

# Applying Machine Learning to Detect Fake News

\* Subhabaha Pal  
\*\* Senthil Kumar T. K.  
\*\*\* Sampa Pal

## Abstract

Spread of fake news has become a big problem in recent times and this has become a source of discomfort in the society in many instances. Application of sophisticated Machine Learning algorithms for detection of fake news has shown some success and it is a new area which shows immense opportunities for practical use and further research. The detection of fake news using machine learning algorithms have been discussed in details in this paper. News items are unstructured data and this paper shows how unstructured news items can be turned into structured form using Count vectorization and TFIDF vectorization techniques. This paper also shows how machine learning models can be developed and the classification can be done using the newly structured data employing sophisticated developed Machine Learning algorithms.

**Keywords:** Count Vectorizer, fake news detection, Machine Learning, NLP, Text Analytics, TFIDF Vectorizer

## I. INTRODUCTION

Information is the most valuable resource today. Data is the new oil [1]. Enormous information is getting generated around the globe every second. Advent of social media has opened the gate for common netizens to publish customized content without any censorship and this has led to appearance of enormous information in the web world. It is not only individuals who are contributing to the generation of information which may be factually incorrect, many small-time internet-based magazines or news portals are also publishing news items whose legitimacy may be questioned. These news items are called 'fake news'. Fake news or junk news is a type of yellow journalism or propaganda that consists of deliberate disinformation or hoaxes spread via traditional print and broadcast news media or online social media [2]. The term is also at times used to cast doubt upon legitimate news from an opposing political standpoint, a tactic known as the lying press [3, 4]. The false information is often caused by reporters paying

sources for stories, an unethical practice called check-book journalism [2]. The news is then often reverberated as misinformation in social media, but occasionally finds its way to the mainstream media as well [5]. Fake news is written and published usually with the intent to mislead in order to damage an agency, entity, or person, and/or gain financially or politically, [6, 7, 8] often using sensational, dishonest, or outright fabricated headlines to increase readership.

These days fake news is creating different issues. These range from sarcastic articles to fabricated news and planted government propaganda in some outlets. Fake news and lack of trust in the media are growing problems with huge ramifications for our society. Obviously, a purposely misleading story is *fake news*, but lately blathering social media discourse is changing its definition. Some of them now use the term to dismiss facts counter to their preferred viewpoints [9]. The importance of disinformation within American political discourse was the subject of weighty attention, particularly following the American president election

---

Manuscript received October 20, 2018; revised November 2, 2018; accepted November 27, 2018. Date of publication January 6, 2019.

\* S. Pal is Senior Faculty, Data Science and Machine Learning with Manipal ProLearn, MAHE South Bangalore Campus (Manipal ProLearn), Electronic City, Bangalore, Karnataka, India - 560100. (e-mail: subhabaha.pal@manipalglobal.com)

\*\* T. K. S. Kumar is Senior Faculty, Data Science and Machine Learning Data Science with MAHE South Bangalore Campus (Manipal ProLearn), Electronic City, Bangalore, Karnataka, India - 560100. (e-mail: senthil.kumar@manipalglobal.com)

\*\*\* S. Pal is Consultant with Arminus Software Private Limited, HSR Layout, Bangalore, Karnataka, India - 560068. (e-mail: sampadas1508@gmail.com)

DOI:10.17010/ijcs/2019/v4/i1/142411

[9]. The term *fake news* became common parlance for the issue, particularly to describe factually incorrect and misleading articles published mostly for the purpose of making money through page views [9]. There are instances, specially, in the Indian context where spreading of 'fake news' through certain communication and Social Networking channels like Facebook and Whatsapp has created disturbances in the society. One well known example is the murder of one bright IT engineer (Google/Accenture employee) by some villagers enraged by the spread of a news in Whatsapp which was utterly false [10]. Several such incidents happened in India and some incidents have resulted in mass disturbances in several parts of India leading to huge loss of life and property.

One of the major reasons for the spread of fake news is the non-availability of proper censorship mechanism before spread of such news items. As internet and social media is accessible to everybody, and each individual has the full right to post anything on the social media, it is very difficult to control the same. Though it is not possible to permanently block the spread of fake news items, machine learning has shown some way to control its spread to some extent by empowering readers with some effective tools which can detect whether a news item is real or fake. The detection of fake news is a classic text analytics prediction problem involving NLP (Natural Language Processing) which needs advanced level of knowledge in application of Machine Learning algorithms.

This paper discusses the methods for detecting fake news from the innumerable news items and successfully flagging them so that readers can understand which is real news and which is propaganda ('fake news'). The current paper discusses different steps involved in building up a Machine Learning model which can be used to segregate (classify) news articles into real or fake news. Detecting fake news involves application of machine learning classification techniques after transforming the textual news items into numerical vectors through different NLP techniques which include text analytics.

## II. TEXT ANALYTICS

News items contain text data and hence, these cannot be processed by machine learning algorithms in the original form. Textual data is unstructured data. The data which is available in databases and which can be readily processed by ML algorithms are called structured data. News items are unstructured data which need to be

transformed into structured form and then these can be used for further processing or analysis. The primary task for text mining or analytics is giving these texts (news items) somewhat structured form. Text mining is similar in nature to data mining. However, focus is on unstructured data, that is, text instead of structured form of data. The first step in text mining or text analytics is organizing and structuring the data. Organizing and structuring the data involves NLP (natural language processing). Text mining includes categorizing the data, clustering, tagging text, summarizing database, creating taxonomies, and extracting information about things like word frequencies, and relationship between data entities.

The text mining process includes:

1. Identification and retrieval of relevant sets of text data to analyze.
2. Running algorithms to categorize, summarize, and organize the data.
3. Use of analytical methods to identify concepts, patterns, and other attributes.
4. Application of the findings for reporting such as sentiment analysis.
5. Preparation of data visualizations and dashboards to share the results.
6. Classifying texts as per different tags in case of classification problem (as in fake news detection).

Text Mining or Text Analytics related tasks can be classified as Descriptive Text Analytics and Predictive Text Analytics. Descriptive text analytics includes getting higher level descriptions of text data like word count, terms, phrases, parts of speech, meaningful associations, and also reporting like word cloud and frequency table. Predictive text analytics includes text categorization (classification), information extraction (predicting chunk boundaries), summarization of text, work prediction, search engine building, and emoji prediction.

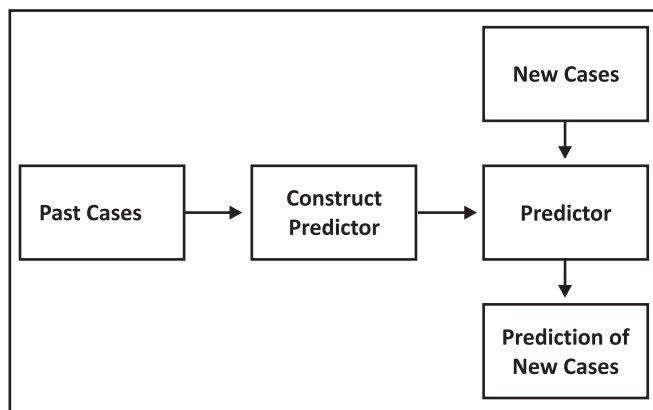


Fig. 1. Predictive Text Analytics

Predictive text analytics includes the scenario shown in fig. 1.

Steps in data mining include (1) Data collection (2) Creating corpus and (3) Data processing.

(1) Data Collection: The data is generally collected through web-scraping, using social media APIs for Twitter and Facebook, conducting surveys and databases. In case of fake news generation, the news items should be downloaded from the web using relevant APIs or if already existing database of news is there, a portion of the dataset should be labeled as 'Fake' or 'Real' as training set for building the model. General data formats are text, csv, html, xml, or pdf.

(2) Creating Corpus: Corpus is database of documents. In this case, it will be database of news items. A corpus is the set which contains all labeled news items for further processing. Table I shows two news items in a corpus.

**TABLE I.  
NEWS ITEMS**

News_no	Title	Text	Label
8476	You Can Smell Hillary's Fear	Daniel Greenfield, a Shillman Journalism Fellow at the Freedom Center, is a New York writer focusing on radical Islam. In the final stretch of the election, Hillary Rodham Clinton has gone to war with the FBI. (Continued .....)	FAKE
10294	Watch The Exact Moment Paul Ryan Committed Political Suicide At A Trump Rally (VIDEO)	There are two fundamental truths in this world: Paul Ryan desperately wants to be president and Paul Ryan will never be president. Today proved it. In a particularly staggering example of political cowardice, Paul Ryan reversed course and announced that he was back on the Trump Train after all. (Continued .....)	FAKE

The two labeled news items shown partially in Table I were collected from a huge dataset of more than 6500 news items. Interested readers can find and download the whole dataset from [11].

### (1) Data Processing

Steps in data processing include:

❖ Splitting / Tokenization: It involves converting the sentence into words. In the fake news detection scenario, it involves converting (breaking) the whole news items into words.

Example – 'Daniel Greenfield, a Shillman Journalism Fellow at the Freedom Center, is a New York writer focusing on radical Islam.'

After Tokenization: Daniel, Greenfield, a, Shillman, Journalism, Fellow, at, the, Freedom, Center, is, a, New, York, write, focusing, on, radical, Islam.

❖ Identification and Replacement of Synonymous words, Abbreviations, Types, and Standardization of cases (converting to lower case).

Example: Daniel, greenfield, a, shillman, journalism, fellow, at, the, freedom, center, is, a, new, york, write, focusing, on, radical, Islam.

❖ Removal of duplicates, special characters, and stop words/commonly used words: Stop words are the words which can be avoided (like the words 'a', 'the', 'in', 'at' etc.). We can also define some words which may not be relevant for the study.

Example: Daniel, greenfield, shillman, journalism, fellow, freedom, new, York, write, focusing, radical, Islam.

❖ Making parts of speech uniform for all words (stemming and lemmatization) : It involves making parts of speech uniform for all words as shown in the following example: focusing : focus; studying : study; boys : boy ; writing : write etc.

Example: daniel, greenfield, shillman, journalism, fellow, freedom, new, york, write, focus, radical, Islam.

❖ Conversion of Text Data into Structured Data: It involves bringing the text data into structured form. It involves the creation of Document Term Matrix using the frequency of each word available in each document. Example: Suppose there are two documents (news items). For understanding purpose, we are considering the first line of each news item mentioned above in the documents.

**TABLE II.  
NEW ITEMS**

News Items	Text
Doc1	Daniel Greenfield, a Shillman Journalism Fellow at the Freedom Center, is a New York writer focusing on radical Islam.
Doc2	There are two fundamental truths in this world: Paul Ryan desperately wants to be president.

The Document Term matrix for these two news items can be shown as follows:

**TABLE III.  
DOCUMENT TERM MATRIX**

	daniel	greenfield	shillman	journalism	fellow	freedom	new	york	write	focus	Radical	islam	fundamental	truth	world	paul	ryan	desperate	want	president
Doc1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
Doc2	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1

This document matrix which contains the count is the input independent X-matrix for the fake news detection for machine learning algorithm. Generally, for a full-length news item there may be 1000+ words (which means 1000+ columns) and when several news items are considered, 50,000 columns are available in the input matrix. Each row will be one news item. Hence, we are giving a structured form to the unstructured text data and we can use it for running machine learning algorithm. Hence, the final matrix looks somewhat like table IV.

**TABLE IV.  
DOCUMENT TERM MATRIX**

X-Matrix																	Y-Matrix (Label)		
1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	FAKE (1)		
0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	FAKE (1)		

We are supposed to run machine learning algorithm on this type of combination when number of rows are more than 6500 and number of columns are more than 50000 as more words are there in all the news items combined.

Each word like 'fundamental', 'truth' is called unigram. Sometimes, the words 'new', 'york' don't have meaning individually and they are combined to make a meaningful entity (phrase) 'new york'. This is called a bi-gram. Similarly, it is possible to create N-grams.

There are several ways to create X-Matrix. We used count vectorization method to create the X-Matrix which has been mentioned earlier. In the count vectorization method, the word-wise count of each document is updated to form the X-matrix. This X-matrix is used as the combination of independent variables for building the machine learning model where Y-matrix is the vector which contains the details of each document, whether it is 'fake' or 'real'.

There is another way to build matrix using Term Frequency and Inverse Document Frequency score. The approach is as follows :

**Term Frequency(TF)** is the number of times each word appears in each document.

TF = (Number of times the word occurred/ Total number of words in each document)

**Inverse Document Frequency (IDF)** is used to calculate the weight of rare words across all documents in the corpus. The words that occur rarely in the corpus have a high IDF score.

IDF = Log of N, where N is (Number of documents in question/Number of documents that contain a particular word)

**TABLE V.  
TERM FREQUENCY (TF) CALCULATION**

	daniel	greenfield	shillman	journalism	Fellow	freedom	new	york	write	focus	radical	islam	fundamental	truth	world	paul	ryan	desperate	want	president
Doc1	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	1/12	0/12	0/12	0/12	0/12	0/12	0/12	0/12	0/12
Doc2	0/8	0/8	0/8	0/8	0/8	0/8	0/8	0/8	0/8	0/8	0/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

**Table VI.  
INVERSE DOCUMENT FREQUENCY (IDF) CALCULATION**

	daniel	greenfield	shillman	journalism	Fellow	freedom	new	york	write	focus	radical	islam	fundamental	truth	world	paul	ryan	desperate	want	president
IDF score	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)	Log (2)

**TABLE VII.**  
**TF-IDF MATRIX**

	daniel	greenfield	shillman	journalism	fellow	freedom	new	york	write	focus	Radical	islam	fundamental	truth	world	paul	ryan	desperate	want	president
Doc1	0.02508583	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0	0	0	0	0	0	0	0
Doc2	0	0	0	0	0	0	0	0	0	0	0	0	0.03762875	0.03763	0.03763	0.03763	0.03763	0.03763	0.03763	0.03763

TF-IDF Score is calculated as TF (Document and word-wise) \* IDF (Word-wise). Hence, the X-matrix will be as follows for these 2 documents (news items containing only 1 line) :

Count vectorizer or TFIDF vectorizer, developing a model based on train dataset using any ML algorithm, fitting the model in the test dataset, and finding the accuracy. The Python code is available in [12].

Doc1	0.02508583	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0.02509	0	0	0	0	0	0	0	0
Doc2	0	0	0	0	0	0	0	0	0	0	0	0	0.03762875	0.03763	0.03763	0.03763	0.03763	0.03763	0.03763	0.03763

In real-life case, we can use X-matrix based on Count or X-matrix based on TD-IDF matrix for developing machine learning model.

The ML algorithms passive aggressive and MLP classifier show most accuracy in detection of fake news from the test dataset.

### III. MACHINE LEARNING MODEL DEVELOPMENT

We can develop machine learning models involving the X-Matrix (developed either using Count Vectorization method or TF-IDF vectorization method) and Y-Matrix using several machine learning algorithms. The machine learning algorithms which can be used include:

- ❖ Decision Tree
- ❖ Random Forest
- ❖ K-NN algorithm
- ❖ Naïve Bayes Multinomial algorithm
- ❖ Support Vector Machine
- ❖ Logistic
- ❖ XGBoost
- ❖ Passive Aggressive Classifier
- ❖ Multiple Layers Perceptron Classifier (Artificial Neural Network – Feed Forward Model)

The models can be developed involving the X-Matrix and the Y-Matrix using different machine learning models.

The process of developing a model and getting accuracy of the model has been demonstrated in the attached Python code. Interested readers can go through the Python code to get a glimpse of the same. The Python code discusses in detail importing of the data, bifurcating the data in train and test set, creating X-matrix using

### IV. CONCLUSION

Spread of fake news has become a perennial problem in recent times and this has become a source of discomfort in the society in many instances. Application of Machine Learning in detection of fake news is a new area which shows immense opportunities for practical use and further research. The detection of fake news using machine learning algorithms have been discussed in detail in this paper. News items are unstructured data. This paper shows how unstructured news items data can be turned into structured form using count vectorization and TFIDF vectorization techniques. This paper also shows how the machine learning models can be developed using the newly structured data.

### REFERENCES

- [1] The Economist, "The world's most valuable resource is no longer oil, but data," 2019. [Online]. Available : <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
- [2] Wikipedia, "Fake News," [Online]. Available: [https://en.wikipedia.org/wiki/Fake\\_news](https://en.wikipedia.org/wiki/Fake_news).
- [3] P. Mihailidis, and S. Viotty, "Spreadable spectacle in digital culture: Civic expression, fake news, and the role

- of media literacies in "Post-Fact" society," *Amer. Behavioral Scientist*, vol. 61, no. 4, pp. 441–454, 2017. doi:10.1177/0002764217701217.
- [4] J. Habgood-Coote, "Stop talking about fake news!", *Inquiry*, 2018, pp. 1–33. doi:10.1080/0020174x.2018.1508363
- [5] M. Himma-Kadakas, "Alternative facts and fake news entering journalistic content production cycle," *Cosmopolitan Civil Societies: An Interdisciplinary J.*, vol. 9, no. 2, pp. 25–41, 2017. doi:http://dx.doi.org.10.5130/ccs.v9i2.5469.
- [6] E. Hunt, "What is fake news? How to spot it and what you can do to stop it," December 17, 2016. *The Guardian*. [Online]. Available: <https://www.theguardian.com/media/2016/dec/18/what-is-fake-news-pizzagate> Accessed on :January 15, 2017.
- [7] R. Schlesinger, "Fake News in Reality," U.S. News, 2017. [Online]. Available: <https://www.usnews.com/opinion/thomas-jefferson-street/articles/2017-04-14/what-is-fake-news-maybe-not-what-you-think>
- [8] "The Real Story of 'Fake News': The term seems to have emerged around the end of the 19th century," Merriam-Webster. [Online]. Available: <https://www.merriam-webster.com/words-at-play/the-real-story-of0-fake-news>. Accessed on:October 13, 2017.
- [9] Pantech Solutions, "Fake news detection using Mach. Learning," [Online]. Available: <https://www.pantechsolutions.net/fake-news-detection-using-machine-learning>. Accessed on:January 19, 2019
- [10] S. Ganjoo, "WhatsApp rumours lead to death again? Google says lynched engineer was working with Accenture," January 16, 2018. [Online]. Available:<https://www.indiatoday.in/technology/feature/s/story/whatsapp-rumours-leads-to-death-again-google-says-lynched-engineer-was-working-with-accenture-1287102-2018-07-16>. Accessed on January 19, 2019.
- [11] <https://drive.google.com/file/d/13BokgGTW04x6wCxqb4m7sKF-VvKtitJ-/view?usp=sharing>
- [12] <https://drive.google.com/file/d/18uR-VZN-6YAR2pumNt8PCcW8-RrTAekZ/view?usp=sharing>

## About the Authors



**Dr. S. Pal** is Senior Faculty (Data Science and Machine Learning) with Manipal ProLearn. MAHE South Bangalore Campus (Manipal ProLearn) Electronic City, Bangalore, Karnataka - 560100



**T. K. Senthil Kumar** is Senior Faculty (Data Science and Machine Learning) with Data Science MAHE South Bangalore Campus (Manipal ProLearn).



**Sampa Pal** is Consultant Arminus Software Private Limited, HSR Layout, Bangalore, Karnataka.