

Fake News Detection: A Survey

Dr. Rachna Somkunwar¹, Anil Kumar Gupta², Faizan Shikalgar³, Shubham Maral⁴, Arpit Paithankar⁵,
Rohan Kapdi⁶

¹Associate Professor, Dr. D. Y. Patil Institute of Technology, Pimpri Pune, Maharashtra, India

²CDAC, Senior member IEEE

^{3,4,5,6}BE, Computer Engineering, Dr. D. Y. Patil Institute of Technology, Pimpri, Pune, Maharashtra,
India

Abstract- To address the modern society's need for reliable News from media has been a challenge since the growth of recent technologies and advancements made in it. With the technological advancements, online news is more exposed to users all around the world and facilitates to increase in spreading disinformation online. Various artificial intelligence tools have been used to solve this problem of fake news identification. Fake news leads to convince the reader to believe fake information which proves these articles difficult to read. Fake news detection problem can be solved by the help of Artificial Intelligence algorithms which includes Machine Learning algorithms. Our system is used for detecting wrong information is based on stance detection which we have implemented using LSTM. We aim for a way to distinguish articles from unknown sources such as agreeing or disagreeing with known sources.

Index Terms - Fake News, LSTM, Stance detection, TensorFlow.

1.INTRODUCTION

Fake news is regarded as one of the biggest threats to society. It has decreased common people's trust in governments. The speed at which news is being produced is huge and fast, that's why it is difficult to detect fake news. The uncontrollable propagation of fake news through the net is not controllable in this digital era. Fake news propagation cannot be tolerated as the bad impacts of it to the society and individuals is worry-some. Moreover, this will lead us to more problems and potential threats and luring users to share lies made from false news through their social media. Within the Indian context, there is a lack of a platform for fake news detection in different language articles and most Indians received news through their social messaging applications. Many sites and social media have devoted efforts to identify disinformation.

For example, Facebook and other social media handles encourage users to report any fake news they encounter on their media and also report the user. Manual fact-checking is also used by fact-checking websites such as Snopes, Fact Check, Politi Fact, and Full Fact. Although social media can be accessed from anywhere and at any time and is also free, it provides anonymity while expressing our opinion, therefore, leading to a lack of accountability which greatly reduces the authenticity of data received from them as compared to a newspaper or any other trusted news source. Lack of constant supervision and an overseeing authority has allowed the wrongdoers to run amok and spread false information [1]. Fake news affects every corner of society. Presidential elections in 2016 were notable for incidents of fake news, including propaganda created to mislead readers or to generate views on websites, or steer public opinion [2]. The fast propagation of false information has the potential for very negative impacts on both individuals & society. Fake news creates the way for duping others and promoting ideologies and narratives. Some people who generate fake information profit by earning money with the number of interactions on their article or their website. Propagating fake news has various intentions, such as, to gain favor in politics, for business and products, it may also be done to gain revenge. But the fake news issue may be tackled with the help of various machine learning algorithms and artificial intelligence. The reason behind this is that the news articles contain a large number of words that are difficult to process but with the help of Natural Language Processing the important keywords are extracted and continue the processing on the extracted keywords. Even the simulation of human behavior like decision making can be achieved by machine learning

and the decision between fake and fact is the basic objective of the fake news detection problem [3].

The main advantage of our system is that we can accept user input as any statement or link to the article. We also search the event register database containing millions of news articles and update their details regularly with news headlines, to remove news articles based on a claim or link provided by a user. In this way, our system acts as a search engine that finds the truth and returns links to relevant articles and the nature of that text in that claim. Identifying fake news is still a challenge in deep learning methods such as the Convolutional Neural Network (CNN), Recurrent Neural Network (RNN).

2.LITERATURE REVIEW

Van-Hoang Nguyen [4], they have proposed FANG (Factual News Graph) in which they represent the social context with the help of a graph. Their focus is on representation learning. FANG got an accuracy of 75% while testing their module.

One advantage of FANG is that it is scalable and it does not have to maintain all nodes and it is fast as compared to other models that the authors studied.

However, they were constructing entity and interaction features before they were passed to FANG, this created errors from the upstream task.

V. Mazzeo [5], authors aim to find fake and misleading content in web search engines by analysing textual information. They have used a dataset of covid-19 news and firstly apply re-sampling techniques then they train their ML model which is built using different classification techniques. URLs associated with news articles have lexical and host-based features which need to be extracted. The efficiency of classification models can be increased by detecting malicious URLs. They plan to construct more discriminative features to detect fake news, by profiling malicious sources of information based on domains and they can blacklist domains that regularly propagate fake news. Their URL analyser was very helpful for the classifier to detect phishing and clickbait URLs. One disadvantage is that they are dealing with imbalanced data and that may be a reason why accuracy may represent a misleading indicator, and it would make it difficult for the classifier to predict the result accurately.

Bhattacharyya, P. [6], authors are proposing a deep transfer learning approach where they find the relationship between headline and body. Their idea is to check whether the body agrees with the claim made. In the future, they would: incorporate the external knowledge into the existing system, take into account, the length difference between the headlines and bodies and enrich the best performing model by incorporating the relevance score between the headline and body texts. The authors have created 2 models one based on SNLI trained model and the other is based on Bi-LSTM and the use of max-pooling increases accuracy by almost 15%. One disadvantage to their model is that they have a huge imbalance in their data as 60% of data is termed unrelated. This may cause the classifier to predict unrelated even though the claim belonged to another label.

Gaglani, J. [7], they have used WhatsApp chats database for classification. They have used NLP techniques for analyzing the messages in the chats. They have trained their model which is based on LSTM. Their model is trained on the Switchboard Dialogue Act Corpus used for the Dialogue Act classification. They have also done feature extraction which is carried out using the Gensim Python package which is based on the TextRank algorithm which in turn helps to find out the summary from the main model. In future work, they would like to extend this to other social media platforms such as Facebook and Twitter. One problem with their system is that they cannot accurately detect fake news in a language other than English if it is extensively used.

Dhruv Khattar [8], Authors have used Encoder- It converts text and image input into latent feature vectors, Decoder – It produces image and text from latent feature vectors, and Fake news Detector (Neural Net) - uses the latent features to predict Fake or real news. It has the advantage that it can learn correlations across the modalities to extract joint feature representation. The disadvantage is that it doesn't use user characteristics and tweet propagation data as features.

S. Singhal [9], Authors have created a model where features are extracted using the language transformer model (bert) and ImageNet Model (VGG). Fusion Model (Neural Net) classifies the features as real or fake. Its advantage is that without using an event discriminator or attention mechanism it gives an

additional 6% accuracy to the baseline models. The disadvantage is that the SPOTFAKE model might be dataset-specific the authors have used different hyper parameters for the two datasets.

Dong Xishuang [10], Two-path deep semi-supervised learning: One path is for supervised learning and the other one is for unsupervised learning. Shared CNN is built to extract low-level features for feeding in two paths. Both labeled and unlabeled data are used for mean squared error loss. Only labeled data is used for cross-entropy loss. The framework is optimized using the weighted sum of these two losses. Implemented models could detect fake news from these two benchmarks, LIAR and PHEME effectively using very limited labeled data and a large amount of unlabelled data. This paper framework doesn't examine other NLP tasks such as sentiment analysis and dependency analysis.

C. Helwe [11], Deep co-learning, a semi-supervised end-to-end deep learning approach: A small training data is used initially to train the classifiers in a supervised fashion. Then some unlabeled data is labeled by using each classifier and other classifiers are re-trained using this label in a semi-supervised approach. This model performed significantly better than multiple baselines, e.g. ensemble models and deep learning models which are fully supervised. Not many linguistic features are considered in this work. Lack of datasets for the Arabic language as well as lack of annotated datasets in Arabic.

3.MATERIALS AND METHODOLOGY

A. Data Collection:

In our system, we have gathered data from various reputed news sources and we also have used event registry API which contains more than 1 million articles from 150000 different sources. Our dataset contains news articles from around the globe. We also perform web scraping to extract the text body from the URLs to predict the fake news. The local database that we are using is has a column stance value associated with it to determine the reputation of the source where that article came from.

B. Data Pre-Processing

In our system, to parse the article and to perform keyword extraction we used Microsoft's Azure cognitive and IBM's Natural Language Processing.

The extracted articles from event registry's database were also passed to ML model. In our system to pre-process the data we use NLP techniques such as Text Embeddings, Bag of Words, Machine Translation, etc. to remove stop words, to lower sentences, to remove punctuations and to transform the data into model understandable form. We also use googles word2vec In Word2vec, we have a large collection of text in which every word is represented by a vector which in turn helps improve model accuracy and training time. Articles which belonged to different languages were converted to English using google translate API. Null values are handles by removing the corresponding values. We have also transformed our articles into features on which model can be trained.

C. Model Training

To train our model we have used supervised learning techniques. We split our data into 70-30 i.e 70% would be our training data and 30% would be the training data. In our model we also perform feature extraction of the news articles so as to help the final model predict accurately.

In our model, we have used stance detection technique to predict the fake news. The classifier is a MLP with one hidden layer of 100 units with rectified liner unit (ReLU) activation. Connected layer and a softmax activation to generate the output. We are comparing the body of text with the claim and our ML model outputs whether text is related or unrelated to the claim. In our model the stance plays an important role to determine the prediction and more reputed sources have more importance.

4.METHODS

A. Stance Detection:

Stance detection has been defined as automatically detecting whether the author of a piece of text is in favor of the given target or against it. In stance detection, systems are to determine author's favorability towards a given target and the target even may not be explicitly mentioned in the text. Moreover, the text may express positive opinion about an entity contained in the text, but one can also infer that the author is against the defined target.

B. TensorFlow:

TensorFlow is developed by Google. Model or program development in TensorFlow consists of two stages: static computation graph definition and running a computational session of this graph. Based on this model representation, it is easy to evaluate data in parallel in pipelines, especially on GPUs (Graphics Processing Units), allowing, creating, and learning deep NN (Neural Network). The core of TensorFlow calculations creates directed graphs structure used for computation, where a node can be, for example, value or math function, and edges are tensors. Each node has zero to N inputs and outputs. TensorFlow assigns kernels to the node representing math function, which

contains the implementation of that function on a particular device [12].

C. Long Short-Term Memory:

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can not only process single

data points but also entire sequences of data for example, LSTM is applicable to tasks such as unsegmented, connected handwriting recognition speech recognition and anomaly detection in network traffic or IDSs (intrusion detection systems)[13].

5.COMPARISON OF CLASSIFICATION METHODS

Comparison Table							
Paper Id	Method	Accuracy	Precision	Recall	F1 Score	Advantages	Disadvantages
[2]	Logistic Regression	0.94	0.94	0.94	0.94	The logistic regression uses a more complex cost function	Their model only works for twitter news analysis. Their model is not trained to extract real news published on verified sources.
[5]	Naive Bayes	0.96	0.76	0.86	0.81	They also perform malicious URL detection which prevents clickbait and also prevents model to run unnecessary check.	The dataset that they have used to train the model is imbalanced. Therefore, accuracy may represent a misleading character.
[14]	Ensemble Model + Heuristic Post-Processing	0.9283	0.9883	0.9883	0.9883	By using ensemble method with soft voting by using prediction vectors from vectors has increased the accuracy	The authors have only focused on data from twitter and their method works better if user is verified but does not work well if given a random claim to be checked.
[15]	K-Nearest Neighbor Classifier	0.79	0.75	0.79	0.785	KNN algorithm using data mining is simple to implement and is efficient.	They have trained their model with very less number of articles. This decreases the accuracy of the model and makes it difficult to understand the pattern in fake news.
[16]	Deep Net	0.952	0.909	0.952	0.930	They have used 7 hidden layers which removes the body which does not match with claim and thereby increases accuracy	The model has performed well on training data but does not perform well on testing data
[17]	Random Forest	0.78	0.68	0.78	0.73	They have used NewsAPI and GraphAPI to collect news stories from social media and to perform synonym-based feature extraction.	The accuracy of their model is very low as compared to other models in the practise.
[18]	LR + RF + ADA	0.9194	0.91	0.93	0.92	Their model does not require any metadata	The classifiers do not perform well on word vector features

						about users, and it prevents user privacy.	when skip-gram and CBOW is used as input to classifiers
--	--	--	--	--	--	--	---

6.CONCLUSION

Fake news remains an open topic for research. More contributions are needed, especially to automate the growing growth of Web knowledge. Our work has tried to get closer to the realization of non-news stories using a simple model based on Stance Detection using LSTM which uses NLP techniques to pre-process data using BOW, TF-IDF, and Google's word-2-vec. This work is the first step towards an ongoing investigation of factors to improve the availability of status in non-existent issues. We have demonstrated how stance detection can help in detecting fake news. We also demonstrate that our system helps as a discovery tool rather than anything else.

7.FUTURE WORK

As future work, we plan to use our system as an extension in the web browser so that the user when reading the article receives information about the authenticity of the news. We also plan to use Gradient Boosted Decision Tree and XG-Boost to have much greater accuracy and so that our model performs much faster than our current loading time (~15s). We plan to bring it down to 5s using those methods.

REFERENCE

- [1] Bhutani, Bhavika, et al. "Fake news detection using sentiment analysis." 2019 Twelfth International Conference on Contemporary Computing (IC3). IEEE, 2019.
- [2] Kudarvalli, Harika, and Jinan Fiaidhi. "Experiments on Detecting Fake News over Social Media using Machine Learning Algorithms." (2020).
- [3] Thakur, Ashokkumar, et al. "MYTHYA: Fake News Detector, Real Time News Extractor and Classifier." 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184). IEEE, 2020.
- [4] Nguyen, Van-Hoang, et al. "FANG: Leveraging social context for fake news detection using graph representation." Proceedings of the 29th ACM International Conference on Information & Knowledge Management. 2020.
- [5] V. Mazzeo, A. Rapisarda, G. Giuffrida, "Detection of Fake news on covid-19 on web search engines", arXiv:2103.11804v, 2021.
- [6] Saikh, Tanik, et al. "A deep learning approach for automatic detection of fake news." arXiv preprint arXiv:2005.04938 (2020).
- [7] Gaglani, Jaynil, et al. "Unsupervised WhatsApp Fake News Detection using Semantic Search." 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2020.
- [8] Khattar, Dhruv, et al. "Mvae: Multimodal variational autoencoder for fake news detection." The World Wide Web Conference. 2019.
- [9] Singhal, Shivangi, et al. "Spotfake: A multi-modal framework for fake news detection." 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM). IEEE, 2019.
- [10] Dong, Xishuang, Uboho Victor, and Lijun Qian. "Two-Path Deep Semisupervised Learning for Timely Fake News Detection." IEEE Transactions on Computational Social Systems (2020).
- [11] Helwe, Chadi, et al. "Assessing arabic weblog credibility via deep co-learning." Proceedings of the Fourth Arabic Natural Language Processing Workshop. 2019.
- [12] Oujezsky, Vaclav, et al. "GPON Traffic Analysis with TensorFlow." 2020 43rd International Conference on Telecommunications and Signal Processing (TSP). IEEE, 2020.
- [13] Chakraborty, Shayak, et al. "Study of Dependency on number of LSTM units for Character based Text Generation models." 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA). IEEE, 2020.
- [14] Das, Sourya Dipta, Ayan Basak, and Saikat Dutta. "A Heuristic-driven Ensemble Framework for COVID-19 Fake News Detection." arXiv preprint arXiv:2101.03545 (2021).
- [15] Kesarwani, Ankit, Sudakar Singh Chauhan, and Anil Ramachandran Nair. "Fake News Detection on Social Media using K-Nearest Neighbor Classifier." 2020 International Conference on

Advances in Computing and Communication Engineering (ICACCE). IEEE, 2020.

- [16] Kaliyar, Rohit Kumar, et al. "DeepNet: An Efficient Neural Network for Fake News Detection using News-User Engagements." 2020 5th International Conference on Computing, Communication and Security (ICCCS). IEEE, 2020.
- [17] Hlaing, May Me Me, and Nang Saing Moon Kham. "Defining News Authenticity on Social Media Using Machine Learning Approach." 2020 IEEE Conference on Computer Applications (ICCA). IEEE, 2020.
- [18] Reddy, Harita, et al. "Text-mining-based fake news detection using ensemble methods." International Journal of Automation and Computing ,2020.