Detection of Fake News on Social Media: A Review

Vivek Jaiswal#1, Akshat Suri\*2, Akanksha Vishvakarma#3, Suruchi Sabherwal#4

*#Computer Science Department, JSS Academy of Technical Education, Noida, Uttar Pradesh, India*

1vinayjaiswal57@gmail.com

2suriakshat2001@gmail.com

3akankshavsh@gmail.com

4suruchi@jssaten.ac.in

***Abstract*—**

**News Utilization from Virtual Entertainment has such countless ramifications for Individual, society and association as far as notoriety, perspectives and convictions, wrongdoing, mental and actual wellbeing to give some examples. Accordingly, inspecting the impacts of Phony News via web-based entertainment is fundamental. Analysts endeavors to list down the issues and discoveries of existing examination in area of Phony News Identification. This paper gives the foundation to future examination and association to basically research the effect of "counterfeit news" on networks. The examination has zeroed in on various strategies for moderating the spread of phony news and spotting it over the media. Specialists tended to that multimodality highlights overwhelms the single methodology highlights with regards to recognize the Phony News via Web- based Entertainment. Context oriented data is expected to expand the presentation of the phony news ID framework. The creator profound jumped into social affair reasonable check-commendable articulations and client remarks for counterfeit data ID. Specialists zeroed in on examining the example of getting out misleading word in informal community and the connection among spreader. The specialists have inspected the modalities of information to investigate the difficulties of the current models for a programmed framework that can perceive the phony news. The barbecued survey functions as a scaffold for completing a successful and effective mechanized counterfeit news recognizable proof framework**

**Keywords: Fake news detection, Online Social Media (OSM), Multimodality, Contextual information.**

***Keywords***— **hate-speech, natural language processing, artificial intelligence**

1. Introduction

The progression of innovation and broad accessibility of the web have reshaped the computerized world and the scattering of data as of late. The most well-known justification behind associating with the web is virtual entertainment. A few web-based virtual entertainment destinations like WhatsApp, Facebook, Twitter, Instagram, YouTube, and numerous others, have acquired prominence because of minimal expense, helpfully use, and viral way of behaving. The quantity of web clients has expanded significantly, and they use it for various goals.

The advanced world is evolving emphatically. OSM has beated all others with regards to news utilization. Conventional media, for example, papers are being cleared out by involving web for online news utilization and overflow of data on web is basically accessible because of developing ubiquity of OSM. Because of the pervasiveness of online entertainment and nonattendance of the necessity of PC proficiency offers a space for cybercrime by dispersing unauthenticated data across the web. Word gets out quickly and it very well may be underhanded or real. Individuals are not sufficiently shrewd to recognize whether news is genuine. Counterfeit word gets out far and wide. News might be spread by listening in on others' conversations and virtual

entertainment. Counterfeit news is characterized as news that is purposely made to misdirect

individuals. It has an unfavorable effect as far as reputational harm to an individual, society, or foundation. It originally

showed up in the 2016 US official races, when manufactured news underwriting one of the two competitors was recognized and shared north of 37 million times on Facebook [1]. Nonetheless, it has as of late drawn in significant consideration. Spotting counterfeit news is a basic work, that guarantees that clients get dependable data, yet additionally safeguards a reliable news environment. Due of the enormous measure of individuals, time length, cost, and different variables, difficult work isn't by any stretch of the imagination supportive. To manage the tremendous amount of information accessible through OSM, a programmed counterfeit news recognition framework is required. Man-made consciousness techniques ended up being very advantageous in the field of programmed counterfeit news recognition. The forthcoming segment tosses light in the work done by the scientists in the field of phony news discovery and break down hole that exist to robotize the Phony news recognition system around web-based entertainment.

1. related work

A few instruments for recognizing counterfeit news have as of late been created. This part talks about applicable examination work on recognizing counterfeit news on web-based virtual entertainment. As per the accessible writing study, AI models were frequently used to distinguish counterfeit news, then, at that point, profound learning models came into picture and these days move learning and pre-prepared models are likewise performing great in this space.

From[2], introduced a device that can be introduced by clients into their own program to distinguish and sift through plausible misleading content sources and the sites\* containing bogus and misdirecting data.

Counterfeit news is for the most part made by changing video, text, picture, and sound, inciting the need of a multi-modular identification framework. The Multi-Modular Component Extractor has been used in different examination, including the accompanying:

[3], proposed EANN (Occasion Ill- disposed Brain Organization) for multi-modular phony news discovery. The current models are confronting inconvenience in recognizing counterfeit stories on recently arisen and time-basic occasions, since they just get occasion explicit attributes that can't be moved on to unseen occasions. Be that as it may, this EANN can learn occasion invariant qualities; offering it a chance to distinguish counterfeit reports progressively occasions. This model comprises of basically three sections: the occasion discriminator, multi-modular component extractor and the phony news locator. This examination depends on sight and sound datasets like Weibo and Twitter. The proposed framework beats the current benchmark methods by utilizing adaptable highlights portrayal.

[4], played out a customary trial to gauge the productivity of different AI calculation on three major and various datasets. Standard models contribute for eight of the 19 AI techniques, while customary profound learning calculations represent six, and unique pre-prepared language models like BERT represent five. It has been seen that the BERT based frameworks beat different methodologies as far as execution and potential on all datasets. BERT-based approaches are likewise sufficiently vigorous to work well even with a minuscule example set. Innocent Bayes with N-Gram has accomplished similar outcomes as brain network-put together models with respect to reasonable enormous informational indexes. The length of the dataset and data provided in the news has likewise been seen to influence the adequacy of LSTM-based frameworks. It works better when the report has adequate data.

[5], introduced a system for distinguishing manufactured reports depending on happy based elements and AI. In this review, the creator broke down various capabilities proposed for trickery identification as well as word implanting to get the best precise model and furthermore gave the "Unbiased"(UNB) dataset, a clever text corpus which adheres to a bunch of rules and standards to get solid results in characterization errands. Over all datasets contemplated, the proposed highlights matched with ML calculations accomplished exactness up to 95%, with AdaBoost positioning top and SVM and Stowing strategies positioning second, yet without measurably massive distinction.

[6], presented a model named as MVAE (Multimodal Variational Auto Encoder) for distinguishing created news. This model comprises of three principal parts:

1) Counterfeit News identifier Module,

2) Encoder and

3) Decoder.

The Multimodal Variational Auto Encoder is utilized to extricate highlights from the two modalities (Printed and Visual), and the phony news identifier module uses these modalities to decide whether the news is created or real. The adequacy of this proposed model has been assessed on two genuine informational collections got from the absolute most visited sites, Twitter, and Weibo. This proposed model outscored the past multimodal procedures by 5% in F1 score and 6% in exactness.

[7], proposed a robotized strategy for distinguishing Twitter news and Count-Vectors, TF-IDF, and Word Installing are utilized to process the dataset. This handled information is utilized by the model to recognize in the event that the substance is counterfeit or certified. The adequacy of

five notable AI procedures like NB, SVM, LR, and RNN models has been analysed in this review, and it has been accounted for that SVM gives 74% precision in every one of the highlighted vectors.

[8], presented a model utilizing Bi-directional LSTM-repetitive brain organization (Bi-LSTM-RNN) and tried on two unstructured news stories datasets which are freely accessible. The in the middle of between the news title and assortment of report is estimated utilizing GloVeppp  
 word implanting in this model. For steady and unsound high layered news informational collections, the proposed approach beats other profound learning models like CNN, vanilla RNN and unidirectional LSTM.

9], zeroed in on the meaning of client profiles in distinguishing counterfeit stories. In this paper, first objective is to gauge the post sharing way of behaving of the clients and to figure out the gathering chiefs who used to share the news all the more much of the time, second is to focus on the attributes of the clients who shares the phony/genuine news lastly examination of these client profiles highlights helps in counterfeit news discovery.

10], presented a system for examining the believability of the data accessible as pictures on different virtual entertainment stages. It comprises of a calculation that approves the veracity of picture texts by actually looking at it on the web and afterward assessing the truth boundary (Rp) in view of the dependability of the main 15 Google look. At the point when the worth surpasses a specific limit, a movement is sorted as practical or deceitful.

[11], presented a Spot fake - a structure that is multi-modular for counterfeit news recognizable proof. It thinks about both printed and visual parts of an article. Dissimilar to earlier multi- modular recognition frameworks, this model has been acquainted with distinguish bogus news without incorporating any extra subtasks. In this work, literary elements have been advanced by utilizing language models like BERT, while visual highlights have been advanced by utilizing the VGG-19 pre-prepared on ImageNet dataset. The review has been made on openly accessible Twitter and Weibo datasets and saw that recommended model beats Twitter and Weibo's current state by 3.27% and 6.83% individually.

[12], proposed Spotfake, a multimodal approach took benefits of the trade sorting out some way to get the setting focused information and semantic from these reports and their photos for the distinguishing proof of fake news. Spotfake+ is the general variation of Spotfake. This study has been done on the FakeNewsNet storage facility and it is the essential work till date which performs multimodal approach on the dataset which involves full-length articles close by the associated pictures. This proposed model beats the work on single system and multimodality both.

[13], cultivated a stunned projecting a polling form model on three specific datasets (Late trends, Kaggle, Reuter) by using the twelve simulated intelligence classifiers. The strength of one model is used to compensate the inadequacy of another model by looking at these twelve classifiers. Utilizing counterfeit estimate (FP) extent, a planned staggered model has been acquainted with achieve further developed results. It is seen that the proposed system has achieved the most raised precision interestingly, with other existing artificial intelligence estimations by using the part extraction strategies like CV, HV and TF-IDF. The major target of this proposed model isn't simply to sort out the best classifier, yet likewise the conditions under which that classifier beats the others for the given task. It is found that the classifiers Key Backslide (LR), Withdrew Intense (Father), and Straight Assistance Vector Machine(Linear SVM) performs best autonomously by using TF-IDF, CV, HV, (Components Extraction Techniques) independently founded on their evaluation rules, while the made model for instance Staggered Model (took a stab at different datasets by using all of the three procedures, TF-IDF, CV, HV) performs better contrasted with the Determined Backslide by 1.3%, Direct SVM by 0.4%, Uninvolved Powerful by 0.8%.

[14], proposed a model named as Safeguarded (Similarity Careful Multi-Particular Fake News Recognizable proof Structure) which essentially drew in the association between the printed and visual components of any report. This model involved three modules:

1). Multi-Secluded Features Extraction

2). Inside Measured Fake News Conjecture

3). Cross-Particular Closeness Extraction.

This proposed model sees the fake news in light of printed,

visual and botch and a great deal of genuine data has been utilized to survey the model's feasibility.

Paper (2017: Deep Learning for Hate Speech Detection in Tweets), focuses on the problem of classifying a tweet as racist, sexist or neither. For this objective, text processing algorithms like (1) Char n-grams (2) TF-IDF (3) Bow were put to use with CNN and LSTM based architectures. The dataset consisted of 16K annotated tweets from the Waseem and Hovey(2016) dataset. Of the 16K tweets, 3383 were labelled as sexist, 1972 as racist, and the remaining were marked as neither sexist nor racist. It was found that among the baseline methods, the word TF-IDF method was better than the character n-gram method. CNN performed better than LSTM which was better than Fast Text. Best results were obtained from “LSTM + Random Embedding + GBDT” where tweet embeddings were initialized to random vectors, LSTM was trained using back-propagation, and then learned embeddings were used to train a GBDT classifier.

Paper (2018: Detecting Online Hate Speech Using Context Aware Models), explored two types of models, feature-based logistic regression models and neural network models, in order to incorporate context information in automatic speech detection. The dataset used consisted of 1528 Fox News user comments, which were taken from 10 complete discussion threads for 10 widely read Fox News The investigation shows that Significant Learning Procedures beats current artificial intelligence Methods in spotting fake news.

[15], proposed a two-step model by consolidating text mining and 23 oversaw man-made thinking computations. Execution of overseen mechanized thinking estimations occurs directly following changing the unstructured dataset into the coordinated from by using the text mining. This joined model has been taken a stab at three veritable world datasets and the viability of the model has been assessed to the extent that precision, exactness, survey, F-measure values. [16], separated PC based knowledge gadgets using simulated intelligence computations, Customary Language Dealing with (NLP) and played out a relationship among open fake news area datasets like BuzzFeedNews, LIAR, BS Finder, CREDBANK, Buzz Face, and Facebook Duplicity, and proposed a broad dataset FakeNewsNet to counter the dispersal of fake news. [17], proposed a model named as pathetically oversaw fake news area structure (WeFEND) which involves clients' reports as weak oversight to raise how much planning data for distinguishing the misleading news. The annotator, the supported selector and the phony news finder were the three primary parts of the proposed work. By utilizing these parts, the clever methodology heightens the preparation information and furthermore the quality for the powerful idea of the news and got the exactness of 82.4% on enormous pool of news stories from true records of WeChat. [18], presented a new dataset FACTDRIL (Reality Actually taking a look at Dataset for Provincial Indian Dialects) and the fundamental spotlight is on the low asset Indian Dialects like Marathi, Bangla, Telugu, Malayalam, Oriya, Tamil, Punjabi, Assamese, Urdu, Burmese and Sinhala. This proposed dataset FACTDRIL contains the 22,435 examples from 11 Indian Low asset dialects with IFCN (Worldwide Reality Actually looking at Organization) certificate. This FACTDRIL is the primary enormous scope multilingual dataset utilized for giving the insights about the veracity of the unconfirmed cases for the low asset Indian Dialects. In this review, another component named as Examination thinking through manual impedance is presented. This component makes sense of the different strategies taken by reality checkers for reach at a choice on unconfirmed news.

[19], planned a managed AI approach for believability keep an eye on twitter news based on two fundamental elements like substance based and client based. The creators have used seven managed AI methods for example Guileless Bayes (NB), Most extreme Entropy (ME), Backing Vector Machine (SVM), Irregular woods (RF), K-Closest Neighbor (KNN), Calculated Relapse (LR), and Contingent Arbitrary Timberland (CRF) on Pheme dataset and this has been separated in this proportion, 80% for preparing, 10% for testing and 10% for approving the informational index. Discoveries of this work are:

1) Arbitrary Woods (RF) gives the best exhibition with precision 82.2% on client put together elements and furthermore with respect to consolidated highlights (content based elements and client based highlights) with exactness 83.4%.

2) Calculated Relapse (LR) performs best with exactness of 73.2% on satisfied based highlights.

3) Additionally dissected that the effect of content-based highlights are not exactly the client based highlights.

[20], acquainted a mechanized methodology with distinguish the bogus news based on various properties for Facebook information using AI and profound learning strategies in chrome climate. For spotting counterfeit stories, this proposed model utilized some extra data connected to the client's Facebook account and its news content. When contrasted with AI techniques, the Long Momentary Memory (LSTM) calculation, which is a profound learning calculation, got a remarkable exhibition of 99.4 %.

[21], conducted a survey on trends in curbing fraudulent news in social media. The author has discussed about:

1) the two major ways i.e., misinformation and disinformation in which fakenews has been proliferated through social media.

2) Also, about the classification or types of fake news whichincludes click baits, hoaxes, propaganda, satire and parody and some others like name-theft, framing, and journalism deception.

3) and various fake news detection models such as Experts or professionals Fact- Checker approach, Crowd sourced approach, Machine Learning approach, Natural Language processing technique, Hybrid technique, Human-Machine approach, Deep Learning, Graph-based method.

[22], proposed a FakeBERT which is deep learning approach that relies on BERT (Bidirectional Encoder Representations from Transformers) by merging distinct parallel blocks of a single- layer deep Convolutional Neural Network (CNN) with various kernel sizes and filters with the BERT. This combination is beneficial for dealing with ambiguity, which is the most difficult aspect of natural language comprehension. The upcoming section aggregates the information grilled from the above literature and presents a comparative analysis with timeline.

1. Comparative Analysis

The rise of social media has opened with respect to fake news and its automatic detection. It is an area that affects views and beliefs, business, mental and physical health. Table 1 throws light on the work of eminent researchers around fake news detection and a comparative view of the outcome and accuracy observed in each of the work.

|  |  |  |  |
| --- | --- | --- | --- |
| **Author Name** | **Approach** | **Efficiency/Accurac y (in %)** | **Findings/outcomes** |
| Y. Wang et al. 2018[3] | EANN (Event Adversarial Neural Networks for multi- modal Framework) | 71.5% on Twitter dataset and 82.7% on Weibo dataset. | Proposed EANN Model outperforms the other models and learns transferable features representations. |
| J. Y. Khan et al. 2021[4] | 19 models including 8 traditional learning models, 6 conventional deep learning models and 5 newfangled models are pre-trained language models likeBERT  (Bidirectional Encoder  Representations from Transformers) are used. | Roberta (Robustly  optimized BERT approach), achieved 96% accuracy in case of fake or real news dataset and 98% in the combined corpus dataset. | Out of 19 models, Roberta performs best in case of Real or fake news dataset and combined corpus dataset, while HAN performs better in case of LIAR dataset. |
| Dhruv Khatter et al. 2019[6] | MVAE (Multimodal Variational AutoEncoder) | 74,5% on Twitter dataset and 82.4% on Weibo dataset. | MVAE outperforms the other deep learning models by ~6% in accuracy and ~5% in F1- scores. |
| Abdullah-All- Tanvir et al. 2019[7] | Machine Learning Approaches: Support Vector Machine (SVM) Naïve Bayes (NB)  Logistic Regression (LR) Deep learning Approaches:  Recurrent Neural Network (RNN) Long-Short Term Memory (LSTM) | 69.47  89.02  89.34  74  78 | SVM and Naïve Bayes outperform other algorithms used on Twitter dataset. |
| Bahad, P. et al. 2019 [8] | Bi-directional LSTM-recurrent neural network (Bi-LSTM- RNN) | ---- | Tested on two unstructured news articles datasets which are publicly available. The correlation in between the news title and body of news story is measured using GloVe word embedding in this model. For stable and unstable high dimensional news data sets, the proposed approach outperforms other deep learning models such as CNN, vanilla RNN and unidirectional LSTM. |
| S. Singhal et al. 2019[11] | SPOTFAKE  (Multi-modal framework) | 77.77% accuracy achieved in case of Twitter Dataset and 89.23% in case of Weibo  Dataset. | The performance of the proposed SPOTFAKE system is higher as compare to other existing models by ~3.27% and ~6.83% respectively. |
| S. Singhal et al. 2020[12] | SPOTFAKE+ | 84.6% on Politifact and  85.6 on Gossipcop of FakeNewsNet dataset. | Proposed SPOTFAKE+, a multi-modal framework which outperforms other multi- modal frameworks like EANN, MVAE and  SPOTFAKE. |

|  |  |  |  |
| --- | --- | --- | --- |
| Xinyi Zhou et  al. 2020[14] | SAFE (Similarity-Aware Multi-  Modal FakE news detection system) | 87.4% accuracy on  PolitiFact Dataset and 83.8% on Gossipcop dataset. | The falsity of news articles based on their text,  images or their ‘mismatches’ have been recognized in the proposed system SAFE. |
| Wang, Y. et al. 2020 [17] | Weakly Supervised Fake News Detection Framework  (WeFEND) | 82.4% accuracy on large pool of news articles  from official accounts of WeChat | This model utilizes users’ reports as weak supervision to raise the quantity of training data  for identifying the false news. The annotator, the reinforced selector and the fake news detector were the three main components of the proposed work. By using these components, the novel approach intensifies the training data and also the quality for the dynamic nature of the news and got the accuracy of 82.4% on large  pool of news articles from official accounts of we chat. |
| Marina Azer et  al. 2021[19] | Random Forest (RF)  Support Vector Machine (SVM) K-Nearest Neighbor (KNN) Naïve Bayes (NB)  Logistic Regression (LR) | LR gives accuracy of  73.2% on content-based features, RF gives 82.2% on user-based features and RF gives 83.4% on overall features set. | LR is best in using content-based features and  RF is best in using user-based features and in overall features (contains both content-based and user-based features). Also, user-based features have enhanced performance than content-based features. |
| S. R. Sahoo et  al. 2020[20] | Machine Learning Classifiers:  K-Nearest Neighbor (K-NN) Support Vector Machine (SVM) Logistic Regression (LR) Decision Tree  Naïve Bayes (NB)  Deep Learning Classifier:  Long Short-Term Memory (LSTM) | Accuracy achieved on  News-Content Features  + Users Profile Features: 99.3  99.3  99.0  99.1  98.6  99.4 | Deep learning model LSTM performs better  than other classifiers with accuracy 99.4% on combined features of user profile and news content. |
| S.Aphiwongsop  hon et al. 2018[23] | Naïve Bayes (NB)  Neural Network (NN)  Support Vector Machine (SVM) | 96.08  99.90  99.90 | NN and SVM perform better than other  methods. |
| A.Kesarwani et  al. 2020[24] | K-Nearest Neighbor  (K-NN) | 79 | 79% accuracy achieved while tested against  Facebook news posts dataset. |
| I. Y. R. Pratiwi  et al. 2017[25] | Naïve Bayes (NB) | 78.6 | 70:30 ratios of training dataset and testing  dataset performs better with accuracy 78.6%. |
| M.Granik et al. 2017[26] | Naïve Bayes (NB) | 74 | Implemented as software system and tested against dataset of Facebook news posts. |

Figure 1 presents the timeline effect of the work done around Fake News Detection approaches.

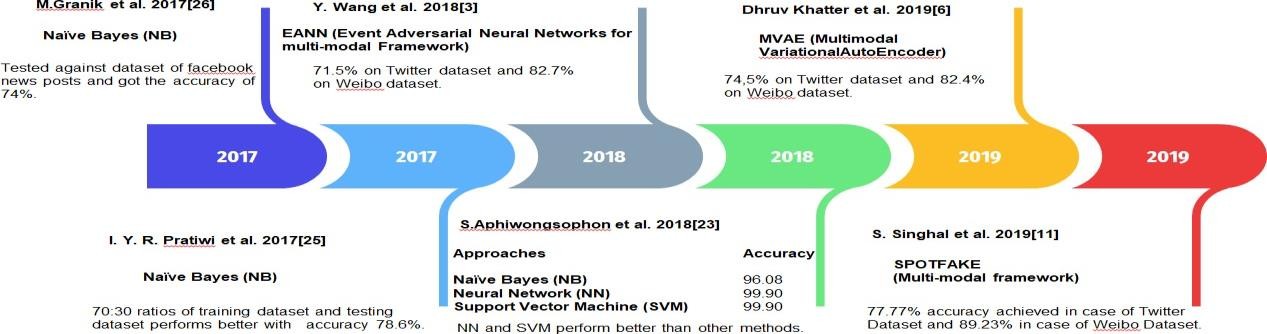
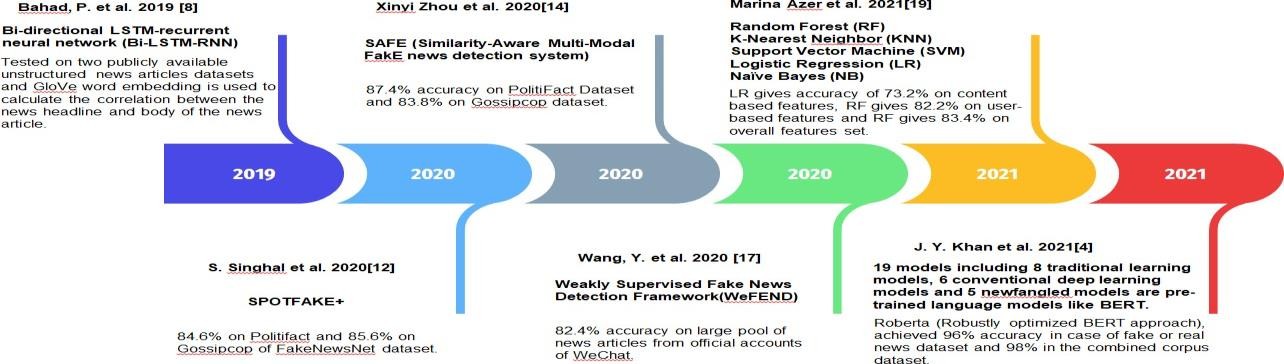


Fig. 1 Timeline of Fake News Detection

The broad examination contemplated has arisen the holes to connect for a successful and proficient phony news identification framework. The impending segment features the holes and difficulties for the computerized counterfeit news location framework.

1. Research Gaps

Single modality feature is a hurdle to effectively identify the fake news. By using Linguistic approach numbers of effective methods have been designed to detect Fake News. But very less work has been done in Visual-based verification. Source verification is observed as a missing piece in the existing models. Limited size of dataset has been observed in the existing literature. The existing approaches have paid less attention towards newly occurring and time critical events. Many of the researchers focused on a particular sort of news (such as political news), which raises the issues of dataset bias. This open the opportunity to take advantage of comprehensive dataset with Multimodal feature for an effective and efficient Automatic Fake News Detection system. Following are the challenge for the potentially benefitting from the system:

As non-manipulated images are jumbled in with the fake news content, it is difficult to tell whether it’s legitimate or not. The identification of fake news is further hindered. As non-manipulated images are jumbled in with the fake news content, it's difficult to tell if it's legitimate or not. The identification of fake news is further hindered by the lack of editorial strictness. A comprehensive dataset comprising of contextual information and a complete multimodal collection of Fake News data types is required. Verifying sources and establishing author credibility is also a hurdle for researchers. There are no ways ofidentifying and blocking the spreaders of fake news on social media. Contextual information is required in order to increase the model's efficiency. Collecting explainable check-worthy phrases, user comments, the patterns of disseminating Fake News as well as the links between the spreaders are an exemplary feature to distinguish fake and real news. The relationship between the title and the body text of news, correlation between the modalities can help in achieving the desired efficiency.

# Conclusion

The exhaustive and basic assessment presents a thorough examination recommendation to recognize counterfeit news by tending to the ramifications for people, society, and associations. In light of the immense measure of data accessible, web-based entertainment has turned into a focal point for internet perusing. Thus, manual grouping of news stories isn't practical because of the extra labor supply, cost, time and aptitude required. Thus, mechanized counterfeit news ID is required. It has been seen that in both single methodology (text just) and multi-methodology, the accentuation is on highlight extraction procedures (text and picture).

Literary element extraction is achieved by utilizing Text-Convolutional Brain Organization (Text-CNN), Term Recurrence Backwards Archive Recurrence (TF-IDF), Hashing-Vectorizer (HV), and Count- Vectorizer (CV), while visual component extraction is achieved by utilizing VGG-19. The different AI, Profound Learning, Move Learning, and Pre-prepared models are looked into to more readily get a handle on the quintessence of recently executed counterfeit news recognition draws near. It has been seen that pre- prepared and profound learning models perform best in recognizing counterfeit news.

# References

1. Ahmed, A. A. A., Aljabouh, A., Donepudi, P. K., & Choi, M. S. (2021). Detecting Fake News using Machine Learning: A Systematic Literature Review. arXiv preprint arXiv:2102.04458. https://arxiv.org/abs/2102.04458
2. 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](https://doi.org/10.1016/j.procs.2018.10.171)
3. Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, GuangxuXun, KishlayJha, Lu Su, and Jing Gao.2018. 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, United Kingdom) (KDD ’18). Association for Computing Machinery, New York, NY, USA, 849–857. https://doi.org/10.1145/3219819.3219903

1. Khan, J. Y., Khondaker, M., Islam, T., Iqbal, A., &Afroz, S. (2021). A benchmark study on machine learning methods for fake news detection. Computation and Language.https://arxiv.org/abs/1905.04749
2. Gravanis, G., Vakali, A., Diamantaras, K., & Karadais, P. (2019). Behind the cues: A benchmark study for fake news detection. *Expert Systems with Applications*, *128*, 201-213. https://doi.org/10.1016/j.eswa.2019.03.036
3. Dhruv Khattar, Jaipal Singh Goud, Manish Gupta, and VasudevaVarma. 2019. MVAE: Multimodal VariationalAutoencoder for Fake News Detection. In The World Wide Web Conference (San Francisco, CA, USA) (WWW ’19). Association for Computing Machinery, New York, NY, USA, 2915–2921. https://doi.org/10.1145/3308558.331355
4. Abdullah-All-Tanvir, Mahir, E. M., Akhter S., &Huq, M. R. (2019). Detecting Fake News using Machine Learning and Deep Learning Algorithms. 7th International Conference on Smart Computing & Communications (ICSCC), Sarawak, Malaysia,
5. Bahad, P., Saxena, P., & Kamal, R. (2019). Fake news detection using bi-directional LSTM- recurrent neural network. *Procedia Computer Science*, *165*, 74-82.<https://doi.org/10.1016/j.procs.2020.01.072>
6. Shu, K., Zhou, X., Wang, S., Zafarani, R., & Liu, H. (2019, August). The role of user profiles for fake news detection.

In *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining* (pp. 436- 439)[.https://doi.org/10.1145/3341161.3342927](https://doi.org/10.1145/3341161.3342927)

1. Vishwakarma, D. K., Varshney, D., &Yadav, A. (2019). Detection and veracity analysis of fake news via scrapping and authenticating the web search. *Cognitive Systems Research*, *58*, 217-229. https://doi.org/10.1016/j.cogsys.2019.07.004
2. Singhal, S., Shah, R. R., Chakraborty, T., Kumaraguru, P., & Satoh, S. (2019). SpotFake: A Multi-modal Framework for Fake News Detection. 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM).doi:10.1109/bigmm.2019.0044.
3. Shivangi Singhal, Anubha Kabra, Mohit Sharma, Rajiv Ratn Shah, TanmoyChakraborty, and Ponnurangam Kumaraguru. 2020. SpotFake+: A Multimodal Framework for Fake News Detection via Transfer Learning (Student Abstract). Proceedings of the AAAI Conference on Artificial Intelligence 34, 10 (Apr. 2020), 13915–13916. https://doi.org/10.1609/ aaai.v34i10.7230
4. Kaur, S., Kumar, P. & Kumaraguru, P. (2020). Automating fake news detection system using multi-level voting model. *Soft Computing, 24*(12), 9049–9069. https://doi.org/10.1007/s00500-019-04436-y
5. Xinyi Zhou, Jindi Wu, and Reza Zafarani. 2020. SAFE: Similarity-Aware Multi-Modal Fake News Detection. arXiv:2003.04981 [cs.CL]
6. Ozbay, F. A., & Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and its Applications*, *540*, 123174. <https://doi.org/10.1016/j.physa.2019.123174>
7. Salazar, A. P. (2020). AI Tools on Fake News Detection: An Overview and Comparative Study.
8. Wang, Y., Yang, W., Ma, F., Xu, J., Zhong, B., Deng, Q., &Gao, J. (2020, April). Weak
9. supervision for fake news detection via reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 01, pp. 516- 523). https://doi.org/10.1609/aaai.v34i01.5389 Springer, Cham. DOI: 10.1007/978-3-319-69155-8\_9
10. Singhal, S., Shah, R. R., & Kumaraguru, P. (2021). Factorization of Fact-Checks for Low Resource Indian Languages. arXiv preprint arXiv:2102.11276.
11. Azer, M., Taha, M., Zayed, H. H., &Gadallah, M. (2021). Credibility Detection on Twitter News Using Machine Learning Approach. International Journal of Intelligent Systems & Applications, 13(3).
12. Sahoo, S. R., & Gupta, B. B. (2021). Multiple features based approach for automatic fake news Detection on social networks using deep learning. Applied Soft Computing, 100, 106983. https://[www.sciencedirect.com/science/article/abs/pii/S1568494620309224](http://www.sciencedirect.com/science/article/abs/pii/S1568494620309224)
13. Collins, B., Hoang, D. T., Nguyen, N. T., & Hwang, D. (2021). Trends in combating fake news on social media–a survey. *Journal of Information and Telecommunication*, *5*(2), 247-266. <https://doi.org/10.1080/24751839.2020.1847379>
14. Kaliyar, R. K., Goswami, A., & Narang, P. (2021). FakeBERT: Fake news detection in Social media with a BERT-based deep learning approach. *Multimedia tools and applications*, *80*(8), 11765-11788. ‘ doi: 10.1007/s11042-020-10183-2
15. 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)*, 528-531. https://doi.org/10.1109/ECTICon.2018.8620051
16. 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)*, Las Vegas, NV, USA,pp.1-4, https://doi.org/10.1109/ICACCE49060.2020.9154997
17. Pratiwi, I. Y. R., Asmara, R. A., & Rahutomo, F. (2017). Study of hoax news detection using naïve bayes classifier in Indonesian language. 2017 11th International Conference on Information & Communication Technology and System (ICTS),

Surabaya, pp.73-78. https://doi.org/10.1109/ICTS.2017.8265649

Malaysia, 2019, pp.1-5, https://doi.org/10.1109/ICSCC.2019.8843612

1. Granik, M., &Mesyura, V. (2017). Fake news detection using naive Bayes classifier. 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON). doi:10.1109/ukrcon.2017.8100379