International Journal of Engineering and Applied Physics (IJEAP)
Vol. 1, No. 2, May 2021, pp. 205~209
ISSN: 2737-8071

Detection of Fake News Using Machine Learning

Mayank Singh Rawat[1], Aviral Srivastava[2], Shubh Aggarwal[3]
[1,2,3]Galgotias University, Noida, Uttar Pradesh

Article Info

Article history:

Received Nov 9, 2020
Revised Apr 01, 2021
Accepted Apr 30, 2021

Keywords:

News
Fake news detection
Machine learning
Passive aggressive classifier

ABSTRACT

For some past recent years, largely since people started obtaining quick access to social media, fake news has become a serious downside and are spreading a lot of and quicker than the true news. As incontestable by the widespread effects of the big onset of fake news, humans are incapable of detecting whether the news is genuine or fake. With this, efforts have been made to research the method of fake news detection. The most popular and well-liked of such efforts is "blacklists" of sources and authors that do not seem to be trustworthy. Whereas these tools area helpful, so as to form a more complete end to end resolution, we also account for tougher cases wherever reliable sources and authors unharnessed false news. The motive of this project is to form a tool for investigation the language patterns that characterize wrong and right news through machine learning. The results of this project represent the flexibility for machine learning to be helpful during this task. We have made a model that detects several instinctive indicators of right and wrong news.

Corresponding Author:

Mayank Singh Rawat,
Galgotias University, Noida,
Uttar Pradesh. India.
Email: mayank.singh.rawat97@gmail.com

1. INTRODUCTION

Fake news has become threat towards journalism and public discourse which has spread hoaxes and misinformation. With the help of famous and widely used social media platforms like Facebook. It has become easier to propagate any data to lots of people in split minute time duration. The propagation of data is proportional to growth of social media and electronic messaging applications; there has been associate degree aggravation within the impact of those news articles. Now days it has become plenty easier to mislead people by posting fake news on twitter etc.

The cause of this issue lays in the fact that among these social media sites none of these uses any automatic process that can identify the truthfulness of news being shared across these platforms.

In the present time, significant amount of study have done in this area with adequate output. With the profit and growth of AIML, methods have influence many people for hard work in this field of research

2. FAKE NEWS DETECTION STRATEGIES

Knowledge primarily based Detection:

The main motive of this is to use the outside process to correct the claims created within the news area. Two hard origins are clearly opening on the internet or information graph. unlocked internet origins is distinguished between the two parties in terms of importance [1][2], but information graph is employed to examine whether or not the claims may be reason from existing facts in graph or not [3]. Several quick checking sites are exploitation domain specialists to see manually the news truthfulness. Drag relating the technique are machine-driven fact, seeing that this is related to category of words into literal, meaningless true and correct worthy true sentences.

Style primarily based Detection:

Vogue primarily observations pivot the means that the sentence has conferred by the people. False news is mostly not provided by the reporter, same fashion of copying would possibly disagree [4]. Within the author has enforced deep syntax models exploitation PCFG to remodel group of words into protocols like adding words production protocol and forebear protocol that elaborates syntax model of fraud observation. Next page enforced thick web model - Convolution nervous web (CNN) to examine truthfulness reports. The above stated kind of technique is termed as hyper partisan designs. Linguistic primarily ruled options may be applied to examine this type of style. there's merely ample info to get the readers eager to wander to definite webpage or clip or link. Such kind of attention-lurking showman or internet URL is known as click bait URL that might result in a supply of made-up chaos [5].

Visual primarily sourced Detection on Social Networking Platforms:

Virtually morphed pictures may be omnipotent currently on the various social networking platforms, sort of the holocaust. Photoshop is well known to be very simply performed broadly to pacify pictures satisfactorily enough to deceive people into cosmologically believing that they are viewing a $64000 image. The aura of multimedia system forensics has gone a space route in making a determinant variation of plans of action for denial of state detection in videos [6] and pictures but, to rename various cults on why these plans are not topologically used to resolve the various pictures of social networking platforms. Among the many rudimentary formations on internet for the common non-tech savvy users to recognize the morphed, edited or simply put photoshopped pictures. To put it out there, Google's reverse image search and Get image data [7] have passed and fulfilled visual, optical, and applied mathematics primarily sourced options that may be used in acting the credibleness of the multimedia system.

3. METHODOLOGY AND ALGORITHMS USED FOR CLASSIFICATION

The implementation of these algorithms was done by using Sci-Kit Learn which is a python library.

IDF-TF vectors:

TF-IDF vectors represent relative significance of a term in the record and the whole work.

TF (TERM FREQUENCY):

Term Frequency is defined as how many times a word repeats in the file. A greater value denotes that the term repeats more frequent than other words, so the document is a fair counterpart when the word is fragment of the searched word [8].

$$TF(t,d) = \frac{\text{Number of times } t \text{ occurs in a document } 'd'}{\text{Total word count of document } 'd'}$$

Figure 1. TF (Term Frequency) Formula

IDF (INVERSE DOCUMENT FREQUENCY)

Words which repeat more often in document and also repeats more times in many different documents might be not to the point. IDF measures how notable a term is in the entire script [8].

$$IDF(t) = \log_e\left(\frac{\text{Total number of documents}}{\text{Number of documents with term } t \text{ in it}}\right)$$

Figure 2. IDF (Inverse Document Frequency) formula

IDF-TF (INVERSE DOCUMENT FREQUENCY -TERM FREQUENCY):

TF-IDF functions by discipline the very often repeating words by allocating them low weight age and assigning high weight age to terms that are there in a genuine subset of a script, and that has greater repetition in certain script. This is equal to Inverse Document Frequency multiplied by Term Frequency [8].

$$\boldsymbol{TFIDF(t,d) = TF(t,d) * IDF(t)}$$

Figure. 3: IDF-TF (Inverse Document Frequency -Term Frequency) formula

PASSIVE AGGRESSIVE CLASSIFIER: A Passive Aggressive classifier algorithm is an on-line algorithm, which is used to categorize huge collection of information (e.g. twitter for micro blogging, facebook). It is pretty handy to use and works very quick. It functions by picking any sample or example, getting knowledge from it and after that moving to next [9].

4. EXPERIMENT RESULT AND ANALYSIS

We performed experiments with the help of Vector features which are algorithms mentioned above. Accuracy was observed. We used text categorization on the report body in the datasets which we will use.

The Dataset:

We will be using dataset for this project which is named as news.csv [10]. The dimension of this dataset is 7796×4. The very first column of the dataset contains news id, the second and third column contains title and text, and the fourth column consists of information about the news whether it is real or fake.

Procedure for detecting fake news:

Step 1)

Make necessary imports of the different libraries (numpy, pandas, itertools) as shown in figure 4.

```
import numpy as np
import pandas as pd
import itertools
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import PassiveAggressiveClassifier
from sklearn.metrics import accuracy_score, confusion_matrix
```

Figure 4. Importing Libraries

Step 2)

Then read the data set into a Dataset and get the first five records. Then get the labels (Fake or real) as represented in fig. 5 and fig. 6.

```
#Read the data
df=pd.read_csv('C:\\Users\\Mayank\\Desktop\\Fake_News_Detection\\news.csv')

#Get shape and head
df.shape
df.head()
```

	Unnamed: 0	title	text	label
0	8476	You Can Smell Hillary's Fear	Daniel Greenfield, a Shillman Journalism Fello...	FAKE
1	10294	Watch The Exact Moment Paul Ryan Committed Pol...	Google Pinterest Digg Linkedin Reddit Stumbleu...	FAKE
2	3608	Kerry to go to Paris in gesture of sympathy	U.S. Secretary of State John F. Kerry said Mon...	REAL
3	10142	Bernie supporters on Twitter erupt in anger ag...	— Kaydee King (@KaydeeKing) November 9, 2016 T...	FAKE
4	875	The Battle of New York: Why This Primary Matters	It's primary day in New York and front-runners...	REAL

Figure 5. Data from dataset

```
#Get the labels
labels=df.label
labels.head()

0    FAKE
1    FAKE
2    REAL
3    FAKE
4    REAL
Name: label, dtype: object
```

Figure 6. Labels

Step 3)

Now divide the dataset and form one set for training and another set for testing. Now, initiate a Tfidf Vectorizer for word stop from English and an upper document with frequency of 7/10.

```
#Split the dataset
x_train,x_test,y_train,y_test=train_test_split(df['text'], labels, test_size=0.2, random_state=7)
```

```
#Initialize a TfidfVectorizer
tfidf_vectorizer=TfidfVectorizer(stop_words='english', max_df=0.7)
```

Figure 7. Splitting the data set and initializing the tfidv vectorizer

Before processing natural language, the word stop is the most common repetitive word that is to be removed out. A group of raw documents is converted into matrix of TF-IDF features using TfidfVectorizer Now, the vectorizer in the training set will be fitted and converted. The test set vectorizer will be transformed.

Step 4)

Next, we will initiate a PassiveAggressiveClassifier. We are going to put this on tfidf train and y train.

```
#Initialize a PassiveAggressiveClassifier
pac=PassiveAggressiveClassifier(max_iter=50)
pac.fit(tfidf_train,y_train)
```

Figure 8. Initializing Passive aggressive classifier

Step 5)

Then, we will make guess on the test set from the Tfidf Vectorizer and using accuracy score() from sklearn.metrics we will find the accuracy.

```
#Predict on the test set and calculate accuracy
y_pred=pac.predict(tfidf_test)
score=accuracy_score(y_test,y_pred)
print(f'Accuracy: {round(score*100,2)}%')
```

```
Accuracy: 92.82%
```

Figure 9. Finding the accuracy

The result we got is an accuracy of 92.82% with this system. In the end, we will take out a confusion matrix to get an idea about the number of false and true negatives and positives.

```
#Build confusion matrix
confusion_matrix(y_test,y_pred, labels=['FAKE','REAL'])
```

```
array([[591,  47],
       [ 44, 585]], dtype=int64)
```

Figure 10. Confusion matrix

And with this system, we resulting values are 591 true positives, 585 true negatives, 44 false positives, and 47 false negatives.

5. CONCLUSION

Detecting fake news using machine learning in python. We used a straightforward dataset, used Tfidf Vectorizer, and initialized a Passive Aggressive Classifier in the project. We finished up by getting an accuracy of 92.82%. If we add on more information in the dataset, it will then analyze the stability of the system therefore putting more faith of people on this fake news detection system. Moreover, collecting real data that somewhat looks like fake data will improve the learning of the system. Additional semantic based traits can be used and put in on reaction to find out the news accuracy.

Social media is rich source of news and has a huge part in the information affirmation procedure, moreover if the information is fresh and is brought out in various news only then, Details from social media cannot be fetched. The changes from conventional media to social and online media and rapid roll out of the news, confirms the clampdown. Hence, by going through more social media characteristic in this project, aggregating them we can make successful and trustworthy system for identifying false news.

REFRENCES

[1] Michele Banko, Michael J Cafarella, Stephen Soder-land, Matthew Broadhead, and Oren Etzioni. Open information extraction from the web. InIJCAI'07.

[2] Amr Magdy and NayerWanas. Web-based statistical fact checking of textual documents. In Proceedings of the 2nd international workshop on Search and mining user-generated contents, pages 103{110. ACM,2010.

[3] Giovanni Luca Ciampaglia, Prashant Shiralkar, Luis M Rocha, Johan Bollen, Filippo Menczer, andAlessandroFlammini. Computational fact checking from knowledge networks. PloS one,10(6):e0128193,2015.

[4] https://www.cjr.org/analysis/facebook-rohingya-myanmar-fake-news.php

[5] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, Huan Liu. "Fake News Detection on Social Media", ACM SIGKDD Explorations Newsletter,2017

[6] Local tampering detection in video sequences Paolo Bestagini, Simone Milani, Marco Tagliasacchi, StefanoTubaro

[7] Andrew Ward, L Ross, E Reed, E Turiel, and T Brown. Naive realism in everyday life: Implications for social conict and misunderstanding. Values and knowledge, pages 103{135,1997.

[8] Wikipedia tf-idf https://en.wikipedia.org/wiki/Tf%E2%80%93idf

[9] Passive aggressive algorithm https://www.geeksforgeeks.org/passive-aggressive-classifiers/

[10] "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection, William Yang Wang.