxes in these two media sources. For example, Aldwairi and Alwahedi [4] researched false information on websites and clickbait on social media (Facebook, Forex, and Reddit). The experimental results showed that a logistic classifier could yield 99.4% accuracy in detecting false or hoax information. Other social media platforms, such as Twitter and YouTube, have also been studied to predict their association with false information. For example, a study by Dhawan et al. [5] showed that the 'likes' ratio has a predictive accuracy value of 92% and 37% for YouTube and Twitter, respectively. When using the engagement rate as a measure, an accuracy of 86% for YouTube and 41% for Twitter was recorded.

Other platforms, such as news portal websites, are also used to identify false information. For instance, Bahad et al. [6] used a dataset from Kaggle containing real news and fake news from various news portal websites. Their experiments showed that the Bidirectional LSTM-RNN model can produce validation accuracy of between 89.74% and 98.25%. Similar research has also been conducted using a dataset from Kaggle, of as many as 18,285 news articles consisting of real and fake news [7]. The results showed that LSTM was the best classifier, yielding an accuracy of 91.05% for detecting hoaxes. In addition, Lin et al. [8] analysed 16 news websites in the Urdu language containing real and fake news on topics related to business, health, showbiz, sports, and technology. The research used several RoBERTa models which achieved accuracy rates from 89% to 90%. Furthermore, Samadi et al. [9] implemented RoBERTa methods as classifiers on a COVID-19 news dataset with binary classification and achieved an accuracy rate of over 97%. Therefore, it is clear that various classifiers, including LSTM and RoBERTa, can be used in fake news detection with a high rate of accuracy.

In addition to using a single source, as has been done by several studies as mentioned above, several studies have used multiple sources to detect hoaxes. For example, Davoudi et al. [10] used a dataset called FakeNewsNet, which consisted of several tweets obtained from Twitter. The researchers applied several classification models, and the proposed method – a hybrid deep model – showed the best results with an accuracy of 98.4% for detecting hoaxes from portal websites and social media. Deepak and Chitturi [11] identified false information using the Feedforward Neural Network (FNN) and LSTM, utilising the George McIntires Fake News Dataset, which consists of various sources including website portals and social media. The research found that LSTM was the better method for detecting hoaxes, with an accuracy of 91.32%.

Existing studies generally focus on information presented in English, but there has been some previous research related to hoax detection in Bahasa. For example, Nayoga et al. [2] identified false information in Bahasa within general news themes. The study collected 1,000 pieces of data, all of which came from news portals in Indonesia. Of the seven methods used, the 1D-Convolutional Neural Network method was able to predict hoax news in Indonesia with an accuracy of approximately 97%. Similar research has also used website portal sources, such as cnnindonesian.com, cekfakta.com, tunbackhoax.id, and others [12,13]. Hoax information in Bahasa can be detected using LSTM [12] and SGD modified-hurber [13], achieving f1-score values of 80.7% and 86% accuracy, respectively.

False information that appears specifically in Indonesia, such as from Twitter, has also been studied [14]. One study focused on false information using hashtag-based keywords that became trending topics in Indonesia, such as government issues, COVID-19, and natural disasters such as floods. The detection results using Neural Networks based on TF-IDF (Term Frequency-Inverse Document Frequency) succeeded in identifying false information, with an accuracy of 78.76%. Essential information was obtained from the study by Kencana et al. [14]. The detection of false information can also be used to identify fake news about COVID-19, which has been a 'hot topic' for the past few years. Several studies have identified the problem of false information related to COVID-19; thus, the detection of false information on this topic should be investigated further.

The K-Nearest Neighbour (KNN) method, based on the Jaccard Space, has previously been used to detect false information about COVID-19 in Indonesia [3]. The information was obtained from traditional media sources, Jabar Saber Hoaks and Jala Hoaks. The study identified false information about COVID-19 with an accuracy of 75.89%. Another study has focused on identifying false information about COVID-19 in English, using 2,084 URLs [15]. Of the various classifier methods used, the research found that the Naïve Bayes method, based on BoW (Bag of Words), could predict false information with an accuracy value of 96%.

In addition to single source studies, several studies have used multiple sources to identify false or hoax information related to COVID-19. Khan et al. [16] identified hoax news about COVID-19 sourced from social media (Facebook and Twitter) and several website portals, in English. Of the four algorithms, the Random Forest classifier was the best algorithm for predicting hoax information related to COVID-19, with an accuracy of 88.50%. However, the number of studies that use multiple sources to identify false information remains limited. The previous study frequently used a single source to identify false information, especially news related to COVID-19. However, no research has been found that identifies fake news related to COVID-19 in Indonesia using multiple sources [16].

False or hoax information has proven very dangerous [17]. Therefore, there is a need for further research on detect false or hoax information related to COVID-19, especially in Bahasa, where the research is limited. In addition, this research should focus on using multiple sources, consisting of traditional media (i.e., news portals and government websites) and social media (i.e., Twitter and Facebook). The reason for using multiple sources is because social media is a common way to get information: 45% of the world's population spends 2 hours and 23 minutes each day using social media [18]. Thus, this study focused on detecting hoaxes and false information related to COVID-19 in Indonesia by using both traditional and social media sources. The algorithms used to detect hoaxes were Long Short-Term Memory (LSTM) and Robustly Optimised Bidirectional Encoder Representations from Transformers Pretraining Approach (RoBERTa). The reason for selecting LSTM and RoBERTa was that they have been proven to be effective in several similar studies and shown a reasonably high rate of accuracy in detecting false information.

Finally, this paper is structured as follows. Section 2 presents the methodology to develop detection models. Section 3 provides the experimental results and discussion of the comparison between the methods and past literatures in detail. Finally, Section 4 provides the conclusions of this study.

2.0 METHODOLOGY

This section describes the methodology in detail. The methodology of this study consisted of four stages: data collection, data pre-processing, model development, and evaluation.

2.1. Data Collection

The first stage was data collection. The experiment was initiated by collecting data, consisting of hoax news and real news articles from various sources. Certain sources were deliberately selected, coming both from traditional media (e.g., Indonesian government websites and news portals websites in Indonesia) and social media (e.g., Twitter and Facebook). The dataset in this study was in the Bahasa language.

The data collection was carried out using Python Scraper (Tweepy and Google News). We focused on articles about COVID-19 in Indonesia, spanning the two years from January 2020 through April 2022. The index used in the dataset consisted of the publication date, news title, full-text news, and news links. All news articles collected were original, as written by each writer, nothing was changed or modified.

2.2. Data Pre-Processing

The second stage was data pre-processing. Data pre-processing was carried out as the initial step of hoax detection. A new index consisting of the ID, title, author, text, and label columns was created for the collected datasets. We took the title and text columns from the entire index; thus, any missing values for the title were replaced with text, and vice versa. Therefore, there were no missing values in the existing dataset for the training sets.

The next stage was text pre-processing. This process focused on replacing punctuation, lowercase letters, split-by words, stemming, and removing stop words. The next stage was to carry out the 'one-hot word representation' process and sequence creation with a maximum sentence length of 50 and 1,000 words for the title and text column, respectively. This stage is essential for creating the correct form of dataset used in neural networks.

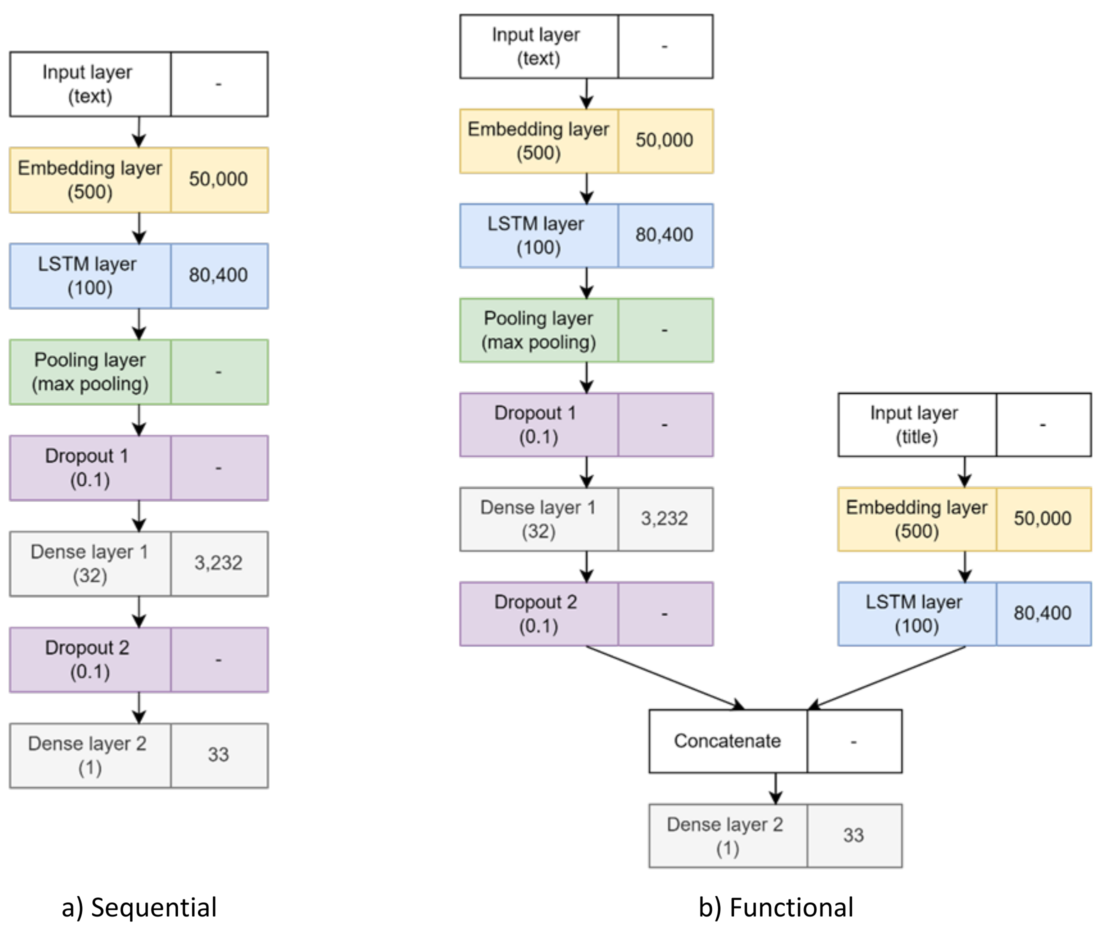


Figure 1 Architecture of LSTM models

2.3. Classification using LSTM

Following the data and text pre-processing, the next stage was to build an architectural model for LSTM. LSTM is a type of recurrent neural network (RNN) architecture that is specifically designed to address the vanishing gradient problem, which is a challenge faced by traditional RNNs. LSTM networks were introduced by Sepp Hochreiter and Jürgen Schmidhuber [19] to address the vanishing gradient problem. The LSTM architecture includes specialized memory cells that allow the network to selectively retain or forget information over long sequences. It achieves this through the use of three gating mechanisms:

- Forget Gate: Determines which information from the previous cell state should be forgotten.
- Input Gate: Determines which new information from the current input should be stored in the cell state.
- Output Gate: Determines which part of the cell state should be output as the current hidden state.

These gates, controlled by learnable sigmoid activation functions, enable LSTM networks to remember or forget information over long time steps, effectively capturing long-term dependencies. The cell state acts as a conveyor belt, carrying information through time while the gates regulate the flow of information.

As in RNN, the LSTM networks consist of several layers. The first layer used was embedding. This layer was initialised with random weights and focused on learning the embedding for all words in the training datasets. Then, the LSTM layer was applied as the second layer. The next layer was the max pooling layer, for downsampling the input representation. Then, the dense layers were added. This layer resembles the ordinary hidden layer of the artificial neural network. Then, the dropout layer was introduced to prevent overfitting. At the end of the network, a dense layer with one neuron and a sigmoid activation function acted as the output layer for binary classification.

We propose four different architectures of LSTM models in this study. The first and second models are sequential models. The first model uses two dense layers, one with a rectified linear unit (ReLU) and the other with sigmoid activation. The second model uses three dense layers, two with ReLU and one with sigmoid activation. Both the first and second models use an embedding vector feature of 100.

The third and fourth models use combinations of the title and text; these are functional models. We propose a multi-input model which merges two previous architectures. These models are close to the first and second models. Similar to the first model, the third model uses two dense layers, one with single ReLU and the other with sigmoid activation. Correspondingly, the fourth model has a similar number of dense layers as the second model, with two dense layers with ReLU and a single dense layer with sigmoid activation. Figure 1 illustrates the architecture, both for sequential models and functional models.

The column on the left indicates the type of layers including their properties – i.e., dimensionality of the output space for the LSTM layer, the number of neurons for a dense layer, the drop rate for the dropout layer. The right-hand column indicates the number of parameters in the layer, which are the weight and bias.

Table 1 Configuration of the LSTM models

Model	Input Data	Maximum sentence length	No. of LSTM layers	No. of dense layers	No. of parameters
1	Text	1,000	1	2	133,665
2	Text	1,000	1	3	138,977
3	Title, Text	50 (title), 1,000 (text)	2	2	221,001
4	Title, Text	50 (title), 1,000 (text)	2	3	232,889

The distinction between functional models and sequential models is in the number of embedding feature vectors. Functional models use embedding vector features of 50 and 1,000 for title and text, respectively. All models use binary cross-entropy as the loss function and Adam as the optimiser during the learning process. Table 1 summarises the configuration details of the four models.

2.4. Classification using RoBERTa

Robustly Optimised Bidirectional Encoder Representations from Transformers Pre-Training Approach (RoBERTa) is a natural language processing (NLP) model that is based on the Transformer architecture, introduced by Liu et all. [20] with the aim of further refining the pretraining process and achieving better performance on various NLP tasks. It is an extension and improvement upon the popular BERT (Bidirectional Encoder Representations from Transformers) model.

The model was trained on a large corpus of publicly available text from the internet, similar to BERT. However, compared to BERT, RoBERTa was trained with a larger number of training steps, a larger batch size, and more diverse data. In addition, RoBERTa is trained with full sentences without Next Sentence Prediction loss, large mini-batches, and a larger byte-level Byte-Pair Encoding [21], thus improving the performance of its predecessor, BERT. It also incorporates other optimization techniques, such as dynamic masking during training, which helps it to better understand the context and relationships within sentences.

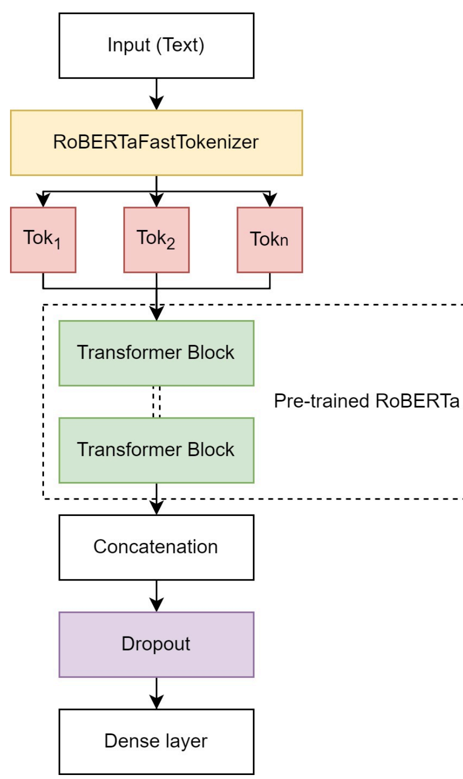


Figure 2 Framework of RoBERTa model

RoBERTa is designed to learn rich representations of text by training a deep neural network on large-scale datasets. It learns to predict masked words in sentences and perform various other auxiliary tasks during pretraining. After pretraining, the model can be fine-tuned on specific downstream tasks, such as text classification, named entity recognition, question answering, and more.

RoBERTa has achieved state-of-the-art results on several NLP benchmarks and tasks, demonstrating its effectiveness in understanding and generating human language. Its architecture and training methodology have served as the basis for further advancements in the field of transformer-based models for NLP.

In this study, RoBERTa models provided sequence output created by pre-trained models for an input dataset (text). The framework is illustrated in Figure 2. Then, RoBERTa models with simple transformer classification were used as classifiers to detect fake news datasets.

The next step utilises simple transformer classification and provides the training arguments of RoBERTa models. Table 2 summarises the details of the RoBERTa training arguments used in this study.

Table 2 Training Arguments of RoBERTa

No.	Training Arguments	Input Data
1	max_seq_length	300
2	train_batch_size	32
3	gradient_accumulation_steps	1
4	evaluate_during_training	True
5	evaluate_during_training_steps	256
6	num_train_epochs	15
7	use_early_stopping	True

3.0 EXPERIMENT AND RESULTS

This section evaluates and discusses the results of the experiment. First, the dataset used to build the models is described. Then, the models are trained and evaluated using that dataset. All coding used to build the detection model for fake news in the context of COVID-19 in Indonesia was written using Python 3.8.8 in Jupyter Lab 1.1.4. The algorithm used in this study is LSTM with four different architectures and RoBERTa. The hoax detection was run on an Intel® CoreTM i5-10400F CPU @ 2.90GHz (12 CPUs) with 16 GB RAM and Nvidia GeForce GTX1650 Super.

This research focused on news related to COVID-19 in Indonesia, so both real and hoax news items were limited to only those about COVID-19. The Indonesian government has verified the hoax news from various media, confirmed via two website portals: turnbackhoax.id and covid19.go.id. Therefore, in addition to social media, fake news related to COVID-19 was collected from those sources. Meanwhile, the real news was collected from social media (55%), Indonesian government websites (31%), and the rest from credible news portals. The overall dataset collected consisted of 4,831 articles with a proportion of 65% real news and 35% fake news. Table 3 shows examples of the data and their corresponding labels, taken from traditional and social media news sources.

The data was divided into two sets, the training and test, with a ratio of 80% to 20%, respectively. This meant that there were 3,847 articles for training and 984 articles for testing, which were randomly selected. Furthermore, of the training data, 20% was used for hyperparameter tuning during the validation phase in each epoch. Each dataset contained articles from both traditional media and social media as well as their corresponding labels, either real or fake, in similar proportions to the whole dataset. It should be noted that the test dataset consisted of different news articles that had not been learned by the trained model before; thus, the test measures the models' ability to generalise other data. Naïve Bayes and Support Vector Machine (SVM) are used as the benchmark methods. The classification models of Naïve Bayes and SVM only utilize the text of the news without considering the title.

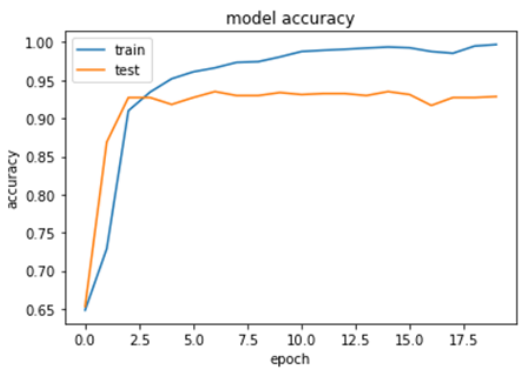
Table 3 Examples of the data and their corresponding labels

Text	Source	Label
RT @detikHealth: Luhut mengingatkan potensi lonjakan kasus COVID-19. Ada 29 persen kabupaten/kota yang mencatat infeksi COVID-19 kembali me…'	Twitter (social media)	Real
Sejumlah pemberitaan lokal UEA mengungkapkan bahwa Putra Mahkota Abu Dhabi tersebut didiagnosis positif virus corona dan saat ini tengah dikarantina………. Menurut sumber di Kementerian Kesehatan UEA mengungkapkan kepada Al-Tawil bahwa jumlah orang yang terinfeksi dua kali lebih banyak dari yang diumumkan oleh pemerintah.(suara.com)"	Beritaterheboh.com (traditional media – news portal)	Fake
Juru Bicara Pemerintah untuk Penanganan COVID-19 Prof. Wiku Adisasmito meminta masyarakat tak perlu mengkhawatirkan kelompok prioritas penerima vaksin………. "Sampai saat ini ilmuwan masih terus mengenali karakteristik penyebab virus baru ini sebagai dasar pengembangan vaksin," papar Prof. Wiku.	COVID19.go.id (traditional media – government website)	Real
Ada Satu Negara yg Pemerintah Tidak Sebutkan??? Mungkin Netizen bisa bantu Jawab Kenapa ya, Pemerintah gak menyebut negara itu??? ………. Mungkin Netizen bisa bantu sebut nama negaranya??? Suudzon aja, mungkin pemerintah kita lupa!!!!	Facebook (social media)	Fake

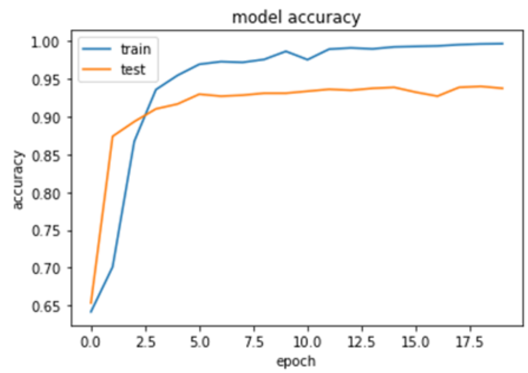
The hyperparameters of LSTM were set to be constant throughout the experiments to ensure a fair comparison of each model. The maximum epoch E was set at 20 epochs, while the batch size B was set at 128. Figure 3 presents the learning progress of the models. The training accuracy in each epoch is highlighted in blue, while the validation accuracy is highlighted in orange.

The four models recorded a similar pattern of learning progress. Figure 3 shows patterns of rapid training in early epochs, which then gradually improve after three epochs, when the training accuracy reaches the value of 0.95. Unlike the training accuracy progress, the validation accuracy in each epoch generally stabilised after three epochs, without any considerable improvement. Thus, there is a gap between the training and validation accuracy at the end of the training process. This indicates that there might be slight overfitting, which is likely caused by the limited amount of training data. Nevertheless, overall, the four models can achieve considerably good learning progress, even using limited data, indicated by the high training and validation accuracy.

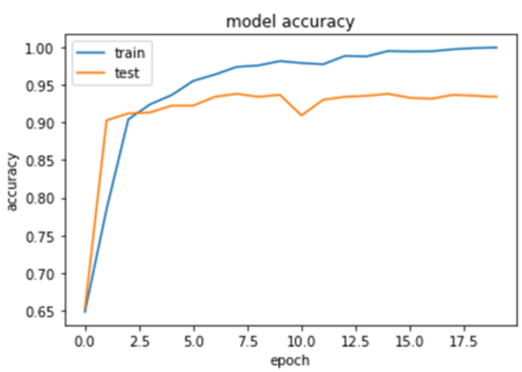
In addition, this study also used RoBERTa classifiers to detect fake news. In this method, the maximum epoch E was set at 15, while the batch size B was set at 32. Moreover, the maximum global steps were set at 1,024.



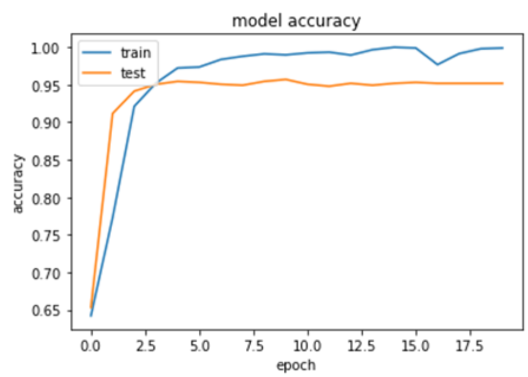
a) LSTM - Architecture 1



b) LSTM - Architecture 2



c) LSTM - Architecture 3



d) LSTM - Architecture 4

Figure 3 Training evaluation of LSTM models

Based on Figure 4, the model loss showed a good improvement for both train and eval loss. This indicates that RoBERTa yielded high accuracy and was more reliable than LSTM.

Figure 4 Model loss evaluation of RoBERTa

After the models had been trained in the LSTM classifier, the next process was testing using the test dataset. First, the news articles underwent data pre-processing with the same procedure as for the training dataset. Afterward, the tests were performed using the trained model in the test dataset with 984 articles.

In the confusion matrix, the True Positive is top-left, the False Positive is bottom-left, the False Negative is top-right, and the True Negative is bottom-right. Due to the negative impact of fake news on society, especially in the context of COVID-19, correctly identifying fake news is critical for the detection model. Therefore, the consequence of a False Negative (mistakenly identifying fake news as real news) is much more severe than a False Positive (mistakenly identifying real news as fake news). Out of the four models, Architecture 2 obtained the least number of False Negatives, at 39 out of 984 articles. Surprisingly, as they had a more complex network, considering the news article title, Architectures 3 and 4 performed worse than the sequential models.

Table 4 summarises the experiment results for the four LSTM models and RoBERTa. The results clearly indicate that the proposed LSTM and RoBERTa models outperform Naïve Bayes and SVM by a significant margin of accuracy. In addition, the proposed models also have favourable performance on precision and recall metrics. However, the proposed LSTM and RoBERTa models took longer computation time since they have relatively more complex architecture than the Naïve Bayes and SVM. Nevertheless, the test times obtained by the proposed models are considerably fast enough (within 3-5 ms per article) to be applied in real world for the hoax news classification.

Among the LSTM models, the best performance was achieved by Architecture 1, as indicated by the highest test accuracy. Architecture 2 followed, with a slight margin, while the two functional models were significantly outperformed with a more than 4% margin. Similar to the previous analysis regarding the number of False Negatives, the test accuracy further confirms that the sequential models perform favourably compared to the functional models. In addition, the value of precision and recall for the models also indicates the same pattern, to a more significant level, whereby the functional models were left behind with a 5–7% margin compared to the sequential models.

As Architectures 3 and 4 have better training accuracy, this confirms that there is overfitting in these two functional models, which was also evident in the learning progress. In addition, the poor performance of Architectures 3 and 4 can be partially attributed to the difficulties in identifying the fake or real news from the news titles in these datasets. The fake news often uses seemingly legitimate and reliable titles, which results in the models identifying this as real news. Thus, it is necessary to further develop a reliable model to identify the fake news using the news title.

The models required a relatively short training and test time. The training process with a dataset of 3,847 took 23–27 minutes, a considerably short amount of time. Although training time is generally overlooked when developing a detection model, it is important to develop an efficient model that requires only a short training time for fake news detection. This is especially the case in the context of COVID-19, where there might be new newsworthy topics in the future, such as new variants, new vaccinations, or new policies. Thus, the detection models might require re-training with additional data. A model with a short training time ensures that the periodic re-training process can be executed efficiently with a larger dataset. In terms of test time, Table 4 lists the overall test times for the test dataset with 984 articles, which translates to a detection time of 3.8–4.1 ms per article. Therefore, the detection models can be implemented for almost real-time identification as soon as news is published.

Furthermore, the testing results achieved using the RoBERTa method significantly outperformed the LSTM models. Based on the best evaluation from RoBERTa, we obtained a confusion matrix consisting of 617 True Positive, 324 True Negative, 12 False Positive, and 14 False Negative classifications for fake news detection. This also indicates that the RoBERTa method performs better than LSTM, especially in regard to accuracy, precision, recall, and training time. Thus, based on the results from the two classifiers, RoBERTa yielded higher accuracy than LSTM. In addition, RoBERTa also significantly outperformed LSTM in regard to training time, by a more than 50% margin.

Table 4 Summary of the results

Model	Training Accuracy	Test Accuracy	Precision	Recall	Training Time (s)	Test Time (s)
Naïve Bayes	0.5069	0.5376	0.41	0.85	0.048	0.0079
SVM	0.8175	0.7327	0.62	0.54	10.92	2.48
LSTM 1	0.9968	0.9217	0.89	0.89	1408	3.74
LSTM 2	0.9968	0.9156	0.87	0.88	1397	4.07
LSTM 3	0.9990	0.8770	0.83	0.80	1596	4.07
LSTM 4	0.9990	0.8790	0.81	0.83	1630	3.87
RoBERTa	0.9782	0.9750	0.97	0.96	700	5.66

The performance of the proposed models are further evaluated and compared with some recent fake news literatures, presented in Table 5. Noted that some of the literatures developing fake news detection for general news covering social, politic, business, and other topics.

The comparison results show that the proposed models achieved generally favourable performance, especially as compared to other models for detecting COVID-19 fake news. Moreover, Indonesian can be considered as a low-resource language in NLP, as compared to Chinese and English which are high-resource languages. It means that it has scarcer available datasets, which limiting the development of prediction model. Considering the accuracy obtained, especially those of RoBERTa model, the proposed models are promising to be implemented for detecting COVID-19 fake news.

Table 5 Comparison of the results with recent literatures

Authors	Topic	News sources	Language	Method	Test Accuracy
Aldwairi and Alwahedi [4]	General news	Social media	Arabic, English	logistic classifier	99.40%
Bahad et al. [6]	General news	Social media	English	Bidirectional LSTM-RNN	98.25%
Yesugade et al. [7]	General news	Social media	English	LSTM	91.05%
Davoudi et al. [10]	General news	Traditional and social media	English	LSTM, Propagation tree, stance network	98.40%
Deepak and Chitturi [11]	General news	Traditional and social media	English	LSTM	91.32%
Lin et al. [8]	General news	Traditional media	Urdu	CharCNN-RoBERTa	90.75%
Nayoga et al. [2]	General news	Traditional media	Indonesian	LSTM, 1D-CNSS	97.90%
Prasetijo et al. [13]	General news	Traditional media	Indonesian	SVM, SGD	86.00%
Mazzeo et al. [15]	COVID-19	Traditional media	English	Naïve Bayes	96%
Khan et al. [16]	COVID-19	Traditional media	English	Random Forest	88.50%.
Samadi et al. [9]	COVID-19 news	Social media	English	RoBERTa	97.43%
Abd Elaziz et al. [22]	COVID-19	Social media	Arabic	AraBERT, FHO	91.00%
Ma et al. [23]	COVID-19	Social media	Chinese	Dual Channel CNN	94.81%
Utami et al. [3]	COVID-19	Social media	Indonesian	kNN	75.89%
This research	COVID-19 news	Traditional and social media	Indonesian	LSTM	92.17%
				RoBERTa	97.50%

4.0 CONCLUSION

The present study developed LSTM and RoBERTa models to detect hoax and fake news from multiple data sources, including traditional and social media, in Indonesian in the specific context of COVID-19. The LSTM models were built from four different architectures, divided into two types: (1) sequential models, where the model uses only text data; (2) functional models, where the model uses text and title data. The architectures also varied in the number of layers and the type of activation functions. In addition, the RoBERTa models utilised sequence output and used simple transformers as classification. Naïve Bayes and SVM are used as benchmark methods to further assess the performance of the prediction models. Some particularly interesting findings from this study are as follows:

The best LSTM configuration, with an accuracy of 92.17%, was the sequential model with an LSTM layer and two dense layers, one with ReLU and another with sigmoid activation. The LSTM method requires a relatively short time of 23 minutes for training and 3.8 ms per article for detection.

Sequential models outperformed functional models, with an accuracy gap of 5–7%. This indicates that the functional models encountered difficulties in classifying fake or real news based on news titles. Fake news is likely to use legitimate and reliable sounding titles, even though the content of the news is fake or incorrect. Therefore, there is room for improvement to develop a more reliable model for detecting hoax and fake news from news titles alone.

The RoBERTa method outperformed other methods, with an accuracy of 97.50%, precision of 97%, and recall of 96%. In addition, this method requires just 12 minutes for training, representing a 50% time saving compared to the LSTM method. Considering that it only requires 5.66 ms for detection, the proposed prediction model using RoBERTa is reliable and can reasonably be implemented for real-time fake news detection.

Acknowledgement

This work is supported by the Study Melbourne Research Partnership program which has been made possible by funding from the Victorian Government through Study Melbourne (Project ID: veski-SMRP#1906).

References

[1] Medioni, G., Cohen, I., BreAˆ mond, F., Hongeng, S. and Nevatia, R. 2001. Event Detection and Analysis from Video Streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 23(8): 873-889. DOI: 10.1109/34.946990

[2] Nayoga, B. P., Adipradana, R., Suryadi, R., and Suhartono, D. 2021. Hoax Analyzer for Indonesian News Using Deep Learning Models, *Procedia Computer Science*, 179: 704-712. DOI: https://doi.org/10.1016/j.procs.2021.01.059

[3] Utami, E., Iskandar, A. F., Hidayat, W., Prasetyo, A. B., Hartanto, A. D. 2021. COVID-19 Hoax Detection Using KNN in Jaccard Space. *Indonesian Journal of Computing and Cybernetics Systems*, 15(3): 255-264. DOI: https://doi.org/10.22146/ijccs.67392

[4] Aldwairi, M., and Alwahedi, A. 2018. Detecting Fake News in Social Media Networks. *Procedia Computer Science*, 141: 215-222. DOI: https://doi.org/10.1016/j.procs.2018.10.171

[5] Dhawan, A., Bhalla, M., Arora, D., Kaushal, R., and Kumaraguru, P. 2022. FakeNewsIndia: A Benchmark Dataset of Fake News Incidents in India, Collection Methodology and Impact Assessment in Social Media, *Computer Communication*, 185: 130-141. DOI: https://doi.org/10.1016/j.comcom.2022.01.003

[6] Bahad, P., Saxena, P., and Kamal, R. 2019. Fake News Detection using Bi-directional LSTM-Recurrent Neural Network, *Procedia Computer Science*, 165: 74-82. DOI: https://doi.org/10.1016/j.procs.2020.01.072

[7] Yesugade, T., Kokate, S., Patil, S., Varma, R., and Pawar, S. 2021. Fake News Detection using LSTM, *International Research Journal of Engineering and Technology*, 8(4): 2500-2507.

[8] Lin, N., Fu, S., and Jiang, S. 2020. Fake News Detection in the Urdu Language using CharCNN-RoBERTa. In *Forum for Information Retrieval Evaluation 2020*

[9] Samadi, M., Mousavian, M., and Momtazi, S. 2021. Deep Contextualized Text Representation and Learning for Fake News Detection, *Information Processing and Management*. 58(6): 1-13. DOI: https://doi.org/10.1016/j.ipm.2021.102723

[10] Davoudi, M., Moosavi, M. R., and Sadreddini, M. H. 2022. DSS: A Hybrid Deep Model for Fake News Detection using Propagation Tree and Stance Network, *Expert Systems with Applications*, 198: 1-21. DOI: https://doi.org/10.1016/j.eswa.2022.116635

[11] Deepak S., and Chitturi, B. 2020. Deep Neural Approach to Fake-News Identification, *Procedia Computer Science*, 167: 2236-2242. DOI: https://doi.org/10.1016/j.procs.2020.03.276

[12] Apriliyanto, A., and Kusumaningrum, R. 2020. Hoax Detection in Indonesia Language using Long Short-Term Memory Model., *Sinergi*. 24(3): 189-196. DOI: 10.22441/sinergi.2020.3.003

[13] Prasetijo, A. B., Isnanto, R. R., Eridani, D., Soetrisno, Y. A. A., Arfan, M., and Sofwan, A. 2017. Hoax Detection System on Indonesian News Sites Based on Text Classification using SVM and SGD. In *4th International Conference on Information Technology, Computer, and Electrical Engineering*. 45-49. IEEE. DOI: 10.1109/ICITACEE.2017.8257673

[14] Kencana, C. W., Setiawan, E. B., and Kurniawan, I. 2020. Hoax Detection on Twitter using Feed-forward and Back-propagation Neural Networks Method, *Jurnal Resti,* 4(4): 648-654. DOI: 10.29207/resti.v4i4.2038

[15] Mazzeo, V., Rapisarda, A., and Giuffrida, G. 2021. Detection of Fake News on COVID-19 on Web Search Engines, *Frontiers in Physics*. 9: 1-14. DOI: 10.3389/fphy.2021.685730

[16] Khan, S., Hakak, S., Deepa, N., Prabadevi, B., Dev, K., and Trelova, S. 2022. Detecting COVID-19-Related Fake News Using Feature Extraction, *Frontiers in Public Health*. 9: 1-9. DOI: 10.3389/fpubh.2021.788074

[17] Bondielli, A., and Marcelloni, F. 2019. A Survey on Fake News and Rumour Detection Techniques, *Information Sciences*, 497: 38-55. DOI: https://doi.org/10.1016/j.ins.2019.05.035

[18] Asano, E., 2017, *How Much Time Do People Spend on Social Media?* [Infographic], [Online, accessed July 24th. 2022] URL: https://www.socialmediatoday.com/marketing/how-much-time-do-people-spend-social-media-infographic

[19] Hochreiter, S., and Schmidhuber, J. 1997. Long short-term memory. *Neural computation*, 9(8): 1735-1780.

[20] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. 2019. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.

[21] Kumar, S., and Singh, T., D. 2022. Fake News Detection on Hindi News Dataset, *Global Transitions Proceedings* 2022, 3(1): 289-297. DOI: https://doi.org/10.1007/s10389-021-01658-z.

[22] Abd Elaziz, M., Dahou, A., Orabi, D. A., Alshathri, S., Soliman, E. M., and Ewees, A. A. 2023. A Hybrid Multitask Learning Framework with a Fire Hawk Optimizer for Arabic Fake News Detection. *Mathematics*, 11(2): 258.

[23] Ma, K., Tang, C., Zhang, W., Cui, B., Ji, K., Chen, Z., & Abraham, A. 2023. DC-CNN: Dual-channel Convolutional Neural Networks with attention-pooling for fake news detection. *Applied Intelligence*, 53(7): 8354-8369.