iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** IV **Month of publication:** April 2021

DOI: https://doi.org/10.22214/ijraset.2021.33633

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue IV Apr 2021- Available at www.ijraset.com

A Review Paper on Fake News Detection Techniques

Dinkal Kewlani[1], Akshita Jasoria[2]
[1, 2]*Computer Engineering Department, NMIMS University*

Abstract: ***Fake news is rapidly eroding the trust of people on journalism and media and also misleading the audience. Therefore discovering fake news online is of utmost importance in today's world where such diversified news content is produced and shared amongst millions of people within the minimum time. These days a lot of information is shared on social media and it is not able to distinguish the information as real or fake manually. People immediately begin to express their concerns or share their thoughts as soon as they encounter a post on any social media platform, without verifying its authenticity and credibility. Fake stories and rumours are the most popular types of lies to unauthorized information and affect the audience's decision making. In order to detect non-linguistic issues, linguistic and non-linguistic analysis can be analysed using a number of methods. While most of these methods of detecting real and fake news are often effective, they have some limitations. This paper reviews and analyses in detail about the different methodologies such as machine learning algorithms, neural network methods, traditional methods and semantic analysis which helps in detecting fake news. By conducting an effective analysis on literature review we compared the most efficient approaches available for detecting fake news.***

Keywords: Fake News Detection, Social Media, Online Social Networks, Machine learning, False positive rate, Misleading, Malicious

I. INTRODUCTION

Fake news are those news stories which are fake with fabricated content and no verifiable facts. These are published with the sole purpose of misleading the readers with the misinformation. The increasing use of social media in our daily lives has decreased the use of traditional sources of news such as Newspapers, Television, Radio, News channels, etc. Therefore the share of False News on various digital platforms have proliferated because of the ease to share it online [11]. The fake news is written with an additional personal opinion of the author about a particular news rather than objective content so the viewers may gain information with the misconception. Political observers have named false information as one of the fundamental threats to democracy [2]. This fake news creates social polarization in the society and leads to many other problems related to politics and therefore creates an unnecessary ambush or rage among the audience for a particular community of the society.

The importance of fake news detection was clearly demonstrated during the current COVID-19 pandemic where the misconceptions of the people about the virus, it's spread and the vaccine as well has created a lot of problems for the government in their effective working. A lot of social media networks are stepping up to detect the fake news getting published on their sites through digital fake news detection or some machine learning algorithms to identify false claims. Only a machine could keep up with the massive spread of diversifying fake news all over social meia. The purpose of this review paper is to gather all the efficient approaches made so far to detect the fake news and to rebuild the trust of people on media and journalism that have been eroded by the spread of fake news. Our review analyses how to detect fake news using different machine learning and traditional algorithms to overcome the problem.

This paper is further explained as per the following sections: Section I talks about the Introduction, Section II tells us about the background of fake news, Section III describes our literature review, Section IV describes the various approaches used for fake news detection, Section V is a comparison of various methods used, Section VI states the application of fake news detection and Section VII, the last section concludes our research and states the future scope.

II. BACKGROUND

Initially, we define the fake news and the factors leading to fake news. After which we discussed the history of fake news and it's evolution with time.

A. What is Fake News?

Even in journalism there hasn't been any universal definition for fake news, it has always been dependent on an individual's opinion. Though a clear and accurate explanation could help to set a strong foundation to detect real and fake news and evaluate their related studies.

In this section we (I) differentiated between different concepts that were frequently encountered by fake news, (II) introduced a broader and accurate meaning of fake news, and (III) continued to discuss the possible research complication raised by those definitions[21].

Till date no explanation is provided for fake news, provided as a “false ”,“ a news story or published message spread through the different media, handling fake information regardless of its methods and motives ”is far from false news, anonymous information, false information, sarcastic news, or even stories that one does not like (they are considered inappropriate).

In addition, what news has is difficult to define as it may move from the account of the latest, most interesting, and important event, to the official account of being specific or different; as for, “the digitization of news has challenged traditional definitions of news. Online platforms provide space for non-journalists to reach a mass audience.”

Broader definition: False or misleading information presented as news. It is usually intended to damage a person's or business reputation, or to monetize advertising profits. Journalist Nolan Higdon provided a broad definition of non-legal news as "false or misleading content presented as news and conveyed in ways that take verbal, written, printed, electronic, and digital communication."

B. History of Fake News

False and distorted news items are not a new concept, it has been a part of our lives and the news history long before the invention of the printing press. On the Internet, popular forms called clickbait sequences entice people to click to learn more, in an attempt to intimidate and surprise us. What annoys you more, to read about something or the lies that never really happened?

There are a lot of examples and case studies of false news and misinformation in history. It was first noticed in Nazi propaganda to build anti-Semitic favour. After that, in 1800 in the United States, racism and its sentiments led to the spread of false news and stories about “African Americans” supposed deficiencies and crime. In 1890, the newspaper's rivals Joseph and William competed with the audience in applause and reported the rumours as if they were real, this practice known at the time as "yellow journalism." This news led to Spanish-American War of 1898.

Fig. 1: Example of Fake News

Fake news is also used to try to undermine popular media (especially President Trump), to suggest that they are fabricated or that they are blowing up a certain amount to be minimal (even if some other sources can confirm its true accuracy). While talking with Lou Dobbs (of FOX BUSINESS in 2017), Donald Trump (ex- president of US) claimed that he stated that whole fake news story. But Ironically, Hillary Clinton used the same term in a speech she made two days before Trump's first use of that phrase. While it is possible that Donald Trump assigned the name in a completely new way, the name itself has been in many stories over many years.

C. Contemporary Fake News

No matter who started this "fake news or stories" but they are different from the nonsense stories of journalists in traditional media. The speed at which the fake news is spreading and the magnitude of its influence on the audience, puts it in a different category from its old cousins. There are three distinct parts of today's fake news story that makes it different from the traditional forms of over-reporting or false reporting: who, what, and how.

1) *Who:* Fake stories are made and circulated by those with interests such as computer programming who wants to make money out of spreading fake news, or the people who spread news to manipulate the market for their personal needs. They are not newspaper publishers at this time.
2) *What:* It usually involves knowing the distortions and deceptions of the media and not just the context. In another example, a video that was initially a part of a public service announcement to help people recognize child trafficking in Pakistan, "was designed to look like a real kidnapping spread like anything" which led to the deadly attacks on innocent people accused of kidnapping.

Fig. 2: Example of Fake News

3) *How:* There can be many perspectives of fake news and the way it is presented makes people more likely to fall for it like many news stories are passed on to people on social media by their friends or the people who follow them. In social media apps, "Most messages are shared in groups, and when they are transmitted, there is no indication of where they are coming from. (False News) It seems to come from family and friends. "

D. How to Spot Fake News

1) *Consider the Source:* Do some basic research about the actual source from which the news was published.
2) *Supporting Sources:* See the resources mentioned in the news and their credibility.
3) *Multiple Sources:* Never rely on one source or one article, read different articles from different authors from different backgrounds.
4) *Comments:* Even if the source is real, be aware of the comments posted on the response because they can be computer generated by the bots or by people hired to put out false stories or and can be confusing or false informative.
5) *Repost Carefully:* False news or articles relies on how many believers to retweet, repost or forward that false informative message. If you're not sure about that article or news, think thrice or hold off sharing it with others

III.LITERATURE REVIEW

With the outburst of spread of misleading information via online platforms, it is very important for people and society to judge the authenticity of the news that they are reading. A variety of machine learning, linguistic and traditional approaches exist to detect fake news. Through this review paper we aim to identify the most efficient models that are available for fake news detection.

A. Methodologies

The author proposed a human-machine interaction case study to distinguish fake news. He suggested a cognitive system that applies backtracking to detect false news [5]. The analysis was done based on the words used in the news article such that the important keywords are prioritized for the web search. Reverse tracking of the published articles helps to reveal features that helps to find the identity of the author. The detection rate of the proposed system was found to be 85% (average) which was a considerate amount keeping in mind the diversification of the news.

Malicious social bots are used to do wrong events to hurt the sentiments of the audience and for negative influencing. The author proposed a novel method that includes transition probability of clickstream sequences and semi-supervised clustering which was used to detect malicious bots by analysing user behaviour clickstreams and by identifying the transition probability features of the user [13]. Through this method, the time of artificial making is reduced and the robustness of the transition probability and other user features is increased thus improving the accuracy of the model to detect fake news. Experimental analysis observed that the accuracy of identification can be as maximum as 93.1% and the recall rate was found to be 97.5%.

The rapid growth of Internet Technology is leading to the rising popularity of Online Social Networks (OSNs). OSN is the most efficient way to broadcast any form of news or content over the digital platforms. The author proposed a model which describes how fake news gets segregated among people with the influence of different misinformation spreading techniques [4]. To detect and control the spread of fake news on Online Social Networks, a susceptible-verified-infected-recovered (SVIR) model has been proposed. In this approach, a basic reproduction number (Ro) is obtained which is an important aspect for the analysis of message spreading within the Online Social Networks.

In 2020, the author proposed a method to recognize fake news through characterizing a bunch of popular fake and real news measured by the shares, retweets, forwards, comments and reaction from two perspectives: Content understanding and Domain reputations [8]. It analyses the similarities and dissimilarities of the fake news and articles based on the tf-idf & LDA topic modelling. It systematically characterizes the published articles and the publishers of the fake news based on their registration patterns, domain popularity, domain ranking and domain reputation. It explores domain similarities between fake news, real news and Jaccard method articles on Hybrid News to differentiate the fake and real news.

The author proposed a system which includes a Hybrid Neural Network (HNN) that merge the proficiencies of both Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) that are used with 2 different proportionate reduction approaches, PCA and Chi-square [10]. The author employed the dimensionality reduction methods and techniques to reduce the proportions of the feature vectors prior to forwarding them to the classifier. The experimental analysis showed that the PCA outperformed state of the art methods and chi-square with an accuracy of 97.8%.

The author proposed a model that extracted all the necessary features from the datasets consisting of fake news articles, and classified those extracted features using the ensemble method consisting three favourable machine learning algorithms: Decision Tree Algorithm, Random Forest Algorithm and Extra Tree Classifier Algorithm [14]. The experimental analysis showed a training of 99.8% and testing accuracy of 44.15% on the Liar dataset and on the ISOT dataset, we attained 100% training and testing accuracy.

The author uses a deep neural network approach for fake and real news detection using google colab platform [2]. Deep learning models based on combined techniques of NLP and live data stage mining are created. LSTM and GRU models are combined with the data mining and Word2Vec models for more effective results. The LSTM model when combined with the additional mined features showed a significant improvement in the accuracy that was in the range of 91-94%. The mining stage was added into the solution approach to find the domain names of the news articles and to upgrade the process of detecting fake news.

The author proposed a computational-stylistic analysis approach which was based on natural language processing [11]. The author introduced a minimal overhead, while it has the potential to provide a high level of confidence in differentiating fraud from real news. Through these automatic approaches we were able to find four different action fronts which helped in extracting the sensitive stylistic approach. The result obtained under the curve was 0.82 by matrix transformation derived from k-mean clustering. In this an accuracy of 86% and precision of 94% stand out in quality of detection.

The author used the present body of knowledge with the application of intelligent tools to fight against the misleading information and Research & Development project related to the problem [9]. In this paper the author proves that apart from surveys related to fake news text based analysis and state of art approach is done and also described the overview of fake and real news and how it is dangerous. The analysis can be divided into 5 parts: linguistic features, deception modelling, clustering, predictive modelling and content cues [9]. The author also represented style based, network based, pattern based approach detection methods. Recently, malware domain names have emerged within large ISP networks effectively with True Positive accuracy up to 85% and False Positive accuracy less than 0.1%. It accurately detected fake portraits and videos with an accuracy of 97.29%.

The author described that the deep investigation of the features was more predictive from automated as well as human point of view for the identification of social media handles that were responsible for spreading fake news [3]. He mentioned that the majority (approx 57%) of people on social media found that news is inaccurate. He proposed a model based on deep learning architecture with the properties of a long short-term memory neural network and CNN. By using KNN classifiers, the accuracy went up from 79% to 84% and by using the cross validation technique it went to 83.13% from 78.61%. Result on the social context data set shows 91.74% accuracy for the first method, and for the second method its accuracy is 92.89%, although the loss value was not exceptionally low.

The author illustrated that the various aspects of communication on digital platforms can be used to find weak social supervision [7]. It opens the door to learning about socially sensitive monitoring of similar activities that arise when labelled data is limited.to illustrate fake news using weak social supervision, we used TriFN that helped to do the empirical comparison of the results in public

benchmark Politifact dataset from FakeNewsNet consisting of 120 true news and 120 false with 23,865 users and 91 publishers. The weak supervision technique proposed by TriFN can achieve an accuracy of 0.75. The author has laid the groundwork for media ideas by using user-friendly information that will help to integrate their ideas in a well-structured unsupervised way to make our rating results.

The author talks about the blast of web-based media that permitted people to share data without any credibility and with very little examination [6]. It intensified the previous issue of phony news, which turned into a significant concern these days because of the negative effect it is bringing to the online social networks. The author also investigated the speculation of the suggested approach across diverse datasets and found promising outcomes. The FA-KES dataset consists of 804 articles on news. For training ISOT is chosen because it is large and has little room for advancement because most models perform phase accuracy limits above 0.9

The author talks about the ease of distribution of user's current data between various people over the same or different online social networks (OSNs) like Facebook, Instagram, Whatsapp etc [15]. He proposed a n-gram model to detect false news by analysing the false reviews. He used explicit and latent features and transformed it to a unified feature space. This method was known as TI-CNN which was used to analyse text and image for false news detection by using the news and social context and other relative information. The proposed method analysed the user profile as well as the news content features based on Naïve Bayes, Decision tree and Deep learning algorithms.

The author pointed out the lack of expertise in traditional newspaper forums nowadays and proposed a model based on sequential neural networks that was less expensive and capable of handling the rich linguistic aspects [1].The neural network was able to observe the complex relationships between the reviews and identify the required details from the given inputs. The author also introduced an efficient linguistic feature based on fake and real news detection model and Readability based model where it evaluates the pattern of textual difficulty and determines the level of text comprehension. The comparison of these models was done based on the Epoch, batch size and linguistic feature parameters to identify the most efficient model for detecting false news. The integrated model based on the sequential neural network achieves 86% accuracy.

The author proposed a model for verifying the news at the event-level based on the patterns of images that were distributed in news events [16]. Here the review of features were given and based on which the finest features from the non - image features as baselines were selected. These features undergo various performance comparisons and feature analysis to give an overview of the image features for news verification. Random Forest classification algorithm was used to do the above analysis and the accuracy came out to be 83.6%. The author concluded that the existing feature-based approaches take no notice of the image content in the news and therefore were incompatible. This model included the image feature analysis from visual content and statistics point of view to differentiate the patterns for fake news and real news both visually and statistically [16].

The author proposed a model for memory network information retrieval which was used for the recognition of misleading news [12]. The author evaluated multiple measures to study and analyse the results of the proposed model with the state-of-the-art stance detection model and finally proposed a novel stance detection model to solve the issue of fake news spread over FNC. The model was a combination of an IR system with a revolutionary MemN2 network.

B. Issues

Cascading nature of fake news was one dispute that the majority of the publications were addressing. The profile of things marked as fake news would move over the long haul on the grounds that the propagators of such false news are aware of different ways through which the advanced automated systems could identify them. Certain attributes of the fake news being spread are changed in order to try and sidestep their information from being marked as fake news.

To perform detection or prediction assignments various scientists and system architectures utilize deep learning methodologies. Whatsoever, the result of such algorithms are given with no interpretation. Usage of proper explanation of algorithms is essential in order for the ML based false news and real news detection approach to be strongly and broadly believed by various groups.

Machine Learning systems can exceed the limits of canonical learning algorithms that need a large group of training and testing samples and they are ready to learn one different task [9]. Key features that need to be developed in this type of program are to use the previously learned information that contains feature modelling, saving lessons from the past, transferring information to future learning activities, updating formerly read and user feedback. Furthermore, the concept of ‘work’ present in many common definitions of Machine Learning lifetime models, proves that it is hard to define in many real setups.

With the proliferation of social media and the increase in time spent on it, users are inevitably using such tools for personal needs [3]. Those requirements include searching and reading articles of your favourite news.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue IV Apr 2021- Available at www.ijraset.com

IV. APPROACHES FOR FAKE NEWS DETECTION

.In this section of the paper we will discuss the methods used to distinguish fake stories and how these methods come together to be used to achieve a more effective outcome. Figure 1 shows the relationship between the 5 approaches for fake news detection, the ratio of the ellipses are based on the number of news articles found referring to that approach.

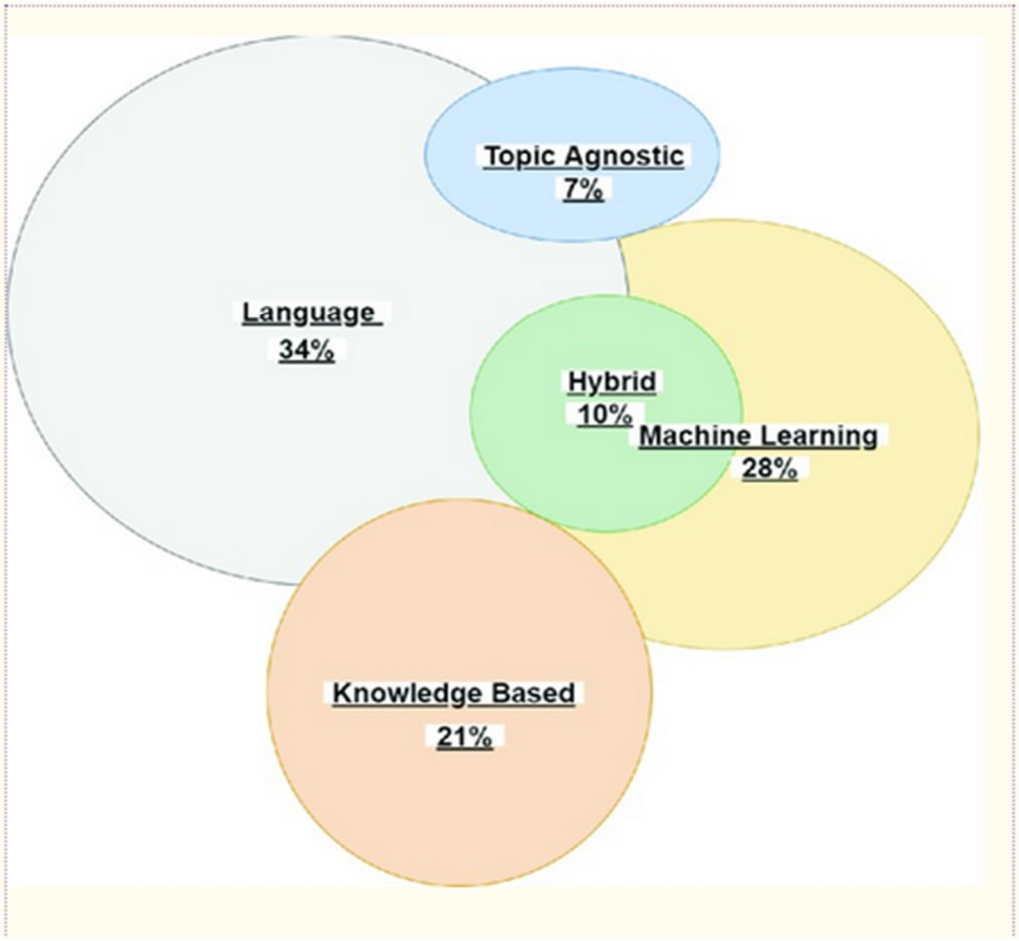


Fig. 1: 5 Approaches to Fake News Detection

A. Language Based Approach

This approach concentrates on the work of linguistic features in a software or personal program to find the malicious information. The people who spread fake news have a different and common slang on language from which they can be caught [17]. This approach looks at all the letters in the sentence word by word and how they fit together in a paragraph to verify the news. There are 3 methods in the language based approach:

1) *Bag of Words (BOW):* In this method each word is contemplated as an individual entity and based on the individual word frequency, signs of malicious information are found. This will eventually assist to recognize word usage design and patterns and by exploring these patterns , misleading and false information can be recognized . The word bag is not valid many times because the context is not considered when the text is converted into numerical representations and the word shape is not always considered [18].
2) *Semantic Analysis:* This approach explains that the truth can be set on by comparing personal information with a profile on a topic based on the same articles. An honest writer will be more likely to say the same thing about a topic than other truthful writers.
3) *Deep Syntax:* In this way the Probability Context Free Grammars performs extensive syntax functions using analytical trees that make Context Free Grammar analysis convenient. Words are changed into a set of rewritten rules and these rules are used to analyse various syntax structures and design. Syntax can be compared to known structures or false patterns and design and can eventually lead to telling the dissimilarity between false stories and true stories.

B. Topic Agnostic based Approach

This type of approach finds stories that are more about ignoring the content of the articles than they are about the features of the articles. This category of approach uses language features and web tagging capabilities to detect non-specific issues [19]. Other examples of features of agnostic topics are (1) an ample number of advertisements, (2) long story articles with conspicuous phrases, (3) distinctive text design and pattern from familiar stories to create emotional reaction, (4) the appearance of the author name [19].

C. Machine learning

Machine learning algorithms can be used to identify misinformation by using different types of training data stocks to improve the algorithms. Datasets enable computer scientists to develop new methods and techniques for machine learning and these datasets are used to train algorithms to identify non-existent stories.

A machine learning system called a rumour detection framework has been developed that verifies the signals of anonymous posts so that one can easily identify false stories [20]. The framework will inform people of potentially fraudulent posts and is built to combat fraudulent tweets on Twitter and focuses on four main areas; tweets metadata, tweet source; date and location of the tweet, compiled where and when.

To support this framework, gossip distribution is compiled to create data sets using the Twitter Streaming API.

Twitter has developed a possible solution to identify and prevent the spread of misleading information about fraudulent accounts, likes and comments by collecting tweets and entering them into a database, creating comparisons between various tweets are possible.

D. Knowledge based approach

Recent studies say that a combination of machine learning and information engineering to get the wrong information. A challenging problem with some of these ways of looking at facts is the speed at which false news spreads on social media. Microblogging platforms such as Twitter cause small fragments of false information to spread more quickly than large populations.

This knowledge-based approach aims to use external sources to verify whether or not stories are false and to identify stories before they spread quickly. There are three main categories; (1) Expert Oriented Fact Checking, (2) Computational Oriented Fact checking, (3) Crowdsourcing Oriented Fact Checking.

1) *Expert Oriented Fact Checking:* Expert Oriented fact checking is necessary to analyse and evaluate the data carefully and it : requires experts to assess the accuracy of the news by conducting research and other studies on a particular claim. Realism is the process of assigning certainty to something by comparing the accuracy of the text with something that has already been tested for truth.
2) *Computational Oriented Fact Checking:* The purpose of computer-based checking is to manage users through an automated authentication process that is able to determine whether a particular piece of news is true or false. A recent tool called ClaimBuster has been developed and is an example of how truth analysis can expose false stories. It uses machine learning techniques integrated with natural language processing and a variety of queries. It analyses the context in forums, interviews and speeches in real time to find 'facts' and compares them to a repository containing verified facts and presents them to the reader. [22]
3) *Crowdsourcing Oriented Fact Checking:* Crowdsourcing provides an opportunity for a group of people to make a collective decision by checking the accuracy of the news. The accuracy of the news is based entirely on the ingenuity of the crowd.

E. Hybrid Approach

Recent research has suggested that a hybrid model that helps to detect false news on social media using a combination of human and machine learning to help detect inaccurate information. To make the hybrid model work more effectively it combines social media issues with machine learning and neural network approach. The purpose of this type is to identify the possibility that the news is false.

Another hybrid model called CSI (catching, earning points, merging) is built and operates on key elements; (1) capture - the process of issuing article submissions through the Recurrent Neural Network (RNN), (2) Points - to create a point vector and representation, (3) Combine - to combine photographic results and obtain the result in a vector used for segmentation [23].

V. COMPARISON OF APPROACHES AND METHODS

Table I summarizes various aspects such as the objectives, the methods, datasets and the experimental analysis, used in this research. Datasets from various social media sites and Buzzfeed were used to propose the different models for detecting false news. All papers showed different accuracies and precision based on the efficiency of the model build.

Table I. Comparison of Approaches And Methods

SR NO	OBJECTIVE	YEAR	DATASET	METHODOLOGY	ACCURACY AND EVALUATION
1.	Detecting Fake News using Domain Reputations and Content Understanding [8]	2020	Facebook data from BuzzFeed News	Domain Reputation and Content Understanding based on tf-idf & LDA Modelling	--
2.	Fake news detection using Defensive Modeling [4]	2020	Online Social Networks (OSNs)	SVIR (Susceptible verified infected recovered) Model	R0(Basic reproduction number)
3.	Fake news Detection Using Deep Learning [10]	2020	Fake News Challenges (FNC) dataset	Hybrid Neural Network with CNN and LSTM	Accuracy: 97.8%
4.	Detecting fake news using Ensemble machine learning approach [14].	2020	Liar and ISOT dataset	Decision Tree Algorithm, Random Forest algorithm and Extra Tree Classifier algorithm	Accuracy in training: 99.8% Accuracy in testing: 44.15%
5.	Fake news detection using backtracking [5]	2019	Database from Social media sites.	FNDS using reverse tracking	Detection Rate: 85%
6.	Detecting Malicious Social Bots using Clickstream Sequences [13].	2019	user data	Transition probability, semi supervised clustering	Accuracy: 93.1% Recall rate: 97.5%
7.	Fake news detection using a sensitive stylistic approach [11].	2020	Images, text and network data from different sources and social media	NLP, Computational-stylistic analysis, LSA	Accuracy: 86% Precision: 94%
8.	Detecting fake news using advanced machine learning techniques [9].	2020	ISO and FA-KES	ML based Approach	Accuracy: 97.29% Precision: 65.18% True Positive Rate: 85% False Positive Rate: 0.1%
9.	Detecting fake news with weak social supervision [7].	2020	Social Media Dataset	Credibility-based, Bias-based, Sentiment-based	Accuracy: 75%
10.	Unreliable users detection using deep learning [3].	2020	Extracted data from Twitter	LSTM,CNN	Accuracy value of the social context dataset: 92.98% Accuracy value of the KNN classifier: 84%
11.	Detecting fake news using Deep neural approach [2].	2020	George McIntires Fake News Dataset , metadata from users	LSTM Neural Network ,GRU, Word2Vec, NLP problem based on Al Deep Learning Models,	Accuracy: LSTM(without mixed 83.7% , mixed 91.3%) Accuracy: Glove Feed Model (FNN) (without mixed 83.7%, mixed 84.3%)

12.	Fake news detection using CNN-RNN [6].	2021	George McIntire dataset, FA-KES dataset	Natural Language processing, convolutional and recurrent neural networks Approaches.	Precision0.48 Recall 0.48 F10.46 Accuracy 0.50
13.	Fake news detection using linguistic features [1].	2021	Zhang & Ghorbani, Bondioli & Marcelloni, satiric dataset	linguistic model, a memory-based learning model	Accuracy: 70%
14.	News verification using image featuring [16].	2021	dataset collected from Twitter, Sina Weibo	classification-based methods, tree building method	Accuracy: 83.6%
15.	Identification of fake news using memory network model [12].	2020	Fake News Challenge dataset	memory network model	Accuracy: 82.1%
16.	Fake news detection using multiple feature based models [15].	2020	Facebook dataset (Fake News NET)	n-gram model, ML based approach	Accuracy in K-Nearest Neighbour: 99.3 % Accuracy in Support Vector Machine : 99.3% Accuracy in Logistic Regression: 99%. Accuracy in Decision tree: 99.1% Accuracy in Long Short Term Memory: 99.4%

VI. APPLICATION

The Readers deserve to know the truth. To make up your mind on the basis of the facts with all the legitimacy is what you are capable of. Feeling insulted on being treated like an idiot on reading fake news is completely equitable.

Credibility is destroyed by fake news. Making an argument on the basis of the facts collected from fake news, deteriorates your credibility and makes it difficult for people to believe you in future instances.

Fake news can hurt the sentiments of the readers. Fake news can be harmful to you along with lots of other people. Source of phony and deceiving medical advice like Mercola.com and NaturalNews.com help propagate myths like HIV and AIDS aren't connected, or that antibodies cause mental imbalance. People tend to visit such sites often and such false information can mislead them and affect their lives dangerously.

Real news can benefit us. On the contrary real news can be beneficial for you. It can help you in fields such as if you want to buy stocks in a particular company, the real and factual articles about the company would help you lead to a wise investment. Another example could be casting a vote for the elections, substantial information on the candidate would help you to decide on considering which candidate you should vote for. Real news can help you to transform this world for a better place as well as help you earn money contrary to the fake news.

VII. CONCLUSIONS

Online social networks enable users to connect with each other and share information and therefore has become a massive hub where fake news has been shared and posted regularly. Through this research we explored the significance of identifying fake news and how this misinformation affects the audience's decision making. This paper presents the collective survey of fake news detection on various online social media platforms. The task of identifying fake news manually needs a deep knowledge of the domain and manpower therefore we needed more efficient methods for distinguishing fake and real news.

In this review paper we discussed the various traditional, deep learning and machine learning approaches available for detection of fake news and extensively analysed the concept of trust and identified the user behaviour on various posts on social media to distinguish fake news from real news. Various methodologies like LDA Modelling, SVIR Model, HNN with CNN and LSTM, Decision Tree, Random Forest, Extra Tree Classifier, reverse tracking, NLP, Computational-stylistic analysis and neural network models are used to predict the content and identify fake news. Some of the above stated models have acquired relatively higher accuracy than the other models. We used different performance measures to compare the experimental analysis of each model and identified the best suitable and efficient model for fake news detection.

REFERENCES

[1] Anshika Choudhary, Anuja Arora, Linguistic feature based learning model for fake news detection and classification, Expert Systems with Applications,Volume 169, 2021, 114171, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2020.114171.

[2] Deepak S, Bhadrachalam Chitturi, Deep neural approach to Fake-News identification, Procedia Computer Science, Volume 167, 2020, Pages 2236-2243, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.276.

[3] G. Sansonetti, F. Gasparetti, G. D'aniello and A. Micarelli, "Unreliable Users Detection in Social Media: Deep Learning Techniques for Automatic Detection," in IEEE Access, vol. 8, pp. 213154-213167, 2020, doi: 10.1109/ACCESS.2020.3040604.

[4] G. Shrivastava, P. Kumar, R. P. Ojha, P. K. Srivastava, S. Mohan and G. Srivastava, "Defensive Modeling of Fake News Through Online Social Networks," in IEEE Transactions on Computational Social Systems, vol. 7, no. 5, pp. 1159-1167, Oct. 2020, doi: 10.1109/TCSS.2020.3014135.

[5] Hoon Ko, Jong Youl Hong, Sangheon Kim, Libor Mesicek, In Seop Na, Human-machine interaction: A case study on fake news detection using a backtracking based on a cognitive system, Cognitive Systems Research, Volume 55, 2019, Pages 77-81, ISSN 1389-0417, https://doi.org/10.1016/j.cogsys.2018.12.018.

[6] Jamal Abdul Nasir, Osama Subhani Khan, Iraklis Varlamis, Fake news detection: A hybrid CNN-RNN based deep learning approach, International Journal of Information Management Data Insights, Volume 1, Issue 1, 2021, 100007, ISSN 2667-0968, https://doi.org/10.1016/j.jjimei.2020.100007.

[7] Shu, Kai & Hassan, Ahmed & Dumais, Susan & Liu, Huan. (2020). Detecting Fake News with Weak Social Supervision. IEEE Intelligent Systems. PP. 1-1. 10.1109/MIS.2020.2997781.

[8] K. Xu, F. Wang, H. Wang and B. Yang, "Detecting fake news over online social media via domain reputations and content understanding," in Tsinghua Science and Technology, vol. 25, no. 1, pp. 20-27, Feb. 2020, doi: 10.26599/TST.2018.9010139.

[9] Choraś, Michał & Demestichas, Konstantinos & Giełczyk, Agata & Herrero, Álvaro & Ksieniewicz, Paweł & Remoundou, Konstantina & Urda, Daniel & Wozniak, Michal. (2020). Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study. Applied Soft Computing. 101. 107050. 10.1016/j.asoc.2020.107050.

[10] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi and B. -W. On, "Fake News Stance Detection Using Deep Learning Architecture (CNN-LSTM)," in IEEE Access, vol. 8, pp. 156695-156706, 2020, doi: 10.1109/ACCESS.2020.3019735.

[11] N. R. de Oliveira, D. S. V. Medeiros and D. M. F. Mattos, "A Sensitive Stylistic Approach to Identify Fake News on Social Networking," in IEEE Signal Processing Letters, vol. 27, pp. 1250-1254, 2020, doi: 10.1109/LSP.2020.3008087.

[12] N. Ebadi, M. Jozani, K. -K. R. Choo and P. Rad, "A Memory Network Information Retrieval Model for Identification of News Misinformation," in IEEE Transactions on Big Data, doi: 10.1109/TBDATA.2020.3048961.

[13] P. Shi, Z. Zhang and K. R. Choo, "Detecting Malicious Social Bots Based on Clickstream Sequences," in IEEE Access, vol. 7, pp. 28855-28862, 2019, doi: 10.1109/ACCESS.2019.2901864.

[14] Saqib Hakak, Mamoun Alazab, Suleman Khan, Thippa Reddy Gadekallu, Praveen Kumar Reddy Maddikunta, Wazir Zada Khan, An ensemble machine learning approach through effective feature extraction to classify fake news, Future Generation Computer Systems, Volume 117, 2021, Pages 47-58,ISSN 0167-739X, https://doi.org/10.1016/j.future.2020.11.022.

[15] Somya Ranjan Sahoo, B.B. Gupta, Multiple features based approach for automatic fake news detection on social networks using deep learning, Applied Soft Computing, Volume 100, 2021, 106983, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2020.106983.

[16] Z. Jin, J. Cao, Y. Zhang, J. Zhou and Q. Tian, "Novel Visual and Statistical Image Features for Microblogs News Verification," in IEEE Transactions on Multimedia, vol. 19, no. 3, pp. 598-608, March 2017, doi: 10.1109/TMM.2016.2617078.

[17] Yang, Y., Zheng, L., Zhang, J., Cui, Q., Li, Z., Yu, P.S.: TI-CNN: convolutional neural networks for fake news detection (2018). arXiv preprint arXiv:1806.00749

[18] Potthast, M., Kiesel, J., Reinartz, K., Bevendorff, J., Stein, B.: A stylometric inquiry into hyperpartisan and fake news (2017). arXiv Preprint arXiv:1702.05638

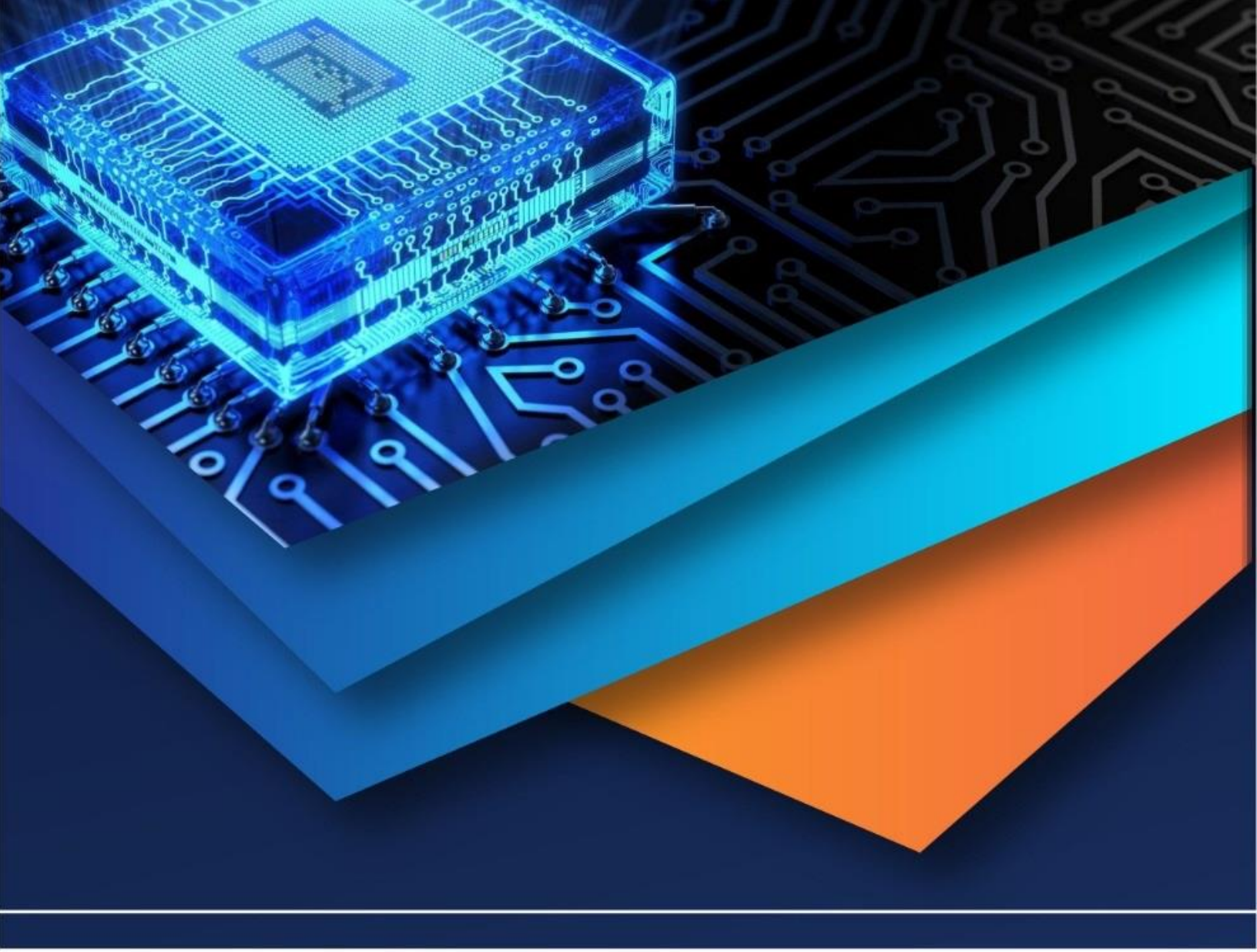
[19] Castelo, S., Almeida, T., Elghafari, A., Santos, A., Pham, K., Nakamura, E., Freire, J.: A topic-agnostic approach for identifying fake news pages. In: Companion Proceedings of the 2019 World Wide Web Conference on - WWW 2019, pp. 975–980 (2019). 10.1145/3308560.3316739

[20] Sivasangari V, Anand PV, Santhya R. A modern approach to identify the fake news using machine learning. Int. J. Pure Appl. Math. 2018;118(20):10. [Google Scholar]

[21] XINYI ZHOU, REZA ZAFARANI, Syracuse University, USA : A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities (2020)

[22] Hassan, N., Arslan, F., Li, C., Tremayne, M.: Toward automated fact-checking: detecting check-worthy factual claims by claimbuster. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 2017, Halifax, NS, Canada, pp. 1803–1812. ACM Press (2017). 10.1145/3097983.3098131

[23] Ruchansky, N., Seo, S., Liu, Y.: CSI: a hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM 2017, pp. 797–806 (2017). 10.1145/3132847.3132877



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY