JRSSEM 2023, Vol. 03, No. 02, 370 – 381
E-ISSN: 2807 - 6311, P-ISSN: 2807 - 6494

IMPLEMENTATION OF DEEP LEARNING FOR FAKE NEWS CLASSIFICATION IN BAHASA INDONESIA

Eko Prasetio Widhi[1]
Dhomas Hatta Fudholi[2]
Syarif Hidayat[3]
Universitas Islam Indonesia
Email: 19917025@students.uii.ac.id[1], hatta.fudholi@uii.ac.id[2], syarif@uii.ac.id[3]
*Corespondence: 19917025@students.uii.ac.id

Abstract: Fake news has become a serious threat in the digital information era. This research aims to develop a model for detecting fake news in Bahasa Indonesia using a deep learning approach, combining the Long Short-Term Memory (LSTM) method with word representations from Word2vec Continuous Bag of Words (CBOW) to achieve optimal results. Our main model is LSTM, optimized through hyperparameter tuning. This model can process information sequentially from both directions, allowing for a better understanding of the news context. The integration of Word2vec CBOW enriches the model's understanding of word relationships in news text, enabling the identification of important patterns for news classification. The evaluation results show that our model performs very well in detecting fake news. After the tuning process, we achieved an F1-Score of 97.30% and an Accuracy of 98.38%. 10-fold cross-validation yielded even better results, with an F1-Score and Accuracy reaching 99%.

Keywords: fake news, deep learning, LSTM, Model Tuning, Word2vec, CBOW

INTRODUCTION

The spread of fake news has become more rampant after the invention of the printing press in 1439 (Anonymous, 2017). Research by (Shu et al., 2017) indicates that the proliferation of social media platforms such as Orkut, Facebook, WhatsApp, Twitter, and Telegram in the late 1990s has made it easier for people to rapidly and widely share information, including the dissemination of fake news. Studies by (Vora et al., 2017) and (Weedon et al., 2017) have shown that Facebook has reported on the distribution of false information by malicious actors, accounting for less than 0.1% of the content posted on the platform.

The adverse effects of fake news cannot be underestimated as it can lead to confusion, influence public opinion, and even trigger overreactions. Often, the public lacks reliable sources of information to verify the accuracy of news, making them susceptible to the rapid spread of false information (Allcott & Gentzkow, 2017). This can severely impact a country's social, political, and economic stability (Pennycook et al., 2020). Therefore, detecting fake news becomes crucial in addressing this challenge.

Although identifying fake news is generally a challenging endeavor, (Ogbuju et al., 2023) research reveals that researchers often rely on fourteen (14) techniques, as noted in recent literature. Machine learning-based fake news detection is gaining popularity, with XGBoost, K-Nearest Neighbor, Naïve Bayes, Support Vector Machine (SVM), Logistic Regression, Decision Trees, and Random Forest being the most commonly used techniques. Meanwhile, deep learning techniques for identifying fake news include Convolutional Neural Network (CNN), Bi-Convolutional Neural Network (BCNN), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (Bi-LSTM), LSTM + BCNN, Bi-LSTM + BCNN, and Multilayer Perception.

In research conducted by (Apriliyanto &; Kusumaningrum, 2020), has evaluated the Long Short-Term Memory (LSTM) model using word2vec feature extract, evaluated using 1000 Indonesian news data. The CBOW and skipgram architectural models were tested in this study, and several modeling scenarios with different parameters were carried out on vector dimensions 100, 200, and 300. Detection model performance is measured using precision, recall, and f1-measure matrices. The study resulted in an outstanding average precision score of 0.819, recall of 0.809, and f1 measure of 0.807.

In the research conducted by (Pakpahan et al., 2022), the Indonesian fake news classification process has been carried out using the BiDirectional LSTM technique, with the Word2vec CBOW architecture used to create a word vectorization model for the classification process. The results showed that models with windows size 3, embedding size 200, and units 128 had the best performance for title data with an accuracy of 79.18%. As for the content data, models with windows size 5, embedding size 300, and units 256 have the best performance with an accuracy of 92.8%.

Another study conducted by (Yusuf & Suyanto, 2022), has conducted an Indonesian hoax detection model that uses short-term memory (LSTM) and is supported by Word2Vec Skip-gram trained and 100-dimensional vectors. The corpus used to design this model consists of 4800 Indonesian-language news articles categorized into two groups: Valid and Hoax. A 10-fold cross-validation method was used to evaluate the model. Based on the results obtained from a 10-fold cross-validation experiment, LSTM combined with the pre-trained Word2Vec corpus on Indonesian Wikipedia resulted in an average accuracy of 89.4%, superior to the average accuracy achieved by Word2Vec of 84.8%. Pre-trained on the corpus of case studies.

Based on the description above, deep learning is superior in learning the representation of hidden input in both context and content. This is different from traditional learning approaches, where feature extraction modeling is done manually (Bondielli &; Marcelloni, 2019). The research revealed to detect fake news with deep learning methods has been widely developed using various models (Ogbuju et al., 2023).

Convolutional Neural Network (CNN) is a deep learning method that can overcome the shortcomings of classical learning, but still has shortcomings in processing sequential data. In contrast, Recurrent Neural Networks (RNNs) are explicitly designed to handle sequential data. However, RNN has limitations in capturing long dependencies even though it is designed to work sequentially. The LSTM model is one variant of the RNN method. The LSTM model is able to overcome long-term dependence by remembering long-term information and is well applied in the case of sentiment analysis or classification, such as hoax news detection (Tin, 2018).

Based on the explanation above, this research chooses to use the LSTM method, Word2vec CBOW method, perform model tuning, and adopt a 10-fold cross-validation method with a strong scientific basis for several reasons:

First, tuning the LSTM model to maximize its performance in the context of Bahasa Indonesia is necessary. Bahasa Indonesia has its own distinctive characteristics in sentence structure, vocabulary, and grammar. Therefore, tuning the LSTM model is required to adapt it to Bahasa Indonesia. This includes adjusting hyperparameters such as the number of LSTM units, the number of LSTM layers, batch size, etc., which are scientifically necessary to enhance the model's performance in processing complex Bahasa Indonesia.

Second, using the CBOW technique to generate better word representations. The CBOW Word2vec technique has proven effective in generating rich and meaningful word representations. In the context of Bahasa Indonesia, where word meanings and relationships between words are often more complex, CBOW can assist in understanding and representing words better compared to other methods that may be less sensitive to language context.

Third, integrating CBOW word representations into the optimized LSTM model can enhance the understanding of context and semantic relationships in Bahasa Indonesia text. This aims to leverage the strengths of each method (LSTM and CBOW) in text processing.

Fourth, improving the accuracy and performance of text classification in Bahasa Indonesia. Text classification in Bahasa Indonesia is a challenging task, especially when distinguishing between fake and genuine news. By using the optimized LSTM, strong word representations from CBOW, and 10-fold cross-validation, this research aims to enhance the accuracy and performance of the model in the task of fake news classification. 10-fold cross-validation helps avoid overfitting and measures the model's performance more accurately.

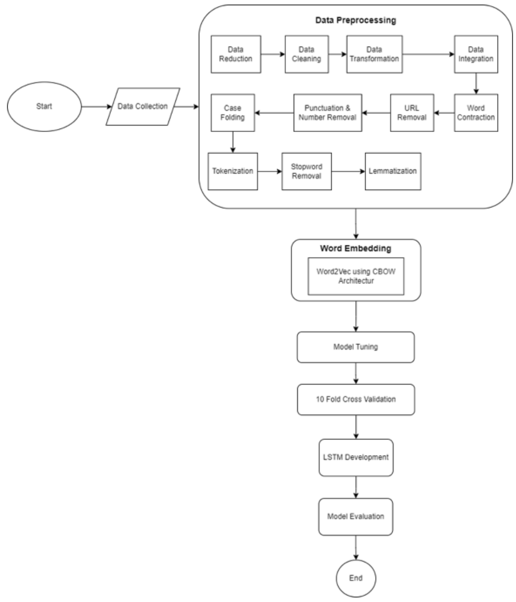
Thus, this research combines important scientific aspects in the development of a fake news classification model in Bahasa Indonesia. By using the right technology and proven scientific steps, this research endeavors to produce a better and more accurate model for this important text classification task.

RESEARCH METHODOLOGY

This research was conducted to obtain the model that has the best performance in classifying news, and attempted to test several parameters on the model to see the effect of its application. The research methodology covers various stages, from data collection to model evaluation. The following are details about the research methodology that contains the purpose of the activity: Modeling flowchart design, data collection, data preprocessing, Word2vec embedding, Tuning model, 10-fold separation, model development and model evaluation.

Modeling Flowchart Show

This stage of modelling flowchart design has a key role in our research methodology. This flowchart visually illustrates the key steps we will take in developing a fake news detection model. Here is the modeling flowchart design (Figure 1):



Dataset Collection

The second step in the study was dataset collection, which was the main foundation of the entire study. Proper dataset collection is an important step to ensure that the models we develop have enough data to train and test effectively. Data collection was carried out by taking samples of Indonesian-language news from various online sources that included fake news and real news.

Pre-processing of Data

The next step in the study is data preprocessing. Data preprocessing is an important stage in data preparation before it is used in modeling. This involves a series of steps to clean and prepare the data to fit the needs of the fake news classification model. The steps include the removal of special characters, conversion of text to lowercase, tokenization of text (separation of words), and removal of stop words.

Word2vec embedding

The next step in the study is 'Word2vec Embedding'. This process involves using Word2vec techniques, specifically the CBOW (Continuous Bag of Words) model, to generate vector representations of words from the researcher's news dataset. The Word2vec model was trained using preprocessed Indonesian news data This vector representation will be used as a feature in the researcher classification model.

Model Tuning

The next step in our research is 'Model Tuning'. At this stage, we made adjustments and experiments with the LSTM model to optimize its performance in classifying fake news and genuine news in Indonesian. The LSTM model will be initiated with initial parameters and then go through a tuning process. It involves experimenting with various hyperparameters such as the number of LSTM units, the number of LSTM layers, batch size, and learning rate. The model will be updated and optimized for the highest accuracy classification of fake news.

Separation of 10 Fold

The next stage in our research is '10 Fold Separation'. This is an important step in ensuring a robust and valid evaluation of the model we are developing. With 10-fold separation, we split the dataset into 10 different subsets to measure the overall performance of the model. This ensures that the model is assessed based on a large amount of disparate data and reduces the risk of overfitting.

Model Development

The next step in our research is 'Model Development'. At this stage, we use the training dataset from each 10-fold iteration to develop and train a pre-optimized LSTM model. An optimized LSTM model will be developed using training data from each fold in cross-validation. This process will result in a model that has a better understanding of the news in Indonesian.

Model Evaluation

The final step in our research is 'Model Evaluation'. At this stage, we measure the performance of the developed model using predefined evaluation metrics. Evaluation metrics used include accuracy, precision, recall, F1-score, and confusion matrix. The results of this evaluation will provide an idea of the extent to which our model is effective in classifying fake news and genuine news in Indonesian

RESULTS AND DISCUSSION

The last stage in this research is the 'Research Results and Discussion', which is the conclusion of the entire research and discussion of the findings that have been found.

Research Results

Data Collection

This study used an Indonesian-language dataset consisting of 40,039 news samples, consisting of 29,618 genuine (reliable) news and 10,521 fake news (hoaxes). This dataset was obtained from various sources, including leading news websites such as Kompas, Medcom, Sindonews, Antaranews, and Kaggle. The dataset has been divided into three main sections: training, testing, and validation, with a balanced composition of real and fake news in each section. The training section consisted of 15,841 real news and 6,693 fake news, while the testing section had 4,890 real news and 2,152 fake news. The validation section contained 3,958 original news and 1,676 fake news. With this powerful and diverse dataset, this study aims to develop a fake news detection model that is effective in Indonesian, focusing on the accuracy and performance of the model in distinguishing fake and genuine news content.

Pre-processing of Data

The data preprocessing process is an important stage in this research. We have taken a series of steps to ensure that our news datasets are ready for use in the development of fake news detection models. First of all, we removed special characters, punctuation, and irrelevant elements from news text to reduce distractions in the dataset. Next, we perform tokenization, that is, separate the text into words or tokens, to understand the structure of sentences and words better. We also removed stop words, which often have no significance in text analysis, in order to focus the model more on informative words. In addition, all text is converted to lowercase in the case folding process for consistency. Finally, we applied lemmatization to convert words into their base word forms, thus reducing the variety of words that have the same meaning. With this careful data preprocessing process, we ensure that our datasets are clean and ready to use for the development of effective fake news detection models in Indonesian.

Word Representation (Word2Vec)

In this study, we used the Word2vec Continuous Bag of Words (CBOW) model to generate vector representations of words from news text in Indonesian. The Word2vec CBOW model is one of the popular methods in natural language processing used to convert words into vectors with rich meanings. The dataset we used for Word2vec CBOW model training comes from the Indonesian Wikipedia corpus, which is a rich source with text in Indonesian.

Model LSTM

At the stage of setting up and configuring the optimized LSTM model, this study has carefully selected several hyperparameters that play an important role in achieving the best performance in detecting fake news in Indonesian. One is the use of the loss function "binary_crossentropy," which is a suitable choice for binary classification tasks such as fake news detection. As an optimizer, the "Adam" method is used with a learning rate of 0.001, which gives the model the ability to effectively adjust the weight of the neural network.

In addition, this study utilizes several important evaluation metrics, namely accuracy, recall, precision, and

F1-Score, to measure model performance. Accuracy gives an idea of how well the model correctly classifies news, while recall measures the model's ability to identify real fake news. Precision assesses the extent to which news classified as fake news is really fake news, while F1-Score combines precision and recall to provide a comprehensive picture of model performance.

Not only that, the batch size setting of 64 is used to determine the number of training samples that should be processed before updating model parameters. Meanwhile, the model is trained for 50 epochs, which determine how many times the learning algorithm will iterate through the entire training dataset. By carefully selecting these hyperparameters, the study has successfully developed an LSTM model optimized to detect fake news in Indonesian with excellent performance.

Model Evaluation

Model evaluation is an important stage in this study, and we compare our model's performance results with a variety of different approaches and methods. Here are the results of model evaluation in several different scenarios:

1. Optimized LSTM Models:
   researchers start by optimizing our basic LSTM model with optimized parameters. The evaluation results show that this model has an F1-Score of 96.70% and an Accuracy of 98.02%. This indicates that with careful tuning, the LSTM model can distinguish well between fake news and real news in Indonesian.
2. LSTM Models That Have Been Optimized by Tuning:
   In an effort to improve model performance, we performed further tuning on optimized LSTM models. The evaluation results show that this model produces an F1-Score of 97.30% and an Accuracy of 98.38%. This suggests that further tuning provides a significant improvement in the model's ability to classify fake news.
3. Word2vec CBOW Integration with Optimized LSTM Models
   Researchers took it a step further by incorporating word representations from Word2vec CBOW into an optimized LSTM model. The result is a more sophisticated model architecture, which results in an F1-Score of 93.06% and an Accuracy of 96.16%. This suggests that Word2vec CBOW can help the model understand the context of the news better, although it needs additional tuning to improve precision and recall.
4. Word2vec CBOW Integration with Optimized LSTM Model and 10 Fold Cross-Validation
   In this step, we run an experiment using 10 Fold cross-validation to measure the performance of models that have been integrated with Word2vec CBOW. The results are remarkable, with an F1-Score of 99% and Accuracy of 99%. This shows that the merging of Word2vec CBOW with an optimized LSTM model, along with cross-validation, resulted

in a model that excels greatly in the task of classifying fake news.

In this series of studies, researchers continuously strive to improve the performance of our models by tuning and integrating different methods. The evaluation results show that Word2vec CBOW integration and 10 Fold cross-validation are important steps in achieving the highest performance in classifying fake news in Indonesian. All of this demonstrates our commitment in developing effective solutions to deal with the spread of fake news in Indonesian.

Results Discussion

Model Evaluation

The model's performance in classifying fake and genuine news in Indonesian was impressive, with an F1-Score of 97.30% and an accuracy of 98.38% after tuning and optimization. This shows the model's ability to distinguish well between the two types of news. However, despite being generally a high-performance model, there are still some situations where the model can be less accurate, especially in classifying genuine news as false (False Positive) or vice versa (False Negative). The integration of Word2vec CBOW and additional tuning has helped in improving performance, but there is still room for further improvement. The most impressive result was when we applied 10 Fold cross-validation, which resulted in a 99% F1-Score and 99% Accuracy, showing that the cross-validation technique can provide significant improvements. In the context of the increasingly sophisticated spread of fake news, the development and improvement of this model continues to be our focus to better meet this challenge.

Effects of Word Representation (Word2Vec)

The influence of Word2vec word representations from the Word2vec CBOW model in this research has a significant impact on the model's performance. Analysis indicates that the use of these word representations aids the model in better understanding the context of news.

Models that integrate Word2vec CBOW in text processing have better capabilities in capturing the meanings and relationships between words in news articles. The word representations obtained from Word2vec CBOW enable the model to understand how words are related to each other within sentences and paragraphs. This helps in recognizing important contextual patterns and relationships crucial in determining whether a news article is fake or genuine.

In other words, the use of Word2vec CBOW enhances the model's ability to identify key words, phrases, or sentence patterns that frequently appear in fake news. This allows the model to more accurately classify news based on a broader context, not just individual words. Therefore, the integration of Word2vec CBOW is a crucial step in improving the model's understanding of the news context, as reflected in its better performance in distinguishing between fake and genuine news.

Advantages and Limitations

The model that has been developed in this study offers a number of significant advantages. The performance of the model reached an excellent level with an F1-Score of 97.30% and an accuracy of 98.38% after going through the tuning process. This demonstrates the model's ability to classify fake and genuine news with high accuracy. The integration of word representations from Word2vec CBOW is also one of the main advantages, helping the model to understand the context of the news better and identify contextual patterns that are important in determining classification. The use of 10 Fold cross-validation also gave outstanding results with an F1-Score and Accuracy of up to 99%, indicating the model's excellent ability to classify fake news.

However, there are some limitations to note. There are certain news stories that are difficult to classify, which can be due to the complexity of the structure of the language and information in the news. In addition, although Word2vec CBOW improves performance, word representation is not always perfect in describing the context of the news. Another challenge is the class imbalance in the dataset, where there is less amount of fake news, which can lead to bias in the model. Finally, some types of news such as those that use ambiguous or controversial language can be more difficult to classify. Overall, this model has strong advantages in classifying fake news, however, complex classification challenges remain a focus for subsequent model development.

Practical Implications

The model that has been developed in this study has significant practical implications in combating the spread of fake news in Indonesian. With the ability to detect fake news automatically and accurately, this model can be used as an important tool in tackling the problem of fake news in today's digital world. The potential applications of this model are very diverse, including use as an automated news filter in online news platforms, a support tool for journalists in news verification, and even as a public education tool to improve media literacy. In addition, this model can be used to analyze sentiment and truth in news content published online. With this model, it is expected to make a positive contribution in maintaining information integrity, improving the quality of news presented to the public, and assisting the public in identifying fake news, thereby minimizing its negative impact on society and encouraging the use of more accurate and trustworthy information in the digital environment.

Recommendations for future research

For future research, there are several suggestions that can be considered to develop understanding and application in the detection of fake news in Indonesian. First, it is recommended to expand the dataset with more news from different sources and topics in order for the model to deal with wider language variation. Second, it is necessary to address the problem of class imbalance in the dataset, for example by using oversampling or undersampling techniques. Third, the potential for developing better word representations such as FastText or BERT needs to be explored to improve context understanding. Fourth, considering the recognition of named entities, custom phrases, and more sophisticated natural language processing can improve model performance. Fifth, evaluation in real cases, such as on social media, will help gauge the effectiveness of the model in the real world. Sixth, the development of a friendly user interface can facilitate the use of the model by journalists and the public. Finally, collaboration with the media, news factors, and platform providers can facilitate the integration of models in a joint effort against fake news. These recommendations are expected to help develop more sophisticated and effective models in combating fake news in Indonesian as well as contribute to efforts to combat disinformation in an increasingly complex digital world.

CONCLUSION

In conclusion, this study has succeeded in developing a deep learning-based fake news detection model, specifically by using the LSTM method and Word2vec CBOW integration in Indonesian. This model has shown excellent performance, with F1-Score reaching 97.30% and Accuracy reaching 98.38% after tuning process and 10 Fold cross-validation resulting in F1-Score and Accuracy reaching 99%. Word2vec CBOW integration also makes a significant contribution to understanding the context of news. However, there are still challenges that need to be addressed, such as difficult news classification and class imbalances in the dataset.

The research has significant practical implications in combating the spread of fake news in Indonesian. This model can be used as an automatic detection tool for fake news, a news filter, a support tool for journalists, and as a public education tool to improve media literacy. Nonetheless, the study also provides suggestions for future research, including dataset expansion, word representation improvement, addressing class imbalances, and real-case evaluation.

Overall, this study shows that the use of deep learning methods and proper word representation can be very effective in identifying fake news in Indonesian. With a wide range of potential real-world applications, this model could play a role in maintaining information integrity and combating the spread of disinformation in an increasingly complex digital age.

REFERENCES

Aceto, G., Ciuonzo, D., Member, S., Montieri, A., Member, G. S., Pescapé, A., & Member, S. (2019). Using Deep Learning : Experimental Evaluation , Lessons Learned , and Challenges. *IEEE Transactions on Network and Service Management*, *PP*(DL), 1.

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Anonymous. (2017). *Fake News : Historical Timeline*. 1–2.

Fernández-Reyes, F. C., & Shinde, S. (2018). Evaluating Deep Neural Networks for Automatic Fake News Detection in Political Domain. In *Advances in Artificial Intelligence - IBERAMIA 2018* (pp. 206–216). Springer, Cham. https://doi.org/10.1007/978-3-030-03928-8_17

Haumahu, J. P., Permana, S. D. H., & Yaddarabullah, Y. (2021). Fake news classification for Indonesian news using Extreme Gradient Boosting (XGBoost). *IOP Conference Series: Materials Science and Engineering*, *1098*(5), 052081. https://doi.org/10.1088/1757-899x/1098/5/052081

Kamath, U., Liu, J., & Whitaker, J. (2019). *Deep Learning for NLP and Speech Recognition*. Springer International Publishing. https://doi.org/10.1007/978-3-030-14596-5

Khanam, Z., Alwasel, B. N., Sirafi, H., & Rashid, M. (2021). Fake News Detection Using Machine Learning Approaches. *IOP Conference Series: Materials Science and Engineering*, *1099*(1), 012040. https://doi.org/10.1088/1757-899x/1099/1/012040

Mridha, M. F., Keya, A. J., Hamid, M. A., Monowar, M. M., & Rahman, M. S. (2021). A Comprehensive Review on Fake News Detection with Deep Learning. *IEEE Access*, *9*, 156151–156170. https://doi.org/10.1109/ACCESS.2021.3129329

Ogbuju, E., Abiodun, T., & Oladipo, F. (2023). *Text Analytics Solutions for the Control of Fake News : Materials and Methods*. *11*(3), 69–74.

O'Shea, T., & Hoydis, J. (2017). An Introduction to Deep Learning for the Physical Layer. *IEEE Transactions on Cognitive Communications and Networking*, *3*(4), 563–575. https://doi.org/10.1109/TCCN.2017.2758370

Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings. *Management Science*, *66*(11), 4944–4957. https://doi.org/10.1287/mnsc.2019.3478

Qawasmeh, E., Tawalbeh, M., & Abdullah, M. (2019). Automatic Identification of Fake News Using

Deep Learning. *2019 6th International Conference on Social Networks Analysis, Management and Security, SNAMS 2019*, *October*, 383–388. https://doi.org/10.1109/SNAMS.2019.8931873

Rohman, M. A., Khairani, D., Hulliyah, K., Arini, Riswandi, P., & Lakoni, I. (2021). Systematic Literature Review on Methods used in Classification and Fake News Detection in Indonesian. *2021 9th International Conference on Cyber and IT Service Management, CITSM 2021*, *Ml*, 1–4. https://doi.org/10.1109/CITSM52892.2021.9589004

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). *Fake News Detection on Social Media: A Data Mining Perspective*. *i*.

Vora, J., Tanwar, S., Tyagi, S., Kumar, N., & Rodrigues, J. J. P. C. (2017). Home-based exercise system for patients using IoT enabled smart speaker. *2017 IEEE 19th International Conference on E-Health Networking, Applications and Services, Healthcom 2017*, *2017-Decem* (October), 1–6. https://doi.org/10.1109/HealthCom.2017.8210826

Weedon, J., Nuland, W., & Stamos, A. (2017). Information Operations and Facebook. *Facebook*, 1–13.

Yang, Y., Zheng, L., Zhang, J., Cui, Q., Li, Z., & Yu, P. S. (2018). *TI-CNN: Convolutional Neural Networks for Fake News Detection*.

Zhang, Y., Gorriz, J. M., & Dong, Z. (2021). Deep learning in medical image analysis. *Journal of Imaging*, *7*(4), NA. https://doi.org/10.3390/jimaging7040074