Rivista Italiana di Filosofia Analitica Junior **Vol 14, No. 2 (2023)**
ISSN: 2037-4445

The Reveal of Fake News on Twitter Using Credibility Analysis and Multimodal Approach

Mr. S. Uthayashangar [1], B. Devibala[2], B. Kalyani Varshini[3], P. Mounishaa[4]

[1] Assistant Professor, Department of Information Technology
Manakula Vinayagar Institute of Technology,Puducherry,India.
[2,3,4] UG Student, Department of Information Technology
Manakula Vinayagar Institute of Technology,Puducherry,India.

Abstract

Misleading news is becoming a big threat to news accuracy since it is so easy to communicate and spread across various social media platforms. Fake news has a tremendous impact on society as a result of the proliferation of online social media. Twitter is one of the social media platforms that is regularly used to propagate misleading information during political campaigns. In this study, we propose a system for detecting bogus news that incorporates text and image analysis techniques, as well as credibility analysis. The system is divided into three modules: text processing, image processing, and credibility analysis**.** We pre-process textual data in the text processing module using the word embedding technique, which allows us to extract significant features and identify potential indicators of fake news. We focus on convolutional neural networks (CNN) to extract useful characteristics from images. The credibility analysis module examines Twitter data, gathering real-time tweets and employing various analysis approaches such as hashtag analysis, user analysis, and retweet analysis. We assess the credibility of the tweets by considering these aspects, providing additional insights into the overall credibility of the news events. The system has a 97% accuracy rate in categorising news item as fake or true.

Keywords: Social-Media, Deep Learning, Fake news, Text Analysis, Image analysis, Credibility Analysis.

1. Introduction

Online social networking sites have surpassed traditional media and television in the age of digital computing when it comes to spreading information or breaking news about any ongoing event or occurrence. One can upload any kind of stuff on social networking sites because they provide users the ability. These contents may be true information or a false rumor. The people who are audience to such fake content through their devices might considered it to be as true and trustworthy information. The fake/ rumored content often includes false spread of the information which diminishes the true facts and misled the society on a larger scale which might bring grave threats for any person and even end up into disastrous condition. These days Twitter, Facebook, Instagram, are the most popular social networking sites with millions of user's databases.

These platforms were created with society in mind and help people communicate with each other in meaningful ways virtually all over the world. Contrary to their intentions, these platforms are frequently exploited to propagate false information, influence others to think incorrectly, and spread rumors. Therefore, it is vitally necessary to develop a model that aids in assessing the reliability of the content submitted on online social networking sites. Many researchers have put in a lot of effort to develop a model that not only help in finding the reported content but also in determining its credibility score. This allows for the free exchange of information in a variety of settings, including those related to regular life, local and international news, and the advancement of new technology. Social media and other online platforms enable users to connect, exchange, and create information without formally giving a source. Although there is still much room for improvement in this area of research, our work focuses

on building a model utilizing machine learning methods for classification and clustering in order to categorize tweets according to their credibility.

A significant number of well-known people utilise the trendy social networking site Twitter to communicate with their fans and followers, regardless of the area in which they work. They simply tweet about a variety of topics on Twitter. According to audience figures, Twitter has 396.5 million users worldwide. With 206 million daily active users, Twitter is sure to keep people interested in what others are tweeting. According to statistics, there were around 27.3 million Twitter users in India as of March 2023.According to a study, misleading information spreads more quickly on Twitter than actual information. For example, misleading news reports are 70% more likely to be retweeted than accurate news items. Furthermore, the time it takes for accurate news to reach 1,500 individuals is roughly six times longer than the time it takes for fraudulent stories to do the same. Consequently, utilising the latest deep learning approaches, it is necessary to automatically identify the fake news that is propagating on Twitter.

2. Literature review

In recent years, the detection of fake news on social media networks has become a major study topic. This review of the literature covers seven separate research that investigated various strategies for detecting fake news on social media sites.

M.Sreedevi, G.Vijay Kumar, B. Aruna, Arvind Yadav [1] (2021) created data mining models to detect fake news in datasets, and tested the accuracy of about five different categorization algorithms in a machine learning algorithm.

Yudith Cardinale, Irvin Dongo, German Robayo, David Cabeza, Ana Aguilera, Sergi Medina [2] (2021) proposed a framework for performing credibility analysis of posts on social media automatically and in real-time, called T-CREo (Twitter Credibility Analysis Framework), that is based on three levels of credibility: text, user, and social. Their system was assessed for functionality and scalability and demonstrated potential for real-world applications.

Uma Sharma, Saran Sidarth & Patil Shankar [3] (2020) developed a fake news detection system that uses data from the user to determine if the news is real or fake. They utilised various NLP and machine learning techniques, with logistic regression achieving the best accuracy rate of 75% after applying grid search parameter optimisation.

YendheRutuja, Kasturi Anusha K, Jatar Rajshri S, Prof.Aparana Patil [4] (2020) developed a web application that uses a knowledge base of tweet patterns and various machine learning and knowledge-based methodologies to determine whether a tweet is real or not and the sentiment it expresses.

Emma Cueva, Grace Ee Akshat Lyer, akexandra Pereia, Alexander Roseman, Dayrene Martinez [5] (2020) compared the predictive abilities of the Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Natural Language Processing (NLP) networks for fake news detection on Twitter. They found that all three models achieved high accuracy in detecting fake news, but only the NLP model could identify satire as fake news.

Anastasia Giachanou, Guobiao Zhang, Paolo Rosso [6] (2020) used a multimodal method that combined textual, visual, and semantic information to distinguish between fake and real posts. Their system was based on a neural network and achieved high accuracy rates.

Monther Aldwairi and Ali Alwahedi [7] (2018) proposed a way to recognise and filter out websites that contain inaccurate and misleading content. Utilising logistic regression, they were able to identify false postings with 99.4% accuracy utilising a few, simple aspects of the title and content.

Overall, these studies highlight the importance of developing effective methods for detecting fake news on social media platforms that use a variety of methodologies such as machine learning, natural language processing, and multimodal methodology. The findings of these studies could aid in the development of more accurate and efficient algorithms for detecting false news and increasing the reliability of information posted on social media networks.

3. Methodology

3.1 Existing System

The textual, visual, and semantic components make up the multimodal multi-image false news detecting system. To represent the post's content, the textual component employs a Bidirectional Encoder Representations from Transformers (BERT) model. The visual component uses the VGG-16 model to extract features from a sequence of images and then applies a Long Short-Term Memory (LSTM) model to obtain a hidden vector sequence. Utilising word2vec embeddings and cosine similarity, the semantic component computes text-image similarity. The representations from the three components are concatenated and fused with an attention mechanism, and the result is passed through a SoftMax layer to produce a probability representation for each feature. A Multi-layer Perceptron (MLP) classifier is used to determine whether the item is fraudulent or not.

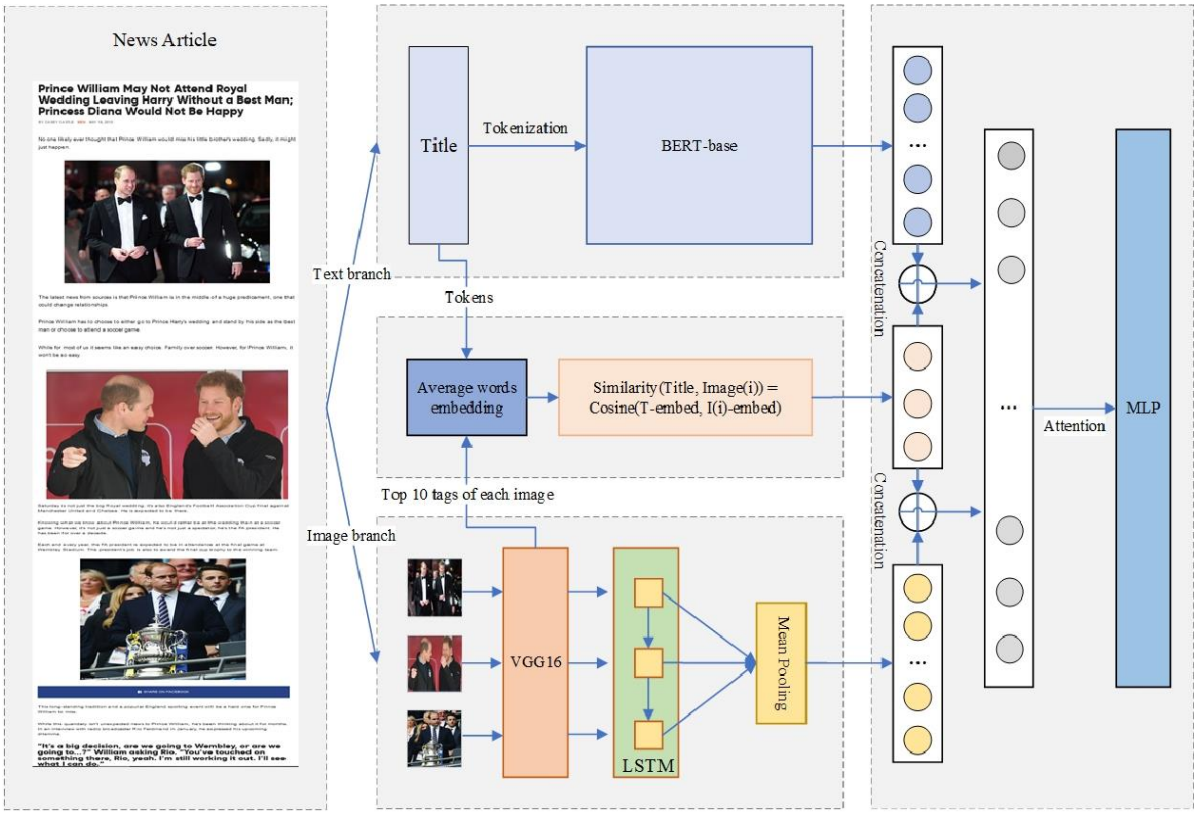

Fig 1: Architecture diagram of multimodal fake news detection

In existing system, they compared the models and find out the best approach. They claimed that BERT performed better than all other unimodal techniques, achieving an accuracy of 78% across all classes. Furthermore, the multimodal technique achieves an accuracy of 87%, which is higher than that of the

other unimodal, leading them to believe that Multimodal yields better results. However, they only utilised this methodology to identify false news in a smaller dataset, the Fakeddit dataset. They didn't employ any more sophisticated datasets. Since Twitter is a medium used by the majority of people worldwide and each person's slang is unique, we are leveraging Twitter's dataset, which is more complicated than the Fakeddit dataset, to construct a fake news identification system for the social media site. Therefore, it is difficult for the current method to identify fake news in the Twitter dataset. Additionally, in the multimodal method, pre-trained networks were not used to provide visual representations.

3.2 Proposed System

The proposed effort attempts to create a complete Fake News Detection System that integrates both text and picture analysis approaches. The system consists of three main modules: a Text Processing Module, an Image Processing Module, and a Credibility Analysis Module.

In the Text Processing Module, the system will gather textual data from various sources, such as online news articles and social media posts. The data will undergo preprocessing steps to remove noise, tokenize the text, and perform other necessary text cleaning tasks. Feature extraction techniques, including word embedding, will be applied to capture the semantic meaning and context of the text.

Similarly, the Image Processing Module will focus on analyzing the visual content associated with the news articles or social media posts. The system will extract relevant features from the images using Convolutional Neural Networks (CNNs) to identify patterns, objects, and potential manipulations or alterations in the images.

The extracted features from both text and image processing modules will be used as inputs to train a machine learning model. The proposed work suggests using techniques such as Adaptive Boosting and Support Vector Machines (SVM) for text analysis, while CNNs will be employed for image analysis. These models will learn from labeled datasets, where fake and genuine news articles or social media posts are appropriately classified.

Various assessment criteria, including as F1 score, precision, and recall, will be utilised to assess the system's performance. By taking into account both textual and visual clues, the suggested study seeks to achieve high accuracy in differentiating between fake and authentic news.

A Credibility Analysis Module will also be created to evaluate the credibility of news sources and social media accounts. This module will leverage Natural Language Processing (NLP) techniques to analyze user profiles, hashtags, and retweet patterns to identify potential indicators of credibility or lack thereof. The credibility analysis module gets live data from the Twitter API, which is then pre-processed and clustered. By using RNN and NLP techniques, the credibility of tweets is analysed through hashtag analysis, user analysis, and retweet analysis. Hashtags in the tweets give an indication of the topics being discussed and the communities targeted. User analysis considers factors like the number of followers, consistency of posting, and engagement levels to assess the user's credibility. Retweet analysis evaluates how many times a tweet has been retweeted and by whom to determine its credibility.

The proposed work has the potential to make a significant contribution to ongoing efforts to combat fake news. The technology intends to deliver more robust and accurate identification of fake news by merging text and picture analysis techniques and incorporating credibility assessment, allowing users to make informed decisions based on reputable information sources.

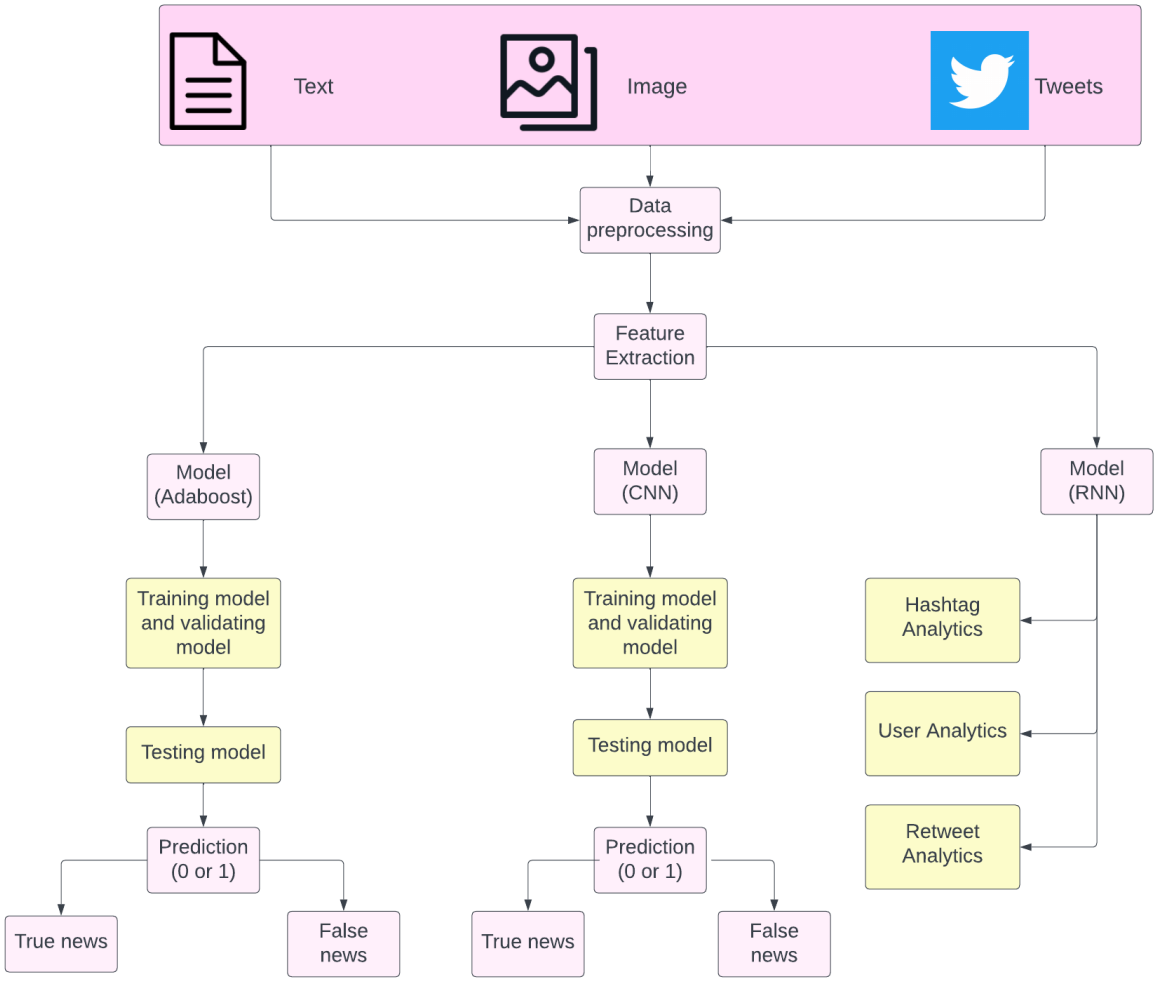

Figure 2: Architecture diagram of reveal of fake news on Twitter using credibility analysis and multimodal approaches

4. RESULT AND DISCUSSION

The results of our proposed system for fake news detection are highly promising, with an impressive accuracy rate of 97%. This indicates that our system is effective in distinguishing between true and fake news. The text processing module and image processing module play crucial roles in extracting relevant features from the datasets. By utilizing word embedding and error level analysis, we were able to capture important linguistic and visual cues that contribute to the identification of fake news. The training phase involved the utilization of Adaptive Boosting, SVM, and CNN algorithms. These algorithms have shown their efficacy in learning from the provided data and making accurate predictions. By combining these algorithms, we were able to leverage their individual strengths and enhance the overall performance of our system. The credibility analysis module, which focused on Twitter data, provided valuable insights into the credibility of tweets. By analyzing hashtags, users, and retweets, we gained a deeper understanding of the context and trustworthiness of the information being shared on the platform.

Overall, our proposed system demonstrates a high level of accuracy in detecting fake news, indicating its potential for real-world applications. It can serve as a powerful tool in combating the spread of misinformation and ensuring the credibility of news sources. However, further evaluation and testing on diverse datasets are necessary to validate the robustness and generalizability of our system.

Additionally, continuous updates and improvements to the system will be crucial to keep up with evolving tactics employed by purveyors of fake news.

Here are some key differences between the proposed system and the existing system: Credibility Analysis: The proposed system incorporates a credibility analysis module that assesses the reliability and trustworthiness of news sources and content. It utilizes techniques such as sentiment analysis, user profiling, and network analysis to evaluate credibility. In contrast, the existing system primarily focuses on the classification and detection of fake news but does not explicitly consider credibility assessment. Enhanced Feature Extraction: The proposed system incorporates advanced feature extraction methods, such as word embedding and error level analysis for textual data, and CNNs for image analysis. These techniques capture more nuanced patterns and characteristics within the data, potentially improving the system's detection capabilities. The existing system does not mention these specific techniques or advancements in feature extraction.

5. CONCLUSION

In conclusion, our proposed system that consists of text processing module, image processing module, and credibility analysis module has achieved an accuracy of 97% in detecting fake news. We achieved this by preprocessing the data, extracting features using word embedding and error level analysis, and training the model using Adaptive Boosting, SVM, and CNN algorithms. The credibility analysis module provided additional insights by analyzing the hashtags, users, and retweets of live Twitter data. We believe that our method can help detect and counteract the spread of fake news in both text and image-based media. In future the accuracy of this model can be increased by training it with much more data. In future the tampered part of the image can also be detected. We can also include other media formats such as video and audio data. We can also Expand the credibility analysis module to include more granular factors such as source reputation, fact-checking references, and cross-referencing information from multiple reliable sources. This will provide users with a more comprehensive credibility assessment of tweets and news articles.

REFERENCE

[1] Marriboyina, Sreedevi & Gogulamudi, Vijay & Kumar, K. & Aruna, B. & Yadav, Arvind. (2021). Prediction of Fake Tweets Using Machine Learning Algorithms. 10.3233/APC210195.

[2] Cardinale, Yudith & Dongo, Irvin & Robayo, German & Cabeza, David & Aguilera, Ana & Medina, Sergio. (2021). T-CREo: A Twitter Credibility Analysis Framework. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.30.

[3] Sharma, Uma & Saran, Sidarth & Patil, Shankar. (2020). Fake News Detection Using Machine Learning Algorithms. 2320-2882.

[4] YendheRutuja, Kasturi Anusha K,Jatar Rajshri S, Prof.Aparana Patil, Fake News Detection and Sentiment analysis in Twitter,Volume 5,Issue 9,September 2020,ISSN (Online) 2456-0774.

[5] E. Cueva, G. Ee, A. Iyer, A. Pereira, A. Roseman and D. Martinez, "Detecting Fake News on Twitter Using Machine Learning Models," 2020 IEEE MIT Undergraduate Research Technology Conference (URTC), 2020, pp. 1-5, doi: 10.1109/URTC51696.2020.9668872.

[6] Giachanou, A., Zhang, G., Rosso, P. (2020). Multimodal Fake News Detection with Textual, Visual and Semantic Information. In: Sojka, P., Kopeček, I., Pala, K., Horák, A. (eds) Text, Speech, and Dialogue. TSD 2020. Lecture Notes in Computer Science (), vol 12284. Springer, Cham.

[7] Monther Aldwairi, Ali Alwahedi, Detecting Fake News in Social Media Networks, Procedia Computer Science, Volume 141,2018, Pages 215-222,ISSN 18770509,https://doi.org/10.1016/j.procs.2018.10.171.

[8] Segura-Bedmar, I.; Alonso-Bartolome, S. Multimodal Fake News Detection. Information **2022**, 13, 284. https://doi.org/10.3390/info13060284

[9] Singhal, S.; Shah, R.R.; Chakraborty, T.; Kumaraguru, P.; Satoh, S. SpotFake: A multi-modal framework for fake news detection. In Proceedings of the 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), Singapore, 11–13 September 2019; pp. 39–47.

[10] Wang, Y.; Ma, F.; Jin, Z.; Yuan, Y.; Xun, G.; Jha, K.; Su, L.; Gao, J. EANN: Event adversarial neural networks for multi-modal fake news detection. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018; pp. 849–857.

[11] D. Westerman, P. R. Spence and B. Van Der Heide, "Social media as information source: Recency of updates and credibility of information", J. Comput.-Med. Commun., vol. 19, no. 2, pp. 171-183, Jan. 2014.

[12] S. M. Shariff, X. Zhang and M. Sanderson, "On the credibility perception of news on Twitter: Readers topics and features", Comput. Hum. Behav., vol. 75, pp. 785-796, Oct. 2017.

[13] Ramkissoon, A.N., Goodridge, W. Enhancing the Predictive Performance of Credibility-Based Fake News Detection Using Ensemble Learning. Rev Socionetwork Strat **16**, 259–289 (2022). https://doi.org/10.1007/s12626-022-00127-7