

# Evaluation of Machine Learning Algorithms for Fake News Detection Based on Accuracy, Using Word2vec as Feature Extraction Method

Aminu Malam Ya'u<sup>1\*</sup>, Abubakar Muhammad Miyim<sup>2</sup>

<sup>1</sup>Department of Computer Science, Federal University Dutse, Dutse, Nigeria.

<sup>2</sup>Department of Information Technology, Federal University Dutse, Dutse, Nigeria.  
amyaujara@gmail.com<sup>1\*</sup>, abubakar.miyim@fud.edu.ng<sup>2</sup>

\*Corresponding Author

## Abstract

*Fake news refers to false information presented as news, circulated over the internet or other communication media for either political purposes or joking. Fake news is a serious problem that needs to be addressed because it causes a lot of issues in human lives such as political crises, fraud, murder and so 'on, to resolve such problems, there is a need to develop a machine learning model that can be classified whether the news is fake or real. Much research works on fake news detection has been carried out, literature review shows that researchers employed many machine-learning techniques for the detection of fake news. In this research, the word2vec feature extraction technique was adopted in converting the news text into numerical values so that it can be fed into machine learning models. Four Machine learning (ML) classifiers were selected, trained, and evaluated to choose the best classifiers for fake news detection. An experimental result shows that Multi-Layer Perceptron (MLP) outperformed the other algorithms with accuracy-98%.*

**Keywords:** Machine learning, fake news, word2vec, MLP, SVM.

## 1. Introduction

According to Wang (2020), Fake news is news that consists of inaccurate information but the inaccuracy is not highlighted in the news. Waldrop (2017) divides Fake News into seven types: (1) Satire or Parody: news that is not initiated to hurt anyone, but can fool someone. (2) False connection: a type of news in which the headlines and content are not related to each other. (3) Misleading content: news that leads an individual or organization to the wrong action. (4) False context: a mixture of false and true news. (5) Imposter content: when a true source of information is impersonated with a false source. (6) Manipulated content: when true information is manipulated to cheat someone. (7) Fabricated content: News that is completely not true created intentionally to deceive someone.

Historically, Rumors and false stories exist as long as people are living in groups. In the pre-printing press era, information is communicated by word of mouth. However, in the post-printing press era, the spread of information has increased widely. Literate people utilize this printing press technology to manipulate information to illiterate people, this is where the fake news come into existence since the introduction of radio broadcasting. (Burkhardt, 2017).

Machine learning is a widely used technique for the detection of fake news thereby classifying the news content as either fake or real. The accuracy of the classifier is the major priority because incorrect detection of fake news by the classifier can negatively affect many people. (Ahmed et al., 2021). In this paper, four machine learning algorithms are trained and evaluated to identify the best classifier for fake news detection.

The rest of the paper is structured as follows: Section two analyses related works of other researchers. Section three discusses the methodology used including data collection and feature selection processes. Section four

discusses the analysis of the simulation results including the performance evaluation. The concluding part and future works are found in section five.

## 2. Related Works

Vijayaraghavan et al., (2020) in their paper titled "Fake news detection with different models", developed a machine-learning model for the detection of fake news. They trained and tested five machine learning algorithms namely; Support Vector Machine (SVM), Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), Logistic Regression (LR), and Random Forest (RF). They used three feature extraction models for text (count vectorizer, Term Frequency-Inverse Document Frequency vectorizer, and Word2vec) and their performance on all the classifiers was evaluated. With the count vectorizer feature extraction model, LSTM stood out to be the best-performing algorithm with an accuracy of 94.88%, with the TF-IDF vectorizer model, Logistic Regression with an accuracy of 94.79%. Lastly, ANN performed best with the word2vec feature extraction method with an accuracy of 93.06%.

Oriola` (2020) in his paper titled "Exploring N-gram, Word Embedding, and Topic Models for Content-based Fake News Detection in FakeNewsNet Evaluation", Compares the performance of three feature extraction methods (N-gram, Word2vec, and topic model) using Logistic Regression algorithm. Two datasets (PolitiFact and GossipCop) were used to train and test the Logistic Regression classifier. The result shows that; in the PolitiFact dataset, Logistic Regression achieved an accuracy of 80% with N-gram, 60% with the topic model, and 73% with word2vec. In the GossipCop dataset, Logistic Regression achieved an accuracy of 82% with N-gram, 51% with the topic model, and 78% with word2vec. From the above results it could be deduced that based on this research, the algorithm achieved the highest accuracy only with the N-gram feature extraction method.

Jain et al., (2019) in their paper titled "A smart system for fake news detection using machine learning" detected fake news using two classifiers namely SVM and Naïve Bayes. They used a count vectorizer for feature extraction to convert the news text into numerical values to be fed into the machine learning classifiers. The experimental result shows that SVM and Naïve Bayes achieved the same accuracy of 93.5% each. However, the researchers did not consider precision and recall in their evaluation. In trying to detect fake news, the authors choose the WELFake dataset and used six machine learning algorithms to train and evaluate their performances. They used Count Vectorizer for feature extraction; results show that SVM was found to outperform the other five algorithms with an accuracy of 96.73% (Verma et al., 2021)

According to Fang et al., (2019), datasets collected from [fakenews.mid.edu](http://fakenews.mid.edu), were trained and tested using Convolutional Neural Network (CNN) with two feature extraction methods (word2vec and Glove). The result obtained from their work shows that the feature extraction with word2vec did achieve an accuracy of 95.90%, while that with Glove was able to achieve an accuracy of 94.67%, indicating that CNN performed best with word2vec feature extraction.

According to Jehad and Yousif (2021) in their paper titled "Classification of Fake news using Multi-layer Perceptron" trained and tested Multi-layer Perceptron (MLP) using TF-IDF as feature extraction method. Their result shows that MLP was able achieved the accuracy of 95.47%

Nurhasanah et al., (2022) conducted fake news detection with the dataset downloaded from Kaggle. They have trained two machine learning classifiers namely SVM and Naïve Bayes using TF-IDF vectorizer as feature extraction technique. In their result, SVM outperformed Naïve Bayes with an accuracy of 96%.

Kesarwani et al., (2020) in their paper titled "Fake news detection on social media using K-Nearest Neighbor Classifier." Employed KNN classifiers to detect fake news using the dataset collected from Buzz feed news organization. Result shows that KNN was able to achieve the accuracy of 79%.

Tanvir et al., (2019) conducted a research using Chile earthquake 2010 dataset to train five classifiers. Four feature extraction methods were used and evaluated separately on all the classifiers. Result shows that SVM with TF-IDF feature extraction outperformed the algorithms with an accuracy of 89.34%.

From the sampled literature considered, it is clearer now that the accuracy of the best-performing machine learning classifiers in fake news detection is about 96% for all the models used by various researchers. The goal of this paper, therefore, is to improve the performance of the machine learning model in terms of accuracy to detect fake news correctly. The researchers felt that there is room to improve the accuracy when word2vec feature extraction techniques and good preprocessing methods are considered with four classifiers for the detection of fake news.

### 3. Methodology

The method employed in this research begins with data collection, preprocessing, feature extraction, data splitting, and model training and testing. If the accuracy has been achieved, then the performance of the classifiers would be evaluated, and choose the best model, otherwise should go back to preprocessing for additional preprocessing activities.

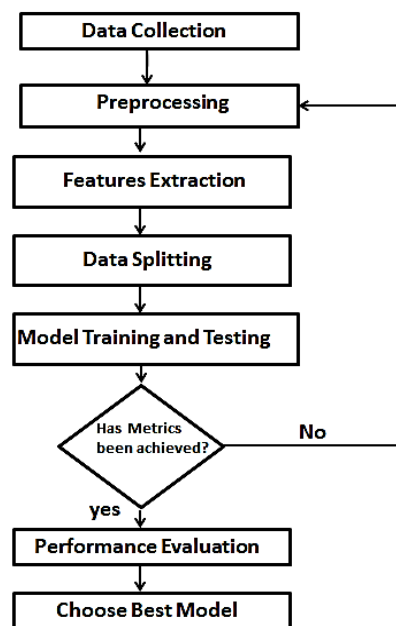


Fig. 1 Flow Diagram of Proposed Method

#### 3.1 Data Collection

A Publicly accessible news dataset was used in this research collected from the University of Victoria, Canada machine learning repository. The dataset consists of two types of news articles; fake and real. The dataset was gathered from real-world sources; the Real news was collected from true articles from reuters.com (News website). While the fake news articles were collected from untrusted websites that were weakened by Politifact (a fact-checking organization in the USA) and Wikipedia. The dataset consists of many articles on different topics, moreover, the majority of the news is concentrated on political and world news topics.

The dataset contains two CSV file; "True.csv" (which consist of more than 12,600 articles from reuters.com) and "fake.csv" (which consist of more than 12,600 news articles from untrusted websites). Each article has the following details: "title" (that shows the headline of the news), text (the actual content of the news), type (the category in which the news belongs), and date (the date on which the news was published).

### 3.2 Preprocessing

After the dataset has been collected, the next step was preprocessing, in which the raw data was prepared for feeding into the machine learning model. It is the first and most important step in developing a machine learning model, the step was necessary to convert raw data into a clean dataset. The dataset was preprocessed with punctuation and special characters removal, conversion to lowercase, stop words removal, and stemming.

### 3.3 Feature Extraction

After preprocessing, the next step is feature extraction. The performances of machine learning classifiers greatly depend on feature extraction techniques (Khan et al., 2019). Fake news detection falls under the domain of Natural Language Processing (NLP; which is the branch of artificial intelligence that analyzes processes and efficiently retrieves text data). Therefore, the dataset to be used for fake news detection is in the form of text, and before feeding text data into the machine learning model, it must be converted into numeric values. As explained earlier, the research aims at improving the performance of machine learning classifiers in fake news detection using the word2vec feature extraction technique.

### 3.4 Word2vec

The feature extraction technique employed in this research was word2vec where each word in the dataset was replaced with its corresponding pre-trained vector. The text column was used in the dataset as the independent variable (X) (which is the actual news content). However, the label column in the dataset was taken as the dependent variable (Y).

To use word2vec one has to follow either of the two ways:

- Use pre-trained models which can download online
- Use deep learning to train your data and model

In this research the first option was chosen; which is a pre-trained model. A spacy NLP library model was downloaded of about 827.9MB which contains built-in words with their corresponding pre-trained vectors. This spacy model was used in converting the text column in the dataset into vectors so that it can be fed into machine learning models.

### 3.5 Data Splitting

The dataset was split into 80:20 ratio which is 80% for training and 20% for testing. This is because of the size of the dataset which is large enough (44,898 news articles) and if the dataset is large 80:20 ratio is preferable.

### 3.6 Model Training and Testing

There exist several algorithms for training and testing dataset and the machine learning classifiers that were used for fake news detection in this research are explained below:

#### 3.6.1 Support Vector Machine (SVM)

According to Jain et al., (2019), SVM is the best classifier for binary classification with an accuracy of 93.5%. It is a supervised learning classifier that can solve both regression and classification problems. This algorithm classifies the data points into two classes with the help of hyper-planes. In this research, the news articles are classified as either fake or real.

#### 3.6.2 Multi-Layer Perceptron (MLP)

Multilayer Perceptron (MLP) is a feed-forward artificial neural network algorithm, in which the input data is mapped to the respective outputs. It consists of three layers; the input layer (which is used for receiving the signal), the output layer (which is used to perform the actual task such as prediction and classification), the hidden layer (which is located between the input and output layers, it is the true processing engine of MLP). In MLP, data move from the input layer to the output layers in the forward direction. It is created for solving problems that cannot separate linearly. Some applications of MLP are; pattern description, identification, estimation, etc. In their results, MLP was able to achieve an accuracy of 94.6%. (Fattoh and Mousa, 2021).

### 3.6.3 Logistic Regression (LR)

Logistic regression: This is a machine learning model that can be used to analyze the association between independent variable and dependent variable, it also fits a logistic curve for calculating the probability of occurrence of an event. Logistic regression has two models, Binary logistic regression and multinomial logistic regression. Binary logistic regression is used when the dependent variable has only two categories (such as true or false, 0 or 1, yes or no, fake or real, etc.). However, if the dependent variable has more than two categories then multinomial logistic regression is used (Hyeoun, 2013).

### 3.6.4 Random Forest (RF)

A Random Forest is a machine-learning model that contains many decision trees working together as an ensemble. In addition, each tree is formed by using baggage and randomness to create a forest of negatively correlated trees which is better than a single tree. The most voted class in the random forest is the predictor of the model. The random forest produces a good result due to the negatively correlated trees that work together as an ensemble and also because there is a low association between the models. In their results, Random Forest was able to achieve an accuracy of 90.9%. (Fattoh and Mousa, 2021).

## 4. Result and Discussion

The result of this research was at the first instance visualized using a confusion matrix to determine the performance of all the trained classifiers. Figure 2 below shows the confusion matrix of MLP where 4633 news article was correctly predicted as fake and 4210 news was correctly predicted as real.

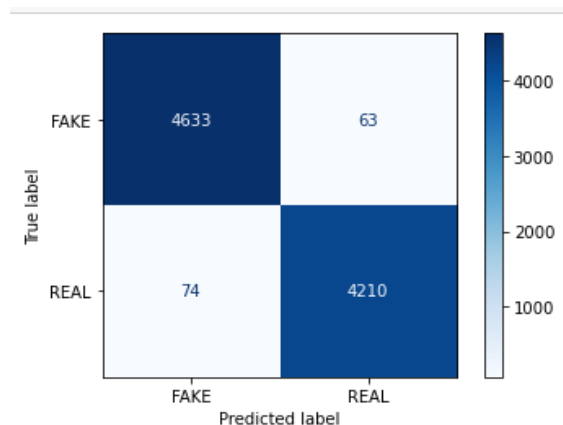


Fig. 2 Confusion Matrix for MLP

From Figure 3 below the total number of news correctly predicted as fake is 4528, whereas 4178 news articles are correctly predicted as real.

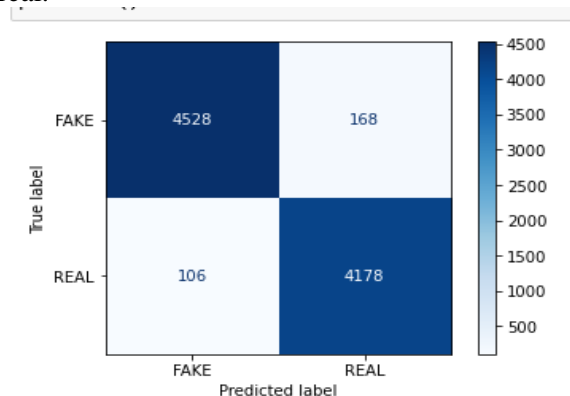


Fig. 3 Confusion Matrix for SVM

Figure 4 below shows the confusion Matrix of logistic regression in which it's correctly predicted 4489 news as fake. However, 4125 news was correctly predicted as real.

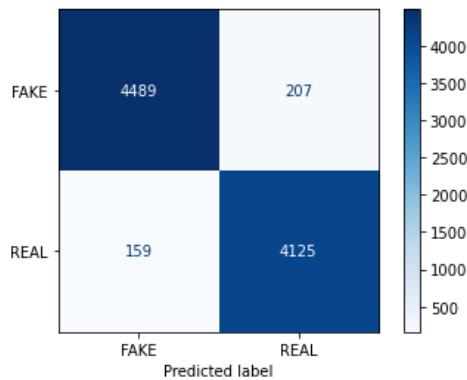


Fig. 4 Confusion Matrix for Logistic Regression

The result of the Random Forest in Figure 5 below indicates that 4491 news was correctly predicted as fake. Furthermore, 4027 real news were correctly predicted.

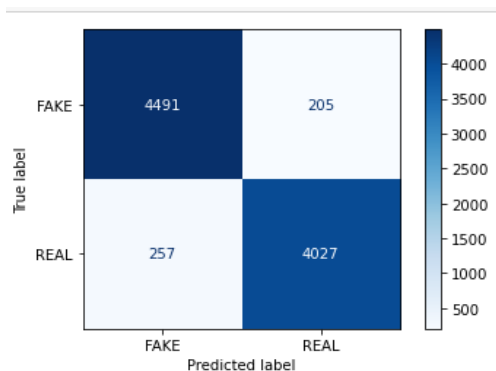


Fig. 5 Confusion Matrix for Random Forest

After the results of the classifiers were visualized using a confusion matrix, the next is to evaluate their performance in terms of Accuracy. Table 1 below shows the performance result of the models.

S/N	Models	Accuracy (%)
1	MLP	98
2	SVM	97
3	LR	96
4	RF	95

From the result of Figure 6 below it could be seen that using word2vec as a feature extraction technique yields a better result due to its strengths in the extraction process. The result shows that MLP is the best-performing algorithm with an Accuracy of 98%; this was due to the ability of MLP to model non-linearly separable datasets and the ability to also learn quickly. The second performing classifier is the SVM with an Accuracy of 97%; for the simple reason that SVM works well with even unstructured and semi-structured data such as text, images, and trees. In general, the kernel trick (which is the method of projecting nonlinear data onto a higher dimensional space to make it linearly separable by a plane) is the strength of SVM. Logistic regression appeared to be the third-best-performing algorithm with an Accuracy of 96%. It is well known, however, that Logistic Regression has the

capability of solving binary classification problems and so produced such wonderful results. The last classifier which is the Random Forest recorded the worst performance with an Accuracy of 95%. Even though the algorithm has the fastest method of training data, however, it is quite slow when it comes to prediction.

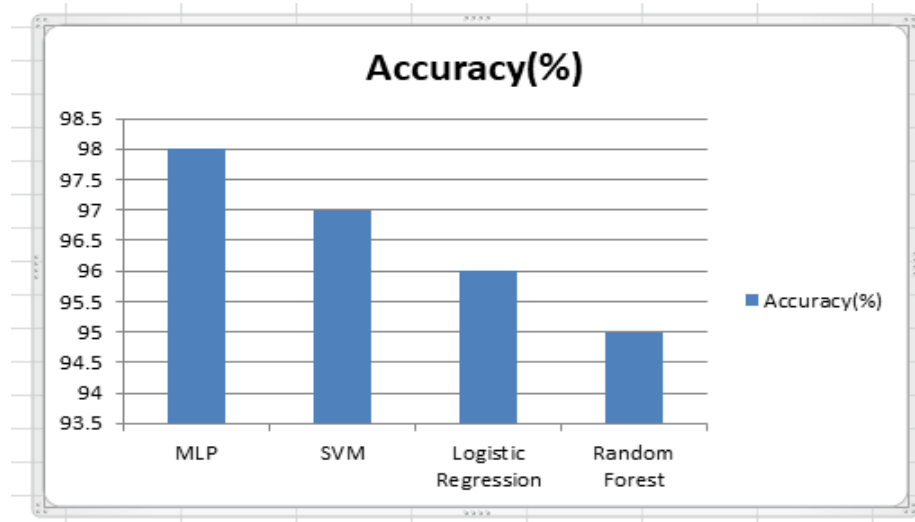


Fig. 6: Performance Comparison between Classifiers

## 5. Conclusion

The result of this research shows the importance of the feature extraction method in developing a machine-learning model for fake news detection. In this research, the text column in the dataset; which is the actual content of the news was used as the independent variable (X) to detect the fake news. In future work, the following are recommended:

1. Though the actual content of news for the detection of fake news did find usage as an independent variable, it will therefore be of interest to merge the title and text columns as one data content to be used as the independent variable (X). The essence is to create additional features in the content that transforms into a binary task where each of the content is labeled true or false through Machine Learning techniques.
2. Given a multi-source news dataset and the social context of news consumers (social media users), the task of detecting fake news is to determine if a news item is fake or real. It is suggested that different architectures of machine learning techniques could be applied to a wide range of scenarios, languages, and platforms as interesting research.
3. In recent times, transfer learning has been applied to detect fake news which could be explored in fake news detection because of the delicate nature of fake news as it deals with semantics, hidden meanings, and context from fake news data. It will be interesting to take further research on fake news detection by considering a transfer learning-based scheme while paying great attention to syntax, semantics, and meanings in fake news data.
4. The use of the factorization method to model the relationships among publishers, news stories, and social media users for fake news detection could be considered in future research. If handled appropriately, some breakthroughs/novelties might surface from such research.

## References

- Ahmed, A. A. A., Aljarbouh, A., Donepudi, P. K., Choi, M. S. (2021). Detecting Fake News using Machine Learning: A Systematic Literature Review. *Computer and Society*, 1, 6. <https://doi.org/10.48550/arXiv.2102.04458>
- Burkhardt, J. M. (2017). *Combating Fake News in the Digital Age*. Library Technology Reports.
- Fang, Y., Gao, J., Huang, C., Peng, H., Wu, R. (2019). Self Multi-Head Attention-based Convolutional Neural Networks for fake news detection. *PLOS ONE*, 14(9), 1-13. <https://doi.org/10.1371/journal.pone.0222713>

- Fattah, I. E. and Mousa, F. A. (2021). Fake News Detection Based on Word and Document Embedding Using Machine Learning Classifiers. *Journal of Theoretical and Applied Information Technology*, 99(8), 1891-1901.
- Jain, A., Khatter, H., Shakya, A., (2019). A Smart System For Fake News Detection Using Machine Learning. *IEEE 2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*. <https://doi.org/10.1109/ICICT46931.2019.8977659>
- Jehad, R., Yousif, S. A. (2021). Classification of fake news using Multi-Layer Perceptron. *AIP Conference Proceedings*, 2334(1), 1-5. <https://doi.org/10.1063/5.0042264>
- Kesarwani, A., Chauhan, S. S., Nair, A. R. (2020). Fake News Detection on Social Media Using K-Nearest Neighbor Classifier. *2020 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, 1-4. <https://doi.org/10.1109/ICACCE49060.2020.9154997>
- Khan, J. Y., Khondaker, M. T. I., Iqbal, A., Afroz, S. (2019). A Benchmark Study on Machine Learning Methods for Fake News Detection Using Machine Learning. *Computation and Language (cs.CL)*, 1, 1-14.
- Nurhasanah, N., Sumarly, D. E., Pratama, J., Heng, I. T.K., Irwansyah, E. (2022). Comparing SVM and Naïve Bayes Classifier for Fake News Detection. *Engineering, MAThematics and Computer Science (EMACS) Journal*, 4(3), 103-107. <https://doi.org/10.21512/emacsjournal.v4i3.8670>
- Oriola, O. (2020). Exploring N-gram, Word Embedding and Topic models for content-based Fake news Detection In FakeNewsNet Evaluation. *International Journal of Computer Applications*, 176(39), 24-29.
- Park, H. A. (2013). An Introduction to Logistic Regression: From Basic Concepts to Interpretation with Particular Attention to Nursing Domain. *J Korean Acad Nurs*, 43(2), 155. <https://dx.doi.org/10.4040/jkan.2013.43.2.154>
- Tanvir, A. A., Mahir, E. M., Akhter, S., Huq, M. R. (2019). Detecting Fake News Using Machine Learning and Deep Learning Algorithms. *2019 7th International Conference on Smart Computing & Communications (ICSCC)*, 1-5. <https://doi.org/10.1109/ICSCC.2019.8843612>
- Verma, P. K., Agrawal, P., Amorim, I., Prodan, R. (2021). WELFake: Word Embedding over Linguistic Features for Fake News Detection. *IEEE Transactions on Computational Social Systems*, 8(4), 1-12. <https://doi.org/10.1109/TCSS.2021.3068519>
- Vijayaraghavan, S., Guo, Z., Wang, Y., Voong, J., Xu, W., Nasser, A., Cai, J., Li, L., Vuong, K., Wadhwa, E. (2020). Fake News Detection with Different Models. *Computation and Language (cs.CL)*, 1, 1-15. <https://doi.org/10.48550/arXiv.2003.04978>
- Waldrop, M. M. (2017). News Feature: The genuine problem of fake news. *Proc Natl Acad Sci USA*, 114(48), 12632. <https://doi.org/10.1073/pnas.1719005114>
- Wang, C. (2020). Fake News and Related Concepts: Definitions and Recent Research Development. *Contemporary Management Research*, 16(3), 149. <https://doi.org/10.7903/cmr.20677>