Postmodern Openings
ISSN: 2068-0236 | e-ISSN: 2069-9387
Covered in: Web of Science (WOS); EBSCO; ERIH+; Google Scholar; Index Copernicus; Ideas RePeC; Econpapers; Socionet; CEEOL; Ulrich ProQuest; Cabell, Journalseek; Scipio; Philpapers; SHERPA/RoMEO repositories; KVK; WorldCat; CrossRef; CrossCheck

2022, Volume 13, Issue 1Sup1, pages: 01-14 | https://doi.org/10.18662/po/13.1Sup1/411
Submitted: August 19th, 2021 | *Accepted for publication:* January 15th, 2022

Multi-scale Machine Learning Prediction of the Spread of Arabic Online Fake News

Fatima ALJWARI[1],
Wahaj ALKABERI[2],
Areej ALSHUTAYRI[3],
Eman ALDHAHRI[4],
Nahla ALJOJO[5],
Omar ABOUOLA[6]

[1] Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia, faljwary.stu@uj.edu.sa
[2] Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia, walkaberi.stu@uj.edu.sa
[3] Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia, aoalshutayri@uj.edu.sa
[4] Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia, eaal-dhahery@uj.edu.sa
[5] College of Computer Science and Engineering, Information system and Technology Department University of Jeddah, Jeddah, Saudi Arabia, nmaljojo@uj.edu.sa
[6] College of Computer Science and Engineering, Information system and Technology Department University of Jeddah, Jeddah, Saudi Arabia, oaboulola@uj.edu.sa

Abstract: There are a lot of research studies that look at "fake news" from an Arabic online source, but they don't look at what makes those fake news spread. The threat grows, and at some point, it gets out of hand. That's why this paper is trying to figure out how to predict the features that make Arabic online fake news spread. It's using Naive Bayes, Logistic Regression, and Random forest of Machine Learning to do this. Online news stories that were made up were used. They are found by using Term Frequency-Inverse Document Frequency (TF-IDF). The best partition for testing and validating the prediction was chosen at random and used in the analysis. So, all three machine learning classifications for predicting fake news in Arabic online were done. The results of the experiment show that Random Forest Classifier outperformed the other two algorithms. It had the best TF-IDF with an accuracy of 86 percent. Naive Bayes had an accuracy rate of 84%, and Logistic Regression had an accuracy rate of 85%, so they all did well. As such, the model shows that the features in TF-IDF are the most essential point about the content of an online Arabic fake news.

Keywords: *arabic fake news, machine learning, naive bayes, logistic regression, random forest, TF-IDF.*

How to cite: Aljwari, F., Alkaberi, W., Alshutayri, A., Aldhahri, E., Aljojo, N., & Abouola, O. (2022). Multi-scale Machine Learning Prediction of the Spread of Arabic Online Fake News. *Postmodern Openings, 13*(1Sup1), 01-14. https://doi.org/10.18662/po/13.1Sup1/411

1. Introduction

Given the ease with which information can be shared on social media platforms, as well as the fact that they allow users to express their opinions, they have grown in popularity as a source of information (Kaakinen et al., 2020). That is, anybody and everyone can express their thoughts on a subject matter that is relevant to the issues at hand. People can receive valuable solutions to their problems in some essential circumstances, but sadly, it can also be abused in other cases. That is, people have the ability to propagate fake news on social media, which is information that has been purposely misrepresented. The spread of fake news can have a severe impact on individuals and the entire population, and it is a major problem (Islam et al., 2020). Online fake news is typically spread through social media, with the result that both the general population and the government bear a significant amount of the consequences. The issue of fake news becomes increasingly important. Fake news can take advantage of multimedia material in order to deceive readers and gain publication (Creech, 2020). This can have harmful ramifications or even be used to control public gatherings. One of the most difficult components of the job is spotting bogus news regarding current events, which can be tough to do.

The majority of the information that tends to be fake news is derived from stories that are purposely manufactured and disseminated with the intent of misleading and harming Internet users. It is common practise for social media and news websites to use clickbait, which entices users to click links with eye-catching headlines or designs in order to raise ad revenue, increase readership, or profit (Aldwairi & Alwahedi, 2018). As a result, some misinformation or rumours are generated and spread via the internet, leading to other internet users believing and propagating these falsehoods themselves. When people are given incorrect information, it can lead to erroneous beliefs and actions. Detecting fake news where it first appears for bad reasons, whether political, economic, or social ones, requires a lot of attention. Machine learning and other techniques have been used extensively to identify fake news. In the context of "fake news," any information generated with the goal of contradicting the truth was deemed to be false and thus labelled phoney (Saadany et al., 2020). Fake news on the internet has two primary characteristics: first, it is not true; and second, it is not credible. Fake news in traditional media refers to publications such as newspapers, radio, and television, among other things. When attempting to determine the legitimacy of something, it is important to be cautious.

Despite the fact that they receive the news from a variety of sources, readers tend to believe what they perceive to be true (Creech, 2020).

There are many ways to tell if something isn't real in today's technologically advanced world, but fake news is still going strong because there are so many ways to tell (Fernandez, 2017). The English language has long used ways to tell if something on the Internet isn't real, but the Arabic language community has lagged far behind. That is why this study proposed to modelled the features for that facilitates the spreading of Arabic fake news. Prior to building a model that can solve the problem of spreading fake news, it is crucial to comprehend the problem associated with it. Specifically, it's important to understand how Internet users are misled and harmed by Arabic fake news items. Typically, this involves spreading of false information on social media platforms such as Facebook, Twitter, and Instagram. It is easy to distribute Arabic fake news on social media by creating an account only for the aim of sharing incorrect information and keeping it active. In part, this is owing to the ease with which malevolent accounts, such as those used to disseminate spam and harass others on social media, may be set up and maintained on the platform (Nyow & Chua, 2019). This is the most crucial motivation for this study.

2. Related Work

The study of fake news and fact-checking has taken years of dedication and work on the part of researchers, crucial to this, is in a study published by Diehl & Lee (2022) has shown that panel survey data from the United States provides strong evidence that the NFMP is related with erroneous ratings of both apolitical and pro-conservative fake news items. Aphiwongsophon and Chongstitvatana (2018) proposed the use of machine learning approaches to detect fake news, which was supported by other researchers. Using the dataset of 948,373 tweets acquired from Twitter, and through the study of the features on Twitter's news, they were able to extract 22 features from Twitter. Following the removal of duplicate data, there were 327,784 tweets that were categorised using machine learning methods. The data set was subjected to the use of the Nave Bayes, Neural Network, and support vector machine algorithms. The results demonstrate that the Nave Bayes algorithm has an accuracy of 96.08 percent. While the accuracy of the Neural Network and Support Vector Machine was 99.90 percent, the accuracy of the Support Vector Machine was 99.90 percent Similarly, this help in identifying bogus news stories in the English language (Aphiwongsophon & Chongstitvatana 2018). The system accepts the user's input and categorises it as true or false based on the answers. This is

accomplished by the creation of a database that contains the most well-known websites that deliver accurate news, into which the user can enter keywords to verify the news or a URL to validate the validity of the website. On three distinct classifiers (Naive Bayes, Logistic Regression, and Random Forest), the system was used to make predictions. The most accurate model was obtained by logistic regression, which had a 65 percent accuracy. Python and its Scikit libraries were utilised in the development of this system. Due to the fact that it is the most reliable source for machine learning methods. The model was made available online through the usage of Django. Following that, "(Sharma et al., 2020)" employ the parameter improvement network search to improve the logistic regression performance, resulting in an accuracy of up to 75%.

According to Jo et al. (2021), an estimation of the cost of battling fake news during catastrophic conditions is connected with those who have psychological harm and a strong reliance on news, and who are willing to pay an average of "9 USD" per individual. Girgis et al. (2018) proposed the use of Deep Learning Algorithms for the detection of fake news. A Recurrent Neural Network (RNN) model (both Vanilla and Gated Recurrent Unit) as well as the Long Short-Term Memory (LSTM) technique were employed in the study. Similarly, on all of the datasets studied, Sastrawan et al. (2021) developed a false news detection strategy based on deep learning CNN-RNN methodologies, and discovered that the Bidirectional LSTM architecture beat CNN and ResNet. In their assessment of numerous connected issues from a variety of sources, Sastrawan et al. (2021) discovered that there are two ways of news disclosure that have traditionally been used: content analysis (i.e., news content analysis) and social media context models (which have been more popular recently). Following the findings of this study, the researcher proposed that the ML approach be used to detect false news by integrating the characteristics of two methods: the news content and the social environment, respectively. Some Facebook likes and posts in the public English language were included in the dataset. This approach exceeded the others in terms of accuracy, which was already high at 4.8 percent when it was introduced.

Della Vedova et al. (2018) proposed a method that has been implemented in a chatbot Facebook Messenger and validated with a realistic application that has an accuracy of 81.7 percent in detecting fake news. Similarly, A temporal evolving graph neural network for false news detection has been suggested by Song et al. (2021a; 2021b), and it has been demonstrated that the proposed model outperforms existing fake news detection approaches. A study by Saadany et al. (2020) was published in

which the authors recommended utilising machine learning models to classify the linguistic aspects of Arabic Satirical Fake News in order to determine whether or not the Arabic news is true. It is demonstrated by the results that Convolutional Neural Networks (CNNs) with word embedding may achieve high accuracy of up to 98.6 percent in classification tasks. In a study conducted by Escolà-Gascón et al. (2021), it was discovered that critical thinking predicts reductions in stress levels among Spanish physicians and increases fake news identification. According to Lin et al. (2019), a system was provided that extracts 134 features from news articles using machine learning models and deep learning models that simply rely on textual input. Afterwards, they must compare their models to the baseline model in order to determine which one performs better in the fake news detection task. They gathered the Fake Newsnet dataset, which covers political news, social background, and celebrity news, among other things. The dataset contains 5,222 fraudulent news pieces, accounting for 23.8 percent of the total, and 16,729 legitimate news articles, accounting for 76.2 percent of the total.

Bodaghi and Oliveira (2021) investigated the pattern of distributing fake news on Twitter, and their findings revealed that fake news retweeters have the highest rate of modularity and intra to inter-linking. According to the findings of the research, the suggested model outperforms state-of-the-art false news detection algorithms in the adversarial multidomain multiclassification multimodal context. While investigating the underlying motives for spreading fake news in Malaysia during the COVID-19 outbreak, Balakrishnan et al. (2021) discovered that different motives drive different types of fake news sharing behaviour. The LAIR dataset is a collection of data derived from the publication "Wang (2017)." The LIAR dataset includes 12,836 short statements from POLITIFACT.COM that were manually categorized and depended on television advertisements, online news, Facebook postings, tweets, and other sources. Use of a word embedding to extract characteristics that discriminate relations between syntactic and semantic words will get the best results; that is, the relationship between two words that have the same meaning but differ in syntactic structure will yield the most accurate findings. The results demonstrated that the GRU approach achieved the highest accuracy (0.217), followed by the LSTM technique, which achieved an accuracy equal to (0.217). (0.2166). Finally, Vanilla achieves an accuracy level comparable to (0.215). Kim and colleagues (2021) presented a fake news diffusion dataset that was primarily used during the COVID-19 timeframe. According to Yuan et al. (2021), a

strategy for improving fake news identification by using domain-adversarial and graph-attention neural networks has been developed.

3. Methodology

Specifically, the research method used in this work is a step-by-step process that starts with building a multi-scale machine learning model, then data collection, then preprocessing, and finally prediction analysis (see Figure 1).

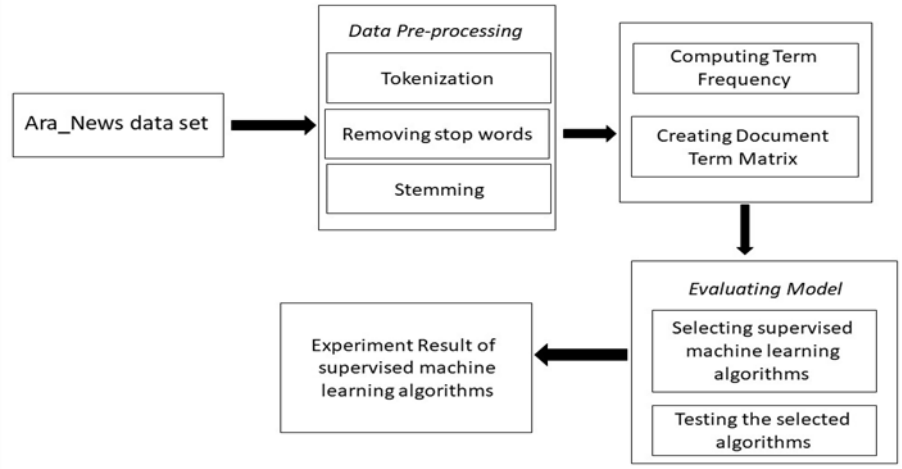


Figure 1. The proposed Methodological Steps
Source: author's own conception

The first step after establishing the model involves dataset collection, Ara_News dataset was used for this current study. This is follows by the preprocessing, prediction analysis and then evaluation.

3.1. The Multi-Scale Machine Learning Model

The multi-scale machine learning model adopted for this study comprises of models such as the "Naive Bayes" method, the "Logistic Regression algorithm", and the "Random Forest" algorithm". They are all well-known machine learning algorithms.

According to Blancquero et al. (2021), Naive Bayes is a good machine learning classification technique that relies on a probabilistic model. It may be viewed as a decent classifier that applies the Bayes theorem for generating a decision rule, as demonstrated by its ability to use prior conditions sets. This has enabled it to achieve strong and tolerable independence among all of the variables set aside for classifications in the first place. The Nave Bayes approach for classification in multivariate

analysis has proven to be a tractable and efficient method for classification. The effectiveness of Naive Bayes classifiers in text classification issues is one of the reasons for their widespread adoption and also the justification for adopting it in this study. Another justification lies with the fact that it contributes to the provision of a precise supervised learning analysis, this is also supported by Singh et al. (2020) who have discovered an extended Naive Bayes algorithm that is suitable for the detection of fake news.

Logistic regression is a good machine learning algorithm, according to Maherolia et al. (2021), Logistic regression is another important machine learning classification algorithm used for predictions of the likelihood of a categorical dependent variable, where it will be either fake or not fake, according to Maherolia and colleagues. An example of how it can be used is to describe the relationship between a collection of categorical dependent variables and a set of independent variables (De Cock et al., 2021). The approach of logistic regression is essentially employed with dichotomous dependent variables; however, it can be expanded to situations involving outcome variables with three or more categories by using the technique. In a recent study, it was discovered that logistic regression optimised data associated with the heterogeneous situation at hand (Hosmer et al., 2013). The reason why this study adopted the use of Logistic Regression as a model-training technique is justified by the fact that it is a suitable approach for the task at hand. That is predicting Arabic fake news from some targeted data sources. The dataset is fitted into the algorithm for the purposes of training and testing. It is capable of predicting the likelihood of fake news in the news datset that was used in the training model. Furthermore, Zhang et al. (2022) also revealed that logistic regression has an optimal regularisation term for identifying groups and picking relevant features in high dimensional data.

Random Forests is a great machine learning algorithm that falls under the supervised classification approach category of ensemble machine learning algorithm that works by constructing multiple decision trees on random samples of the dataset (Chen et al., 2021). Random Forests is one of the great machine learning algorithms that falls under the supervised classification approach category of ensemble machine learning algorithm that works by constructing multiple decision trees on random samples of the dataset. The reason for choosing this method is due to the fact that it makes use of the combination of results produced by a decision tree output, which allows for better confirmation of trained model (Bai et al., 2022 Another reason to use this technique is that it is part of a class of decision tree algorithms, and their expanded form for building trees, which takes into

account attributes that have been randomly chosen for each of their features, to produces series on optimized option.

Working with a variety of machine learning algorithms is essential, and it provides a strong explanation for this research's goal of developing a better prediction model. This is why every algorithm in this study makes use of multi-scales, resulting in an overall more accurate classifier.

3.2. The Dataset and Feature Extraction

The dataset used in this study was adopted from "The Ara News" dataset, which was made publicly available (Elmadany et al., 2020), and served as the basis for the dataset used in testing the proposed model of this research. It is intended for general Arabic disinformation and has been compiled from various publications covering a wide range of subjects in 15 Arabic countries, Great Britain, and the United States. It was created to be used for research purposes and is intended for general Arabic disinformation. Each sentence in the dataset is assigned a label, which can be either fake or not fake depending on the context of the utterance. Following validation, the dataset was classified as either "FAKE" or "NOT-FAKE," depending on the authenticity of the data for analysis.

The news that is considered fake has features that can be extracted and labelled in order to be trained by the machine learning multi-scale algorithm, which is used to detect fake news. Those characteristics that determine whether news is fake or not are derived from their primary purpose, which is the "intent," which was then characterised as the input text and made suitable for processing by the machine learning classifier by a machine learning classifier. The extracted features that will be used to train the classifier are defined in accordance with this attribute. Using the information from each news item, the researchers extracted features under the "Term Frequency (TF)" and "Inverse Document Frequency (IDF)" features. When used in conjunction with a dataset, TF-IDF is typically the representation of the descriptive statistic for the entire dataset, which indicates the frequencies of terms deemed to be fake and those that are not fake throughout the entire document. Based on this, a label was created to assign a value to each of the attributes that represent the presence or absence of fake or non-fake content. As a result, the widely used TF-IDF was used to develop the proposed concept for analysing the news dataset that was obtained. Even though the weight of terms or the frequency of occurrences can be used to assess how important a term is by descriptive statistics, this may not be the case for the prediction analysis, which is why the machine learning algorithm was implemented.

3.3. The Experimental Analysis

The experiment was carried out in Python because it has a large number of efficient packages that can deal with any type of data, including text. The machine learning was carried out with the help of the following libraries: Scikit-learn, Pandas, Numpy, Keras, and matplotlib.

Following the acquisition of the "Ara News dataset," pre-processing is carried out to ensure that the data is cleaned and suitable for use in the prediction model analysis. Pre-processing techniques such as text normalisation and stop words, as well as punctuation removal from the data are being used, and it has been observed that pre-processing the data results in better performance.

Following the reading of the dataset, the dataset is divided into two parts: training data and testing data. The TF-IDF method was used.

This data is gathered based on the results of the TF-IDF applied to Logistic Regression (LR), Naive Bayes, and Random Forest (RF) models to determine whether or not the news is fake. A classifier's prediction accuracy was calculated with the help of the metrics class from the Scikit-learn library, which was implemented in Python. The dataset which were partitioned into training data and testing data has been able to allow the study to determine the accuracy of the model. Similarly, when the testing data was passed to the classifier, it was examined to see what predictions were made. The predictions made by the classifier were compared.

4. Presentation of the Results and Discussion

The three different algorithms are selected for training in a similar environment based on their performance in previous training. Data has been cleaned and is ready for use. Many partitioned in a different ratio were analyzed with "Naive Bayes," "Logistics Regression," and "Random Forest" in each of the three cases. The results of the analysis were gathered in order to be able to evaluate the models more effectively. First, the train-test split procedure was carried out, and then all of the algorithms were executed one by one. When predictions are made on data that has not been used to train the model, the results indicate that an estimate of how well the Machine Learning algorithms (Naive Bayers, Logistics Regression, and Random Forest) performed is obtained. After that, the algorithms were run again to determine the most efficient method of performing the prediction; the results of this experiment allowed the researchers to evaluate the performance of the algorithms they had chosen for their predictive modelling problem.

At this point, different partitioning's of the dataset were included in the analyses to determine whether the model was well-suited to the data set in question. This method allows for more partitioning of the train and test datasets than is possible with other methods.

Furthermore, another partitioning method that was used in this study was with a "Train dataset" that contained 70% of the data and a "Test dataset" and "Validation dataset" that contained 15% of the data each are reported (see Table 1). Based on this partitioning, the Naive Bayes model on the word level was tested using the TF IDF in the first experiment, and it was found to be accurate to within 84 percent. This was followed by the application of a Logistic Regression model to the same partitioning, with the results being superior to those obtained by Naive Bayers. This time, the performance has been slightly improved, with predictions of 85 percent on word-level vectors this time around. Finally, using the same partitioning, it was discovered that Random Forest performed significantly better, achieving an accuracy of 86 percent. This represents an improvement over the previous results that has been observed.

Table 1. Experiment Result

Classifiers	Accuracy
Naive Bayes	0.844
Logistic Regression	0.859
Random Forest	0.866

Source: author's own conception

Considering how well Random forest exceeds all other algorithms in terms of performance, it is reasonable to conclude that this algorithm is particularly effective for predicting Arabic fake news. Consider the fact that the random forest strategy is advantageous when dealing with predictions among datasets that are obviously well specified and well organised as a rationale for this. Furthermore, it is an ensemble learning technique, in which multiple separate decision trees are utilised to generate a huge number of decisions during training, and which results in a class that is within the mean forecast of all the individual variables, among other things. Random Forest, which was determined to be adaptive, has become a popular choice for solving challenges involving prediction and categorization because of its versatility. Random Forest's superior performance in dealing with the prediction of fake news was recognized by the researchers in this study, which was supported by the data

5. Conclusions

A significant issue in the Arab world is the proliferation of fake news and the fabrication of stories about them. Due to the rapid advancement of technology and communication systems, individuals can distribute information without verifying it first. Detecting false news has become critical, which is why researcher began looking for a realistic answer to the rapid spread of fake news and other forms of misinformation. The purpose of this article is to examine the features of Arabic fake news and determine the model for predicting the spread of fake news using machine learning. The machine learning approach utilised was extremely accurate. Additionally, the paper demonstrates how machine learning uses multi-scale algorithms for the prediction of the spread of Arabic fake news. Classifiers of three types were used: logistic regression, naive bayes, and random forest. In the experimental analysis, a Random Forest was utilized, and it produced the best results with an accuracy of 86 percent. Both Logistic Regression and Naive Bayes performed admirably. Logistic Regression, on the other hand, outperformed Naive Bayes in forecasting the spread of fake news. While these findings are significant for predicting fake news, future research can take a different approach. To be more specific, this study recommends that in the future, feature extraction approaches be used with deep learning algorithms to achieve the greatest outcomes.

References

Aldwairi, M., & Alwahedi, A. (2018). Detecting fake news in social media networks. *Procedia Computer Science, 141*, 215-222. https://doi.org/10.1016/j.procs.2018.10.171

Aphiwongsophon, S., & Chongstitvatana, P. (2018). Detecting fake news with machine learning method. *2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)* (pp. 528-531). IEEE. https://doi.org/10.1109/ECTICon.2018.8620051

Bai, J., Li, Y., Li, J., Yang, X., Jiang, Y., & Xia, S. T. (2022). Multinomial random forest. *Pattern Recognition, 122*, 108331. https://doi.org/10.1016/j.patcog.2021.108331

Balakrishnan, V., Ng, K. S., & Rahim, H. A. (2021). To share or not to share–The underlying motives of sharing fake news amidst the COVID-19 pandemic in Malaysia. *Technology in Society, 66*, 101676. https://doi.org/10.1016/j.techsoc.2021.101676

Blanquero, R., Carrizosa, E., Ramírez-Cobo, P., & Sillero-Denamiel, M. R. (2021). Variable selection for Naïve Bayes classification. *Computers & Operations Research, 135*, 105456. https://doi.org/10.1016/j.cor.2021.105456

Bodaghi, A., & Oliveira, J. (2021). The theater of fake news spreading, who plays which role? a study on real graphs of spreading on twitter. *Expert Systems with Applications, 116110*. https://doi.org/10.1016/j.eswa.2021.116110

Chen, Y., Zheng, W., Li, W., & Huang, Y. (2021). Large group activity security risk assessment and risk early warning based on random forest algorithm. *Pattern Recognition Letters, 144*, 1-5. https://doi.org/10.1016/j.patrec.2021.01.008

Creech, B. (2020). Fake news and the discursive construction of technology companies' social power. *Media, Culture & Society, 42*(6), 952-968. https://doi.org/10.1177/0163443719899801

De Cock, M., Dowsley, R., Nascimento, A. C., Railsback, D., Shen, J., & Todoki, A. (2021). High performance logistic regression for privacy-preserving genome analysis. *BMC Medical Genomics, 14*(1), 1-18. https://doi.org/10.1186/s12920-020-00869-9

Della Vedova, M. L., Tacchini, E., Moret, S., Ballarin, G., DiPierro, M., & de Alfaro, L. (2018). Automatic online fake news detection combining content and social signals. In *2018 22nd Conference of Open Innovations Association (FRUCT)* (pp. 272-279). FRUCT. https://doi.org/10.23919/FRUCT.2018.8468301

Diehl, T., & Lee, S. (2022). Testing the cognitive involvement hypothesis on social media: 'News finds me' perceptions, partisanship, and fake news credibility. *Computers in Human Behavior, 128*(107121), 345-354. https://doi.org/10.1016/j.chb.2021.107121

Elmadany, A., Abdul-Mageed, M., & Alhindi, T. (2020). Machine generation and detection of Arabic manipulated and fake news. In *Proceedings of the Fifth Arabic Natural Language Processing Workshop* (pp. 69-84). Association for Computational Linguistics https://aclanthology.org/2020.wanlp-1.7

Escolà-Gascón, Á., Dagnall, N., & Gallifa, J. (2021). Critical thinking predicts reductions in Spanish physicians' stress levels and promotes fake news detection. *Thinking Skills and Creativity, 42*, 100934. https://doi.org/10.1016/j.tsc.2021.100934

Fernandez, P. (2017). The technology behind fake news. *Library Hi Tech News, 34* (7), 1-5.https://doi.org/10.1108/LHTN-07-2017-0054

Girgis, S., Amer, E., & Gadallah, M. (2018). Deep learning algorithms for detecting fake news in online text. In *2018 13th International Conference on Computer Engineering and Systems (ICCES)* (pp. 93-97). IEEE. https://doi.org/10.1109/ICCES.2018.8639198

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*,Vol. 398. John Wiley & Sons.

Islam, F., Alam, M. M., Hossain, S. S., Motaleb, A., Yeasmin, S., Hasan, M., & Rahman, R. M. (2020). Bengali Fake News Detection. In *2020 IEEE 10th International Conference on Intelligent Systems (IS)* (pp. 281-287). IEEE. https://doi.org/10.1109/IS48319.2020.9199931

Jo, H., Park, S., Shin, D., Shin, J., & Lee, C. (2021). Estimating Cost of Fighting against Fake News during Catastrophic Situations. *Telematics and Informatics*, 101734. https://doi.org/10.1016/j.tele.2021.101734

Kaakinen, M., Sirola, A., Savolainen, I., & Oksanen, A. (2020). Shared identity and shared information in social media: development and validation of the identity bubble reinforcement scale. *Media Psychology, 23*(1), 25-51. https://doi.org/10.1080/15213269.2018.1544910

Kim, J., Aum, J., Lee, S., Jang, Y., Park, E., & Choi, D. (2021). FibVID: Comprehensive fake news diffusion dataset during the COVID-19 period. *Telematics and Informatics, 64*(1), 101688. https://doi.org/10.1016/j.tele.2021.101688

Lin, J., Tremblay-Taylor, G., Mou, G., You, D., & Lee, K. (2019). Detecting fake news articles. In *2019 IEEE International Conference on Big Data (Big Data)* (pp. 3021-3025). IEEE. https://doi.org/10.1109/BigData47090.2019.9005980

Mehrolia, S., Alagarsamy, S., & Solaikutty, V. M. (2021). Customers response to online food delivery services during COVID-19 outbreak using binary logistic regression. *International journal of consumer studies, 45*(3), 396-408. https://doi.org/10.1111/ijcs.12630

Nyow, N. X., & Chua, H. N. (2019). Detecting fake news with tweets' properties. In 2019 IEEE Conference on Application, Information and Network Security (AINS) (pp. 24-29). IEEE. https://doi.org/10.1109/AINS47559.2019.8968706

Saadany, H., Mohamed, E., & Orasan, C. (2020). Fake or real? A study of Arabic satirical fake news. *Proceedings of the 3rd International Workshop on Rumours and Deception in Social Media (RDSM),* (pp. 70-80). Association for Computational Linguistics. https://aclanthology.org/2020.rdsm-1.0

Sastrawan, I. K., Bayupati, I. P. A., & Arsa, D. M. S. (2021). Detection of fake news using deep learning CNN-RNN based methods. *ICT Express.* In Press https://doi.org/10.1016/j.icte.2021.10.003

Sharma, U., Saran, S., & Patil, S. M. (2020). Fake News Detection using Machine Learning Algorithms. *International Journal of Engineering Research & Technology (IJERT) NTASU, 9*(03), 509-518, https://www.ijert.org/fake-news-detection-using-machine-learning-algorithms

Singh, M., Bhatt, M. W., Bedi, H. S., & Mishra, U. (2020). Performance of Bernoulli's naive bayes classifier in the detection of fake news. *Materials Today: Material Today Proceedings, 49*(10), 1865-1870. https://doi.org/10.1016/j.matpr.2020.10.896

Song, C., Ning, N., Zhang, Y., & Wu, B. (2021b). Knowledge augmented transformer for adversarial multidomain multiclassification multimodal fake news detection. *Neurocomputing, 462*, 88-100. https://doi.org/10.1016/j.neucom.2021.07.077

Song, C., Shu, K., & Wu, B. (2021a). Temporally evolving graph neural network for fake news detection. *Information Processing & Management, 58*(6), 102712. https://doi.org/10.1016/j.ipm.2021.102712

Wang, W. Y. (2017). "Liar, liar pants on fire": A new benchmark dataset for fake news detection. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics 2(P17-2067)* (pp. 422–426). Association for Computational Linguistics https://doi.org/10.18653/v1/P17-2067

Yuan, H., Zheng, J., Ye, Q., Qian, Y., & Zhang, Y. (2021). Improving fake news detection with domain-adversarial and graph-attention neural network. *Decision Support Systems, 151*, 113633. https://doi.org/10.1016/j.dss.2021.113633

Zhang, P., Wang, R., & Xiu, N. (2022). Multinomial logistic regression classifier via lq, 0-proximal Newton algorithm. *Neurocomputing, 468*, 148-164. https://doi.org/10.1016/j.neucom.2021.10.005