A Review: Fake News Detection using Hierarchical Attention Network and Hypergraph Attention

Ms. Alpana A. Borse, Prof., Dr.G.K.Kharate

[1] *Asst. Prof., Information Technology, PCCOE, Pune;Research Scholar, MCOERC, Nasik, India*
[2]*Principal, MCOERC, Nasik, India*

alpana.borse@pccoepune.org, gkkharate@rediffmail.com

Received ## Mon.20##, Revised ## Mon. 20##, Accepted ## Mon. 20##, Published ## Mon. 20##

Abstract: This paper aims a survey of some of the detection of Fake 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 – Generative Adversarial Networks (GAN), Long Short Term Memory (LSTM), Recurrent Neural Network (RNN), 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, Hierarchical Attention Network, Hypergraph Attention

1. Introduction

Fake News has proliferated significantly in the area thanks to the expanding internet and rising use of social media. Both conventional and online techniques of information circulation are possible. This information's contents can be true or false. False content may be accidentally produced or purposely constructed to be perceived as true. Why is this Fake News so problematic? As every one of us is governed by two forces: As we know, during the lockdown in 2021, the news that the trains were going to return tomorrow had an influence on people in such a way that chaos results. As a result, people congregated on a train station. And managing such a dreadful situation in society becomes challenging. Therefore, a prediction mechanism for news authenticity is required, and it must be quick and efficient to prevent the Fake News. We understand recognizing News is a technically challenging subject that has most recently caught the publics' and scholars' attention. The objective of Fake News prediction identifying news is that contains misinformation. Fake News Detection is very crucial identifying Fake News predictions and minimizing its detrimental impacts. Early Fake News Detection techniques sometimes use extensive sets of manually created features according to user profiles, the content, and transmission routes of News, and then classifierstrain to differentiate between news that is true or false. Still, it is challenging to construct features that are all-inclusive because Fake News is typically produced across a variety of subjects, writing techniques, and social media channels. Here, we provide a survey of neural network methodologies based on machine and Deep Learning. Unfortunately, earlier research improved the Fake News Prediction model's performance, but it drastically decreased when the news's contents were brief. Topical news and contextual data, such as author biographies, have some effects on the Fake News prediction job. By establishing relationships between news subjects, author credibility distributions, and news veracity all at once, prediction and classification are accomplished.

2. STATE-OF-THE-ART SURVEY

A. Comparisons and contributions

Our paper offers a survey based on Deep Learning (DL) and Machine Learning (ML), Hierarchical Attention Networks (HAN), and Hypergraph for authenticity detection of News as True/False. Therefore, academics have been focusing on creating ways for categorizing news to get better detection of News as well as the creation of diverse methods relying on various DL approaches. Techniques for improving performance in applications involving news or text classification must be considered. The survey of news classification methods based on ML and DL algorithms that has been developed throughout time is the main topic of the research. In addition, we give the Research Gap Identified, the Motivation, and the Objectives of the Proposed Research after the survey. We highlight the various News Classification approaches that were built utilizing either a single or multiple DL methodology. We present the fundamentals of text classification, Word Embedding, and various Classifiers, and review the state-of-the arts.

Fake News prediction techniques have significantly improved, helping to lessen their detrimental effects on society by distributing real news on social media. Regardless of whether the news is factual or fraudulent, there are various techniques used to identify and categorize it. Models for prediction and classification are helpful for identifying and categorizing news, respectively. Following is a discussion of some prediction and classification techniques for news detection.

B. 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.

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

C. Paper Organization

This article is further organized into following sections.

We provide the Literature Survey for Fake News Authenticity Techniques in Section_II. In Section_III, we provide a methodology that includes word embedding techniques, text classification techniques, numerous classifier techniques, and classifier approaches as a survey. Finally, we wrap up the conclusions in Section_IV.

2. STATE-OF-THE-ART SURVEY

Fake News prediction techniques have significantly improved, helping to lessen their detrimental effects on society by distributing real news on social media. Regardless of whether the news is factual or fraudulent, there are various techniques used to identify and categorize it. Models for prediction and classification are helpful for identifying and categorizing news, respectively. Following is a discussion of some prediction and classification techniques for news detection. Likewise, here literature survey is discussed and given in Table I: Literature survey from reference [1] to [41].

Table I: Literature Survey for Fake News detection

SN	Title	Cited	Year	Author	DOI	Publication	Methods used	Remark and Scope
1	An Integrated Multitask model for Fake News Detection	[1]	January 2021	Qing Liao, et.al	10.1109/TKDE.2021.3054993	IEEE Transaction	▪ N-Graph ▪ LSTM ▪ Glove	▪ Detection and classification done, simultaneously ▪ Short news the performance improved ▪ It learns statement and relation representation
2	Fake News Detection on Social Media- A Data Mining Perspective	[2]	2017	K Shu, et. al.	10.1145/3137597.3137600	SuhangWang, SIGKDD Explorations,	▪ Natural Language Features	▪ Examines the veracity of news reports ▪ Features such as perceived cognitive authority, style content-based characteristics, and source authenticity
3	Unsupervised Fake News Detection on social media: A Generative Approach	[3]	2019	S.Yang, et.al.	10.1609/aaai.v33i01.330156 44	AAAI Conference on Artificial Intelligence	▪ To conclude news authentication, Gibbs sampling approach is used ▪ With unlabeled data, users' credibility is obtained ▪ Bayesian network model	▪ Check to see whether we can spot bogus news without supervision ▪ Determine news veracity, treating the news veracity and users' credibility as random latent variables
4	Detecting Fake News in Social Media Networks	[4]	2018	MontherAldwairi	10.1016/j.procs.2018.10.171	EUSPN Conference	▪ WEKA Classifiers ▪ Pythonscript ▪ InfoGainAtributeEval	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.

							▪ CorrelationAttributeEval	
5	defend: Explainable Fake News Detection	[5]	August 2019	K. Shu	10.1145/3292500.3330935	ACM SIGKDD International Conference	▪ lexical Features ▪ Syntax Features ▪ Semantic Features ▪ Linguistic Features	▪ Created a Sentence and comment-basedCo-Attention subnetwork to take advantage of news items and users comments. ▪ Used side information from other user interactions, likes, to find comments that have explanations. ▪ The credibility of people who leave explicative comments is utilized
6	Automatic Detection of Fake News	[6]	2018	V. P´erez-Rosas, et.al	10.48550/arXiv.1708.07104	27th ICCLC International Conference Linguistics COLING	▪ Used syntactic and semantic level of news	▪ Automatic fake News detection ▪ Accuracy is more if compared with manual system
7	Automatic Deception Detection - Methods for finding Fake News	[7]	2018	N. J. Conroy	10.1002/pra2.2015.145052010082	78th ASIS and Annual Meeting: Information Science	▪ Linguistic cue approacheas ▪ Network analysis approacheas ▪ Deep Syntax ▪ Semantic analysis	Operational guidelines for a workable News detection system Linguistic-basedand 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.	0.1007/s10588-018-09280-3	Computational & Mathematical Organization Theory	▪ Semi-supervised features ▪ Streaming is used	▪ Can use features like favorites, retweets, social network ▪ Can explore in real-time ▪ Automatically collect data for news and visualization
9	Toward automatic Fake News	[9]	2019	S. Ghosh	10.1002/pra2.2018.14505501125	HICSS	▪ TF-IDF ▪ NLP, SVM ▪ RF ▪ Logistic regression	▪ Accuracy 82.4% ▪ Vector space model
10	CSI- A Hybrid Deep Model for Fake News Detection	[10]	2017	N. Ruchansky	10.1145/3132847.3132877	ACM Conference-Information and Knowledge Management	▪ SVM ▪ LSTM ▪ GRU	▪ The first novel model ▪ Capture the three shared properties– Text, response and source whichdetects false information.
11	Multi Source , Multi Class Fake News Detection	[11]	2018	H. Karimi	--	27th International Conference on Computational Linguistics	▪ CNN ▪ LSTM	▪ Use many sources and differentiate between different levels of fakeness. ▪ Automation of Feature extraction ▪ Multiple-source fusion
12	A Comparative Analysis - News Categorization Using Machine Learning Approaches	[12]	JANUARY 2020	Nabamita Deb	10.3390/info13120576	International Journal Of Scientific and Technology Research	▪ Machine Learning based Survey on News Categorization	▪ 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%. ▪ The Random Forest is next, with a 94.1% accuracy rate. ▪ Neural Networks had an accuracy of 96.4%
13	Multi modal fusion with recurrent Neural	[13]	2017	H. Z. Jin	10.1145/3123266.3123454	Proceeding of the ACM on	▪ RNN ▪ Attention mechanism	▪ A unique Attention based Recurrent Neural Network ▪ Combines multi modal signals

	Networks for rumor Detection					Multimedia Conference	▪ LSTM	for efficient rumour detection. ▪ Experiments are done on Twitter and Weibo Multi-media data. ▪ In this, Joint Features are used in Image features
14	Early detection of fake News on Social Media through Propagation Path Classification with Recurrent and Convolutional Networks	[14]	2018	Y. Liu	10.1609/aaai.v32i1.11268	AAAI: Conference on Artificial Intelligence	▪ Geometric Deep Learning approach ▪ Convolution neural network	▪ Detecting of bogus news, based onDeep Learning method of propagation is employed. ▪
15	FAKEDETECTOR: Effective Fake News Detection with Deep Diffusive Neural Network	[15]	2020	J. Zhang	10.1109/ICDE48307.2020.00180	Proceedings of the IEEE 36th International Conference: Data Engineering (ICDE)	▪ Creator-Article Historical Records ▪ Analysis of Subject Credibility ▪ Gated Recurrent Unit is used	▪ Looking into the guiding ideas, approaches, and formulas for spotting Fake News ▪ Content, authors, and issues from online social networks, together with an assessment of how well they perform ▪ The issues posed by the unidentified qualities are addressed in this work.
16	Defensive Modeling of Fake News Through Online Social Networks	[16]	October 2020	Gulshan Shrivastava	10.1109/TCSS.2020.3014135	IEEE Transactions On Computational Social Systems	▪ Jacobian matrix, Matrix Eigenvalues, and Lyapunov function is used	▪ They proposed a mathematical model to explore regulating actions and dynamic spreading of message transmission in OSNs ▪ A model is design that can identify and remove false information from OSNs. ▪ alleviate some OSN users' anxiety about the pandemic
17	Fake News Detection Using Machine Learning Ensemble Methods	[17]	2020	Iftikhar Ahmad	10.1155/2020/8885861	Hindawi, Wiley	▪ Naive Bayes ▪ Decision Tree ▪ Adaboost ▪ SVM ▪ Linguistic Features: n-grams	▪ The explanation of Machine Learning Ensemble Methods ▪ Gradient descent (SGD), with SVM obtaining the maximum accuracy of 92%
18	EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection	[18]	August 2018	Yaqing Wang	10.1145/3219819.3219903	24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining	▪ Feature Extractor ▪ text, visual features	▪ End-to-end framework named EANN ▪ Fake News Detector ▪ Used Twitter dataset
19	Liar, liar pants on fire: A new benchmark dataset for Fake News detection	[19]	2017	W. Y. Wang	10.18653/v1/P17-2067	Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics	▪ Dataset	▪ Largest Fake News dataset ▪ Useful for short news
20	Filter bubbles and Fake News	[20]	2017	D. DiFranzo	10.1145/3055153	ACM Crossroads	▪ Survey	▪ Various surveys are conducted regarding techniques, including social network analysis and machine learning ▪ Researchers are still exploring and looking into them from many technical and sociological angles
21	Syntactic	[21]	2012	S. Feng	--	50th Annual	▪ Feature	▪ Syntactic stylometry is learnt for

	stylometry for deception detection				Meeting of the Association for Computational Linguistics	Encoding ▪ Words ▪ Shallow Syntax ▪ Deep syntax	deception detection ▪ Parse tree Context Features ▪ Enhances the detection performance with the use of Free Grammar
22	The spread of true and false news online	[22]	2018	S. Vosoughi	10.1126/science.aap95	Science, this issue p. 1146	▪ Complementary cumulative distribution functions (CCDFs) ▪ A number of unique Twitter users.	▪ Correlations between the spread of news is found ▪ the sentiments of responses for the news is found
23	Misleading online content: Recognizing click bait as false news	[23]	2015	Y.Chen,	10.1145/2823465.2823467	Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection.	▪ Hybrid approaches	▪ Its primary goal - to draw attention and persuade viewers; also a certain web page is obtained with the help of click on link ▪ It looks at potential tools for automatically spotting clickbait as a kind of deception. ▪ Textual and non-textual Techniques for identifying clickbaiting cuesare reviewed
24	Fake News Detection Tools and Methods – A Review	[24]	Nov 2021	Sakshini Hangloo1	10.48550/arXiv.2112.11185	International Journal of Advance and Innovative Research, Volume 8, Issue 2 (IX) ISSN 2394 - 7780	▪ Review	▪ Machine learning tools and its related methods survey is given in this
25	A Comprehensive Survey on Word Representation Models	[25]	2018	Usman Naseem	10.1145/3434237	ACM Transactions on Asian and Low-Resource Language Information Processing	▪ Text Mining ▪ Natural Language Processing ▪ Word representation ▪ Language Models	▪ Various word representation models are explained and compared here with analysis
26	Identifying Fake News and Fake Users on Twitter	[26]	2018	Costel-Sergiu Atodiresei	10.1016/j.procs.2018.07.279	International Conference on Knowledge Based and Intelligent Information and Engineering	▪ Text based features ▪ N-gram ▪ Propagation based approach	▪ System created with the intention of detecting bogus Twitter users and news sources
27	Deep Learning Algorithms for Detecting Fake News in Online Text	[27]	2018	Sherry Girgis	10.1109/ICCES.2018.8639198	13th International Conference on Computer Engineering and Systems (ICCES)	▪ RNN (vanilla, GRU) ▪ LSTMs ▪ GRU	▪ A classifier determines if a ews nis true/false based on its content ▪ Results indicate that GRU is the best, followed by LSTM, and vanilla
28	Fake News Identification Characteristics	[28]	2018	HerleyShaori Al-Ash	10.1109/ICITEED.2018.853	10th International	▪ Frequency term, ▪ Inverse	▪ Uses the Indonesian language to form vectorwhich handles features of fake News

	Using Named Entity Recognition and Phrase Detection				4898	Conference on Information Technology and Electrical Engineering (ICITEE)	document frequency ▪ Support vector machine ▪ NER	▪ This uses vector space paradigm to differentiateLegitimate and Fake News ▪ With the word frequency, Vector performance representation seems auspicious.
29	Automatic Online Fake News Detection Using Social Signals and Content	[29]	2018	Marco L. Della	10.23919/FRUCT.2018.8468301	Procedding of the 22nd Conference of FructAssociation.	▪ News content features ▪ Social content features ▪ ML methods	▪ Combines social context features with news articles ▪ A Facebook Messenger chatbot uses the diffusion pattern as an implementation strategy
30	Hierarchical Attention Network for Document Classification	[30]	2016	Z. Yang	10.18653/v1/N16-1174	Conference of the North American Chapter of the Association for Computational Linguistics (NACCL)	▪ GRU ▪ Word encoder ▪ Sentence encoder ▪ Attention Mmechanism	▪ 2 levels ▪ A structure that is equivalent to the hierarchy of documents ▪ Second-level: Attention modelis designed at Word and Sentence level
31	3HAN: A Deep Neural Network for Fake News Detection	[31]	2017	S. Singhania	10.1007/978-3-319-70096-0_59	Neural Information Processing - 24th International Conference, ICONIP	▪ GRU ▪ Word, Sentence, and Headline encoder ▪ Attention Mechanism	▪ Article structure having 3 levelsword-encoder, sentence encoder, and headline encoder
32	Nothing stands Alone: Relational fake News Detection with Hypergraph Neural Networks	[32]	2022	UjunJeong	10.48550/arXiv.2212.12621	IEEE Conference BigData	▪ GCN ▪ Hypergraph Attention	▪ Use a hypergraph to depict how news is interacted with collectively
33	Multi Depth Graph Convolution N/k for News detection	[33]	2019	Federico Monti	10.1007/978-3-030-32233-5_54	Natural Language Processing and Chinese Computing Conference paper	▪ Graph ▪ CNN	▪ Employed propagation-based techniques ▪ Utilized information including user profiles, content, activities, Social Network, and news dissemination ▪
3 4	Attention Models in a Graph: A Survey	[34]	2018	John Boaz Lee	doi.org/0000001.0000001	ACM Trans. Knowl. Discov. Data.	▪ Survey	▪ Literature Review on Graph Attention models
35	Fake News Detection on News Oriented heterogeneous information networks through Hierarchical Graph Attention	[35]	July 2021	Y. Ren and J. Zhang	10.1109/IJCNN52387.2021.9534362	International Joint Conference on Neural Networks (IJCNN). IEEE	▪ HGAT ▪ Attention mechanism	▪ A novel Framework ▪ Learn node representation ▪ Hierarchical Attention mechanism is used
36	Hypergraph convolution and hypergraph attention	[36]	2020	Song Bai	10.1016/j.patcog.2020.107637	Internal journal Pattern Recognition	▪ Hypergraph convolution ▪ Hypergraph attention.	▪ Utilises a hierarchical attention technique to carry out node representation learning in HIN ▪ New framework for false news identification.
37	GCAN: Graph aware Co-	[37]	2020	Yi-Ju Lu	10.18653/v1/2020.a	Proceedings of the 58th	▪ Tweets are used	▪ GCAN consistently outperforms cutting-edge techniques by 16%

	Attention Networks for Explainable fake News Detection on Social Media				cl-main.48	Annual Meeting of the Association for Computational Linguistics	▪ Graph CNN ▪ GRU	in accuracy. ▪ Co-attention between sources ▪ Co-attention for Source-Propagation
38	Hierarchical Attention Network for Document Classification	[38]	2016	Z. Yang	10.18653/v1/N16-1174	North American Chapter of the Association for Computational Linguistics (NACCL)	▪ HAN ▪ Attention visualization	▪ 2 distinguishing characteristics (i) Its organizational structure is hierarchical. (ii) Word-and Sentence levels are used as Attention Strategies Attention visualization
39	Identifying Fake News on Social Networks based on natural Language Processing: Trends and Challenges	[39]	2021	Nicollas R	10.3390/info12010038	MDPI Information	▪ Reduction, ▪ Machine learning ▪ NLP ▪ Survey	▪ Natural language, vectorization, and dimensionality survey preprocessing techniques for data ▪ Discuss research objectives and prospects while contextualizing the detection of Fake News ▪ Information retrieval quality evaluation
40	Automatic detection of rumor on Sina-Weibo	[40]	2015	F. Yang	10.1007/978-3-319-25207-0_10	NaturalLanguage Processing and Chinese Computing. NLPCC	▪ Features ▪ Client, Content, Propagation, account andLocation based	▪ First study on Sina Weibo rumour analysis and detection Gather a large collection of microblogs ▪ Used the Twitter Monitor-defined style of keyword-based query
41	Graph Convolution Networks for Text Classification		2019	Liang Yao	10.48550/arXiv.1809.05679	33rd AAAI Conference on Artificial Intelligence	▪ Deep neural network models ▪ CNN, RNN LSTM	▪ 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 ▪ They also introduced the attention mechanism as a crucial component of the models used for text categorization

3. METHODOLOGY

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 synopsisfor these multiple meanings, are listed asSatires and parodies, Rumors, Conspiracy theories, Spams, Scams and hoaxes, Clickbaits, Misinformation, disinformation, propaganda. Table II gives information for various Methods used for Word Representation Models

Table II: Comparison between various Methods for Word Representation Models

SN	Methods	Based on	Purpose	Advantage	Disadvantage
1	One Hot Encoding	▪ Count based ▪ Categorical word representation	The conversion of categorical information into a format that may be fed into machine learning algorithms to improve prediction accuracy	▪ Enables us to apply machine-learning algorithms to data with categorical columns	▪ The problem is that it increases dimensionality so training becomes slower and more complex
2	BoW	▪ Count based ▪ Categorical word	Used to determine the word count of texts and documents	▪ Its simplicity and usability	▪ If the new phrases use new terms, then our vocabulary

		representation	Word order within a document is not considered	▪ It can be used to develop a rough draught model before moving on to more complex word embeddings	would grow, which would also lengthen the vectors. ▪ A sparse matrix would be produced as a result of the vectors' high number of zeros, which is something we want to avoid
3	TF	Frequency based Weighted word representation	In a collection of documents, it evaluates the weight of a phrase	▪ Simple and easy to use it is	▪ Does not consider the significance of phrases that are used sparingly throughout a collection of documents' individual documents
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	▪ It is an easy starting point for similarity computations, is inexpensive to compute, and is straightforward to calculate.	▪ Assign low values to comparatively significant terms, and be excessively sensitive on the extensive margin and excessively resistant on the intensive margin.
5	Word2Vec	▪ Prediction-based method	▪ Representing word embeddings ▪ A fixed vector size	▪ The embedding vector is tiny and versatile ▪ Able to capture relationships between words, including their syntactic and semantic relationships ▪ Since it is unsupervised, less human effort is required to tag the data	▪ 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.
6	GloVe	▪ Prediction-based method ▪ count-based technique ▪	▪ 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	▪ It typically outperforms word2vec in analogy tasks ▪ GloVe requires less training time because it is easier to parallelize than Word2Vec.	▪ Compared to word2vec, GloVe uses a co-occurrence matrix and global information, which increases memory cost. ▪ Because words and vectors are identical, it, too, like word2vec, does not address the problem of polysemous words
7	FastTex	Text based Non Contextual word representation	▪ Achieving scalable solutions for the challenges of text categorization and representation while reliably and quickly processing huge datasets	▪ It is a lightweight, open-source, that makes users to train text classifiers and text representations	▪ High memory use since it embeds words using characters rather than words
8	Context2Vec	Contextual word representation	▪ A bidirectional LSTM-based unsupervised model for learning generic context embedding of broad sentential contexts	▪ It may take into account the context of the full statement	▪ more complicated
9	Continuous Bag of Words	Non-Contextual word representation	▪ Tries to anticipate the middle word based on the words around it ▪ In other words, it attempts to fill the blanks with the word that best fits the context and surrounding words ▪ Less expensive when using smaller datasets. ▪ Shorter training period than with Skip-Gram	▪ several times faster to train than SG with slightly better accuracy for frequent words	▪ Resultant vectors will be sparse since they will have a lot of null values and be of huge dimension.
10	Skip Gram	Non Contextual word representation	▪ Is the inverse of CBOW and aims to predict the words in the surrounding context from a target word. ▪ Performs better with bigger datasets but requires more training time.	▪ Unsupervised Learning ▪ Requires less memory to vector representations	▪ Need a huge number of training data

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

Table III: Machine Learning and Deep Learning based classifiers

SN	Classifier	Reference	Purpose	Advantage	Disadvantage
Deep Learning based classifiers					
1	Naïve Bayes Classifier	[12], [17]	▪ While training, each word will be used by the NB algorithm to determine its likelihood of being categorized.	▪ Less computational effort, simplicity in understanding and application, and ease of training with less data	▪ Depends heavily on the independence of the class characteristics and performs poorly if the requirement is not satisfied. ▪ For Zero frequency features, zero conditional probability-related issues result in Zero likelihood
2	Support vector machine (SVM)	[7], [12], [28]	▪ Non Probabilistic based Binary Linear Classification ▪ Works by projecting Training data in Multi Dimensional space ▪ Hyper-plane is used to classifier into classes ▪ It learns from the labelled data set	▪ Working in higher dimensions ▪ Can nonlinear decision boundaries be modeled ▪ Resistant to the problem of overfitting	▪ More processing time is required for huge datasets. ▪ Kernel selection is challenging and performs poorly when classes overlap.
3	Logistic Regression (LR) classifier	[9], [17]	▪ When a category value has to be anticipated, True or false, for instance ▪ A Statistical Model	▪ Easy and simple to implement; ▪ Less expensive in terms of calculation ▪ Does not require adjusting, and features are dispersed evenly	▪ Fails when dealing with non-linear issues ▪ Need for big datasets ▪ Make predictions based on independent factors.
4	Decision Tree (DT) classifier	[12], [17]	▪ It divides the dataset into several smaller subgroups. ▪ The root, decision, and leaf nodes of the DT classifier architecture represent the dataset, conduct computations, and classification executes respectively ▪ Classifier learns choices, while the training phase ▪ Classifier understands the choice that are used to separate categories	▪ Easy to interpret and comprehend ▪ Less need for pre-processing ▪ Fast tuning with practically no hyper-parameters	▪ High potential for overfitting; lower than average prediction accuracy ▪ Complex computation over several classes
5	Random Forest (RF) Classifier	[12]	▪ The real output of this classifier is a value with more votes that are provided by several random forests. ▪ A strategy for ensemble learning is random forest.	▪ Trains faster ▪ Flexible ▪ Produces very high results ▪ Pre-processing is less	▪ Interpretation is not easy ▪ More Computational resources are required ▪ More time requires to predict
▪ Deep Learning based classifiers					
6	Recurrent Neural Network (RNN)	[13], [14],[27], [41]	▪ RNN –handles information with the last layers/nodes, in a way that their superiority in terms of corpus semantic analysis. ▪ Commons types of RNN – GRU, and LSTM are used for Text Classification	▪ Ability to Handle Variable-Length Sequences. ▪ Memory Of Past Inputs ▪ Parameter Sharing ▪ Non-Linear Mapping ▪ Sequential Processing ▪ Flexibility ▪ Improved Accuracy ▪ Vanishing And Exploding Gradients.	▪ RNNs get trained ▪ RNNs - cannot be fixed up. ▪ Training procedures get slow and complex too ▪ For longer sequence, it finds difficult
7	Long Short-Term Memory (LSTM) Gated	[10], [13], [27], [41]	▪ Ability to understand long-term relationships, particularly in situations involving sequence prediction.	▪ Considerably good at managing Long-term dependencies ▪ This is, a result of the propensity for long-term	▪ Online learning tasks are less effective ▪ Prediction or classification tasks need more training as input data is in a non-

	Recurrent Unit (GRU)		▪ Gradient vanishing is combated by LSTM by discarding redundant data and information in the network.	memory ▪ LSTMs are far less vulnerable tovanishing gradient issue	sequence ▪ More complex ▪ LSTMs can take a long time to train on huge datasets.
8	Convolutional Neural Networks (CNN)	[17, [18]	▪ A Deep Learning network design that gains knowledge directly from data ▪ CNNs are very helpful for recognizing objects, classifications, and categories in photos by looking for patterns in the images ▪ Additionally, they may be quite successful in categorizing audio, time series, and signal data.	▪ For the purpose of finding crucial traits, CNNs do not require human oversight. ▪ They are incredibly proficient in classifying and recognizing images. ▪ Another key benefit of CNNs is weight sharing.	▪ The location and orientation of objects are not encoded, hence a lot of training data is needed for the CNN to work effectively. ▪ CNNs frequently perform much slower because to procedures like the maximum pool advantage.
9	Hierarchical Attention Networks (CNN)	[30], [31], [38]	▪ Uses stack, RNN ▪ It works on word level ▪ Use of attention network ▪ Important words get extracted that are to the meaning of the entire sentence ▪ Summative instructional words to form a sentence vector	▪ Ability to locate the data in a source that is most important for task completion	▪ Increased computation.
10	Graph based Convolutional Neural Networks (H-CNN)	[1], [41] - [48]	▪ LSTM outperforms with the help of GRU and a sort of RNN ▪ LSTM performs better when dealing with datasets including longer sequences, although GRU is faster and uses less memory ▪ GRU's bag has two updated and reset gates	▪ GRU has fewer gates than LSTM, making it less complicated ▪ GRU is preferable for small datasets, whereas LSTM is better for bigger datasets	▪ Fail to distinguish multi-sets with the same distinct elements but with different structure

4. CONCLUSION

Many readers nowfavorto read newsusing social media sources Here, the first is usto 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.

ACKNOWLEDGMENTS

I would want to say thanks to Prof. Dr. G. K. Kharate, who has been my mentor at MCERC in Nasik, for their unwavering leadership. Additionally, we are appreciative to the reviewers whose insightful criticism considerably raised the caliber of this article.

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

[4] 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

[5] MontherAldwairi, 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

[6] 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, 2019, pp. 395– 405, doi: 10.1145/3292500.3330935

[7] V. P´erez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, "Automatic Detection of Fake News," in Proceedings of the 27th InternationalConferenceonComputationalLinguistics,COLING2018, SantaFe,NewMexico,USA,August20-26,2018,2018,pp.3391–3401, doi: 10.48550/arXiv.1708.07104

[8] 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/pra2.2015.145052010082

[9] 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

[10] S. Ghosh and C. Shah, "Toward automatic Fake NewsClassification," in 52nd Hawaii International Conference on System Sciences, HICSS 2019, Grand Wailea, Maui, Hawaii, USA, January 8-11, 2019, 2019, pp. 1–10, doi: :10.1002/pra2.2018.14505501125

[11] 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

[12] 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.

[13] 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

[14] 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

[15] 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

[16] 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

[17] 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

[18] Iftikhar Ahmad, Muhammad Yousaf, Suhail Yousaf , and Muhammad Ovais Ahmad, “Fake News Detection Using Machine Learning Ensemble Methods”, in in Hindawi, Wiley Volume 2020, Article ID 8885861,doi: 10.1155/2020/8885861

[19] Yaqing Wang, Fenglong Ma, ZhiweiJin, 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

[20] 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

[21] D. DiFranzo and M. J. K. Gloria, “Filter bubbles and Fake News,” ACMCrossroads,vol.23,no.3,pp.32–35,2017

[22] 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

[23] S. Vosoughi, D. Roy, and S. Aral, “The spread of true and false news online,” Science, vol. 359, no. 6380, pp. 1146–1151, 2018.

[24] 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.

[25] Sakshini Hangloo1, Bhavna Arora2, “Fake News Detection Tools And Methods – A Review”

[26] 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

[27] Costel-Sergiu Atodiresei, AlexandruTănăselea, Adrian Iftene , “Identifying Fake News and Fake Users on Twitter”, 2018, International Conference on Knowledge Based and Intelligent Information and Engineering Systems,Belgrade, Serbia

[28] 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)

[29] HerleyShaori 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)

[30] Marco L. Della Vedova, Eugenio Tacchini, Stefano Moret, Gabriele Ballarin, “Automatic Online Fake News Detection Combining Content and Social Signals”, 2018, Procedding of the 22nd Conference of FructAssocition.

[31] 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

[32] S. Singhania, 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

[33] UjunJeong, Kaize Ding, Lu Chengy, RuochengGuoz, 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

[34] 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

[35] John Boaz Lee, Ryan A. Rossi, Sungchul Kim, Nesreen K. Ahmed, Eunyee Koh, “Attention Models in Graphs: A Survey”, ACM Trans. Knowl. Discov. Data., Vol. 0, No. 1, August 2018, doi.org/0000001.0000001

[36] 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

[37] 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

[38] 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

[39] 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

[40] 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

[41] F. Yang, Y. Liu, X. Yu, and M. Yang, “Automatic detection of

rumor on sinaweibo," in Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics. ACM, 2012, p. 13

[42] 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

[43] Y. Ren,"Fake News detection on news-oriented heterogeneousinformationnetworks through hierarchical graph attention", International Joint Conference on Neural Networks (IJCNN). July 2021; doi: 10.1109/IJCNN52387.2021.9534362