# FAKE NEWS DETECTION FROM REAL NEWS

Submitted for the Summer Internship

on

## Machine Learning and Deep Learning

(from 8th June, 2021 to 31st July, 2021)

**Organised by**

**DST Centre of Excellence – Artificial Intelligence, IGDTUW**

**IGDTUW-Anveshan Foundation**

**Department of AI and Data Sciences, IGDTUW**

By

**SALONI SETH**
Bachelors in Technology(B. Tech), Final Year

**HARSHA KUMARI**
Bachelor of Computer Applications (BCA), First Year

**CHARU PRABHA**
Bachelors in Technology(B. Tech), Final Year

## INDIRA GANDHI DELHI TECHNICAL UNIVERSITY FOR WOMEN
(Established by Govt. of Delhi vide Act 09 of 2012)
Kashmere Gate, Delhi-110006

# Synopsis

## Introduction

This project aims to detect fake news and identify it and report it to the user. 'Fake news', as it is called, is nothing but misleading news stories that come from non-reputable sources. The online social media is expanding its base and henceforth people have started to believe whatever news or items spread there. One such very prominent problem that is prevalent there is the spread of fake news. It has a great impact on society. There is much information from disparate sources among various users around the world which might or might not be true. In this project thus, by using the techniques of natural language processing we will identify whether the news item is legitimate or not and prevent our target audience from falling into the trap of irrelevant and fake news items. Though the meaning of the term 'Fake news' has been changing a lot due to the widespread cases and incidents that have occurred in the recent past. Hence it becomes necessary for us to be aware of the fact that whatever information we are gathering from and source is legitimate and sound. Untrue and untrustworthy articles and news items are harmful for the populace. They can lead to disastrous resentments by their audience, instigate people's minds and have a negative impact as well. Thus, it is necessary to focus on increasing awareness about information which can be considered as knowledge and discarding the one which is not true.

## Literature Review

Social networks are stepping up in using digital fake news detection tools and educating the public towards spotting fake news. At the time of writing, **Facebook** uses machine learning algorithms to identify false or sensational claims used in advertising for alternative cures, they place potential fake news articles lower in the news feed, and they provide users with tips on how to identify fake news themselves (Sparks and Frishberg 2020). **Twitter** ensures that searches on the virus result in credible articles and **Instagram** redirects anyone searching for information on the virus to a special message with credible information.

For fake news detection, current surveys have mostly limited their scope to reviewing research from a certain perspective (or within a certain research area, e.g., NLP [Oshikawa et al. 2018] and data mining [Shu et al. 2017]). These surveys generally classify fake news detection models by the types of [deep] machine learning methods used [Oshikawa et al. 2018] or by whether they utilize social context information .

The Facebook Fake News Task Team and their approach
In November 2016 it was reported 19, 20 that Facebook had formed an unofficial task force working on the problem of fake news. Through their official channels, Facebook promised to reprioritize fake news on its pages, saying one of their news feed values is "authentic communication" and that it's acting to prevent posts that are "misleading, sensational or spammy". Pages that have been posting fake news have been studied by Facebook's experts and are now expected to be seen less frequently in news feeds, was then reported.
According to facebook -
• "We categorized pages to identify whether or not they were posting spam or trying to game the Feed by doing things, like asking for likes, comments or shares"
• "We then used posts from these pages to train a model that continuously identifies whether posts from other pages are likely to be authentic"
• "For example, if page posts are often being hidden by people reading them, that's a signal that it might not be authentic".
• "If a post is likely to be authentic (...) your relationship to the person who shared [it] will also be taken into consideration along with the number of likes, shares and comments".

Therefore, Facebook uses, on a first instance, a human classification to identify the purpose of the post, eventually to conclude that it was spamming, or cheating the Facebook Feed. Then, it uses some sort of a machine learning algorithm using the classified posts. As a remark, we notice that there isn't a clear similarity between a post that is not spam, nor a cheating-post, and a fake news. Nevertheless, assuming that there is, the third sentence is still confusing because from the best of our experience, fake news tends to spread, and to spread fast. We must say that at the time of writing this paper Facebook changed the share button of posts, only allowing public posts to be shared. Finally, it is very interesting to see that Facebook uses the "social graph" to obtain intelligence about the spreading of fake news.

## Issues/Challenges with the Existing Work

Research into identifying, detecting, and mitigating the spread of fake news is still under development. Nevertheless, it is already possible to identify the main challenges in combating fake news, which are listed below .

• Great interests and the plurality of actors involved
The spread of fake news occurs as a distributed event, and involves multiple entities and technological platforms. Thus, there is an increasing difficulty in studying and designing computational, technological, and business strategies to combat fake news without compromising speed and collaborative access to high-quality information.

• Opponent's malicious intent
The fake news content is designed to make it difficult for humans to identify the fake news, exploiting our cognitive skills, emotions, and ideological prejudices. Moreover, it is challenging for computational methods to detect fake news, as the way fake news is presented is similar to true news, and sometimes fake news uses artifacts to make it difficult to identify the source or falsify the real source of the news.

• Susceptibility and lack of public awareness
The user does not have information about the credibility of the source or patterns of spreading of the news on the network. Thus, to increase public awareness, several articles and advertising campaigns are run to provide tips on how to differentiate between false and legitimate news. For example, the University of Portland in the United States provides a guide for identifying misinformation (fake news).

• Propagation dynamics
The spread of fake news on social media complicates detection and mitigation, as fake information can easily reach and affect large numbers of users in a short time. Verification of veracity must be carried out in an agile way, but it must also consider the patterns of propagation of information throughout the network .Constant changes in the characteristics of fake news.

Developments in the automated identification of fake news also drive the adaptation of the generation of new disinformation content to avoid being classified as such. The detection of fake news based on writing style, differentiating false and legitimate news by an analysis based on Natural Language Processing, is one of the most-used alternatives due to the unsolved challenges in automatic fact verification from pre-defined knowledge bases.

## Objectives of the Proposed Work

The project aims to build a model to detect and identify 'fake' news items and report them to the target audience.

## Methodology

In our project we have fake news and true news dataset. We  have created a word cloud to analyse the most frequent words in our data. The stop words are removed from the data, and the word clouds are generated.

We have applied three classification models namely MultiNomial Naive Bayes,Support Vector Machine, Passive Aggressive Classifier and evaluated the performance. As classification is supervised learning,we have to first split the data into training and test data. We train the model using the train data, and test the performance of our model using test data.

**1. MultiNomial Naive Bayes**
Naive Bayes are mostly used in natural language processing. Naive Bayes classifier algorithm is a family of algorithms which use Bayes Theorem. It uses the naive assumption that all the features are independent of each other. Bayes theorem calculates the probability $P(c|x)$ where c is the class of possible outcomes and x is the given instance which has to be classified.
$P(c|x) = P(x|c) * P(c) / P(x)$
According to our data, the class is 0 or 1, where 0 implies fake news and 1 implies true news. Given a news x, we will compute P(true news|x) as well as P(fake news|x). If P(true news|x) > P(false news|x), the algorithm predicts it is true news. Otherwise, the news will be predicted as fake.

**2. Support Vector Machine**
Support Vector Machine or SVM is a linear model for classification and regression problems. The SVM model takes the data in the training set, and maps it to data points in space so that there is a clear gap between points belonging to different categories. This gap is made as wide as possible to improve the performance of the model. Whenever a new data point is given to the model, it maps the point to the same space, and predicts the category based on the side of the gap on which they fall.

**3. Passive Aggressive Classifier**
Passive aggressive classifier is an online algorithm that learns from massive streams of data. The idea is to get an example, update the classifier, and throw away the example. It is fast and easy to implement, but does not provide global guarantees like SVM.

Before applying these models to our data we have to convert text into numbers.For that we will use a model called bag of words to deal with text words. It finds all the unique words in the document, and creates a vector of size equal to the number of unique words. Each word is assigned an index in the vector.

The main drawback with this approach is that it ignores all the information related to the order of the words, and only takes into account the frequency of the words. We are using CountVectorizer and TfidfTransformer for the transformation.

**Count Vectorizer**

The count vector tokenizer a collection of documents and builds a vocabulary of unique words. It can also encode new documents using this vocabulary.

**Tfidf Transformer or Term Frequency-Inverse Document Frequency**

It gives frequency scores to words by highlighting the ones which are more frequent in a document, but not across documents.It tokenizes documents, learns the vocabulary and inverse document frequency scores. It also encode new documents

We have used Pipeline object, a utility used to automate machine learning workflow. A pipeline allows several transformers to be chained together.

A pipeline has two main methods:

a) fit_transform: this method is called for each transformer and each time the result is fed into the next transformer.
b) fit_predict: if your pipeline ends with an estimator, fit_predict is called on the estimator.

In our project we have used accuracy and confusion matrix as two metrics to evaluate the performance of our model. **Accuracy** is considered as a good metric only if your data is balanced. **Confusion matrix** is a matrix which gives more insights into our model. It compares predicted values and the actual values.

## Conclusion

The news data has been classified using natural language processing techniques as fake and real. The performance of the models is being assessed based on the model accuracy and confusion matrix obtained. Hence, the desirable results have been achieved.

## References

1. https://www.geeksforgeeks.org/applying-multinomial-naive-bayes-to-nlp-problems/

2. https://machinelearningmastery.com/prepare-text-data-machine-learning-scikit-learn/

3. https://signal-to-noise.xyz/post/sklearn-pipeline/

4. https://www.analyticsvidhya.com

5. https://en.wikipedia.org/wiki/Support_vector_machine

6. https://www.ijrte.org/wp-content/uploads/papers/v7i6/F2457037619.pdf

7. https://www.researchgate.net/publication/343916946_Fake_News_Detection_Using_Machine_Learning_Algorithms

8. https://www.analyticsvidhya.com/blog/2019/12/detect-fight-neural-fake-news-nlp/

9. https://www.pantechsolutions.net/fake-news-detection-using-machine-learning

10. https://blog.usejournal.com/using-nlp-to-detect-fake-news-289314fb9198

11. https://iopscience.iop.org/article/10.1088/1757-899X/1099/1/012040/pdf

12. https://pdf.sciencedirectassets.com/280203/1-s2.0-S1877050920X00032/1-s2.0-S1877050920300430/main.pdf

13. https://bit.ly/2Ti7FiO