Fake news detection using deep learning

Submitted By

Vishakha K Ralegankar

20MCED18

DEPARTMENT OF COMPUTER ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2022

Fake news detection using deep learning

Major Project

Submitted in partial fulfillment of the requirements

for the degree of

Master of Technology in Computer Science and Engineering

Submitted By

Vishakha K Ralegankar

(20MCED18)

Guided By

Dr Ankit Thakkar

NIRMA UNIVERSITY

INSTITUTE OF TECHNOLOGY

NAAC ACCREDITED 'A+' GRADE

DEPARTMENT OF COMPUTER SCIENCE

AHMEDABAD-382481

May 2022

Certificate

This is to certify that the major project entitled **"Fake news detection using deep learning"** submitted by **Vishakha K Ralegankar (20MCED18)**, towards the partial fulfillment of the requirements for the award of degree of Master of Technology in Computer Science and Engineering of Nirma University, Ahmedabad, is the record of work carried out by him under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination. The results embodied in this Major Project Part-I, to the best of my knowledge, haven't been submitted to any other university or institution for award of any degree or diploma.

Dr Ankit Thakkar
Internal Guide & Associate Professor
CSE Department
Institute of Technology
Nirma University, Ahmedabad

Dr. Gaurang Raval
Professor & PG Coordinator (M.Tech - DS)
CSE Department
Institute of Technology
Nirma University, Ahmedabad

Dr Madhuri Bhavsar
Professor & Head
CSE Department
Institute of Technology
Nirma University, Ahmedabad

Dr R.N. Patel
Director
Institute of Technology
Nirma University, Ahmedabad

Statement of Originality

I, **Vishakha K Ralegankar**, **20MCED18**, give undertaking that the Major Project entitled **"Fake news detection using deep learning"** submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in **Data science** of Institute of Technology, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made.It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Signature of Student
Date:
Place:

Endorsed by
Dr Ankit Thakkar
(Signature of Guide)

Acknowledgements

It gives me immense pleasure in expressing thanks and profound gratitude to **Dr Ankit Thakkar**, Associate Professor, Computer Science and Engineering Department, Institute of Technology, Nirma University, Ahmedabad for his valuable guidance and continual encouragement throughout this work. The appreciation and continual support he has imparted has been a great motivation to me in reaching a higher goal. His guidance has triggered and nourished my intellectual maturity that I will benefit from, for a long time to come.

It gives me an immense pleasure to thank **Dr Madhuri Bhavsar**, Hon'ble Head of Computer Science And Engineering Department, Institute of Technology, Nirma University, Ahmedabad for her kind support and providing basic infrastructure and healthy research environment.

A special thank you is expressed wholeheartedly to **Dr R.N. Patel**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for the unmentionable motivation she has extended throughout course of this work.

I would also thank the Institution, all faculty members of Computer Science and Engineering Department, Nirma University, Ahmedabad for their special attention and suggestions towards the project work.

Vishakha K Ralegankar
20MCED18

Abstract

The unexpected surge in social media has given rise to distribution of unlimited content over the internet. The social media platforms behaves like a broadcasting medium these days. Increase in social media users over past decade has given these platforms a liberty to own and entertain their audience. Users can access daily news through these platforms. Various channels (news related) have their official accounts which usually post news online. Publicising and forwarding the content online is very easy and free. This is one of the reasons why the fake news is circulated on the web. Platforms like Facebook, Instagram, and twitter are used by millions of people every day and therefore, data is generated with a great velocity. Huge amount of data out of there exist that is difficult to process and classify into false data. Some users intentionally or unintentionally communicate these kind of stuff online. Fake news or fake information can be of any form i.e, post, image, video, audio, etc. Proliferation of these kind of data amongst the people may mislead them to believe something that is not true. The widespread usage of social media plays a vital role in setting up the mindsets of the people and their actions in the real world. The free will provided by these platforms(social media), allows people to upload it easily and hence, it is circulated among a huge number of masses. The credibility of any content being uploaded on social media platforms is still a quest and the major reason for fake news spread. This is the very motivation of this study. This research study is a systematic review of the recent work performed in detecting the fake news using deep learning. Herein, we perform a task for improving the performance of deep learning model and also we implement the graph neural network for the identification of the false news.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

The fake news refers to a false information set forth as a piece of information [4]. The purpose behind it can either be right-minded or wrong-minded. Nowadays, online social media platforms (mostly Facebook, Twitter, Snapchat, YouTube, etc.) have become the chief source of information for masses all over the world [5]. Even if some untruthful opinions may be put out on purpose, these online social media platforms are excellent venues for people to communicate their emotions, tales, worries, and provide a better means of getting immediate responses and feedback on various worldwide topics [6]. In economics, politics, and social development, social media play a critical role. The proliferation of erroneous information or news on the internet and social media could have an impact on financial markets, response times in critical situations, and terrorist strikes, among other things [7]. Fake news is disseminated primarily for governmental (politics) or commercial advantage. They volunteer these well-fabricated news pieces and also hire communicating bots or paid fraudsters to roll out the news more quickly [8]. There are many instances where fake news made people believe into falsely propagated information and act accordingly. For example, as the COVID-19 epidemic progressed, social media platforms grew in importance as a means of socialising as well as finding and sharing disease-related information. As a result, an explosion of unregulated information and the propagation of disinformation occurred. In Italy exclusively, an average of 46000 news tweets on Twitter were false and related to incorrect information regarding the issue every day in March 2020 [9]. Individuals nowadays utilize the social media and diversification of other online news streams as their primary sources of information [10]. The social media platforms are quite easy to utilize in terms of spreading any type of information. People have this flexibility to upload content on social media without ensuring the credibility of the content they upload. The fake news spreading online is mostly in the form of post that can be a written content, image(meme), audio, or video. The sudden internet upsurge is also responsible for propagation of these fake content among the masses. Conventional media such as printing news articles or magazines is steadily diminishing, and everyone who has a social media profile or account has the ability to be a journalist or a news writer [8]. Hence, this issue of fake news is increasing everyday with increased use of internet [11]. Therefore, the issues we face as researchers is figuring out how to create a tool that can assist users of any form of content (for example, a news story) in determining if what they are seeing is false or real. In the social media ecosystem, fake news can take many various structure and shapes, making it much more arduous to discover and contrast them, both manually and automatically. Therefore it is vital to survey and review the state-of-the-art strategies to create a learning ground before devising new solutions.

1.1 Motivation

There are various aspects where fake news can become a bane to those who trust their sources, they read and make up their mind sets or act accordingly. The motivation of this review article comes from the importance of detecting fake news and preventing it. Some of the important aspects of detecting the false news is listed below.

- Individuals need to be well-informed about current events and news, and their social and political actions should not be influenced by outside forces.
- Detecting misleading news is a first step in identifying economic incentives for spreaders who conduct their "business" on social media.
- By proper detection of fake news, people would continue to trust the internet and social media.
- The negativity induced by any fake news will be nullified before it up-rises among the people.
- The voluminous data generated from the social media every single day becomes a tedious content verifying task (detection of fake and real news) if performed manually. Hence, the automating the process in a correct way is very essential.

1.2 Structure of this dissertation

The dissertation is divided into five chapters in all. The chapter 1 consist of the introduction where we begin this research study by establishing the concept of fake news and its repercussions on the mindsets of the people or a group of community which is targeted as a part of propaganda. We then perform a detailed review on the current state of the art in chapter 2. We discuss the models proposed by distinct studies with the results that are achieved by the same. We also review different datasets which are available publicly for analysis in chapter 2. The chapter 3 discusses the uniquely designed architecture that is utilised to develop a small proof of concept using a highly imbalanced dataset for classification of false and real news claims. We discuss the performance of the model designed and achieved good results. In chapter 4, we implement a novel technique graph neural network based models for fake news detection. In chapter 5, we examine the challenges faced while detecting the fake news and the possible research directions that can be considered in future to enhance this area of research. Finally in chapter 6, we conclude this dissertation with the brief review on the performed experiment along with the scope of future work that can be performed for more contributions toward this problem statement.

Chapter 2

Literature Survey

This chapter reviews the recent work and contributions done hitherto for identification of the fake news. The researchers have proposed different solutions by proposing novel work. This section also reviews the datasets created so far for developing these models.

2.1 Review on publicly available datasets

The quality and availability of datasets is a major impediment for automating fake news identification. Herein, we survey the publicly available datasets on fake news that are utilized by various researchers. The datasets available publicly are mostly made by collecting the data from social media APIs. We elucidate the ways and nature of the datasets used in distinct studies. Furthermore, the datasets are analyzed using a data requirements that the datasets must meet in order to be useful for building and testing fake news detection techniques. Components of the review datasets used in the detection of misinformation are as follows and summarized in table 2.1:

- Domain of news: Fake news pieces targeting specific news topics, such as healthcare, academia, travel, entertainment, business, military, technology, and governmental elections, may be found in the collection.

- Language in which the dataset is collected: This refers to the language of the false information in the dataset, which might be written in a variety of languages depending on the sources employed to gather the information.

- Dataset dimensions: The count of news items in the dataset is typically used to determine its size. It can also be expressed in terms of kilobytes/megabytes.

Table 2.1: Review on datasets

Reference	Dataset name	Domain	Language	Size	Platform used for data collection
[12]	PHEME dataset	Political and social	German and English	4842 tweets	Twitter
[13]	LIAR dataset	Political	English	12836 brief lines	Twitter and Facebook
[14]	Fact checking dataset	Social and political	English	221 brief statements	Traditional media
[15]	FakeOrReal News dataset	Social and political	English	33063 articles	Traditional media
[16]	Zheng et al.'s dataset	Social	Chinese	14922 headlines	Traditional and social media
[17]	Yelp dataset	Technological news	English	18912 reviews	Traditional media
[18]	Fake News Net dataset	Political and social	English	422 news articles	Traditional and twitter data
[19]	Spanish fake news corpus	Sciences, Game, Economics, Academia, Entertainment, Government, Healthcare, Security, and Social	Spanish	971 news articles	Traditional media
[20]	Tam et al.'s dataset	Scams, technology, political, and science	English	4 million tweets and 1022 rumours	Twitter
[21]	ISOT dataset	Political and global	English	25200 articles	News website data

2.2 Literature review on fake news detection models

After the data has been collected, labeled, and features extracted, the classification method must be decided. This section will give an overview of various methodologies for detecting fake news using different deep learning algorithms. [22] performed experiments based on CNN and amalgamation of CNN with other ensemble models. In all, they performed seven experiments among which the best accuracy achieved was 88.78% by model 7. There are two datasets used in this study, one was collected from twitter and the other was PolitiFact dataset which is available publicly. Another study based on CNN architecture [23] also used two datasets viz fake_or_real and a kaggle dataset. They used two different variants of CNN i.e., LIWC-CNN(Linguistic Inquiry and Word Count-CNN) and N-gram CNN. The CNN based models are summarized in table 2.2.

[24] created its own dataset from facebook posts and employed LSTM for the detction of fake news. The model used the user profile and content as features for the detection. This model achieved the accuracy of 99.4%. Another article [22] used LSTM on two datasets (twitter scraped posts and PolitiFact) along with some hybrid models and the LSTM model attained 80.62% of accuracy. [23] implemented two types of LSTM (depth LSTM and embedding LSTM) out of which the embedding LSTM performed the best (94% accuracy) on a kaggle dataset.

Another study [22] executed the amalgamation of two deep learning models such as CNN+LSTM ensembled model, CNN+BiLSTM ensembled model with attention mechanism, and Bi-LSTM+LSTM ensembled model. They best model (CNN+LSTM ensembled learning model) pulled off the accuracy of 88.78%. [25] make use of a dataset from kaggle competition. Four distinct experiments are performed here viz hybrid model (LSTM+CNN) with and without pre-processing, hybrid model with PCA(principle component analysis), and hybrid model with chi-square. The finest experiment (CNN+LSTM+PCA) gained the accuracy of 97.8%. [26] presented a hybrid model(LSTM+CNN) for classifying the fake news on a kaggle dataset. This deep learning paradigm attained 97.21% of precision, 91.89% of recall, and 97.44% of specificity. It was also mentioned in this study that this proposed architecture is suffering from the problem of overfitting. We present the conducted literature review of all the above discussed models in table 2.3 and table 2.4.

Table 2.2: Review on CNN based studies

Reference	**Embedding or encoding used**	**CNN variant used**	**Evaluaiton metrics used**	**Dataset used**	**Results achieved**
[22]	GloVe	CNN for text classification	Accuracy	Two datasets are used here. One collected from the twitter scraping and the other is PoliFact dataset	73.28% accuracy achieved
[23]	Word2vec	N-gram CNN and LIWC CNN	Accuracy	$Fake_or_Real_news$ dataset available publicly on github and kaggle	87% accuracy attained by n-gram CNN
[27]	GloVe	Graph CNN	ROC-AUC curve	Twitter data was collected and for labelling the data, journalist fact checking organizations were employed (such as buzzfeed, politifact and snopy)	ROC AUC of 92.70 ± 1.80% and 88.30 ± 2.74%
[1]	GloVe	CNN for text classification	Accuracy	Real-world fake news dataset available on kaggle is used	97.55% accuracy
[28]	Word2Vec	1-D CNN and CNN with more convolutional layer	Accuracy, recall, precision, f-1 score and time taken by the model	Dataset taken from kaggle competition	Both variants of CNN crossed accuracy of 86%
[29]	Word2Vec	CNN and deep CNN	Accuracy, precision, recall and F-1 score	Kaggle dataset was used	Both the models crossed precision of 0.69
[30]	Glove and one hot encoding	CNN	Accuracy	Liar	63% best achieved accuracy
[31]	GloVe	CNN	Accuracy	Two datasets from kaggle are used here (DS1 and DS2)	100% accuracy best achieved
[32]	word2vec, doc2vec, tfidf, and one hot	CNN	Accuracy	Liar and kaggle dataset	96.89% accuracy achieved by CNN

Table 2.3: Review on RNN based studies

References	Embedding or encoding used	Dataset used	RNN variant used	Evaluation metric used	Results achieved
[24]	Not mentioned	Created their own dataset by scraping posts from facebook. Both real and fake news posts were web-scraped	LSTM	Accuracy	99.4% accuracy achieved
[22]	GloVe	Two datasets are used here, one is twitter scrapped data and other is PolitiFAct dataset	LSTM and Bi-LSTM	Accuracy	83.82% accuracy achieved by Bi-LSTM (best case)
[23]	Word2Vec	A publicly available dataset from kaggle "$fake_or_real$" is used here	Depth LSTM and embedding LSTM	Accuracy	94% accuracy achieved by embedding LSTM
[1]	GloVe	A kaggle dataset "Real-world fake news" is used here	LSTM	FPR, FNR, cross-entropy loss, and accuracy	97.55% accuracy achieved by LSTM
[28]	Word2Vec	Kaggle dataset is used here	LSTM	Accuracy, recall, precision, f-1 score and time taken by the model	LSTM crossed the accuracy of 83%
[29]	TF-IDF and Word2Vec	Kaggle dataset is used here	LSTM	Accuracy, precision, recall and F-1 score	97.3% accuracy achieved
[33]	Word2Vec	It is produced using 20015 news articles which are gathered from two online sources	LSTM	accuracy, ROC-AUC, F1-score	91% accuracy achieved

Table 2.4: Review on hybrid and other different deep learning based studies

Reference	Model used	Dataset used	Embedding or encoding used	Evaluation metric used	Results achieved
[34]	Keras neural network model, four models are proposed, Model 1: Fed with N-gram vectors of news title, Model 2: Fed with N-gram vectors of news content, Model 3: Fed with sequence vectors of news title, and Model 4: Fed with sequence vectors of news content	Data collected and combined from two different kaggle datasets	Tf-Idf	Accuracy, Recall and computation time	90% accuracy achieved by model 4
[22]	CNN + LSTM ensembled model, Bidirectional LSTM + LSTM ensembled model, CNN + LSTM ensembled model with attention mechanism, and CNN + bidirectional LSTM ensembled model with attention mechanism	Two datasets are used here. One collected from the twitter scraping and the other is PoliFact dataset	GloVe	Accuracy	88.78% accuracy achieved by CNN + Bidirectional LSTM ensembled model with attention mechanism.
[25]	CNN+LSTM with four types of feature engineering techniques	The Fake news challenge dataset was taken	Word2Vec	Accuracy and F-1 Score	CNN-LSTM with PCA : 97.8 % is the best achieved accuracy
[26]	Hybrid CNN+LSTM	Dataset was taken from kaggle	GloVe	Accuracy, sensitivity, and specificity	Precision: 97.21%, Recall: 91.89%, Specificity: 97.44%
[35]	Hybrid CNN+RNN	FA-KES dataset and ISOT dataset	Keras embedding layer	Accuracy, precision, recall, and F-1 score	FA-KES: 60% accuracy, ISOT: 99% accuracy

Chapter 3

Improving the efficiency of fake news detection using deep learning

This chapter proposes a novel architecture that improves the efficiency of the results of the reproduced results of [3]. We present the details of the dataset used in this experiment along with the unique approach to resolve the lacunae of the existing approach.

3.1 Motivation and objectives

There are several ways in which individuals who read fake news stories, form their opinions based on what they read, and then act in accordance with their convictions put themselves in danger of being misled. The significance of identifying and avoiding the spread of false information served as the inspiration for this proposed approach. Following are the objectives of the proposed model:

1. To prevent the harm that may have been caused by spreading false information will be neutralised before it can spread to the general public by detecting the news which is fake.

2. To address the issue of class imbalance in the dataset which is previously not been addressed by [3].

3. To propose a novel architecture where fake news can be identified by a deep learning model(LSTM).

4. To make internet a safe space for individuals for their daily news reading/sharing.

3.2 Dataset description

The dataset produced by [36] is used in this study. The dataset consist of enumeration of news related to the COVID-19 news broadcasted across the web. The dataset consist of two columns viz., headlines spread over internet and whether these headlines are true or false. In all, there are 10201 headlines (in text format). 474 headlines were labelled true and 9727 were false. The proportion is shown in figure 3.1. Class 1 refers to the correct news claims and class 0 contains the false claims.

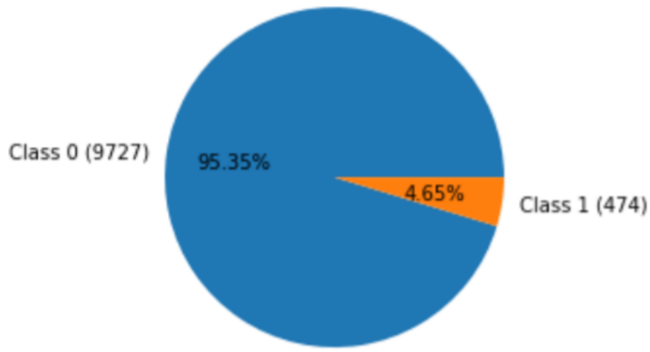

Figure 3.1: Dataset class distribution

3.3 Proposed architecture

We propose a idiosyncratic approach to detect the false claims present in the dataset. The architecture proposed in this work is designed by us for improving the performance of [3]. Figure 3.2 shows the exact process of identifying the false headlines using a unique deep learning approach. First the dataset is pre-processed before it is fed into any deep learning model. Then the encoding/embedding takes place for converting the cleaned data into vector format so that it becomes machine friendly for classification. The embedded data is given to the deep learning model for further classification of the target. The output generated is then evaluated using the suitable evaluation metric. The detailed process is discussed in this chapter. The unique layered system is designed for recognition of false claims and it is explained in the latter subsections of this chapter.

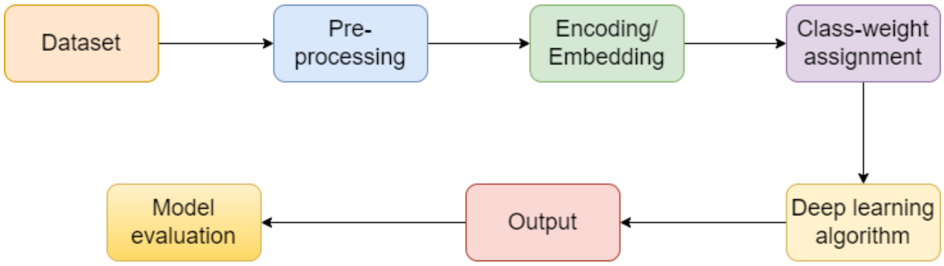

Figure 3.2: System flow diagram

3.4 Explanation of the layered architecture

Herein, we discuss the detailed proposed layered mechanism for fake news classification. The proposed approach contains distinct layers and each layer plays significant role in predicting the target attribute. In all five layers are involved in the architecture and detailed description of each layer is presented in the following sub-sections.

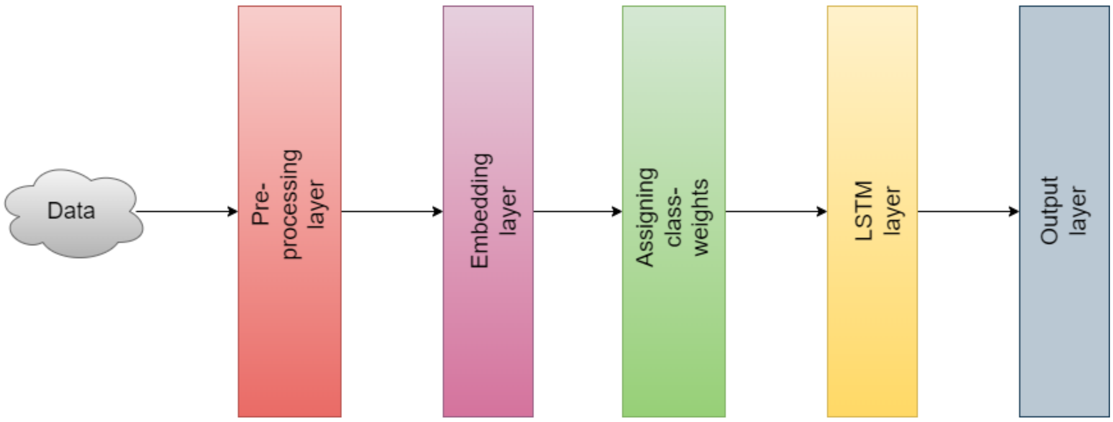

Figure 3.3: Proposed layered architecture

3.4.1 Pre-processing layer

This layer is useful when the data is raw and requires cleaning for better learning of the model. Cleaning raw data is very important because it contains a lot of useless symbols or words which do not help the algorithm in classification. The dataset is pre-processed before it is fed into any deep learning model. The pre-processing involves the dropping the null rows if any, followed by removal of stopwords, stemming, lower-casing the words. The cleaned data is then converted into one-hot encoding represenation.

3.4.2 Embedding layer

Word embedding is a term used for word representation to analyze text, usually in the form of a vector with a real value that includes the meaning of the word as words close to the vector space are expected to match the meaning so that the machine understands it. Herein, the embedding layer is utilised such that it pads the given sequence of texts into equal lengths (sentence length 60 is chosen). The embedding proposed by [3] is utilised in this experiment.

3.4.3 Class-weight assignment layer

The dataset is highly skewed and hence not fit for learning. In order to solve this issue, we introduce the concept of assigning the class-weights before we send the cleaned data into any deep learning algorithm for prediction purposes. The class-weights can be assigned to both majority and minority classes present in the dataset. The weights assigned before training the model to the classes impacts the prediction by giving equal priority to all the available classes irrespective of lower number of samples available in the minority class. The more priority to the minority class in the cost function of the applied algorithm and therefore reduces the error in predicting the target in case of the minority class.

The formula for calculating the class weights for any respective class is given by equation 3.1 .The class weights in this experiment are manually passed using a dictionary data structure with 70% weightage is given to the minority class and remaining 30% weightage is given to the majority class.

$$W_i = Total_{samples}/(target_{classes} * R_i) \tag{3.1}$$

where W_i is weight for class "i", $Total_{samples}$ denote the total rows in the dataset, $target_{classes}$ are the total number labels present in the target class, and R_i is number of rows present in the class "i".

3.4.4 LSTM(Long-short term memory) layer

After assigning the class-weights, the LSTM layer with 128 neurons is added for training followed by a dense layer with sigmoid activaiton function is employed for output purpose. The data was split into two parts, 80% training and 20% testing. The model was trained for 35 epochs with batch size of 32. Figure 3.4 shows the working diagram for LSTM used as the deep learning algorithm for classification purpose.

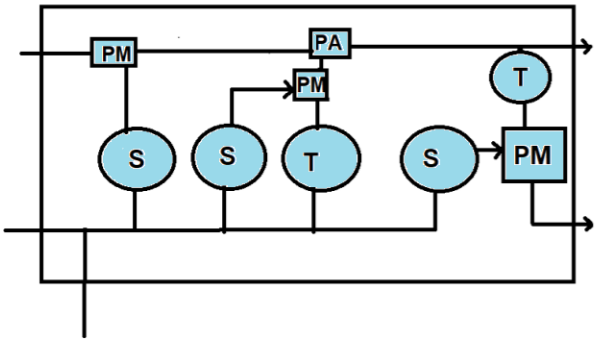

Figure 3.4: LSTM architecture [1]

where "T" denotes the tanh (i.e., activation function), "S" is the sigmoid function, "PA" represents the addition to be performed point-wise, and "PM" denotes the pointwise-multiplication.

3.4.5 Output layer

This layer is responsible for collecting the results that are given by the deep learning algorithm. The output in terms of classification of the claims present in the dataset into fake or real. The further evaluation of the model is analyzed here. For the performance analysis of the proposed model, we calculate f1-score, support, precision, and recall along with the accuracy. The precision for both the classes is calculated to check the performance of both the classes. For more visualization of the results, confusion matrix is also created for the proposed model.

3.5 Experimental setup and motivation for conducting this experiment

The experiment conducted in this research study employed google colab platform and github(for data extraction). The parameters used in this experiment and the simulations are same as [3]. The simulated study utilized a dataset created by [36]. The dataset used for the experiment is highly imbalanced resulting into poor performance into existing paradigms. This research experiment therefore gets its motivation from the fact that there is still need and scope for improvement for handling such sensitive data when used for classification purposes .The recent advancements in the deep learning models are boon to the NLP(Natural language processing) problems, but lack the ability to classify when the data is disproportioned. The prime goal of this experiment is to differentiate the false news from the real one.

3.6 Analysis of the achieved results

This section discusses the results achieved by the proposed approach. This experiment has exclusively taken care of the sensitive data created by [36]. The performance of the reproduced results of [3] were poor for the minority class. The dataset imbalance problem in the domain of false news claims can lead to biased learning of the model which is definitely not a good idea. Whether a news is true or false, the model shall predict both the classes correctly. In the next subsection, we discuss the results achieved by the proposed novel approach.

3.6.1 Results

The accuracy attained by the proposed model is 97% and the original accuracy was 96%. The achieved achieved results are produced by the proposed architecture as an improved version of [3]. The detailed classification reports for both the approaches can be seen in table 3.1 and table 3.2. The recall, precision, and F-1 score of class 1 and 0 are quite promising than the existing approach [3].

Table 3.1: Results achieved by the proposed approach

	Precision	**Recall**	**F-1 score**	**Support**
Class 0	0.98	0.99	0.99	1945
Class 1	0.72	0.65	0.68	96
Macro-avg	0.85	0.82	0.83	2041
Weighted avg	0.97	0.97	0.97	2041

Table 3.2: Results achieved by the existing approach [3]

	Precision	**Recall**	**F-1 score**	**Support**
Class 0	0.98	0.98	0.98	1945
Class 1	0.55	0.49	0.42	96
Macro-avg	0.76	0.74	0.75	2041
Weighted avg	0.96	0.96	0.96	2041

Chapter 4

Fake news detection using graph neural network

Herein, we aim to detect fake news using the concept of graph neural networks. The study [2] is taken into the consideration and the results for the same are reproduced in this work. In this research, we leverage media users' previous postings to determine their endogenous news consumption preferences. We offer User Choice-aware False Detection, a fake news detection approach that integrates intrinsic preference and extrinsic context together.

4.1 Motivation and objectives for implementing graph neural network

The previously implemented models for fake news detection are majorly based on the content of the news present as a dataset. The other data points such as the user's historical data, the user profiles, news propagation patterns, etc are not considered by the models. When it comes to take the other data points in to consideration, the deep learning models requires a structure in which the data can be represented (graph based data depiction). The data structure used to model the dataset for the 'fake news detection' task for GNN is graph which allows us to keep more data points for model learning which leads to better results. The motivation for implementing the graph based model is to take both the intrinsic as well as extrinsic features for more accurate prediction of the fake news which is the ultimate goal of this study. Following are the objectives we aim to attain by utilizing graph based models:

1. Graph based models not only focus on the news content but also other features which are useful for better prediction.

2. The inclusion of endogenous and exogenous feature for better prediction.

3. The massive amounts of data that are produced by social media platforms on a daily basis become a laborious undertaking, and therefore maximum number of data points extraction is performed using graph neural networks.

4.2 Dataset description

The FakeNewsNet database [18] was chosen to explore both user preference and false news distribution patterns. FakeNewsNet contains multi-dimensional data on content of the news, social context, and spatial data. It includes false and actual news via two fact-checking sites, as well as social media participation from Twitter. Table **??** displays the dataset characteristics.

Dataset name	**Total graphs (for fake news)**	**Total graphs (for real news)**	**Total nodes**
PolitiFact	157	314	41,054
GossipCop	2732	2732	314,262

Table 4.1: Dataset description

4.3 Endogenous user preference:

The endogenous preference in this context refers to the intrinsic behavior of the user which can be seen and collected by collecting the user's historical data. Herein, we utilize the last 200 for each account to get historical data for user preference modelling, totaling about 20 million tweets retrieved (from twitter API). We utilise randomly picked tweets from available users engaged in the very same news as the comparable historical postings for inaccessible individuals whose accounts are disabled or deleted. The intrinsic features of users help us to identify the user preference as a feature for fake news detection.We encode news items and user historical postings using different text depiction (Embeddings) learning algorithms to simulate the user endogenous preference. The encoding performed in this experiment make use of word embedding (Bert).

The news material is encoded employing BERT with the longest possible input sequence (i.e., 512 tokens). We couldn't utilise BERT to encode 200 tweets as one chain of sequence due to the input pattern length constraint, so we had to encode every tweet individually and then averaged them to get a user's preference depiction. Because twitter text is often much shorter than news text, we experimentally limited BERT's maximum input sequence length at 16 tokens (this reduces the time for encoding). The encoding strategy discussed above is depicted in the figure 4.1.

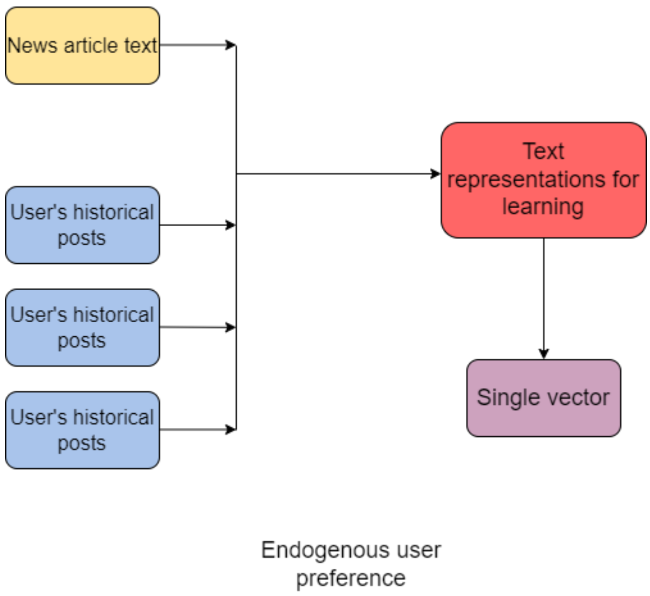

Figure 4.1: Endogenous user preference [2]

4.4 Exogenous user preference:

The individual's exogenous context is made up of all people who interacted with a news story on social media. To create a news propagation graph, we use the retweet data from news articles. We create a news dissemination graph using retweet data from news articles. The information propagation network(graph) is a tree-structured network in which the root node denotes the news item and other nodes indicate people who disseminate the root article, as illustrated in Figure 4.2.

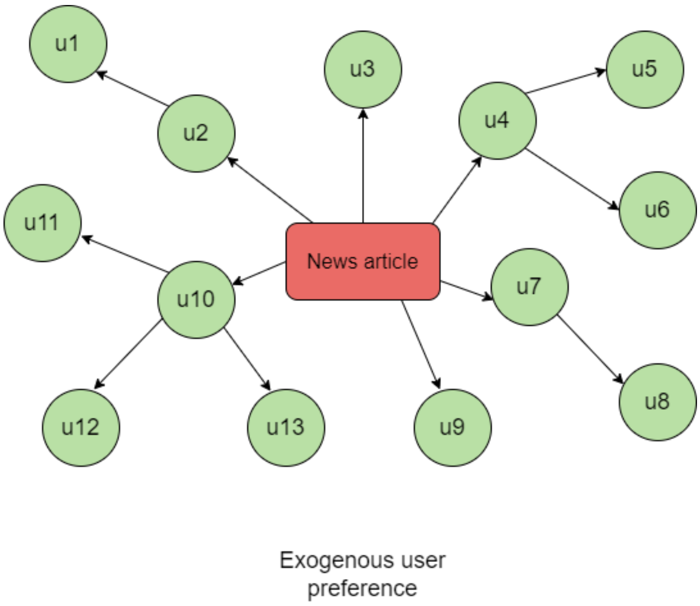

Figure 4.2: News propagation graph [2]

4.5 Fusion of the extracted information

Using the GNN, we basically combine the endogenous and exogenous data. The news text embedding and user interest embedding may both be considered node characteristics in a GNN. Typically GNNs discover the node embedding by aggregating the characteristics of its neighbouring nodes given the news propagation network. To acquire the embedding of an information transmission graph, we employ a readout function across all node embeddings. Second, news material often includes more clear clues about the news' reliability.As the ideal news embedding, we concatenate news textual encoding and user interaction embedding to improve the embedding of the news content.

Later the fused information is given to a multi-layered perceptron for the binary classification (fake vs real). The loss function used here is the binary cross entropy, and the model updation is performed using SGD (Stochastic gradient decent). The dimension of the graph embedding used here is 128 along with the Adam optimizer.

4.6 Experimental setup

This whole experiment is performed on CloudOcean platform. The PyTorch-Geometric is used here to implement all graph based models. The test-train-validation ratio here is 70-20-10 for all the algorithms. The batch size used here is 128 along with the 'Adam' optimizer, and the regularization technique used here is 'L2' with weight 0.001. The mean of five different simulations is carried out here.

4.7 Methodology

There are three primary components to our framework. To understand user endogenous inclination, we crawl the past postings of people interested in the news, provided a news story. By encoding previous postings utilizing text representation learning algorithms, we may infer the opinions of active users. Second, to take advantage of user exogenous knowledge, we construct the news dissemination graph using engagement data from social media sites (for example, Twitter retweets). To merge the user's endogenous preference and external surroundings, we design a tiered information fusion technique. The engagement of the users is presented in the form of embedding along with the news text embedding are combined to create the final news word embedding. The findings of the study are averaged across five separate runs.

In this research, we demonstrate that user endogenous content aggregation preferences are crucial in detecting bogus news. To back up this claim, we gathered user historical postings to simulate their endogenous preferences and used the news dispersion graph from twitter as the users' extrinsic social context. In all, there are five graph convolution operators were employed in this work viz., GCN [37] , GCN-FN [27], Bi-GCN [37], GraphSage [38], and GAT [39]. The text embedding for both the 'news content' and 'historical posts of users' used is BERT.

The aforementioned convolution operators are pre-defined in the Py-Geometric framework. For this experiment, we run 10 simulations for each model.

4.8 Results achieved

The results for both the datasets (PolitiFact and GossipCop) is calculated in this work. The simulations carried out were recorded and then average of all the achieved results (for 10 times) is calculated. We present the accuracy and F1-score of PolitiFact in table 4.2 and GossipCop in table 4.3.

Simulation no.	Accuracy	F-1 score
1.	85.54	85.46
2.	83	83
3.	82.99	82.89
4.	83	83
5.	82.98	82.98
6.	83	83
7.	82.98	82.89
8.	83	83.41
9.	83	83
10.	82.98	82.89

Table 4.2: Results achieved for the dataset PolitiFact

Simulation no.	Accuracy	F-1 score
1.	95.43	95.40
2.	95.41	95.39
3.	95.41	95.38
4.	94.77	94.75
5.	95.45	95.43
6.	95.27	95.24
7.	95.25	95.45
8.	95.22	95.21
9.	95.42	95.40
10.	95.42	96.40

Table 4.3: Results achieved for the dataset GossipCop

4.8.1 Statistical analysis

We conduct a t-test between the original results (stated by [2]) and the results reproduced in this study.

We present the original results in table 4.4 and the reproduced results in table 4.5

Dataset	Accuracy	F-1 score
PolitiFact	84.62	84.65
GossipCop	97.23	97.22

Table 4.4: Originally produced results by [2]

We perform a t-test between the two pair of results and try to check whether the reproduced results are statistically similar or not.

Firstly, we perform the t-test between the original pair of accuracy and the reproduced accuracy on the chosen two datasets.

Dataset	Accuracy	F-1 score
PolitiFact	83.25	83.25
GossipCop	95.32	95.41

Table 4.5: Reproduced results

The p value obtained for the mean vector of accuracy = 0.0001, which is less than 0.05. The reproduced results therefore are very much statistically similar in case of accuracy.

Next, we perform the t-test between the original and the reproduced F-1 score on the chosen two datasets.

The p-value obtained for the mean vector of F1-score = 0.8713, which is greater than 0.05. The reproduced results therefore are not statistically similar in case of F1-score.

Chapter 5

Challenges and future research directions

The bias created by these news posts(false claims) has been a focal point for various research studies so far. Researchers have made use of current developments and simple access to deep learning approaches to detect the falsely generated information, rumour, spamming, and other issues in social networking websites over the last few years [40]. It allows platforms to manage, analyze, and adjust a large amount of data. It's now most commonly used in business intelligence frameworks and predictive analytics. The proposed models for recognition of such false news face from some challenges that are discussed in this section as follows:

- Detecting fake news that has incorrect information itself is sounds confusing. Because no other news pieces have been published, this will require a thorough gathering of evidence as well as a thorough and proper inspection of facts. Although skilled specialists find assessing the validity of a news article to be a complex and challenging task.

- The task of comparing the facts with the various news items presented in various forms (audio, video, image post, or text post) becomes a tedious task if we try to automate it.

- The deep learning algorithms (hybrid models let's say) are computationally exhaustive, and when applied to a large amount of data would definitely add more to the computational and infrastructure cost.

- The data generated on different social media platforms of different languages would require their own separate embedding or encoding techniques which might result into an overhead. The code-mixed data(data containing languages more than one such as hindi+english) is difficult to handle.

All the above mentioned challenges must be addressed in order to enhance this area of research. This give rise to the administration of the future research that has the scope for improvisation of the existing methodologies and their lacunae. We discuss the foreseeable future research orchestration in this section. Though much has been done in recent years to improve the reliability and trustworthiness of internet content, certain crucial areas remain unresolved. This subsection we will be discussing the hitherto research void as well as possible future research directions. Administering the circulation of incorrect news and hereby lowering the negative whack on society requires quick and real-time detection of the originator are some of the unaddressed problems. Real-time data collection, automatic rumour detection, and tracking down the original source are all complicated tasks. These things give birth to some of the possible research opportunities. These research directions for future are discussed below in brief:

5.1 Echo chamber neutralization

When a user's inherent thoughts and perspectives are reinforced on social media, and he is unaware of the alternative viewpoints, result into echo chambers formation. As a result, more study is needed to connect the contradictory echo chambers and effectively communicate contrasting viewpoints to readers for reducing radicalization. It also aids in the finding of truth by requiring users to think critically and rationally across several perspectives.

5.2 Detecting the actual sources of fake news

Individuals usually have accounts on a number of social networking platforms, and they can spread rumours across one or more social networks that they own , making source discovery arduous. Accompanying this, the dissemination of misleading information from one online platform to many others, also known as cross platform propagation and hence the identification for the same has become a substantial hurdle for academics to investigate.

5.3 Dataset framing

Because most research is done on highly personalized datasets, the creation of credible standard datasets in this industry is essential. Datasets that are based and framed on actual real-time scenario of detecting the fake information across the social media platforms .A conventional juxtaposition between different algorithmic techniques is impossible due to a deficiency of availability of large-scale datasets publicly. The need for massive dataset containing all the probable news domain so that a deep learning or machine learning algorithm can use it for learning purposes. Such datasets are much demanding in this field of research.

5.4 Detecting fake news before its proliferation amongst the mass audience

Catching fake news immediately on, before it circulates, is a difficult challenge that must be accomplished in order to implement effective, a timely response, and prevention measures. It's nearly tough to change people's minds after fraudulent news has gotten prevalent and acquired their confidence. This is one of the major challenges where researchers need to focus because this can solve the problem even before it can take place in real time space.

5.5 Complexity in discerning the correctness or truthfulness of news articles

If we perform the veracity classification task before it is properly addressed, it becomes a prediction problem that necessitates a large amount of supporting data. Because of the intricate and variable network organization of social platforms, the problem becomes even more complicated. This issue will always be there and it requires special efforts to address it.

5.6 Language barriers

The majority of the research focuses on grammatical structures in English-language content. Other admired and accepted regional or local languages are not yet evaluated. The linguistic flexibility would add to the major advancements in determining the false news to a complete new audience which is left behind due to language barriers.

Chapter 6

Conclusion

The proliferation of any wrong piece of information amongst the masses can be dangerous. The origins of fake news are very difficult to trace due to vague sources and broadcasting medium. This dissertation begins with a detailed introduction on false news recognition that has arised out of the online platforms. We discuss the impacts these inaccurate content in the web has the ability to drive the mentality of the masses. The substantial impacts of fake news motivates this study to develop a small proof of concept in order to contribute a little into this domain. This study examines various deep learning paradigms used in the recent times by many researchers. The models include both raw and hybrid types that are capable of performing false news identification when implemented. Herein, we propose a unique architecture that has the potential to perform well even when the data has imbalance behaviour using LSTM. We implement various graph based models for fake news identification where not only news content is taken into consideration, but also user's historical posts. We discuss the challenges and possible upcoming research methodologies that can be employed in this domain. The future direction of this research work is to implement a notable prototype aiming for the detection if the fake and real news from a good benchmark dataset. This study aims to contribute towards the implementation of the same using some less applied techniques (such as fusion or hybrid models) as a part of future research work to achieve outstanding results.

Bibliography

[1] R. K. Kaliyar, A. Goswami, and P. Narang, "Deepfake: improving fake news detection using tensor decomposition-based deep neural network," *The Journal of Supercomputing*, vol. 77, no. 2, pp. 1015–1037, 2021.

[2] Y. Dou, K. Shu, C. Xia, P. S. Yu, and L. Sun, "User preference-aware fake news detection," in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 2051–2055, 2021.

[3] M. U. Islam, M. M. Hossain, and M. A. Kashem, "Covfake: A word embedding coupled with lstm approach for covid related fake news detection," *International Journal of Computer Applications*, vol. 975, p. 8887.

[4] T. Quandt, L. Frischlich, S. Boberg, and T. Schatto-Eckrodt, "Fake news," *The international encyclopedia of journalism studies*, pp. 1–6, 2019.

[5] A. Erlich, D. F. Jung, J. D. Long, and C. McIntosh, "The double-edged sword of mobilizing citizens via mobile phone in developing countries," *Development Engineering*, vol. 3, pp. 34–46, 2018.

[6] D. Westerman, P. R. Spence, and B. Van Der Heide, "Social media as information source: Recency of updates and credibility of information," *Journal of computer-mediated communication*, vol. 19, no. 2, pp. 171–183, 2014.

[7] S. Kumar and N. Shah, "False information on web and social media: A survey," *arXiv preprint arXiv:1804.08559*, 2018.

[8] S. B. Parikh and P. K. Atrey, "Media-rich fake news detection: A survey," in *2018 IEEE conference on multimedia information processing and retrieval (MIPR)*, pp. 436–441, IEEE, 2018.

[9] E. Hollowood and A. Mostrous, "Fake news in the time of c-19," *Tortoise, March*, vol. 23, 2020.

[10] T. Saikh, A. De, A. Ekbal, and P. Bhattacharyya, "A deep learning approach for automatic detection of fake news," *arXiv preprint arXiv:2005.04938*, 2020.

[11] B. Narwal, "Fake news in digital media," in *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, pp. 977–981, IEEE, 2018.

[12] A. Zubiaga, M. Liakata, R. Procter, G. Wong Sak Hoi, and P. Tolmie, "Analysing how people orient to and spread rumours in social media by looking at conversational threads," *PloS one*, vol. 11, no. 3, p. e0150989, 2016.

[13] W. Y. Wang, "" liar, liar pants on fire": A new benchmark dataset for fake news detection," *arXiv preprint arXiv:1705.00648*, 2017.

[14] A. Vlachos and S. Riedel, "Fact checking: Task definition and dataset construction," in *Proceedings of the ACL 2014 workshop on language technologies and computational social science*, pp. 18–22, 2014.

[15] P. S. Dutta, M. Das, S. Biswas, M. Bora, and S. S. Saikia, "Fake news prediction: a survey," *International Journal of Scientific Engineering and Science*, vol. 3, no. 3, pp. 1–3, 2019.

[16] H.-T. Zheng, J.-Y. Chen, X. Yao, A. K. Sangaiah, Y. Jiang, and C.-Z. Zhao, "Click-bait convolutional neural network," *Symmetry*, vol. 10, no. 5, p. 138, 2018.

[17] R. Barbado, O. Araque, and C. A. Iglesias, "A framework for fake review detection in online consumer electronics retailers," *Information Processing & Management*, vol. 56, no. 4, pp. 1234–1244, 2019.

[18] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media," *Big data*, vol. 8, no. 3, pp. 171–188, 2020.

[19] J.-P. Posadas-Durán, H. Gómez-Adorno, G. Sidorov, and J. J. M. Escobar, "Detection of fake news in a new corpus for the spanish language," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 5, pp. 4869–4876, 2019.

[20] N. T. Tam, M. Weidlich, B. Zheng, H. Yin, N. Q. V. Hung, and B. Stantic, "From anomaly detection to rumour detection using data streams of social platforms," *Proceedings of the VLDB Endowment*, vol. 12, no. 9, pp. 1016–1029, 2019.

[21] H. Ahmed, I. Traore, and S. Saad, "Detection of online fake news using n-gram analysis and machine learning techniques," in *International conference on intelligent, secure, and dependable systems in distributed and cloud environments*, pp. 127–138, Springer, 2017.

[22] S. Kumar, R. Asthana, S. Upadhyay, N. Upreti, and M. Akbar, "Fake news detection using deep learning models: A novel approach," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 2, p. e3767, 2020.

[23] Y.-F. Huang and P.-H. Chen, "Fake news detection using an ensemble learning model based on self-adaptive harmony search algorithms," *Expert Systems with Applications*, vol. 159, p. 113584, 2020.

[24] S. R. Sahoo and B. B. Gupta, "Multiple features based approach for automatic fake news detection on social networks using deep learning," *Applied Soft Computing*, vol. 100, p. 106983, 2021.

[25] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi, and B.-W. On, "Fake news stance detection using deep learning architecture (cnn-lstm)," *IEEE Access*, vol. 8, pp. 156695–156706, 2020.

[26] A. Agarwal, M. Mittal, A. Pathak, and L. M. Goyal, "Fake news detection using a blend of neural networks: an application of deep learning," *SN Computer Science*, vol. 1, no. 3, pp. 1–9, 2020.

[27] F. Monti, F. Frasca, D. Eynard, D. Mannion, and M. M. Bronstein, "Fake news detection on social media using geometric deep learning," *arXiv preprint arXiv:1902.06673*, 2019.

[28] V. M. Krešňáková, M. Sarnovskỳ, and P. Butka, "Deep learning methods for fake news detection," in *2019 IEEE 19th International Symposium on Computational Intelligence and Informatics and 7th IEEE International Conference on Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics (CINTI-MACRo)*, pp. 000143–000148, IEEE, 2019.

[29] R. K. Kaliyar, "Fake news detection using a deep neural network," in *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, pp. 1–7, IEEE, 2018.

[30] S. Garg and D. K. Sharma, "Phony news detection using machine learning and deep-learning techniques," in *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART)*, pp. 27–32, IEEE, 2020.

[31] P. Bahad, P. Saxena, and R. Kamal, "Fake news detection using bi-directional lstm-recurrent neural network," *Procedia Computer Science*, vol. 165, pp. 74–82, 2019.

[32] L. Singh, "Fake news detection: a comparison between available deep learning techniques in vector space," in *2020 IEEE 4th Conference on Information & Communication Technology (CICT)*, pp. 1–4, IEEE, 2020.

[33] Á. I. Rodríguez and L. L. Iglesias, "Fake news detection using deep learning," *arXiv preprint arXiv:1910.03496*, 2019.

[34] S. H. Kong, L. M. Tan, K. H. Gan, and N. H. Samsudin, "Fake news detection using deep learning," in *2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, pp. 102–107, IEEE, 2020.

[35] J. A. Nasir, O. S. Khan, and I. Varlamis, "Fake news detection: A hybrid cnn-rnn based deep learning approach," *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100007, 2021.

[36] S. Banik, "Covid fake news dataset." https://www.kaggle.com/thesumitbanik/covid-fake-news-dataset/. Accessed: 2021.

[37] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[38] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *Advances in neural information processing systems*, vol. 30, 2017.

[39] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, 2020.

[40] A. Abedalla, A. Al-Sadi, and M. Abdullah, "A closer look at fake news detection: A deep learning perspective," in *Proceedings of the 2019 3rd International Conference on Advances in Artificial Intelligence*, pp. 24–28, 2019.