A Report

on

FAKE NEWS DETECTION USING MACHINE LEARNING

Submitted in partial fulfillment of

the requirement for the award of the

degree of

Bachelor of
Technology in
Computer Science and Engineering

Under The Supervision of
Mr. Tarun Kumar
Assistant Professor

Submitted By
Mohammad Azeem Ansari - 19SCSE1010571
Rishi Singh Sengar – 19SCSE1010330

GALGOTIAS UNIVERSITY
(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY,
GREATER NOIDA INDIA
December 2021

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the project, entitled **"FAKE NEWS DETECTION USING MACHINE LEARNING"** in partial fulfillment of the requirements for the award of the B.tech CSE submittedin the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of October, 2021 to December 2021, under the supervision of Mr. Tarun Kumar - Assistant Professor, Department of Computer Science and Engineering/Computer Application and Information and Science, of School of Computing Science and Engineering , Galgotias University, Greater Noida

The matter presented in the thesis/project/dissertation has not been submitted by me/usfor the award of any other degree of this or any other places.

Mohammad Azeem Ansari - 19SCSE1010571
Rishi Singh Sengar – 19SCSE1010330

This is to certify that the above statement made by the candidates is correct to the best of myknowledge.

Mr. Tarun Kumar

Assistant Professor

CERTIFICATE

The Final Project Viva-Voce examination of Mohammad Azeem Ansari(19SCSE1010571) and Rishi Singh Sengar(19SCSE1010330) has been held on ________________and his work is recommended for the award of Bachelor of Technology.

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Project Coordinator **Signature of Dean**

Date: December , 2021Place: Greater Noida

Abstract

Abstract In our modern era where the internet is ubiquitous, everyone relies on various online resources for news. Along with the increase in the use of social media platforms like Facebook, Twitter, etc. news spread rapidly among millions of users within a very short span of time. The spread of fake news has far-reaching consequences like the creation of biased opinions to swaying election outcomes for the benefit of certain candidates. Moreover, spammers use appealing news headlines to generate revenue using advertisements via click- baits. In this paper, we aim to perform binary classification of various news articles available online with the help of concepts pertaining to Artificial Intelligence, Natural Language Processing and Machine Learning. We aim to provide the user with the ability to classify the news as fake or real and also check the authenticity of the website publishing the news.

Table of Contents

List of Figures

Introduction

1.1 Intoduction

As an increasing amount of our lives is spent interacting online through social media platforms, more and more people tend to hunt out and consume news from social media instead of traditional news organizations. The explanations for this alteration in consumption behaviours are inherent within the nature of those social media platforms: (i) it's often more timely and fewer expensive to consume news on social media compared with traditional journalism , like newspapers or television; and (ii) it's easier to further share, discuss , and discuss the news with friends or other readers on social media. For instance, 62 percent of U.S. adults get news on social media in 2016, while in 2012; only 49 percent reported seeing news on social media.

It had been also found that social media now outperforms television because the major news source. Despite the benefits provided by social media, the standard of stories on social media is less than traditional news organizations. However, because it's inexpensive to supply news online and far faster and easier to propagate through social media, large volumes of faux news, i.e., those news articles with intentionally false information, are produced online for a spread of purposes, like financial and political gain. it had been estimated that over 1 million tweets are associated with fake news “Pizzagate" by the top of the presidential election. Given the prevalence of this new phenomenon, “Fake news" was even named the word of the year by the Macquarie dictionary in 2016.

The extensive spread of faux news can have a significant negative impact on individuals and society. First, fake news can shatter the authenticity equilibrium of the news ecosystem for instance; it's evident that the most popular fake news was even more outspread on Facebook than the most accepted genuine mainstream news during the U.S. 2016 presidential election. Second, fake news intentionally persuades consumers to simply accept biased or false beliefs. Fake news is typically manipulated by propagandists to convey political messages or influence for instance, some report shows that Russia has created fake accounts and social bots to spread false stories. Third, fake news changes the way people interpret and answer real news, for instance, some fake news was just created to trigger people's distrust and make them confused; impeding their abilities to differentiate what's true from what's not. To assist mitigate the negative effects caused by fake news (both to profit the general public and therefore the news ecosystem). It's crucial that we build up methods to automatically detect fake news broadcast on social media.

Internet and social media have made the access to the news information much easier and comfortable. Often Internet users can pursue the events of their concern in online form, and increased number of the mobile devices makes this process even easier. But with great possibilities come great challenges. Mass media have an enormous influence on the society, and because it often happens, there's someone who wants to require advantage of this fact. Sometimes to realize some goals mass-media may manipulate the knowledge in several ways. This result in producing of the news articles that isn't completely true or maybe completely false. There even exist many websites that produce fake news almost exclusively producing of the news articles that isn't completely true or maybe completely false. There even exist many websites that produce fake news almost exclusively.

World is changing rapidly. No doubt we have a number of advantages of this digital world but it also has its disadvantages as well. There are different issues in this digital world. One of them is fake news. Someone can easily spread a fake news. Fake news is spread to harm the reputation of a person or an organization. It can be a propaganda against someone that can be a political party or an organization. There are different online platforms where the person can spread the fake news. This includes the Facebook, Twitter etc. Machine learning is the part of artificial intelligence that helps in making the systems that can learn and perform different actions (Donepudi, 2019). A variety of machine learning algorithms are available that include the supervised, unsupervised, reinforcement machine learning algorithms. The algorithms first have to be trained with a data set called train data set. After the training, these algorithms can be used to perform different tasks. Machine learning is using in different sectors to perform different tasks. Most of the time machine learning algorithms are used for prediction purpose or to detect something that is hidden. Online platforms are helpful for the users because they can easily access a news. But the problem is this gives the opportunity to the cyber criminals to spread a fake news through these platforms. This news can be proved harmful to a person or society. Readers read the news and start believing it without its verification. Detecting the fake news is a big challenge because it is not an easy task (Shu et al., 2017). If the fake news is not detected early then the people can spread it to others and all the people will start believing it. Individuals, organizations or political parties can be effected through the fake news. People opinions and their decisions are affected by the fake news in the US election of 2016 (Dewey, 2016). Different researchers are working for the detection of fake news. The use of Machine learning is proving helpful in this regard. Researchers are using different algorithms to detect the false news. Researchers in (Wang, 2017) said that fake news detection is big challenge. They have used the machine learning for detecting fake news. Researchers of (Zhou et al., 2019) found that the fake news are increasing with the passage of time. That is why there is a need to detect fake news. The algorithms of machine learning are trained to fulfill this purpose. Machine learning algorithms will detect the fake news automatically once they have trained. This literature review will answer the different research questions. The importance of machine learning to detect fake news will be proved in this literature review. It will also be discussed how machine learning can be used for detecting the false news. Machine learning algorithms that are used to detect false news will be discussed in the literature review. The structure of the rest of paper is as Methodology in section two, section three shows the research questions, section four is showing the search process model that is followed for this literature review, result and discussion is given in section five, the conclusion is presented in section six. In the last, references are given for the papers that are discussed in this literature review.

1.2 Aim and Objectives

The main objective behind the development and upgradation of existing projects are the following smart approaches:

- Be Aware of such article while forwarding to others
- Reveal True stories
- Prevent from false crisis events
- Be Informative

1.3 Motivation

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values. Machine learning is an application of AI which provides the ability to system to learn things without being explicitly programmed. Machine learning works on data and it will learn through some data. Machine learning is very different from the traditional approach. In, Machine learning we fed the data, and the machine generates the algorithm. Machine learning has three types of learning.

1. Supervised learning
2. Unsupervised learning
3. Reinforcement learning

Supervised learning means we trained our model with labeled examples so the machine first learns from those examples and then performs the task on unseen data. In this fake news detection project, we are using Supervised learning.

The extensive spread of faux news can have a significant negative impact on individuals and society. First, fake news can shatter the authenticity equilibrium of the news ecosystem for instance.

Understanding the truth of new and message with news detection can create positive impact on the society.

Why machine learning is required to detect the fake news?

Increasing use of internet has made it easy to spread the false news. Different social media platforms can be used to spread fake news to a number of persons. With the share option of these platforms, the news spread in a fast way. Fake news just not only affects an individual but it can also affect an organization or business. So, controlling the fake news is mandatory. A person can know the news is fake only when he knows the complete story of that topic. It is a difficult task because most of the people do not know about the complete story and they just start believing in the fake news without any verification. The question arises here how to control fake news because a person cannot control the fake news. The answer is machine learning. Machine learning can help in detecting the fake news . Through the use of machine learning this fake news can be detected easily and automatically. Once someone will post the fake news, machine learning algorithms will check the contents of the post and will detect it as a fake news. Different researchers are trying to find the best machine learning classifier to detect the fake news. Accuracy of the classifier must be considered because if it failed in detecting the fake news then it can be harmful to different persons. The accuracy of the classifier depends on the training of this classifier. A model that is trained in a good way can give more accuracy. There are different machine learning classifiers are available that can be used for detecting the fake news that will be answered in the next question.

1.4 Scope

The usage of this system greatly reduces the time required to search for a place leading to quicker decision making with respect to places to visit. Used to view the location view (the user can even zoom in and zoom out to get a better view) as well as 360-degree image embedded in the application. The System makes use of weather underground API for fetching the details of weather at accuracy.

The user can also find the paths to follow to reach the final destination in map which gives a better view to the users. It becomes convenient for users to book their tour via website instead of visiting agency ultimately saves time and money.

What is a TfidfVectorizer?

TF (Term Frequency): The number of times a word appears in a document is its Term Frequency. A higher value means a term appears more often than others, and so, the document is a good match when the term is part of the search terms.

IDF (Inverse Document Frequency): Words that occur many times a document, but also occur many times in many others, may be irrelevant. IDF is a measure of how significant a term is in the entire corpus.

The TfidfVectorizer converts a collection of raw documents into a matrix of TF-IDF features.

2. Methodology

This literature review is written for answering some research questions. So the methodology that is used is the systematic literature review. This methodology helps in answering the research questions. The papers were collected from various databases to be discussed in this literature review. To answer the research questions, different research papers are discussed and cited in thisliterature review.

2.1 Exclusion and Inclusion

A number of papers are published every day. So when a string is searched a number of papers arepresented in the result. Not all the papers are relevant to that string. This means there is a need forthe criteria. The criteria for inclusion and exclusion that is followed in this literature review is given in the below table.

Exclusion Criteria	**Inclusion Criteria**
The language of the paper is not the English language.	Papers that are written in the English language.
The complete paper is not accessible.	Paper can be accessed completely.
Paper is not related to machine learning and fakeor false news detection.	Paper showing content related to machinelearning and fake or false news detection.

Table 1: Exclusion and Inclusion Criteria

Papers that fulfilled the above mentioned inclusion criteria were included in the literature review.

2.2 Quality Assessment

Quality of all included papers was assessed on the basis of the research work presented in those papers. The papers in which the researchers have discussed the machine learning use for fake or false news detection were considered as good quality papers to be included in this literature review.

2.3 Research Question

A SLR has to answer some RQs. In this literature review, three research questions will be answered on the basis of valid arguments. These two research questions are given below. RQ1: Why machine learning is required to detect the fake news? RQ2: Which machine learning supervised classifiers can be used for detecting fake news? RQ3: How classifiers of machine learning are trained to detect fake news? These research questions will be answered in the result and discussion section of this literature review.

2.4 Search Process

A search process is followed to collect the papers that can be discussed in this literature review. This search process can easily be understood through the below given diagram.

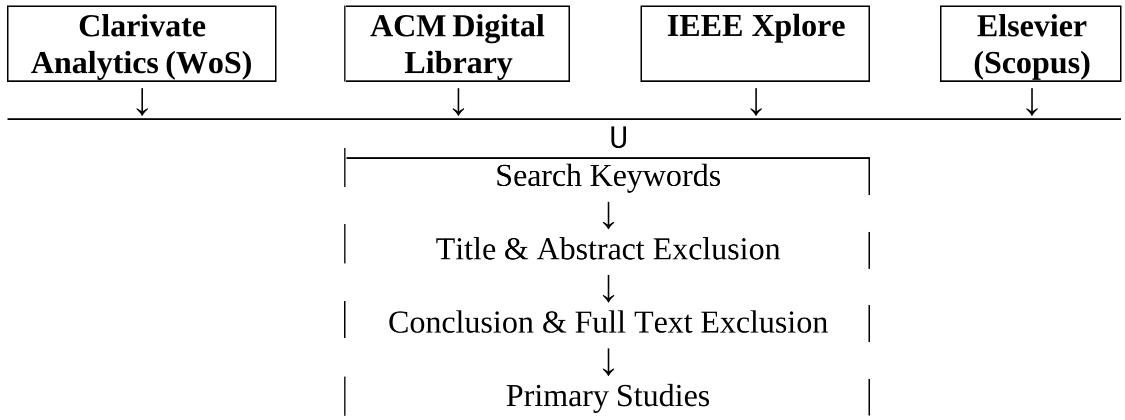


Figure 1: Search Process Model Diagram

Papers were collected from different databases. But not all of them were relevant to the topic. So first of all, the papers were excluded on the basis of their titles and abstract. An abstract is a kind of short summary of the whole paper that can give the idea about the contents presented in the paper. In the next phase, the further part of the papers was studied against the inclusion and exclusion criteria. Seventy three papers were collected from different databases against the searchkeyword. After the exclusion, there were twenty six papers were remaining that are discussed in this literature review.

3. Discussion

Internet is one of the great sources of information for its users (Donepudi, 2020). There are different social media platforms that includes Facebook or Twitter that helps the people to connectwith other people. Different kind of news are also shared on these platforms. People nowadays prefer to access the news from these platforms because these are easy to use and easy to access platforms. Another advantage to the people is that these platforms provide options of comments, reacts etc. These advantages attract people to use these platforms (Donepudi et al., 2020b). But aslike their advantages, these platforms are also used as the best source by the cyber criminals. Thesepersons can spread the fake news through these platforms. There is also a feature of sharing the post or news on these platforms and this feature also proves helpful for spreading such fake news.People start believing in such news as well as shares the news with other peoples. Researchers in (Zubiaga et al., 2018) said that it is difficult to control the false news from spreading on these social media platforms.

Anyone can be registered on these platforms and can start spreading news. A person can create a page as a source of news and can spread the fake news. These platforms do not verify the person whether he is really reputable publisher. In this way, anyone can spread news against a person or an organization. These fake news can also harm a society or a political party. The report shows that it is easy to change people opinions by spreading fake news (Levin, 2017). Therefore, there isa need for detecting these fake news from spreading so that the reputation of a person, political party or an organization can be saved.

RQ1: Why machine learning is required to detect the fake news?

Increasing use of internet has made it easy to spread the false news. Different social media platforms can be used to spread fake news to a number of persons. With the share option of these platforms, the news spread in a fast way. Fake news just not only affects an individual but it can also affect an organization or business (Donepudi et al., 2020a). So controlling the fake news is mandatory. A person can know the news is fake only when he knows the complete story of that topic. It is a difficult task because most of the people do not know about the complete story and they just start believing in the fake news without any verification.

The question arises here how to control fake news because a person cannot control the fake news.The answer is machine learning. Machine learning can help in detecting the fake news (Khan et al., 2019). Through the use of machine learning these fake news can be detected easily and automatically (Della Vedova et al., 2018). Once someone will post the fake news, machine learning algorithms will check the contents of the post and will detect it as a fake news. Differentresearchers are trying to find the best machine learning classifier to detect the fake news (Kurasinski, 2020). Accuracy of the classifier must be considered because if it failed in detecting the fake news then it can be harmful to different persons. The accuracy of the classifier depends on the training of this classifier. A model that is trained in a good way can give more accuracy. There are different machine learning classifiers are available that can be used for detecting the fakenews that will be answered in the next question.

RQ 2: Which machine learning supervised classifiers can be used for detecting fake news?

Detecting the fake news is one of the most difficult tasks for a human being. The fake news can easily be detected through the use of machine learning. There are different machine learning classifiers that can help in detecting the news is true or false. Nowadays, the dataset can easily becollected to train these classifiers. Different researchers used machine learning classifiers for checking the authenticity of news. Researchers in (Abdullah-All-Tanvir et al., 2019) used the machine learning classifiers for detecting the fake news. According to the experiments of the researchers the SVM and Naïve Bayes classifiers are best for detecting fake news.

These two are better than other classifiers on the basis of accuracy they provide. A classifier with more accuracyis considered as a better classifier.

The major thing is the accuracy that is provided by any classifier. Classifier with more accuracy will help in detecting more fake news. Researchers in (Kudarvalli & Fiaidhi, 2020) said that detection of false news is necessary because many personsspread the fake news of social media to mislead the people. To safe the individuals or organizationsfrom losing their reputation because of false news it is necessary to detect it (Rahman et al., 2020). They have said that the machine learning is very helpful in this regard. They used the different machine-learning algorithms and they also found that the Logistic regression is a better classifier because it gives more accuracy.

Researchers in (Aphiwongsophon & Chongstitvatana, 2018) said that the social media produce a large number of posts. Anyone can register on these platforms can do any post. This post can contain false information against a person or business entity. Detecting such false news is an important and also a challenging task. For performing this task the researchers have used the threemachine learning methods. These are the Naïve Bayes, Neural network and the SVM. The accuracy provided by the Naïve Bayes was 96.08%. On the other hand, the other two methods thatare neural network and SVM provided the accuracy of 90.90%.

According to the researchers of (Ahmed et al., 2017), false news has major impact on the political situation of a society. False news on the social media platforms can change opinions of peoples. People change their point of view according to a fake news without verifying it. There is a need for a way that can detect such news. The researchers have used classifiers of machine learning for this purpose. The classifiers that are used by different researchers are the K-Nearest Neighbor, Support Vector Machine, Logistic Regression, Linear Support Vector Machine, Decision tree, Stochastic Gradient Descent. According to results, linear support vector machine provided the good accuracy in detecting the false news.

Researchers (Reis et al., 2019) have used the machine-learning classifiers for the detection of fakenews. They have used different features to train these classifiers. Training of the classifiers is an important task because a trained classifier can give the more accurate results. According to the researchers of (Granik & Mesyura, 2017), artificial intelligence is better to detect the fake news. They have used Naïve Bayes classifier to detect fake news from Facebook posts. This classifier has given them the accuracy of 74% but they said the accuracy can be improved. To improve the accuracy different ways are also described by these researchers in that paper. There are classifiersof machine learning that are used for detecting fake news.

Some of these popular classifiers are given below that are used for this purpose.

Support Vector Machine: This algorithm is mostly used for classification. This is a supervised machine learning algorithm that learns from the labeled data set. Researchers in (Singh et al., 2017) used various classifiers of machine learning and the support vector machine havegiven them the best results in detecting the fake news.

Naïve Bayes: Naïve Bayes is also used for the classification tasks. This can be used to check whether the news is authentic or fake. Researchers in (Pratiwi et al., 2017) used this classifier of machine learning to detect the false news.

Logistic Regression: This classifier is used when the value to be predicted is categorical. For example, it can predict or give the result in true or false. Researchers in (Kaur et al., 2020)have used this classifier to detect the news whether it is true or fake.

Random Forests: In this classifier, there are different random forests that give a value and a valuewith more votes is the actual result of this classifier. In (Ni et al., 2020) researchers have used different machine learning classifiers to detect the fake news. One of these classifiersis the random forest.

Recurrent Neural Network: This classifier is also helpful for detecting the fake news. Researchers in (Jadhav & Thepade, 2019) have used the recurrent neural network to classify the news as true or false.

Neural Network: There are different algorithms of machine learning that are used to help in classification problems. One of these algorithms is the neural network. Researchers in (Kaliyar et al., 2020) have used the neural network to detect the fake news.

K-Nearest Neighbor: This is a supervised algorithm of machine learning that is used for solvingthe classification problems. This stores the data about all the cases to classify the new caseon the base of similarity. Researchers (Kesarwani et al., 2020) have used this classifier to detect fake news on social media.

Decision Tree: This supervised algorithm of machine learning can help to detect the fake news. It breaks down the dataset into different smaller subsets. Researchers in (Kotteti et al., 2018)have used different machine learning classifiers and one of them is the decision tree. Theyhave used these classifiers to detect the fake news.

RQ3: How machine learning classifiers are trained for detecting fake news?

Training of the classifiers of machine learning is an important task. This plays an important role for the accuracy of results of these classifiers. A classifier must have to be trained in a proper waywith proper data set. Different researchers have trained the machine learning classifiers to detect the fake news. The main problem that occurs while training these classifiers is that mostly the training data set in an imbalanced form (Wang et al., 2020). Researchers in (Al Asaad & Erascu, 2018) have used the supervised machine learning classifiers for fake news detection. To train theseclassifiers they have used the three different models for feature extraction. Actually, these featuresare used to train the classifiers.

Dataset

U

Removing Stop Words or Stemming

U

Test dataset & Training Dataset

U

Training Classifier

U

Experimental Results

Figure 2: Training Dataset

These models are the TF-IDF Model, N-Gram Model, Bag of Words Model. These models extract the features from the training data set and then the classifier is trained through these features. Researchers in (Ahmed et al., 2018) has trained some machine learning classifiers to detect the fake news. For the training purpose, they have used a training dataset. They have first removed the unnecessary words and the words are transformed to its single form. So that the training dataset that is given to these classifier should only have the valuable data.

4.LITERATURE SURVEY

4.1 Introduction

Our project is an web application which gives you the guidance of the day to day routine of fake news, spam message in daily news channel, Facebook, Twitter, Instagram and other social media. We have shown some data analysis from our dataset which have retrieve from many online social media and display the main source till now fake news and true news are engaged.

Our project is tangled with multiple models trained by our own and also some pretrained model extracted from Felipe Adachi. The accuracy of the model is around 95% for all the self-made model and 97% for this pretrained model. This model can detect all news and message which are related to covid-19, political news, geology, etc.

4.2 Existing System

We can get online news from different sources like social media websites, search engine, homepage of news agency websites or the factchecking websites. On the Internet, there are a few publicly available datasets for Fake news classification like Buzzfeed News, LIAR [15, BS Detector etc. These datasets have been widely used in different research papers for determining the veracity of news. In the following sections, I have discussed in brief about the sources of the dataset used in this work.This Existing system can help us to trained our model using machine learning technique.

4.3 Need of New System

Currently, many people are using the internet as a central platform to find the information about reality in world and need to be continue. Hence I has mention above we will create fake news and message detection model which detect the reality of the news and message.

Also, whose use our website can see the up to date about main source or keyword are getting most fake news and message and mapped up with chart. After and all everyone want to know how to prevent this hence we are giving some important tips to avoid this fake news of spreading rumour in the world.

4.4 Problems Definition

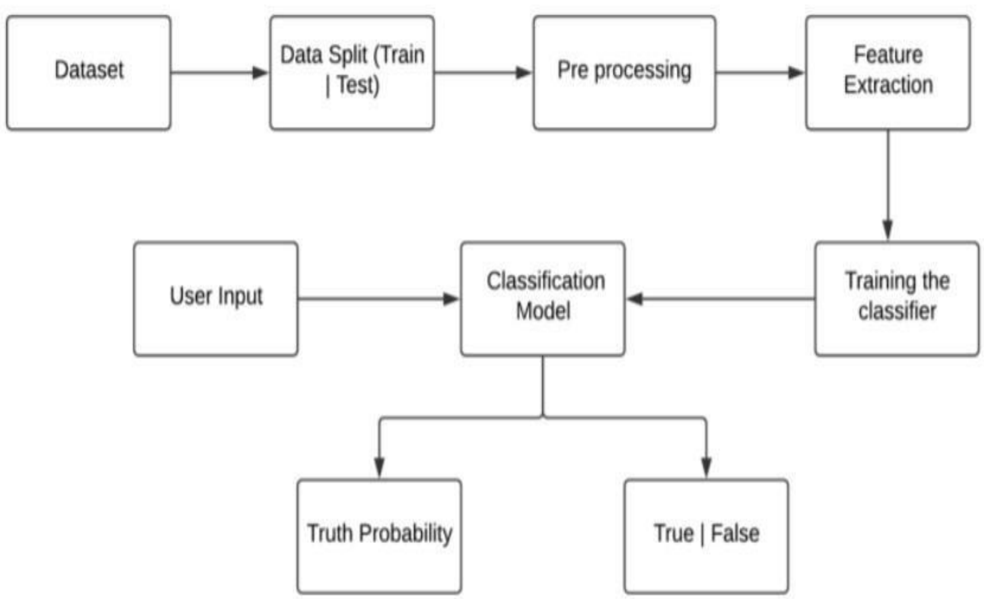
The system is an Web application which help user to detect the fake news. We have given the text box where the user has the option to paste the message or paste the url link of the news and other message link and after that it gives the reality of it. All the user gives data to detector may save for further use in order to update the statue of model, data analysis in future. We also help user by giving some guidance of how to prevent from such false event and how to stop with such event from spreading it.

5. DESIGN AND IMPLEMENTATION

5.1 Proposed system

The system is an Web application which help user to detect the fake news. We have given the text box where the user has the option to paste the message or paste the url link of the news and other message link and after that it gives the reality of it. All the user gives data to detector may save for further use in order to update the statue of model, data analysis in future. We also help user by giving some guidance of how to prevent from such false event and how to stop with such event from spreading it.

5.2 System Architecture Design



- **Random Forest Classifier**:

Random Forest is a trademark term for an ensemble of decision trees. In Random Forest, we’ve collection of decision trees (so known as “Forest”). To classify a new object based on attributes, each tree gives a classification and we say the tree “votes” for that class. The forest chooses the classification having the most votes (over all the trees in the forest). The random forest is a classification algorithm consisting of many decisions trees. It uses bagging and feature randomness when building each individual tree to try to create an uncorrelated forest of trees whose prediction by committee is more accurate than that of any individual tree. Random forest, like its name implies, consists of a large number of individual decision trees that operate as an ensemble. Each individual tree in the random forest spits out a class prediction and the class with the most votes becomes our model’s prediction. The reason that the random forest model works so well is: A large number of relatively uncorrelated models (trees) operating as a committee will outperform any of the individual constituent models.

- **Passive Aggressive Classifier Algorithm** :

Passive-Aggressive algorithms are generally used for largescale learning. It is one of the few '**online-learning algorithms**'. In online machine learning algorithms, the input data comes in sequential order and the machine learning model is updated step-bystep, as opposed to batch learning, where the entire training dataset is used at once. This is very useful in situations where there is a huge amount of data and it is computationally infeasible to train the entire dataset because of the sheer size of the data. We can simply say that an online-learning algorithm will get a training example, update the classifier, and then throw away the example.

6. Related Work

This problem is well-known and several papers have already studied this problem. However, they vary in their choice of data modeling. Ma et al. [9] introduced some important datasets and offered a first solution, using text content only. They framed it a sequential problem by transforming each example to a sequence of chronologically ordered tweets. They transformed each text to a fixed-size vector using word embeddings for the k-most important words of each tweet according to TF-IDF. On top of that, they used a multi-layer RNN on the embedding sequences to compute the label. A lot of other solutions use this sequential format as it has been proved efficient. Liu and Wu [4] use a combination of RNNs and CNNs on user features instead of text to achieve better results. Some methods using more advanced models have recently emerged. For instance, Ma, Gao, and Wong [10] use Tree Recursive Neural Networks and improve on their past results with the exact same data. The same team also experimented with recent deep learning techniques, such as Generative Adversarial Networks [7]. Except for the Tree Recursive Neural Networks, all these solutions do not really make use of the graph structure induced by the tweet history. Yet, there might be a lot of useful information hidden in the structure of these graphs. We observed that propagation graphs' shape depends on their label. Fake news are often striking and people are more inclined to react to them, so the speed and number of tweet.

7. Data overview

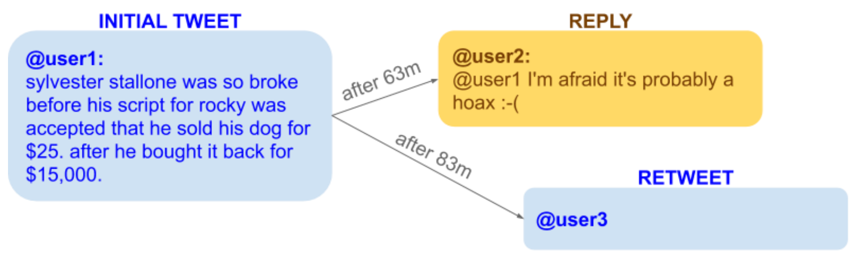


Figure 1: Example of **fake** news sharing tweet and subsequent retweets on Twitter16

The datasets we work with are *Twitter15* and *Twitter16* [6, 9]. These two datasets share the same exact structure. Both of them contain the tweets and retweets from a thousand of news articles published in 2015 and 2016. For each news article, the data contains the first tweet that shared it on Twitter, and a sequence of retweets following this initial post. We show one such data point (initial tweet and first two retweets) on Figure 1. Each event is labeled according to the initial news article, the label is taken out of four possible classes: "true", "false", "unverified", "non-rumor". Labels are evenly distributed in both datasets.

The task we address is *supervised graph classification*: our models label each event between 4 possible classes using the tweet/retweet data at hand. We highlight below the fundamental graphical structure of this data, that our models strive to leverage.

For a given news article, the sequence of tweets and retweets subsequently observed on Twitter is atree:

1 Nodes correspond to pairs of Tweet and Twitter User IDs. A retweet without commentshares ID of the initial Tweet, we use the User ID to distinguish them.

2 We put an edge between node 1 and 2 if node 2 is a retweet (with or without comment) of node 1.

Data cleaning Twitter15/16 were not perfectly collected and we spotted some data quality issues (e.g. negative propagation time, lines that were present twice). We implemented some fixes to get an acceptable level of data quality. For instance, we made sure that all propagation times are positive and that the tree structure always holds.

The dataset already contains the text written with each tweet or retweet. Due to privacy issues however, features of the user behind a tweet are not directly available in the Twitter datasets. We augmented our dataset by retrieving them using the Twitter API via the tweepy library. A major drawback to this is that we could only get the current state of the users (*i.e* as of October 2019). Hence, we had to make the strong assumption that user features did not change too much since 2015, or change in a way that doesn't affect too much our downstream classification task. Twitter15/16 are standard datasets for Fake News detection, and all papers published since 2016 must have made the same assumption with user features.

As it is usually done in papers using Twitter15/16 for Fake News detection, we hold out 10% of the events in each dataset for model tuning (validation set), and the rest of the data is split with a ratio of 3:1 for the train and test set.

Text features To get the best of the text data, we cleaned it before feeding it to a Transformer-based model. We tried BERT [1] and RoBERTa [5], and went with the latter option as it is the one which provided the best accuracy. Transformer-based models take as input a sequence of words, tokenize it, and output a list of highly-dimensional token embeddings. Right after tokenization, some starting and end tokens are added at the beginning and end of each sequence. We extract the final embedding of the starting token to get a fixed-dimension vector out of each sentence, as it is done in [1]. This 768 dimensions vector supposedly gives a representation of the sentence meaning. This is currently the state-of-the-art in word sequences embedding.

User features Some users deleted their accounts between the creation of Twitter15/16 datasets and when we started the project. Hopefully, we still managed to get the data for 90% of the users. We filled in the blanks with sensible values: medians for some numerical features, 0 for others, etc. Aggregated features used to fill out the blanks (e.g. medians) were computed from users of the train set solely to avoid data leakage. The features we extracted are presented on Table 1.

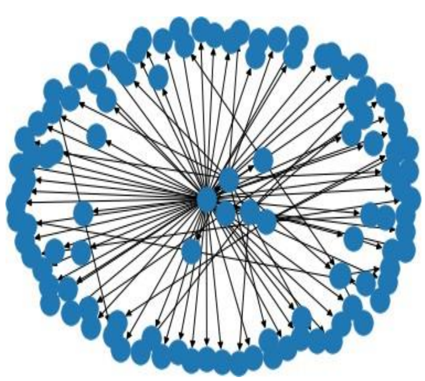
Name	Type	Description
created_at	numeric	Normalized time since account creation.
favourites_count	numeric	User favourites count.
followers_count	numeric	User followers count.
friends_count	numeric	User friends count.
geo_enabled	boolean	User geographical location enabled or not.
has_description	boolean	User has description on Twitter profile or not.
len_name	numeric	Length of the User username on Twitter.
len_screen_name	numeric	Length of the User name on Twitter, as seen by other users.
statuses_count	numeric	User count of statuses.

verified	boolean	Verified Tweeter user or not.

Table 1: Outline of our User features for Fake News classification

Trees extracted from Twitter15/16 The trees we defined (see above) from our news articles and Twitter data do not show a complex structure and are mostly "flat" (with low depth). On Twitter15/16, the maximum depth of the root tweet / retweets trees we build is around 5, and most nodes are simply at depth 1 or 2 from the root node. We illustrate this problem on Figure 2.

News article: a Twitter tree or a forest? We must mention an important limitation of the Twit- ter15/16 datasets we realized. For a given (Fake / True) news article, a *single* "root tweet" sharing this news article on Twitter is retrieved in the data, and we end up with a single tree of retweets. In reality, we expect that any important news article is shared independently by several users on Twitter ("independently" meaning here that they do not retweet one another). In Twitter15/16 we miss all those alternative "root tweets" as only one is selected, and lose a lot of graphical information: we have a single tree instead of a "forest".



The Twitter data collected by Ma et al. [9] to build Twitter15/16 is thus partial.

8. Baseline Models

Here, we describe briefly the baseline models we used to compare to our models leveraging techniques from Machine Learning on Graphs.

8.1 Gradient Boosting with Decision Trees (GBDT)

We first built a baseline using the *user features* of tweets, in the spirit of Liu and Wu [4]. Indeed, their results are "state-of-the-art" for Fake News classification on Twitter15/16 and we would like to see if we can get similar results with our models. However, we note that their code is not publicly available and hard to reproduce.

We opted for a GBDT model for two main reasons: user features are "tabular" in nature and boosted tree models perform well on those; GBDT is quick to implement and tune for good results. We use the Python API of the efficient LightGBM implementation of GBDTs [11].

To train the GBDT model, data is processed very simply: for each news article's series of retweets, we aggregate user features from Table 1 using user IDs for each retweet (mean aggregation for numerical features, sum aggregation for boolean features encoded as 0 or 1), the time (in minutes since the root tweet initial post) at which each retweet occurs is also mean-aggregated and used as a feature.

We improved on the baseline GBDT model by adding additional features from a fitted graphical contagion model (SEIZ: see Section 5.1) for each tweet/retweet's propagation observed on Twitter.

8.2 Long Short Term Memory (LSTM) Network and Multi-Layer Perceptron (MLP)

In their paper, Liu and Wu [3] obtain best results with a combination of Recurrent Neural Networks(RNNs) and Convolutional Neural Networks (CNNs) to classify news articles. The idea being that their RNNs and CNNs capture the "sequential" aspect of the data at hand: from an initial tweet sharing a news article, retweets and reactions from Twitter users are observed sequentially as time elapses and the information propagates on Twitter. Intuitively, their RNN model captures "global" information from this temporal tweets propagation while their CNN model captures "local" information.

In [3], authors employ GRU (Gated Recurrent Unit) cells for their RNNs. We opt for LSTM cells, a more common RNN cell, with arguably better representation power but higher complexity. Following their work, the sequence fed to our baseline LSTM corresponds to the sequence of retweets that follows the root tweet for any of our data point, earliest retweets first. Doing so, the sequence obtained for each data point is of *variable* length. We deal with this characteristic as Liu and Wu: if this length is below a threshold $L = 40$, authors randomly oversample some user features in this sequence to reach threshold size L; if the sequence longer than L, it is truncated to size L.

For the project milestone, we represented each retweet in the sequence via the user feature of the user retweeting. Those features are tabular and hard-to-learn-from for a LSTM (we had to properly scale them to obtain reasonable results). For this final report, we experimented with the Transformer-based features of the retweets text, those *dense* 768-dimensional embeddings are easier to learn from for a LSTM and we obtained better result.

We also train a MLP on the text feature of root tweets. By comparing this bare baseline to the LSTM, we could evaluate nicely the benefits from capturing the "sequential" nature of the retweets data with a LSTM.

The baseline LSTM model is then compared to Graph Neural Network (GNN) models to evaluate the additional benefits from capturing the "sequential" *and* "graphical" nature of our retweets data (trees in our modelling, as exposed above).

9. Graphical Models

9.1 SEIZ Contagion Model for Fake News Detection

The SEIZ model is probabilistic contagion model that has been applied to Tweets by Jin et al. [2]. SEIZ initials stand for the different states in which Twitter user can be over time, with respect to a news article propagating on the social network:

S: Susceptible. In theory, all active Twitter users that can reasonably be in contact with tweets associated with the news article.

9.1.1 E: Exposed. Twitter users exposed to tweets associated with the news article.

9.1.2 I: Infected. Twitter users believing the content of the news article.

9.1.3 Z: Skeptics. Twitter users that do not believe the content of the news article.

A graphical model defines the possible transitions between those intuitive states. The attributed of this graph and the initial populations in each of the SEIZ states govern the dynamics of their populations over time, according to an Ordinary Differential Equation (ODE). The graphical model is reproduced on Figure 3 as seen in class.

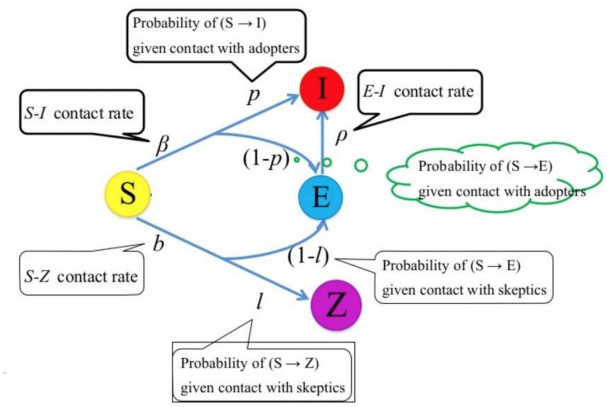


Figure 3: SEIZ graphical model

In our work, we followed the approach of Jin et al. [2] on our specific data. For each news article in Twitter15/16, the input to Jin et al.'s procedure is the number of distinct retweets (distinct (user_ID, tweet_ID) tuples) observed over time. We remind their procedure (for a single data point i.e. number of retweets over time) below:

1. Start from random SEIZ parameters and initial populations.
2. Find the best fit (in the least squares sense) of the obtained I(t) curve to our number of retweets over time. We use the *least_square* function from Scipy optimize for optimization, and implement ourselves a simple Euler method to solve the ODE (following the approach of the authors).

Examples of fit obtained on Twitter15/16 are presented on Figure 4. We restrict the fit of the SEIZ parameters to the first 120 minutes after the inital root tweet is posted on Twitter. It took us several hours to compute the "best" SEIZ parameters and initial populations for the roughly 2000 news articles of Twitter15/16.

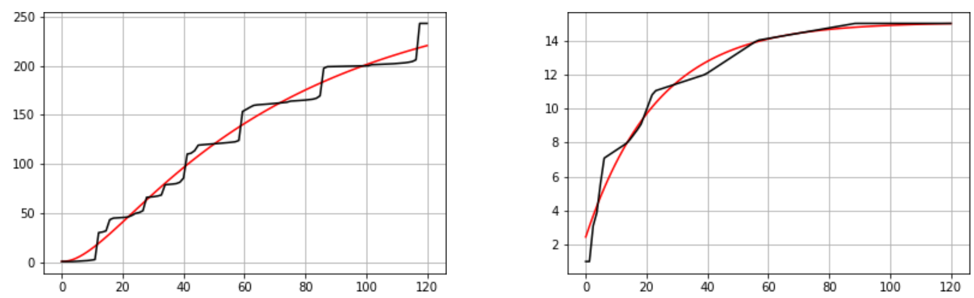


Figure 4: Example SEIZ fit on Twitter15 (left) and Twitter16 (right). The red curve is the SEIZ fit for I(t) and the black curve is the number of retweets observed over time.

The "best" parameters obtained (10 in total) for each news article are used for classification. We employ a GBDT model on this structured tabular data for best results. In the results section, we compare this GBDT SEIZ approach to the baseline GBDT approach using user features only. We also evaluate an *ensemble* of the two models to see if it can compete with Graph Neural Networks (GNNs).

10. Graph Neural Networks

Finally, we experimented with GNNs, our best performing models. GNNs are powerful neural network models to obtain node embeddings in a graph. In our 4-class classification problem, the graphs are the trees described in Section 3 and the task becomes *graph classification*.

We experimented with 3 variants of Graph Neural networks: Graph Convolutional Networks (GCNs), Graph Attention Networks (GATs) and GraphSAGE. We evaluated our model every epoch and selected the model with best validation accuracy.

As node features to any of our GNN models, we concatenate the User Features and (BERT-based) text features detailed in Section 3. For a given news article to classsify, we obtain GNN-based embeddings for all nodes of its retweets tree. For graph classification, we then need to aggregate those node embeddings: common approaches include max-pooling, mean, or sum aggregation. Because each tree has a root node that is directly related to the news we want to classify (see Figure 5), we thought that using the output embedding of this node would be useful. However, if the number of GNN layers is small, this embedding doesn't take into account nodes that are deep in the retweets tree. Thus, we decided to concatenate this root node embedding with the mean and max aggregates of all nodes of the tree, so that we could get a more accurate representation of the tree. We obtained better results with this approach.

11. Experiments and results

11.1 Experimental details

We ran two sets of experiments: a first set of experiments on the multi-class problem ("true", "fake","unverified", "non-rumor") and another set of experiments on the binary classification problem (retaining only news labelled "true" or "fake"). Reference papers usually carry out those two types of experiments.

GNN experiments We compared the three models (GAT, GraphSage, GCN) and did an ablationstudy where we examined the respective benefits of user and text features. We have three different sets of features in our ablation study: `text_only` using only the BERT-based text features, `user_only` using only User features described in Table 1, and `all` that uses both. Results are presented on Figures 5 and 6.

For the experiments, we used 1 to 2 layers, a batch size of 32, a dropout probability ranging from 0.1to 0.7 and an AdamW optimizer. Specific hyperparameters were tuned for each model and dataset based on the *validation* set classification accuracy.

Gradient Boosting experiments We trained three models: a baseline GBDT model leveraging (engineered) User features, a GBDT model leveraging SEIZ features, and an ensemble model of thetwo.

We used 2000 tree learners, a maximum depth of 5 for the decision trees grown, and a learning ratevarying between datasets (Twitter15/16) and models. The best learning rate was selected case by case to optimize accuracy on the *validation* set.

Competitive results on multi-class classification

	Twitter15			Twitter16		
Split	**Train**	**Val**	**Test**	**Train**	**Val**	**Test**
Recursive Tree[8]	NA	NA	0.723	NA	NA	0.737
RNN+CNN[3]*	NA	NA	0.842	NA	NA	0.863
GBDT_user	0.962	0.629	0.628	1.00	0.671	0.647
GBDT_seiz	0.672	0.412	0.360	0.741	0.506	0.377
Ens_GBDT	0.959	0.635	0.577	0.995	0.617	0.618
MLP text	0.931	0.568	0.536	0.882	0.634	0.549
LSTM text	0.899	0.584	0.622	0.922	0.622	0.587
GraphSage text	0.954	0.624	0.622	0.866	0.756	0.712
GCN all (Our best)	1.00	0.719	0.690	0.859	0.841	0.750

Table 2: Final results: Accuracy of baselines and graphical models on 4 classes classification. Results from top reference papers are featured for comparison. *: no code available and not reproducible

Results are presented in Table 2. We divide models in (1) baselines from the litterature [3, 8], (2) Gradient Boosted models, (3) Neural models (4) our best GNN mode: a GCN using both user and text node features. We highlight that, despite their high scores, Liu and Wu [3] did not make their code available.

Overall performance on Twitter15 On this dataset, our performances are competitive with stateof the art (excluding [3]). We can notice heavy overfitting on all variants of our model (see Figure 5) that could maybe be addressed with dropout (we didn't have time to exhaustively grid search the best parameters).

Overall performance on Twitter16 Here, our model outperforms the current state of the art (excluding [3]). Note that the size of this dataset is approximately half of the first one, so results (forevery model) are subject to higher variance. It also induced higher overfitting on the validation set ascan be seen on Figure 6, a trend we didn't notice on the first dataset.

12. Conclusion

We came up with a graphical model to represent Fake News data on Twitter15/16. We used this model to train powerful Graph Neural Networks classifiers, leveraging text features (extracted using the recent BERT transformer-based language model) and user features effectively. Their scores are comparable to the state-of-the-art on our datasets, they outperform methods relying on information aggregates, and a simple sequential representation of the tweets data. This work shows that useful information is hiding in the graphical structure of the news propagation on the Twitter social network, that our GNNs could leverage efficiently.

Following our above discussion, we identify two main axes of improvement on this work. First, to improve the generalization performance of our models and limit overfitting, via clever regularization and hyperparameter tuning. Then, our study highlights the current shortcomings of the Twitter15/16 datasets. To properly track the progress made on those, it would be important to have defined a reference train/validate split of the data, and to have a neutral third party hold a *held-out* test set. Twitter datasets are also very small, which prevents the development of complex models. Finally, the user features are now outdated, and it would be interesting to see if it would be possible to build a new version the Twitter datasets, coming with publicly available user features that correspond of the time of the events. Creating such a dataset is not possible today as it would clearly violate Twitter's privacy policy. It would be interesting to see if user features can be safely anonymized to permit the creation of such datasets.

12.1 Summary

- With the help of Machine Learning we have created 5 prediction model which gives the accuracy above 90% and it cover all latest political covid 19 news. Also, with some pretrained model we have cover news related to history and sport.

- We intent to build our own dataset which will be kept up to data according to the latest news in future.

12.2 Future Scope

This project can be further enhanced to provide greater flexibility and performance with certain modification whenever necessary. Deep fake learning which can be help to detect fake image. Deep learning machine learning to get more accurate result.