

# Deep Learning-Powered Approaches for News Authenticity Verification and Fake News Detection: A Comprehensive Survey

<sup>1</sup>Alpana A. Borse, <sup>2</sup>Gajanan K. Kharate, <sup>3</sup>Namrata G. Kharate

<sup>1</sup>Resrach Scholar, MCERC, Nasik, Assistant Professor, PCCOE, Pune

<sup>2</sup>Director, MCERC, Nasik

<sup>3</sup>Assistant Professor, VIIT, Pune

---

## Article History:

**Received:** 12-12-2024

**Revised:** 28-1-2025

**Accepted:** 9-2-2025

## Abstract:

In today's digital world, increase of Fake News rapidly spreads substantial threats to both the public and individuals, emphasizing the crucial necessity for better-quality News Authenticity confirmation. This study gives a summary of some of the detection of bogus news using a Deep Learning (DL) technique and classification-based authenticity prediction techniques that to a large extent applied in several ways to news detection applications. Fake?True News Detection is currently a hard subject that is attracting investigation due to its detrimental effects on society. Deep Learning is employed in the crucial and often-used field of Fake News classification. Due to its excellent classification accuracy, the DL based approach has been widely used in the classification of news. Neural Network based – Convolutional Neural Networks (CNN), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), Generative Adversarial Networks (GAN), Hierarchical Attention Network (HAN), and Graph-based CNN (GCN) are some of the Deep Learning approaches that have been taken into consideration in our study and are used for developing a variety of News Detection methods. The study also includes comparative studies on a few news detection and classification methods that have been used to various Fake News prediction issues.

**Keywords:** Authenticity, Deep Learning, Fake News, HAN, Hypergraph Attention, Hypergraph Convolution

---

## 1. Introduction

In today's digital era, Fake News prediction has appeared as a serious encounter, aggravated by the quick feast of distortion transversely available display place. Customary procedures principally emphasis on examining manuscript or system structures autonomously, however they scrap to imprisonment the intricacy of the problem. Fake News involves a multi-dimensional interplay of factors—textual semantics, user engagement, and the interactions among articles, users, and

publishers—that cannot be effectively addressed in isolation. To address these constraints, the proposed technique leverages the capabilities of Hierarchical Attention Networks (HAN) and Hypergraph Convolutional Networks (HGNN) to represent the complex semantic and relational processes that underpin Fake News identification. By merging attention mechanisms with Hypergraph convolutions, the algorithm enables a deeper understanding of news content, user interactions, and publication networks for a more robust Fake News prediction system.

### *1.1. Problem Formulation*

Fake News prediction is defined as a detection and classification task, as goal of categorizing items as either Fake or True. The data being processed involves news items with a variety of features, including headlines, bodies, and metadata such as publication date. Additional further information, such as user interactions (comments, shares), source reliability, and topic similarity, help to enrich the input. The system also uses a diverse graph structure to connect users, articles, and publishers. The result is a binary label (fraudulent/Real) or a confidence score indicating the chance of an article being fraudulent. The key issue in this task is to accurately capture both the hierarchical structure of text and the complicated relationships between elements in the graph.

## **2. STATE-OF-THE-ART SURVEY**

### *2.1 Fundamentals of Distributed Learning*

The following researchers were active during this time:

The News Classification and Detection System were developed by Qing Liao et al. to concurrently classify topics and determine if the given news is true or false. In order to learn statement and relation representation, it employs N-Graph. Additionally, a dynamic weighting technique is employed to balance the relevance of the two tasks of topic classification and Fake News detection, using the GloVe tool as an unsupervised learning algorithm to create word vector representations [1].

K. Shu et al. presented the distinctive characteristics of Social media Fake News, such as foundations for psychology, malevolent accounts, and social foundations, in their article “Fake News Detection on Social Media: A data Mining Perspective”. They provide a summary of the various Fake News Prediction methods, currently in use, including stance based, style based, knowledge based, and propagation based. They also discuss a number of identify future directions and address unsolved concerns to identify Fake News on Social Media[2].

S. Yang et. al. suggested Unsupervised Social Media Fake News Detection. News accuracy, user trustworthiness, user interactions on social media, and latent random factors are used to determine opinions on the veracity of a news. Using a Model – Bayesian Network, they depict symbiotic relationships between news realities, user perceptions, and user credibility. To address the inference issue, they propose a productive collapsing Gibbs sampling method without any tagged data, to infer the News veracity and users reliability [3].

The purpose of "Detecting Fake News in Social Media Networks" by Monther Aldwairi et al. is to make money through clickbait, which allure users with attention-grabbing headlines or eye-catching designs to click links in order to boost ad income. This work aims to construct a method that consumers may be used to recognize and exclude websites that present false or misleading information. With the use of a logistic classifier, they accurately identify bogus posts using a few simple, carefully chosen elements of the title and content. [4].

According to Kai Shu et al.'s proposed Explainable Fake News classification, Fake News can be identified using dFEND methods, which build a co-attention sub-network between sentences and comments, using both user comments and news content [5].

Automatic Detection of Fake News was proposed by V. P'erez-Rosas and colleagues. Several sets of linguistic features, including N-grams, readability, syntax, and psycholinguistic features are extracted to accomplish this. They employed an SVM classifier for this [6].

A taxonomy of several veracity evaluation methods deriving from Linguistic Cue Approaches and network based approaches was proposed by N. J. Conroy et. al. in their paper, "Automatic deception detection: Methods for discovering false news. As a result, a novel hybrid approach is developed, fusing linguistic cues with Machine Learning and Network-based behavioral data. N-grams or a text bag of words technique are used to represent text. Three classifiers are used: Support Vector Machine (SVM), Rhetorical Structure (RST), Vector Space Model (VSM)[7].

K. Shu et al. suggested Tracker for bogus News, a tool for gathering, predicting, and visualizing bogus news. The system offers these 3 features: 1) Fake News collection, which involves the automatic gathering of news material and social situations. 2) To categorize Fake News, they introduced model – Social Article Fusion (SAF) which uses news content - Linguistic Characteristics and Social Environment characteristics. 3) Fake News Visualizations In this section, they use powerful visualisation tools to illustrate the characteristics of Fake News distribution [8].

Automatic false news classification approaches based on the discovery of Hidden-based patterns in news. Characteristics like Syntax, Style and content were proposed by S. Ghosh et al. Deep Neural N/k -based factors are used to identify if news is reliable. They employ techniques like the TF-IDF method, NLP, and Deep Learning for information retrieval as well. For classification, they employed SVM, RF, and logistic regression. They classified texts using articles. The features are applied to writing style and word-level features. There is a benchmark dataset [9].

In order to produce more precise and automated predictions, N. Ruchansky et. al. introduced a Hybrid Deep Model for Fake News Detection called CSI. In this, they proposed- a) Capture, b) Score, and c) Integrate methods. The Capture module uses Response and Text. Also, temporal pattern of user activity on a particular article is get captured with the use of RNN. In order to determine if an item is fraudulent or not, the third module integrates the second module's source characteristic learning, which is based on user behavior [10].

“Multi Source, Multi Class Fake News detection” was proposed by H. Karimi et al., which dealt with how to combine several source and how to distinguish classes' of fakeness degrees. They employed the CNN and LSTM models for this. This technique functions as a fakeness discrimination, multi-source fusion, and automated feature extraction [11].

Machine Learning Comparative Analysis of News Categorization Approaches was proposed by Nabamita Deb et al. The automatic news categorization problem is evaluated using some of the most popular machine learning approaches, primarily Neural Networks, random Forest, Naïve Bayes, SVM, and Decision Tree [12].

RNN Att - Attention mechanism (att-RNN) were employed by H. Z. Jin et.al. in their proposal for "Multimodal fusion with recurrent neural networks for rumour prediction on microblogs," which fused multimodal information for efficient rumour identification. In order to create a trustworthy fused classification in a complete network, where features were joined with – Text content and context joint features, and then employed with LSTM. In the rumour detection job, Topic models features- LDA are used[13].

Early Detection of Fake News based on Propagation Path Classification with RNN and CNN is a model that Y. Liu et al. proposed. In order to identify bogus news at an early stage of its spread, this model integrates Propagation Path Classification with RNN and CNN classification. For each altered propagation path, they learn a vector representation using an RNN version known as the Gated Recurrent Unit (GRU) [14].

Deep diffusive network model for Fake News Detection is introduced by J. Zhang et. al. .This is a brand-new FAKEDETECTOR automated Fake News credibility inference model. It makes use of a collection of explicit and latent qualities that were taken from the textual data. For learning the New representations, News creators and subject matters are used to design – “ A Deep Diffusive Network Model”. A GDU model is also used by FAKEDETECTOR for efficient relationship modelling between news pieces, creators, and topics [15].

Defensive Modelling of Fake News through Online Social Networks is a mathematical model that Gulshan Shrivastava and colleagues presented to analyze the Dynamic Fake News spread and controlling methods of message transmission in OSNs. It creates a Model that can identify and remove false information from OSNs, which will make certain OSN users feel less stressed about the epidemic. They employed a Jacobian matrix and applied the Lyapunov function to get the matrix's eigenvalues [16].

Fake News Detection - Using Machine Learning Ensemble Methods, which suggests a Machine Learning ensemble strategy for automatic classification of news articles, was proposed by Iftikhar Ahmad et al. They investigate many linguistic characteristics that can be utilized to tell bogus information from authentic. They train to different ML algorithms utilizing diverse ensemble approaches by making use of such properties. Convolutional Neural Networks were their main area of focus [17].

In "EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection" by Yaqing Wang et al., they suggest an whole framework that may generate characteristics and hence aid in the prediction of Fake News on recently occurring events. a) Multi model Feature Extractor, b) Fake News Detector and c)Event discriminator design its three primary parts. The features – Textual and Visual from postings are extracted by Multi-model Feature Extractor. The fundamental component of their textual feature extractor uses CNN. Utilized operations are the loss function, max-pooling operation, and ReLU activation function. [18].

"Liar, liar trousers on Fire" – is a new Benchmark Dataset for Fake News detection that was developed by W. Y. Wang et al. It is the largest dataset and is useful for short news [19].

In "Filter bubbles and Fake News," D. DiFranzo et al. present several surveys on technologies, such as social network analysis and machine learning, as well as how researchers continue to explore and research from multiple technological and social perspectives [20].

Syntactic stylometry for deception detection examines for deception detection, bringing a slightly unorthodox perspective to prior studies, as demonstrated S. Feng. et. al. They found that Context Free Grammar (CFG) parse trees have properties that are driven by CFG significantly and it improve the performance of identification over plenty of baselines which are based specifically features on "Shallow Lexico Syntactic feature", across four various datasets ranging based on the product evaluation domain to the Essay domain. Their findings increase the best-published result's accuracy by 14.0% while reducing errors by 91.2%. Additionally, they significantly outperform the Essay Data from Mihalcea and Strapparava (2009), achieving up to 85.0% accuracy [21].

The dissemination of true and fraudulent news online was postulated by S. Vosoughi et al. They looked into how each type of verified, factual, and fraudulent news story was disseminated differently on Twitter. They used data from six independent different fact-checking organizations for news classifications as genuine or false, and there was 95%–98% agreement between the classifications [22].

Y. Chen et al. are depicted in False internet information: Recognising click bait as misleading news that spreads quickly online and has been linked to the spread quickly of rumours and misinformation. The main goal of click bait – is to grab attention and get people to click on specific Web page link. They look at possible techniques for automatically spotting clickbait as a sort of fraud. A hybrid strategy may produce the best results, according to an assessment of techniques for spotting both Textual and Non-Textual Click-Baiting indicators. [23].

Fake News identification tools and methods have been demonstrated by Sakshini Hanglloo1 et al. in "Fake News Detection Tools And Methods - A Review"[24].

In "A Comprehensive Survey on Word Representation Models: From Classical to State-Of-The-Art Word Representation Language Models," Usman Naseem et al. demonstrate the various approaches and languages that exist to identify bogus news [25].

The goal of "Identifying Fake News and Fake Users on Twitter" by Costel-Sergiu et al. is to find out Twitters Fake users and Fake News. Twitter crawler component, the Processing module, which lends credibility to new tweets, and an algorithmic component called NER, which separates the text into its constituent pieces, bring out the themes, tweets and social-tags, general and hashtags sentiment. Additionally, we have modules for developing credibility and analyzing data, which analyze the most recent tweet in relation to what is locally recorded in our database [26].

"Deep Learning Algorithms for Detecting Fake News in Online Text" was idea put out by Sherry Girgis et al. The methods used to identify bogus news are numerous, creative, and even thrilling. This paper aims – to design a classifier, to determine News authenticity as True/False based solely on its content. For this, RNN, GRU as well as LSTMs are utilized to address to solve the issue. They demonstrate that GRU is the best, followed by LSTM, and then vanilla [27].

The "Fake News Identification Characteristics Using Named Entity Recognition and Phrase Detection" paper by Herley Shaori et al. tries employing the Indonesian language to design vectors that were taken into account features of misleading news before being further processed. This uses a Vector-Space approach to represent both real news and fraudulent news. A Term Frequency (TF), Inverse Document Frequency (IDF), and Frequency reverse vector model were combined using a Support Vector machine (SVM) technique classifier. In vector representation, Names Recognition Entity-NRE and phrase detection variations are also employed. The vector performance is then evaluated by TF - Term Frequency [28].

In "Automatic Online Fake News Detection Combining Content and Social Signals" by Macro L. Della et. al. et al., are the first to propose a Machine Learning (ML) Fake News Detection model that outperforms other methods in survey and improvements their accuracy by 4.8% by integrating social context characteristics with news articles. The second technique they employ is a Facebook Messenger chatbot [29].

In "Hierarchical attention networks for document classification," Z. Yang et. al. demonstrate how weighting word level-encoder, sentence level-encoder, and headline-encoder may be used to classify Fake News in news content [30].

The 3HAN (Pre-Trained) approach for Fake News identification was developed by S. Singhanian et al. It mostly interprets an article's structure in this. For quick, precise Fake News identification, a three-level hierarchical attention network is deployed. It has three layers, one for the headline and two for the words. It makes an effective representation for an input of news article, with the use of hierarchical bottom-up article processing [31].

To describe group wise interaction among news, Ujun Jeong et al. suggested a hypergraph based "Nothing Stands Alone: Relational Fake News Detection with Hypergraph Neural Networks" to highlight key news interactions relation using its dual – level attention method [32].

For the purpose of identifying bogus news, Federico Monti et al. presented a propagation-based system like Multi-Depth Graph Convolution Networks. It makes use of information from the social

graph, content, user behaviour, and user profile. To learn propagation representations, Gating Recurrent Units (GRU) and Convolutional Neural Networks (CNN) are used[33].

In "Attention Models in Graphs: A Survey," John Boaz Lee and colleagues conducted a thorough and in-depth literature review of the newly emerging topic of graph attention models. Additionally, three intuitive taxonomies based on the job (e.g., Graph classification and Link prediction, etc.), various types of Attention mechanism are used, and the problem setting were introduced. They provided specific instances, using each to compare rival strategies from a different angle [34].

Hierarchical Graph Attention Network (HGAT) was proposed by Y. Ren et. al. that is a novel framework for Fake News Detection - "Fake News Detection on News-oriented Heterogeneous Information Networks through Hierarchical Graph Attention,". This performs node representation learning and then classification of articles node in the HIN with the use of novel Hierarchical Attention Mechanism This compares to text based and other network models which performance on experience do better[35].

Song Bai et al. introduced "Hypergraph convolution and hypergraph attention" which has 2 end-to-end operators which become part of Graph Neural Networks, Hypergraph convolution and Hypergraph Attention respectively. It learns effectively deep embeddings on the high-order graph-structured data. The representation capability gets improved by Hypergraph Attention and fundamentals of formulation of convolution on Hypergraph is done by Hypergraph convolution. Likewise, representation and formulation using a Graph neural network can be easily designed with the help of these 2 operators - Attention and Convolution respectively These operators are applied to a case where non pairwise links are seen [36].

Yi-Ju Lu et al. demonstrate in "GCAN: Graph-aware Co-Attention Networks for Explainable Fake News Detection on Social Media" how to predict whether suspicious tweets are fake or not, by considering related retweets and the terms they concern and by generating information for pieces of evidence To do this, they create the innovative Graph aware Co-Attention Networks (GCAN) which are neural network-based model[37].

Z. Yang et. al. proposed "Hierarchical attention networks for document classification" based on Deep Learning. It has two distinguishing characteristics (i) Document hierarchy and (ii) attention to word and sentence level, finding or weighing differentially more and less important contents while constructing document representation or classification. In this, the attention layer visualization is presented to show how effectively the model selects most important highly informative words and sentences in a hierarchy [38]

According to Nicollas R. de Oliveira et al. in "Fake News on Social Networks Based on Natural Language Processing: Trends and Challenges". Fake News is revealed to be logical truth threat, to democracy, to journalism, and the authority of government-based institutions They look at machine learning, dimensionality reduction, vectorization, quality assessment of information retrieval, and

approaches for preparing data in natural language. We also explore research goals and prospects, contextualize the definition of false news, and [39]

On Sina Weibo, F. Yang et al. presented Automatic Rumour Detection. The analysis and detection of rumours on Sina Weibo has never been done before. They compile a sizable collection of microblogs that have been identified as fake rumors using data with help of rumor busting service from SinaWeibo. They made advantage of the kind of keyword-based query that Twitter Monitor specified. They discovered a collection of traits that may be taken from microblogs in order to classify them. These include numerous platform-specific characteristics, such as those that are location-based, client-based, content-based, propagation-based, account-based, and account-based on the Sina Weibo platform. [40]

Graph Convolution Networks for Text Classification were proposed by Liang Yao et al. This work is the first in which they model the Heterogeneous graph as an entire corpus and jointly train a Graph Neural Network to make document word and embeddings. Convolutional and recursive neural networks, both deep neural network models, were used to classify texts. To learn text representation, they employed LSTM, a particular kind of RNN. Additionally, attention mechanisms had been added as a crucial component of the text classification models [41]

Likewise, here literature survey is discussed and given in Table 1: Literature survey from reference [1] to [41].

### *2.2 Research Gap Identified:*

According to this survey, existing algorithms still require improvement in the prediction and classification of news authenticity. In order to increase accuracy, hierarchical attention networks and hypergraph should be used to create algorithms for Fake News prediction.

### *2.3 Motivation*

A Fake News Prediction system is a need of an era to make social media a credible place for the authenticity of any information and make a robust system for Fake News Prediction. For this, in 2022 ABP News done one agreement with IIM, Indore to make system robust against Fake News detection.

### *2.4 Objectives of the Proposed Research*

An accurate Fake News Prediction system towards predicting Fake News from social media resources using Prediction and Classification methods. This has tremendous amount of applications in surrounding for researchers and students. The prime goal of proposed research work is to predict and draw attention to Fake News. The objectives of this surveys are -

1. To comprehend the fundamentals of machine learning and data mining algorithms, confined to Prediction
2. To undergo through Literature review to understand the trends, challenges in Fake News Prediction

### 3 Methods

In this, we have discussed various fundamentals of word embedding, ML-based classifiers, DL based classifiers, Graph Convolution Networks and Hypergraph. With the idea of Fake News, several essential ideas compete and overlap. A synopsis for these multiple meanings, are listed as Satires and parodies, Rumors, Conspiracy theories, Spams, Scams and hoaxes, Clickbaits, Misinformation, disinformation, propaganda is given in Table 1 and Table 2 gives synthesis of various terms for Fake News Prediction.

**Table 1: Literature Survey for Fake News detection**

SN	Title	Cited	Year	Author	DOI	Publication	Methods used	Remark and Scope
1	An Integrated Multi-Task Model for Fake News Detection	[1]	January 2021	Qing Liao, et.al	<a href="https://doi.org/10.1109/TKDE.2021.3054993">10.1109/TKDE.2021.3054993</a>	IEEE Transaction	<ul style="list-style-type: none"> <li>▪ N-Graph</li> <li>▪ LSTM</li> <li>▪ Glove</li> </ul>	<ul style="list-style-type: none"> <li>▪ Detection and classification done, simultaneously</li> <li>▪ Improves the performance of short news fake detection</li> <li>▪ It learns statement and relation representation</li> </ul>
2	Fake News Detection on social media: A data mining perspective	[2]	2017	K. Shu, et. al.	<a href="https://doi.org/10.1145/3137597.3137600">10.1145/3137597.3137600</a>	Suhang Wang, SIGKDD Explorations,	<ul style="list-style-type: none"> <li>▪ Natural Language Features</li> </ul>	<ul style="list-style-type: none"> <li>▪ Examines the veracity of news reports</li> <li>▪ Features such as perceived cognitive authority, style content-based characteristics, and source authenticity</li> </ul>
3	Unsupervised Fake News Detection on social media: A Generative Approach	[3]	2019	S.Yang, et.al.	<a href="https://doi.org/10.1609/aaai.v33i01.33015644">10.1609/aaai.v33i01.33015644</a>	AAAI Conference on Artificial Intelligence	<ul style="list-style-type: none"> <li>▪ To conclude news authentication, Gibbs sampling approach is used               <ul style="list-style-type: none"> <li>▪ With unlabeled data, users' credibility is obtained</li> </ul> </li> <li>▪ Bayesian network model</li> </ul>	<ul style="list-style-type: none"> <li>▪ Check to see whether we can spot bogus news without supervision.</li> <li>▪ Utilize user interactions on social media to determine their attitudes about the veracity of news, treating the veracity of news and users' credibility as latent random variables.</li> </ul>
4	Detecting Fake News in Social Media Networks	[4]	2018	Monther Aldwairi	<a href="https://doi.org/10.1016/j.procs.2018.10.171">10.1016/j.procs.2018.10.171</a>	EUSPN Conference	<ul style="list-style-type: none"> <li>▪ WEKA Classifiers</li> <li>▪ Pythonscript</li> <li>▪ InfoGainAttributeEval</li> <li>▪ Correlation AttributeEval</li> </ul>	When a person clicks on a link, they are taken to a website with content that is far less than what they were expecting. This is known as clickbait.
5	defend: Explainable Fake News Detection	[5]	August 2019	K. Shu	<a href="https://doi.org/10.1145/3292500.3330935">10.1145/3292500.3330935</a>	ACM SIGKDD International Conference	<ul style="list-style-type: none"> <li>▪ lexical Features</li> <li>▪ Syntax Features</li> <li>▪ Semantic Features</li> <li>▪ Linguistic Features</li> </ul>	<ul style="list-style-type: none"> <li>▪ Created a Sentence and comment-based Co-Attention subnetwork to take advantage of news items and users comments.</li> <li>▪ Used side information from other user interactions, likes, to find comments that have explanations.</li> <li>▪ The credibility of people who leave explicative comments is utilized</li> </ul>
6	Automatic Detection of	[6]	2018	V. P'erez-Rosas, et.al	<a href="https://doi.org/10.48550/arXiv.170">10.48550/arXiv.170</a>	27th ICCLC International	<ul style="list-style-type: none"> <li>▪ Used syntactic</li> </ul>	<ul style="list-style-type: none"> <li>▪ Automatic fake News detection</li> <li>▪ Accuracy is more if compared</li> </ul>

	Fake News				<a href="#">8.07104</a>	Conference Linguistics COLING	and semantic level of news	with manual system
7	Automatic deception Detection: Methods for finding Fake News	[7]	2018	N. J. Conroy	<a href="#">10.1002/p ra2.2015. 14505201 0082</a>	78th ASIS & Tannual Meeting: Information Science	<ul style="list-style-type: none"> <li>▪ Linguistic cue approach</li> <li>▪ Network analysis approach</li> <li>▪ Deep Syntax</li> <li>▪ Semantic analysis</li> </ul>	Operational guidelines for a workable News detection system Linguistic-based and Network-based approaches have produced classification results with high accuracy in a few specific fields.
8	FakeNewstracker: a tool for Fake News collection, Detection, and visualization	[8]	2019	K. Shu, et.al.	<a href="#">0.1007/s1 0588- 018- 09280-3</a>	Computational & Mathematical Organization Theory	<ul style="list-style-type: none"> <li>▪ Semi-supervised features</li> <li>▪ Streaming is used</li> </ul>	<ul style="list-style-type: none"> <li>▪ Can use features like favorites, retweets, social network</li> <li>▪ Can explore in real-time</li> <li>▪ Automatically collect data for news and visualization</li> </ul>
9	Toward automatic Fake News	[9]	2019	S. Ghosh	<a href="#">10.1002/p ra2.2018. 14505501 125</a>	HICSS	<ul style="list-style-type: none"> <li>▪ TF-IDF</li> <li>▪ NLP, SVM</li> <li>▪ RF</li> <li>▪ Logistic regression</li> </ul>	<ul style="list-style-type: none"> <li>▪ Accuracy 82.4%</li> <li>▪ Vector space model</li> </ul>
10	CSI: A hybrid deep model for Fake News Detection	[10]	2017	N. Ruchansky	<a href="#">10.1145/3 132847.3 132877</a>	ACM on Conference on Information and Knowledge Management	<ul style="list-style-type: none"> <li>▪ SVM</li> <li>▪ LSTM</li> <li>▪ GRU</li> </ul>	<ul style="list-style-type: none"> <li>▪ The first novel model</li> <li>▪ Capture the three shared properties – Text, response and source which detects false information.</li> </ul>
11	Multi-source multi- class Fake News detection	[11]	2018	H. Karimi	--	International Conference on Computational Linguistics	<ul style="list-style-type: none"> <li>▪ CNN</li> <li>▪ LSTM</li> </ul>	<ul style="list-style-type: none"> <li>▪ Use many sources and differentiate between different levels of fakeness.</li> <li>▪ Automation of Feature extraction</li> <li>▪ Multiple-source fusion</li> </ul>
12	A Comparative Analysis of News Categorization Using Machine Learning Approaches	[12]	JANUARY 2020	Nabamita Deb	<a href="#">10.3390/i nfo13120 576</a>	International Journal Of Scientific & Technology Research	<ul style="list-style-type: none"> <li>▪ Survey Analysis of News Categorization Using Machine Learning Approaches</li> </ul>	<ul style="list-style-type: none"> <li>▪ Analyzing some of most popular machine learning techniques- naïve Bayes techniques performs better on average than the other four algorithms, with a classification accuracy of 96.8%.</li> <li>▪ The Random Forest is next, with a 94.1% accuracy rate.</li> <li>▪ 96.4% accuracy Support Vector Machine (SVM)</li> <li>▪ The Decision Tree had an accuracy of 83.2% while Neural Networks had an accuracy of 96.4%.</li> </ul>
13	Multimodal fusion with recurrent Neural Networks for rumor Detection	[13]	2017	H. Z. Jin	<a href="#">10.1145/3 123266.3 123454</a>	Proceedings of the ACM on Multimedia Conference	<ul style="list-style-type: none"> <li>▪ RNN</li> <li>▪ Attention mechanism</li> <li>▪ LSTM</li> </ul>	<ul style="list-style-type: none"> <li>▪ A unique Attention-based Recurrent Neural Network (att-RNN) to aggregate multimodal signals for efficient rumour detection.</li> <li>▪ Experiments are done with 2 multimedia datasets collected from Twitter and Weibo.</li> <li>▪ In this, all complete Image features are incorporated into the joint features of Social text and its context</li> </ul>
14	Early Detection of Fake News	[14]	2018	Y. Liu	<a href="#">10.1609/a aai.v32i1.</a>	AAAI Conference	<ul style="list-style-type: none"> <li>▪ Geometric Deep</li> </ul>	<ul style="list-style-type: none"> <li>▪ Detecting of bogus news, based on Deep Learning</li> </ul>

	on social media through propagation path Classification with recurrent and convolutional networks				<a href="#">11268</a>	on Artificial Intelligence	Learning approach <ul style="list-style-type: none"> <li>Convolution neural network</li> </ul>	method of propagation is employed.
15	FAKEDetectOR: Effective Fake News Detection with Deep Diffusive Neural Network	[15]	2020	J. Zhang	<a href="#">10.1109/ICDE4830.7.2020.00180</a>	Proceedings of the IEEE 36th International Conference on Data Engineering (ICDE)	<ul style="list-style-type: none"> <li>Creator-Article Publishing Historical Records</li> <li>Subject Credibility Analysis:                             <ul style="list-style-type: none"> <li>Deep Diffusive Unit Model</li> <li>Gru</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>Looking into the guiding ideas, approaches, and formulas for spotting Fake News</li> <li>Content, authors, and issues from online social networks, together with an assessment of how well they perform</li> <li>The issues posed by the unidentified qualities are addressed in this work.</li> </ul>
16	Defensive Modeling of Fake News Through Online Social Networks	[16]	October 2020	Gulshan Shrivastava	<a href="#">10.1109/TCSS.2020.3014135</a>	IEEE Transactions On Computational Social Systems	<ul style="list-style-type: none"> <li>Jacobian matrix</li> <li>Eigenvalues of the matrix</li> <li>Lyapunov function used</li> </ul>	<ul style="list-style-type: none"> <li>They proposed a mathematical model to explore regulating actions and dynamic spreading of message transmission in OSNs</li> <li>A model is design that can identify and remove false information from OSNs.</li> <li>alleviate some OSN users' anxiety about the pandemic</li> </ul>
17	Fake News Detection Using Machine Learning Ensemble Methods	[17]	2020	Iftikhar Ahmad	<a href="#">10.1155/2020/8885861</a>	Hindawi, Wiley	<ul style="list-style-type: none"> <li>Naive Bayes</li> <li>Decision Tree</li> <li>Adaboost</li> <li>SVM</li> <li>Linguistic features such as n-grams</li> <li>KNN</li> <li>Logistic regression (LR),</li> <li>Stochastic gradient descent (SGD)</li> </ul>	<ul style="list-style-type: none"> <li>The explanation of Machine Learning Ensemble Methods</li> <li>Gradient descent (SGD), with SVM obtaining the maximum accuracy of 92%</li> </ul>
18	EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection	[18]	August 2018	Yaqing Wang	<a href="#">10.1145/3219819.3219903</a>	24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining	<ul style="list-style-type: none"> <li>Feature Extractor</li> <li>text, visual features</li> </ul>	<ul style="list-style-type: none"> <li>End-to-end framework named EANN                             <ul style="list-style-type: none"> <li>Fake News Detector</li> <li>Used Twitter dataset</li> </ul> </li> </ul>
19	Liar, liar pants on fire: A new benchmark dataset for Fake News detection	[19]	2017	W. Y. Wang	<a href="#">10.18653/v1/P17-2067</a>	Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics	<ul style="list-style-type: none"> <li>Dataset</li> </ul>	<ul style="list-style-type: none"> <li>Largest Fake News dataset                             <ul style="list-style-type: none"> <li>Useful for short news</li> </ul> </li> </ul>
20	Filter bubbles and Fake News	[20]	2017	D. DiFranzo	<a href="#">10.1145/3055153</a>	ACM Crossroads	<ul style="list-style-type: none"> <li>Survey</li> </ul>	<ul style="list-style-type: none"> <li>Various surveys are conducted regarding techniques, including</li> </ul>

								social network analysis and machine learning <ul style="list-style-type: none"> <li>▪ Researchers are still exploring and looking into them from many technical and sociological angles</li> </ul>
21	Syntactic stylometry for deception detection	[21]	2012	S. Feng	==	50th Annual Meeting of the Association for Computational Linguistics	<ul style="list-style-type: none"> <li>▪ Feature Encoding</li> <li>▪ Words</li> <li>▪ Shallow Syntax</li> <li>▪ Deep syntax</li> </ul>	<ul style="list-style-type: none"> <li>▪ Syntactic stylometry is learnt for deception detection</li> <li>▪ Parse tree Features from Context Free Grammar are used to consistently enhance the detection performance</li> </ul>
22	The spread of true and false news online	[22]	2018	S. Vosoughi	<a href="https://doi.org/10.1126/science.aap95">10.1126/science.aap95</a>	Science, this issue p. 1146	<ul style="list-style-type: none"> <li>▪ Complementary cumulative distribution functions (CCDFs)</li> <li>▪ Depth, Size, Maximum breadth, depth and (F)</li> <li>▪ A number of unique Twitter users.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Correlations between the spread of news is found</li> <li>▪ the sentiments of responses for the news is found</li> </ul>
23	Misleading online content: Recognizing click bait as false news	[23]	2015	Y.Chen,	<a href="https://doi.org/10.1145/2823465.2823467">10.1145/2823465.2823467</a>	Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection.	<ul style="list-style-type: none"> <li>▪ Hybrid approaches</li> </ul>	<ul style="list-style-type: none"> <li>▪ Its primary goal - to draw attention and persuade viewers; also a certain web page is obtained with the help of click on link</li> <li>▪ It looks at potential tools for automatically spotting clickbait as a kind of deception.</li> <li>▪ Textual and non-textual Techniques for identifying click baiting cues are reviewed</li> </ul>
24	Fake News Detection Tools and Methods – A Review	[24]	Nov 2021	Sakshini Hangloo1	<a href="https://doi.org/10.48550/arXiv.2112.11185">10.48550/arXiv.2112.11185</a>	International Journal of Advance and Innovative Research, Volume 8, Issue 2 (IX) ISSN 2394 - 7780	<ul style="list-style-type: none"> <li>▪ Review</li> </ul>	<ul style="list-style-type: none"> <li>▪ Machine learning tools and its related methods survey is given in this</li> </ul>
25	A Comprehensive Survey on Word Representation Models	[25]	2018	Usman Naseem	<a href="https://doi.org/10.1145/3434237">10.1145/3434237</a>	ACM Transactions on Asian and Low-Resource Language Information Processing	Text Mining, Natural Language Processing, Word representation, Language Models	<ul style="list-style-type: none"> <li>▪ Various word representation models are explained and compared here with analysis</li> </ul>
26	Identifying Fake News and Fake Users on Twitter	[26]	2018	Costel-Sergiu Atodiresei	<a href="https://doi.org/10.1016/j.procs.2018.07.279">10.1016/j.procs.2018.07.279</a>	International Conference on Knowledge Based and Intelligent Information and Engineering	<ul style="list-style-type: none"> <li>▪ Text based features</li> <li>▪ N-gram</li> <li>▪ Propagation based approach</li> </ul>	<ul style="list-style-type: none"> <li>▪ System created with the intention of detecting bogus Twitter users and news sources</li> </ul>

27	Deep Learning Algorithms for Detecting Fake News in Online Text	[27]	2018	Sherry Girgis	<a href="#">10.1109/1CCES.2018.8639198</a>	13th International Conference on Computer Engineering and Systems (ICCES)	<ul style="list-style-type: none"> <li>RNN (vanilla, GRU)</li> <li>LSTMs</li> <li>GRU</li> </ul>	<ul style="list-style-type: none"> <li>Based on its content, a classifier determine if a piece of news is true or false</li> <li>Results indicate that GRU is the best, followed by LSTM, and vanilla</li> </ul>
28	Fake News Identification Characteristics Using Named Entity Recognition and Phrase Detection	[28]	2018	Herley Shaori Al-Ash	<a href="#">10.1109/1CITEED.2018.8534898</a>	10th International Conference on Information Technology and Electrical Engineering (ICITEE)	<ul style="list-style-type: none"> <li>Frequency term,</li> <li>Inverse document frequency</li> <li>Support vector machine</li> <li>NER</li> </ul>	<ul style="list-style-type: none"> <li>Uses the Indonesian language to form vector which handles features of fake News</li> <li>This uses vector space paradigm to differentiate Legitimate and Fake News</li> <li>With the word frequency, Vector performance representation seems auspicious.</li> </ul>
29	Automatic Online Fake News Detection Using Social Signals and Content	[29]	2018	Marco L. Della	<a href="#">10.23919/FRUCT.2018.8468301</a>	Proceeding of the 22nd Conference of Fruct Association.	<ul style="list-style-type: none"> <li>News content features</li> <li>Social content features</li> <li>ML methods</li> </ul>	<ul style="list-style-type: none"> <li>Combines social context features with news articles</li> <li>A Facebook Messenger chatbot uses the diffusion pattern as an implementation strategy</li> </ul>
30	Hierarchical Attention Network for Document Classification	[30]	2016	Z. Yang	<a href="#">10.18653/v1/N16-1174</a>	Conference of the North American Chapter of the Association for Computational Linguistics (NACCL)	<ul style="list-style-type: none"> <li>GRU</li> <li>Word encoder</li> <li>Sentence encoder</li> <li>Attention Mechanism</li> </ul>	<ul style="list-style-type: none"> <li>2 levels</li> <li>A structure that is equivalent to the hierarchy of documents</li> <li>Second-level: Attention model is designed at Word and Sentence level</li> </ul>
31	3HAN: A Deep Neural Network for Fake News Detection	[31]	2017	S. Singhania	<a href="#">10.1007/978-3-319-70096-0_59</a>	Neural Information Processing - 24th International Conference, ICONIP	<ul style="list-style-type: none"> <li>GRU</li> <li>Word, Sentence, and Headline encoder</li> <li>Attention Mechanism</li> </ul>	<ul style="list-style-type: none"> <li>Article structure having 3 levels word-encoder, sentence encoder, and headline encoder</li> </ul>
32	Nothing stands Alone: Relational fake News Detection with Hypergraph Neural Networks	[32]	2022	Ujun Jeong	<a href="#">10.48550/arXiv.2212.12621</a>	IEEE Conference BigData	<ul style="list-style-type: none"> <li>GCN</li> <li>Hypergraph Attention</li> </ul>	<ul style="list-style-type: none"> <li>Use a hypergraph to depict how news is interacted with collectively</li> </ul>
33	Multi Depth Graph Convolution N/k for News detection	[33]	2019	Federico Monti	<a href="#">10.1007/978-3-030-32233-5_54</a>	Natural Language Processing and Chinese Computing Conference paper	<ul style="list-style-type: none"> <li>Graph</li> <li>CNN</li> </ul>	<ul style="list-style-type: none"> <li>Employed propagation-based techniques</li> <li>Utilized information including user profiles, content, activities, Social Network, and news dissemination</li> </ul>
34	Attention Models in a Graph: A Survey	[34]	2018	John Boaz Lee	<a href="#">doi.org/10.00001.000001</a>	ACM Trans. Knowl. Discov. Data.	<ul style="list-style-type: none"> <li>Survey</li> </ul>	<ul style="list-style-type: none"> <li>Literature Review on Graph Attention models</li> </ul>
35	Fake News Detection on News Oriented heterogeneous information	[35]	July 2021	Y. Ren and J. Zhang	<a href="#">10.1109/1JCNN52387.2021.9534362</a>	International Joint Conference on Neural Networks	<ul style="list-style-type: none"> <li>HGAT</li> <li>Attention mechanism</li> </ul>	<ul style="list-style-type: none"> <li>A novel Framework – to detect Fake News</li> <li>It uses a Hierarchical Attention mechanism, to learn node representation</li> </ul>

	networks through Hierarchical Graph Attention					(IJCNN). IEEE		
36	Hypergraph convolution and hypergraph attention	[36]	2020	Song Bai	<a href="https://doi.org/10.1016/j.patcog.2020.107637">10.1016/j.patcog.2020.107637</a>	Internal journal Pattern Recognition	<ul style="list-style-type: none"> <li>Hypergraph convolution</li> <li>Hypergraph attention.</li> </ul>	<ul style="list-style-type: none"> <li>Utilises a hierarchical attention technique to carry out node representation learning in HIN</li> <li>New framework for false news identification.</li> </ul>
37	GCAN: Graph aware Co-Attention Networks for Explainable fake News Detection on Social Media	[37]	2020	Yi-Ju Lu	<a href="https://doi.org/10.18653/v1/2020.acl-main.48">10.18653/v1/2020.acl-main.48</a>	Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics	<ul style="list-style-type: none"> <li>Tweets are used</li> <li>Graph CNN</li> <li>GRU</li> </ul>	<ul style="list-style-type: none"> <li>GCAN consistently outperforms cutting-edge techniques by 16% in accuracy.</li> <li>Co-attention between sources</li> <li>Co-attention for Source-Propagation</li> </ul>
38	Hierarchical Attention Network for Document Classification	[38]	2016	Z. Yang	<a href="https://doi.org/10.18653/v1/N16-1174">10.18653/v1/N16-1174</a>	North American Chapter of the Association for Computational Linguistics (NACCL)	<ul style="list-style-type: none"> <li>HAN</li> <li>Attention visualization</li> </ul>	<ul style="list-style-type: none"> <li>2 distinguishing characteristics                             <ul style="list-style-type: none"> <li>(i) Its organizational structure is hierarchical.</li> <li>(ii) Word-and Sentence levels are used as Attention Strategies</li> </ul> </li> <li>Attention visualization</li> </ul>
39	Identifying Fake News on Social Networks based on natural Language Processing: Trends and Challenges	[39]	2021	Nicollas R	<a href="https://doi.org/10.3390/info12010038">10.3390/info12010038</a>	MDPI Information	<ul style="list-style-type: none"> <li>Reduction,</li> <li>Machine learning</li> <li>NLP</li> <li>Survey</li> </ul>	<ul style="list-style-type: none"> <li>Natural language, vectorization, and dimensionality survey preprocessing techniques for data</li> <li>Discuss research objectives and prospects while contextualizing the detection of Fake News</li> <li>Information retrieval quality evaluation</li> </ul>
40	Automatic detection of rumor on Sina-Weibo	[40]	2015	F. Yang	<a href="https://doi.org/10.1007/978-3-319-25207-0_10">10.1007/978-3-319-25207-0_10</a>	Natural Language Processing and Chinese Computing, NLPCC	<ul style="list-style-type: none"> <li>Features</li> <li>Client, Content, Propagation, account and Location based</li> </ul>	<ul style="list-style-type: none"> <li>First study on Sina Weibo rumour analysis and detection</li> <li>Gather a large collection of microblogs</li> <li>Used the Twitter Monitor-defined style of keyword-based query</li> </ul>
41	Graph Convolution Networks for Text Classification		2019	Liang Yao	<a href="https://doi.org/10.48550/arXiv.1809.05679">10.48550/arXiv.1809.05679</a>	33rd AAAI Conference on Artificial Intelligence	<ul style="list-style-type: none"> <li>Deep neural network models</li> <li>CNN, RNN LSTM</li> </ul>	<ul style="list-style-type: none"> <li>In this study, for the first time, they represent the entire corpus like heterogeneous Graph and Graph Neural Networks uses to concurrently train word document and document embeddings</li> <li>They also introduced the attention mechanism as a crucial component of the models used for text categorization</li> </ul>

**Table 2: Synthesis of various terms for Fake News and its meaning**

SN	Attribute	Meaning
1	Satires	Humour that is embedded and uses sarcasm and irony
2	Rumors	It is not based on events of news, but are widely acknowledged
3	Conspiracy	Hard to authenticate as True/False

4	Spams	Also referred - unsolicited emails and messages
5	Scams or Hoaxes	Simply for fun or to fool the people being targeted
6	Clickbaits	Make users want to access and distribute questionable content by using tiny graphics or spectacular headlines, a kind of deceptive marketing
7	Misinformation	Unintentionally produced, without a clear source or aim to deceive the reader;
8	Disinformation	Information elements that have been generated specifically to confuse readers
9	Propaganda	Refers to news articles that have been produced and spread by a political organisation in order to affect public opinion

Significant development has been seen in Fake News Prediction methods to reduce negative impacts on the society by providing authentic News to society and individuals

3.1. Fake News Detection Approaches used for

- Propagation based Approach
- Stance based Approach
- Deep Learning based Approach
- Graph – Neural Network based Approach
- Network Approaches
- Linguistic Approaches

**Table 3: Table for Detection and Classification for Fake News Approaches**

SN	Methods/ Approaches	Used by	Remark
1	Propagation-Based Approach	[14],[18]	<ul style="list-style-type: none"> <li>▪ An efficient method for discovering linkages in a graph</li> <li>▪ It does not need an objective function that has been predefined or knowledge of the communities beforehand in order to detect these communities; instead, it uses network structure alone as its sole guide.</li> </ul>
2	Stance-Based Approach	[6], [18], [21]	<ul style="list-style-type: none"> <li>▪ Rely on manually created Embedding/Linguistic features to identify writing, linguistic hints, or keywords emphases from News articles.</li> </ul>
3	Deep Learning-Based Approach	[1], [9]-[11], [30], [31]-[37], [41], [42]	<ul style="list-style-type: none"> <li>▪ Deep Learning based Neural Networks are designed for detection of any Fake News automatically.</li> </ul>
4	Graph Neural Network-Based Approach	[32]-[37], [41], [42]	<ul style="list-style-type: none"> <li>▪ By utilizing neighborhood aggregation to convey messages between close neighbors in the graph, GNN based models seek to improve representation for node and edge</li> <li>▪ As a mean-pooling neighborhood, GCN can be explained.</li> </ul>
5	Network Approaches	[14], [18]	<ul style="list-style-type: none"> <li>▪ A network strategy needs a lot of people who regard themselves as leaders, are able to reach out to others, organize exploratory projects, and collaborate in novel ways.</li> <li>▪ Transformation frequently occurs as a result of our interactions with others, not just what we accomplish.</li> </ul>
6	Linguistic Approaches	[7]	<ul style="list-style-type: none"> <li>▪ An approach to translation from a linguistic perspective emphasises the meaning and similarity of words and expressions in sentences</li> </ul>

3.2 : News [Text] Classification – Steps

To categorize the given news in a particular category. Its procedure is shown in Figure 3 Text Classification steps

- Text representation

Figure 3 gives the processing to convert raw data into clean data and also it tells us to extract features of news. It consists of two blocks

- Text Preprocessing

It converts text into a clean and consistent format

- Feature Extraction

There are various Feature extraction methods given in Figure 4

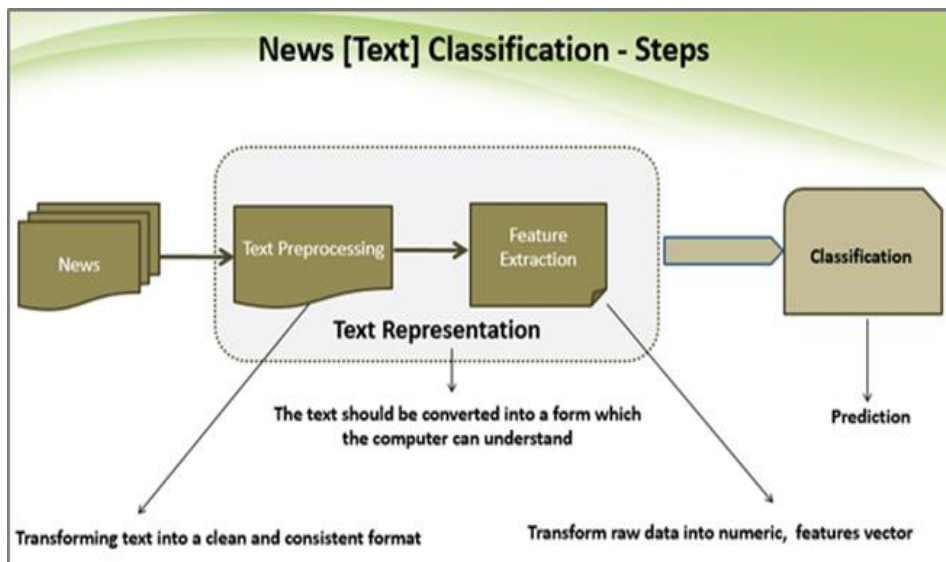


Figure 3: Text Classification steps

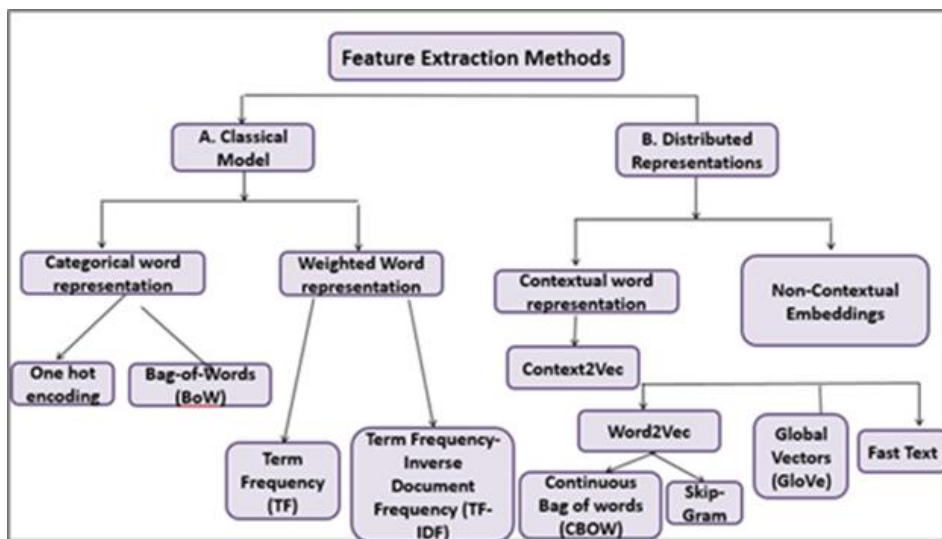


Figure 4 : Feature Extraction Method

### 3.3 Feature Extraction Methods

Figure 4 gives the pictorial representation of Feature extraction methods. Figure 5 gives a Classification example and Figure 6 gives Text preprocessing methods respectively.

**Table 4: Comparison between various Methods for Word Representation Models**

SN	Methods	Based on	Purpose	Advantage	Disadvantage
1	One Hot Encoding	<ul style="list-style-type: none"> <li>▪ Count based</li> <li>▪ Categorical word representation</li> </ul>	The conversion of categorical information into a format that may be fed into machine learning algorithms to improve prediction accuracy	<ul style="list-style-type: none"> <li>▪ Enables us to apply machine-learning algorithms to data with categorical columns</li> </ul>	<ul style="list-style-type: none"> <li>▪ The problem is that it increases dimensionality so training becomes slower and more complex</li> </ul>
2	BoW	<ul style="list-style-type: none"> <li>▪ Count based</li> <li>▪ Categorical word representation</li> </ul>	Used to determine the word count of texts and documents Word order within a document is not taken into account by the model.	<ul style="list-style-type: none"> <li>▪ Its simplicity and usability</li> <li>▪ It can be used to develop a rough draught model before moving on to more complex word embeddings</li> </ul>	<ul style="list-style-type: none"> <li>▪ If the new phrases use new terms, then our vocabulary would grow, which would also lengthen the vectors.</li> <li>▪ A sparse matrix would be produced as a result of the vectors' high number of zeros, which is something we want to avoid</li> </ul>
3	TF	Frequency based Weighted word representation	It evaluates the weight of a phrase in a document in light of a collection of documents	<ul style="list-style-type: none"> <li>▪ Simple and easy to use it is</li> </ul>	<ul style="list-style-type: none"> <li>▪ Does not consider the significance of phrases that are used sparingly throughout a collection of documents' individual documents</li> </ul>
4	TF-IDF	Frequency based Weighted word representation	Metric that uses statistics to determine what relevance a term has to each page in a group of papers	<ul style="list-style-type: none"> <li>▪ It is an easy starting point for similarity computations, is inexpensive to compute, and is straightforward to calculate.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Assign low values to comparatively significant terms, and be excessively sensitive on the extensive margin and excessively resistant on the intensive margin.</li> </ul>
5	Word2Vec	<ul style="list-style-type: none"> <li>▪ Prediction-based method</li> </ul>	<ul style="list-style-type: none"> <li>▪ Representing word embeddings</li> <li>▪ A fixed vector size</li> </ul>	<ul style="list-style-type: none"> <li>▪ The embedding vector is tiny and versatile</li> <li>▪ Able to capture relationships between words, including their syntactic and semantic relationships</li> <li>▪ Since it is unsupervised, less human effort is required to tag the data</li> </ul>	<ul style="list-style-type: none"> <li>▪ Word2Vec can't handle terms that aren't part of its lexicon well; it can't share training parameters for additional languages; and it needs a somewhat larger corpus before the network converges.</li> </ul>
6	GloVe	<ul style="list-style-type: none"> <li>▪ Prediction-based method</li> <li>▪ count-based technique</li> <li>▪</li> </ul>	<ul style="list-style-type: none"> <li>▪ It aims to combine the best of prediction-based technique (Word2Vec) and count-based strategy (co-occurrence matrix), hence it is also known as a Hybrid technique for continuous representation of words</li> </ul>	<ul style="list-style-type: none"> <li>▪ It typically outperforms word2vec in analogy tasks</li> <li>▪ GloVe requires less training time because it is easier to parallelize than Word2Vec.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Compared to word2vec, GloVe uses a co-occurrence matrix and global information, which increases memory cost.</li> <li>▪ Because words and vectors are identical, it, too, like word2vec, does not address the problem of polysemous words</li> </ul>
7	FastTex	Text based Non Contextual word representation	<ul style="list-style-type: none"> <li>▪ Achieving scalable solutions for the challenges of text categorization and representation while reliably and quickly processing huge datasets</li> </ul>	<ul style="list-style-type: none"> <li>▪ It is a lightweight, open-source, that makes users to train text classifiers and text representations</li> </ul>	<ul style="list-style-type: none"> <li>▪ High memory use since it embeds words using characters rather than words</li> </ul>
8	Context2Vec	Contextual word representation	<ul style="list-style-type: none"> <li>▪ A bidirectional LSTM-based unsupervised model for learning generic context embedding of broad sentential contexts</li> </ul>	<ul style="list-style-type: none"> <li>▪ It may take into account the context of the full statement</li> </ul>	<ul style="list-style-type: none"> <li>▪ more complicated</li> </ul>
9	Continuous	Non-Contextual	<ul style="list-style-type: none"> <li>▪ Tries to anticipate the</li> </ul>	<ul style="list-style-type: none"> <li>▪ several times faster to</li> </ul>	<ul style="list-style-type: none"> <li>▪ Resultant vectors will be</li> </ul>

	Bag of Words	word representation	<p>middle word based on the words around it</p> <ul style="list-style-type: none"> <li>▪ In other words, it attempts to fill the blanks with the word that best fits the context and surrounding words</li> <li>▪ Less expensive when using smaller datasets.</li> <li>▪ Shorter training period than with Skip-Gram</li> </ul>	train than SG with slightly better accuracy for frequent words	sparse since they will have a lot of null values and be of huge dimension.
10	Skip Gram	Non Contextual word representation	<ul style="list-style-type: none"> <li>▪ Is the inverse of CBOW and aims to predict the words in the surrounding context from a target word.</li> <li>▪ Performs better with bigger datasets but requires more training time.</li> </ul>	<ul style="list-style-type: none"> <li>▪ It is unsupervised learning hence can work on any raw text given.</li> <li>▪ It requires less memory comparing with other words to vector representations.</li> </ul>	▪ Need a huge number of training data

Example for Text Classification is given in Figure 5.



Figure 5: Example for Text Classification

Following Figure shows text Preprocessing methods

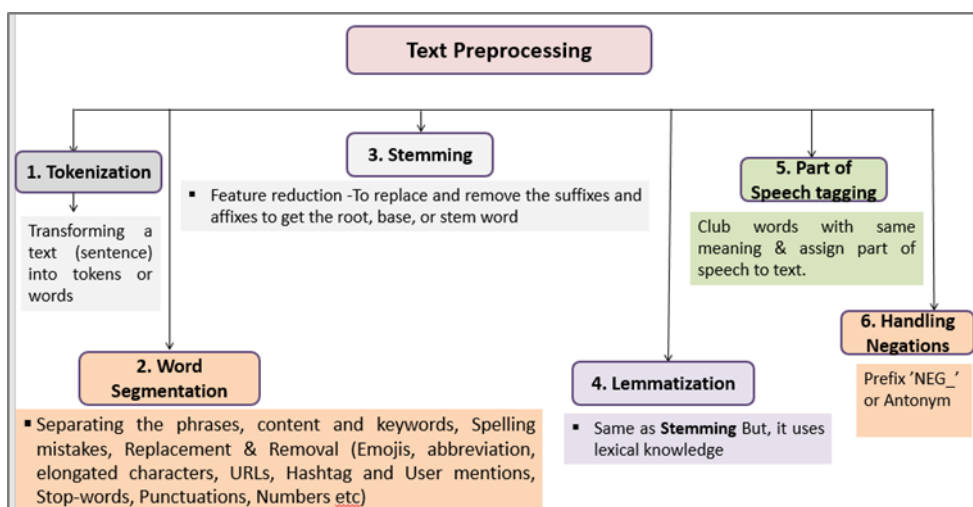


Figure 6: Example for Text Classification

C. Deep Learning and Machine-learning based classifiers

In following table, ML based and DL based classifiers details are given.

**Table 5: Machine Learning and Deep Learning based classifiers**

SN	Classifier	Reference	Purpose	Advantage	Disadvantage
<i>Deep Learning based classifiers</i>					
1	Naive Bayes Classifier	[12], [17]	<ul style="list-style-type: none"> <li>While training, each word will be used by the NB algorithm to determine its likelihood of being categorized.</li> </ul>	<ul style="list-style-type: none"> <li>Less computational effort, simplicity in understanding and application, and ease of training with less data</li> </ul>	<ul style="list-style-type: none"> <li>Relies heavily on the independence of the class characteristics and performs poorly if requirement is not satisfied.</li> <li>Zero frequency features, zero conditional probability-related issues in Zero likelihood</li> </ul>
2	Support vector machine (SVM)	[7], [12], [28]	<ul style="list-style-type: none"> <li>A non-probabilistic Binary Linear Classification technique, that works by projecting Training data in Multi-Dimensional space</li> <li>Hyper-plane is used to classifier into classes</li> <li>It learns from the labelled data set</li> </ul>	<ul style="list-style-type: none"> <li>Working in higher dimensions in nonlinear decision boundaries modeled</li> <li>Resistant to the problem of overfitting</li> </ul>	<ul style="list-style-type: none"> <li>More processing time is required for huge datasets.</li> <li>Kernel selection is challenging and performs poorly when classes overlap.</li> </ul>
3	Logistic Regression (LR) classifier	[9], [17]	<ul style="list-style-type: none"> <li>When a category value has to be anticipated, True or false, for instance</li> <li>A Statistical Model</li> </ul>	<ul style="list-style-type: none"> <li>Easy and simple to implement;</li> <li>Not expensive in terms of computation</li> <li>Does not require adjusting, and classes are dispersed evenly</li> </ul>	<ul style="list-style-type: none"> <li>Struggles when dealing with non-linear</li> <li>Not suitable for big datasets</li> <li>Does not make predictions based on dependent factors.</li> </ul>
4	Decision Tree (DT) classifier	[12], [17]	<ul style="list-style-type: none"> <li>It divides the dataset into several smaller subgroups.</li> <li>The root, decision, and leaf nodes of the DT classifier architecture represent the dataset, conduct computations, and classification executes respectively</li> <li>Classifier learns choices, while the training phase</li> <li>Classifier understands the choice that are used to separate categories</li> </ul>	<ul style="list-style-type: none"> <li>Easy to interpret and comprehend</li> <li>Need for pre-processing</li> <li>Training with practically no hyper-parameters</li> </ul>	<ul style="list-style-type: none"> <li>Potential for overfitting; lower average prediction accuracy</li> <li>Complex computation over several steps</li> </ul>
5	Random Forest (RF) Classifier	[12]	<ul style="list-style-type: none"> <li>The real output of this classifier is a value with more votes that are provided by several random forests.</li> <li>A strategy for ensemble learning is random forest.</li> </ul>	<ul style="list-style-type: none"> <li>Runs faster</li> <li>Stable</li> <li>Produces very high results</li> <li>Overfitting is less</li> </ul>	<ul style="list-style-type: none"> <li>Retrieval is not easy</li> <li>Computational resources are needed</li> <li>More time requires to predict</li> </ul>
<b>Deep Learning based classifiers</b>					
6	Recurrent Neural Network (RNN)	[13], [14],[27], [41]	<ul style="list-style-type: none"> <li>RNN - handles information with the last layers/nodes, in a way that their superiority in terms of corpus semantic analysis.</li> <li>Commons types of RNN – GRU, and LSTM are used for Text Classification</li> </ul>	<ul style="list-style-type: none"> <li>Ability to Handle Variable-Length Sequences.</li> <li>Memory Of Past Inputs</li> <li>Parameter Sharing</li> <li>Non-Linear Mapping</li> <li>Sequential Processing</li> <li>Flexibility</li> <li>Improved Accuracy</li> <li>Vanishing And Exploding Gradients.</li> </ul>	<ul style="list-style-type: none"> <li>RNNs get trained</li> <li>RNNs - cannot be fixed up.</li> <li>Training procedures get slow and complex too</li> <li>For longer sequence, it finds difficult</li> </ul>
7	Long Short-Term Memory (LSTM) Gated Recurrent Unit (GRU)	[10], [13], [27], [41]	<ul style="list-style-type: none"> <li>Ability to understand long-term relationships, particularly in situations involving sequence prediction.</li> <li>Gradient vanishing is combated by LSTM by discarding redundant data and information in the network.</li> </ul>	<ul style="list-style-type: none"> <li>Considerably good at managing Long-term dependencies</li> <li>This is, a result of the propensity for long-term memory</li> <li>LSTMs are far less vulnerable to vanishing gradient issue</li> </ul>	<ul style="list-style-type: none"> <li>They are less effective for online learning tasks, such as prediction or classification tasks when the input data is not a sequence, as they are more complex than regular RNNs and need more training data to learn successfully.</li> <li>LSTMs can take a long</li> </ul>

					time to train on huge datasets.
8	Convolutional Neural Networks (CNN)	[17, [18]	<ul style="list-style-type: none"> <li>▪ A Deep Learning network design that gains knowledge directly from data</li> <li>▪ CNNs are very helpful for recognizing objects, classifications, and categories in photos by looking for patterns in the images</li> <li>▪ Additionally, they may be quite successful in categorizing audio, time series, and signal data.</li> </ul>	<ul style="list-style-type: none"> <li>▪ For the purpose of finding crucial traits, CNNs do not require human oversight.</li> <li>▪ They are incredibly proficient in classifying and recognizing images.</li> <li>▪ Another key benefit of CNNs is weight sharing.</li> </ul>	<ul style="list-style-type: none"> <li>▪ The location and orientation of objects are not encoded, hence a lot of training data is needed for the CNN to work effectively.</li> <li>▪ CNNs frequently perform much slower because to procedures like the maximum pool advantage.</li> </ul>
9	Hierarchical Attention Networks (CNN)	[30], [31], [38]	<ul style="list-style-type: none"> <li>▪ uses stacked recurrent neural networks on word level, followed by an attention network</li> <li>▪ The goal is to extract such words that are important to the meaning of the entire sentence and aggregate these instructional words to form a vector of the sentence</li> </ul>	<ul style="list-style-type: none"> <li>▪ Ability to locate the data in a source that is most important for task completion</li> </ul>	<ul style="list-style-type: none"> <li>▪ Increased computation.</li> </ul>
10	Graph based Convolutional Neural Networks (H-CNN)	[1], [41] - [48]	<ul style="list-style-type: none"> <li>▪ LSTM outperforms with the help of GRU- Gated Recurrent Unit and a sort of RNN- Recurrent Neural Network</li> <li>▪ LSTM performs better when dealing with datasets including longer sequences, although GRU is faster and uses less memory</li> <li>▪ GRU's bag has two updated and reset gates</li> </ul>	<ul style="list-style-type: none"> <li>▪ GRU has fewer gates than LSTM, making it less complicated</li> <li>▪ GRU is preferable for small datasets, whereas LSTM is better for bigger datasets</li> </ul>	<ul style="list-style-type: none"> <li>▪ Fail to distinguish multi-sets with the same distinct elements but with different structure</li> </ul>

*D. Comparison of classifier functions*

Various classifier functions are compared in following Table 6

**Table 6: Comparison of classifiers functions**

SN	Classifier Functions	Purpose	Methods
1	Functions of a Numerical Classifier	Numerical Classifiers and Classifier Languages gives a thorough examination that mostly focuses on classifiers.	Logistic Regression, Support Vector Machine (SVM), Neural network, Linear Classifiers
2	Probabilistic (parametric) functions	For an given input, able to predict a Probability Distribution over collection of multiple classes	Gaussian discriminant analysis, hidden Markov models, probabilistic graphical models, and naive Bayes
3	(Instance-based) non-parametric functions	Classification and Regression group that predicts the category of an object based on how similar it is to its nearest neighbors in the training set.	Kernel regression, kernel density estimation, local regression, and k-nearest neighbors
4	Functional symbols	Trains Artificial Intelligence (AI) the same way human brain learns	CART - classification, regression, and decision trees.
5	(Ensemble) learning through Aggregation	A method for creating new classifier based on different base classifiers which outperform than other classifiers. The method employed, the	Bagging, boosting (Adaboost), random forest

		hyperparameters, the representation, or the training data may vary amongst these base classifiers.	
--	--	--	--

#### 4 Conclusion

Many readers now favor to read news using social media sources Here, the first is us to offer an overview of studies that have focused on classifying Fake News using classifiers based on Machine Learning and Deep Learning. Then, methods are discussed related to Text Representation, word embedding methods, and Classifiers which includes mostly Deep Learning based features.

#### References

- [1] Qing Liao , Heyan Chai, Hao Han, Xiang Zhang, Xuan Wang, Wen Xia, Ye Ding , “An Integrated Multi-Task Model for Fake News Detection” , IEEE Transactions on Knowledge and Data Engineering, DOI 10.1109/TKDE.2021.3054993, March '2021
- [2] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake News Detection on social media: A data mining perspective,” SIGKDD Explorations, vol. 19, no. 1, pp. 22–36, 2017, doi: 10.1145/3137597.3137600
- [3] S.Yang, K.Shu, S.Wang, R.Gu,F.Wu, and H.Liu, “ Unsupervised Fake News Detection on social media: A generative approach,” in Proceedings of 33rd AAAI Conference on Artificial Intelligence, 2019, Vol. 33 No. 01: AAAI-19, IAAI-19, EAAI-20, doi: 10.1609/aaai.v33i01.33015644
- [4] Monther Aldwairi, Ali Alwahedi, “Detecting Fake News in Social Media Networks”, in Proceedings of the 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2018), pp. 215–222, doi: 10.1016/j.procs.2018.10.171
- [5] K. Shu, L. Cui, S. Wang, D. Lee, and H. Liu, “defend: Explainable Fake News Detection,” in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019, Anchorage, AK, USA, August 4-8, 2019, pp. 395– 405, doi: 10.1145/3292500.3330935
- [6] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, “Automatic Detection of Fake News,” in Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018, pp. 3391–3401, doi: 10.48550/arXiv.1708.07104
- [7] N. J. Conroy, V. L. Rubin, and Y. Chen, “Automatic deception Detection: Methods for finding Fake News,” in Proceedings of the 78th ASIS & Tannual Meeting: Information Science with Impact: Research in and for the Community. American Society for Information Science, 2015, p. 82, doi: 10.1002/pr2.2015.145052010082
- [8] K. Shu, D. Mahudeswaran, and H. Liu, “FakeNewstracker: a tool for Fake News collection, Detection, and visualization,” Computational & Mathematical Organization Theory, vol. 25, no. 1, pp. 60–71, 2019, doi: 0.1007/s10588-018-09280-3
- [9] S. Ghosh and C. Shah, “Toward automatic Fake News Classification,” in 52nd Hawaii International Conference on System Sciences, HICSS 2019, Grand Wailea, Maui, Hawaii, USA, January 8-11, 2019, pp. 1–10, doi: :10.1002/pr2.2018.14505501125
- [10] N. Ruchansky, S. Seo, and Y. Liu, “CSI: A hybrid deep model for Fake News Detection,” in Proceedings of the ACM on Conference on Information and Knowledge Management, 2017, pp. 797–806, doi: 10.1145/3132847.3132877
- [11] H. Karimi, P. Roy, S. Saba-Sadiya, and J. Tang, “Multi-source multi- class Fake News detection,” in Proceedings of the 27th International Conference on Computational Linguistics, 2018, pp. 1546–1557.
- [12] Nabamita Deb, Vishesh Jha, Alok K Panjiyar, Roshan Gupta, “A Comparative Analysis of News Categorization Using Machine Learning Machine Learning Approaches”, INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 01, JANUARY 2020, ISSN 2277-8616, doi: 10.3390/info13120576

- [13] H. Z. Jin, J. Cao, H. Guo, Y. Zhang, and J. Luo, "Multimodal fusion with recurrent Neural Networks for rumor Detection on microblogs," in Proceedings of the ACM on Multimedia Conference, 2017, pp. 795–816, doi: 10.1145/3123266.3123454
- [14] Y. Liu and Y. B. Wu, "Early Detection of Fake News on social media through propagation path Classification with recurrent and convolutional networks," in Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, 2018, Vol. 32 No. 1, doi: 10.1609/aaai.v32i1.11268
- [15] J. Zhang, L. Cui, Y. Fu, and F. B. Gouza, "FAKEDETECTOR: Effective Fake News Detection with Deep Diffusive Neural Network," in Proceedings of the 2020 IEEE 36th International Conference on Data Engineering (ICDE), doi: 10.1109/ICDE48307.2020.00180
- [16] Gulshan Shrivastava, Prabhat Kumar, Rudra Pratap Ojha, "Defensive Modeling of Fake News Through Online Social Networks", IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, VOL. 7, NO. 5, PP:1-9, OCTOBER 2020, doi: 10.1109/TCSS.2020.3014135
- [17] Iftikhar Ahmad, Muhammad Yousaf, Suhail Yousaf, and Muhammad Ovais Ahmad, "Fake News Detection Using Machine Learning Ensemble Methods", in Hindawi, Wiley Volume 2020, Article ID 8885861, doi: 10.1155/2020/8885861
- [18] Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, "EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection", KDD '18: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, August, 2018, pp. 849–857, doi: 10.1145/3219819.3219903
- [19] W. Y. Wang, "liar, liar pants on fire": A new benchmark dataset for Fake News Detection," in Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Volume 2, 2017, pp. 422–426, doi: 10.18653/v1/P17-2067
- [20] D. DiFranzo and M. J. K. Gloria, "Filter bubbles and Fake News," ACM Crossroads, vol.23, no.3, pp.32–35, 2017
- [21] S. Feng, R. Banerjee, and Y. Choi, "Syntactic stylometry for deception detection," in The 50th Annual Meeting of the Association for Computational Linguistics, Volume 2, 2012, pp. 171–175
- [22] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," Science, vol. 359, no. 6380, pp. 1146–1151, 2018.
- [23] Y. Chen, N.J. Conroy, and V.L. Rubin, "Misleading online content: Recognizing click bait as false news," in Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection. ACM, 2015, pp. 15–19.
- [24] Sakshini Hanglool, Bhavna Arora, "Fake News Detection Tools And Methods – A Review"
- [25] Usman Naseem, Imran Razzak, Shah Khalid Khan, Mukesh Prasad, "A Comprehensive Survey on Word Representation Models: From Classical to State-Of-The-Art Word Representation Language Models", Woodstock '18, June 03–05, 2018, Woodstock, NY
- [26] Costel-Sergiu Atodiresei, Alexandru Tănăselea, Adrian Iftene, "Identifying Fake News and Fake Users on Twitter", 2018, International Conference on Knowledge Based and Intelligent Information and Engineering Systems, Belgrade, Serbia
- [27] Sherry Girgis, Eslam Amer, Mahmoud Gadallah, "Deep Learning Algorithms for Detecting Fake News in Online Text", 2018, 13th International Conference on Computer Engineering and Systems (ICCES)
- [28] Herley Shaori Al-Ash, Wahyu Catur Wibowo, "Fake News Identification Characteristics Using Named Entity Recognition and Phrase Detection", 2018, 10th International Conference on Information Technology and Electrical Engineering (ICITEE)
- [29] Marco L. Della Vedova, Eugenio Tacchini, Stefano Moret, Gabriele Ballarin, "Automatic Online Fake News Detection Combining Content and Social Signals", 2018, Proceeding of the 22nd Conference of Fruct Association.
- [30] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document Classification," in Conference of the North American Chapter of the Association for Computational Linguistics (NACCL), 2016, pp. 1480–1489, doi:10.18653/v1/N16-1174
- [31] S. Singhanian, N. Fernandez, and S. Rao, "3HAN: A deep Neural Network for Fake News Detection," in Neural Information Processing - 24th International Conference, ICONIP 2017, Guangzhou, China, November 14-18, 2017, Proceedings, Part II, 2017, pp. 572–581, doi: 10.1007/978-3-319-70096-0\_59

- [32] Ujun Jeong, Kaize Ding, Lu Chengy, Ruocheng Guoz, Kai Shux and Huan Liu, “Nothing Stands Alone: Relational Fake News Detection with Hypergraph Neural Networks,” in Proceedings of the IEEE Conference BigData 2022, Osaka, Japan, doi: 10.48550/arXiv.2212.12621
- [33] Federico Monti, Fabrizio Frasca, Davide Eynard, Damon Mannion, “Multi-Depth Graph Convolution Networks for Fake News Detection” , natural language Processing and Chinese Computing Conference paper pp 698-710], September 2019, doi: 10.1007/978-3-030-32233-5\_54
- [34] John Boaz Lee, Ryan A. Rossi, Sungchul Kim, Nesreen K. Ahmed, Eunye Koh, “Attention Models in Graphs: A Survey”, ACM Trans. Knowl. Discov. Data., Vol. 0, No. 1, August 2018, doi.org/0000001.0000001
- [35] Y. Ren and J. Zhang, “Fake News Detection on News-oriented heterogeneous information networks through hierarchical graph attention,” in International Joint Conference on Neural Networks (IJCNN). IEEE, July 2021, doi: 10.1109/IJCNN52387.2021.9534362
- [36] Song Bai, Feihu Zhang, Philip H.S. Torr, “Hypergraph convolution and hypergraph attention”, in Internal journal Pattern Recognition Volume 110, February 2021, doi: 10.1016/j.patcog.2020.107637
- [37] Yi-Ju Lu, Cheng-Te Li, “GCAN: Graph-aware Co-Attention Networks for Explainable Fake News Detection on Social Media” in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 505–514, doi: 10.18653/v1/2020.acl-main.48
- [38] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, “Hierarchical attention networks for document classification,” in Conference of the North American Chapter of the Association for Computational Linguistics (NACCL), 2016, pp 1480–1489, doi:10.18653/v1/N16-1174
- [39] Nicollas R. de Oliveira, Pedro S. Pisa, “Identifying Fake News on Social Networks Based on Natural Language Processing: Trends and Challenges”, MDPI Information 2021, 12, 38. <https://doi.org/10.3390/info12010038>
- [40] F. Yang, Y. Liu, X. Yu, and M. Yang, “Automatic detection of rumor on sina weibo,” in Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics. ACM, 2012, p. 13
- [41] Liang Yao “Graph Convolution Networks for Text Classification” in Proceedings of 33rd AAAI Conference on Artificial Intelligence (AAAI 2019); <https://doi.org/10.48550/arXiv.1809.0567>