International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2020 IJSRCSEIT | Volume 6 | Issue 3 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT2063219

Fake News Detection on Social Media

Prof. Dipti Chaudhari, Krina Rana, Radhika Tannu, Snehal Yadav
Computer Engineering, D. Y. Patil Institute of Technology, Pune, Maharashtra, India

ABSTRACT

Most of the smart phone users prefer to read the news via social media over internet. The news websites are publishing the news and provide the source of authentication. The question is how to authenticate the news and the articles which are circulated among the social media like WhatsApp groups, Facebook Pages, Twitter and other micro blogs and social networking sites. It can be considered that social media has replaced the traditional media and become one of the main platforms for spreading news. News on social media trends to travel faster and easier than traditional news sources due to the internet accessibility and convenience. It is harmful for the society to believe on the rumors and pretend to be a news. The basic need of an hour is to stop the rumors especially in the developing countries like India, and focus on the correct, authenticated news articles. This paper demonstrates a model and methodology for fake news detection. With the help of Machine Learning, we tried to aggregate the news and later determine whether the news is real or fake using Support Vector Machine. Even we have presented the mechanism to identify the significant Tweet's attribute and application architecture to systematically automate the classification of the online news.

Keywords : social media, authentication

I. INTRODUCTION

Traditionally people used to get news from trusted sources that were required to keep up strict codes of practice. However, the presence of internet has encouraged an advanced method to distribute, share data and news with almost no guideline or article benchmark.

The huge amounts of information are generated on the social network with various social media's formats. When some event has occurred, people discuss it on the web through the social network. They search or retrieve and discuss the news events as the routine of daily life. Without even getting the idea of whether the provided news on the social media is reliable or not they keep on spreading it by sharing it on various social media platforms. Unreliable sources of information expose people to a dose of fake news, hoaxes, rumors, conspiracy theories and misleading news.

The huge spread of false news brings negative impact to individual and society. The fake news comes from the misinformation, misunderstanding or the unbelievable contents which the creditability sources. These are difficult to detect whether to believe or not when they receive the news information. A recent event (2020) in India was the misinformation related to the virus is in the social media messages related to home remedies that have not been verified, fake advisories and conspiracy theories. At least two people have been arrested for spreading fake news about the coronavirus pandemic. The Press Information Bureau brought out a fact check on 24th

March that the stories about a financial emergency being imposed in India are fake.

In this paper we present the method to detect fake news. Our model classifies unreliable news into real or fake news after computing a score and will be able to distinguish between real and fake news based on various parameters obtained from text entered, Uniform Resource Locator (URL) and the tweet's Id obtained from the specific tweet. The proposed model uses Machine Learning techniques and Support Vector Machine (SVM).

The rest of the paper is organized as follows: Section II outlines the related works on fake news detection and Support Vector Machine (SVM) are briefly discussed. Section III presents the proposed approach where the methodology used in the proposed model is discussed. Section IV concludes the paper.

II. RESEARCH WORK

A. Fake News

The creditability of information can be defined by many different words such as trustworthiness, believability, reliability, accuracy, fairness, objectivity, and other with the same concepts and definitions. There are several researches use the machine learning approach to calculate the creditability of tweet's message, text entered to check etc.

Fake news is the content that makes people believe the falsification, sometimes it is the sensitive messages. When the messages were received, they will rapidly dispersed it to other. The dissemination of fake news in today's digital world has affected beyond a specific group. Mixing both believable and unbelievable information on social media has made the ×confusion of truth. That is the truth will be hardly classified. However, the appearance of fake news causes great threat on the safety of people's lives and property.

B. Support Vector Mechanism

SVM is a supervised machine learning algorithm that can be employed for both classification and regression purpose. It uses the hyperplane to split two classes data point with the maximum margin.

There are four evaluation results. They are precision, recall, F-measure and accuracy. They are used to compute four measures: True Positive, True Negative, False Positive, False Negative.

True Positive is the number of messages that is correctly classified by believable messages.

True Negative is the number of messages that is correctly classified by unbelievable messages.

False Positive is the number of messages that is incorrectly classified by believable messages.

False Negative is the number of messages that is incorrectly classified by unbelievable messages.

The precision, recall, F-measure, and accuracy are calculated by following equations:

$$Precision = \frac{\text{True Positive}}{(\text{Truve Positive} + \text{False Positive})}$$

$$Recall = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Negative})}$$

$$\text{F}-\text{measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$Accuracy = \frac{True\ Positive + True\ Negative}{(\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative})}$$

The data is collected from the database with selected topics. After that raw data is retrieved, we use the normalization rule to manipulate them. Next, the process of replication

data removing is being used. The last process of this work is the machine learning method for data classification that we use Support Vector Machine (SVM).

III. Methodology

A. Text Collection

The dataset was taken from Kaggle and has often considered to be the first step towards classification of fake news. It consists of almost 13000 posts over a period of 30 days. Research on this dataset using language processing tools has already been carried out by Kaggle users. The dataset was generated by Andrew Thompson to create document term matrices using the articles and analyze connections between articles using common political affiliations, medium or subject matter. It contains articles from top 15 American publications and the articles were mostly published between the years of 2016 and 2017. It consists of around 150000 articles that were collected by scraping news website homepages and RSS feeds. However, we will randomly select only 13000 articles from this dataset and merge it with the fake news dataset for more accurate predictions and for avoiding a skewed dataset.

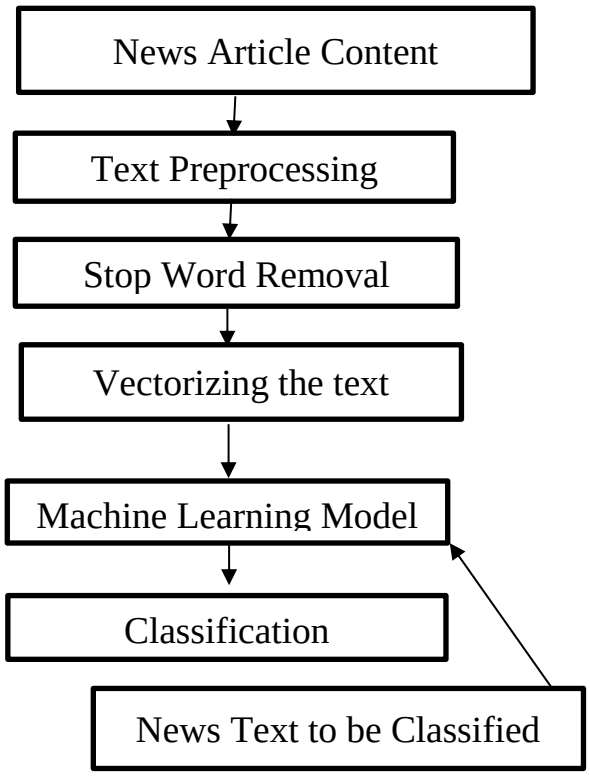

Fig 1 : Flow chart representing the methodology flow.

B. Text Pre-processing

Text needs to be converted into numbers before it is used with a machine learning algorithm. For classification of documents, documents are taken as input and a class label is generated as output by the predictive algorithm. The documents need to be converted into fixed-length vectors of numbers for the algorithm to take them as input. A process called tokenization, where certain words are removed after parsing the text is used. The input for the machine learning algorithm is the words encoded as integers or floating point values (feature extraction or vectorization) [7]. The Bag of Words model is a simple and effective model used for processing text documents in machine learning. The model discards all the order information in the words and only considers the occurrence of the words in the document. There are two different ways of implementing this model

- Count Vectorizer
- TF-IDF Vectorizer

The Count Vectorizer generates an encoded vector that contains the length of the entire vocabulary coupled with the frequency of each word by which it appears in the document. A popular alternative is to calculate word frequencies with a method called TF-IDF (Term Frequency Inverse Document). The TF - IDF score is then attached to each document. • Term Frequency: A summary for the frequency of each word in the document. (Eqn. 1) • Inverse Document Frequency: Reduces the importance of words that are repeated frequently across documents. (Eqn. 2)

$$tf(t,d) = 0.5 + 0.5 * \frac{f_{t,d}}{max_{\{t'od\}\ f_{t,d}}}$$

$$idf(t,d) = log\frac{N}{n_t}$$

$$tfidf(t,d,D) = \text{tf}(\text{t},\text{d}) * \text{idf}(\text{t},\text{D})$$

C. Classification

Classification could be achieved using five main machine learning algorithms - Decision Trees, Naive Bayes, Logistic Regression, Support Vector Machines and Neural Networks

IV. MODELING AND EVALUATION

A. Our Pipeline

After cleaning the data and generating features, we execute a 90/10 random test-train split on the dataset and feed it into a modeling pipeline. This pipeline iteratively fits models varying the tuning parameters with which they are executed up to 50 times, depending on the number of possible permutations for that model. These models are then tested on the 10% holdout data to understand their performance.

B. Baseline Models for Comparison

As a baseline comparison for understanding the performance of our models, we look at two methods. First, a Naive Bayes model that predicts all majority class; in this case, all articles are from reliable news sources. Second, a model that randomly selects a classification for each article as either reliable or unreliable based on the posterior probability of that class in the training set. These are the Naïve, Random models and Support vector machine respectively.

C. Combining PCFG and TF-IDF bi-gram features

Combining both these features our model performs well.

We note that our best models tended to be Stochastic Gradient Descent (SGD) models, which, given that they tend to perform well with sparse and highly dimensional data, is not surprising. In particular, SGDs far outperform on precision while retaining a high recall, meaning that these models would work well both as identification of high priority articles in addition as 'fake news' filters.

D. TF-IDF Bi-gram only model performance

Removing the PCFG features allows us to understand in more depth the value of those features in achieving these combined feature results. The results from this more limited feature run are displayed in Table IV. The removal of PCFGs improves most of the metrics across our models. This is surprising, indicating that the PCFG features add little predictive value to the models. Indeed, the only noticeable decrease in performance is in our recall figures for Decision Trees and SGDs.

E. PCFG model performance

The removal of TF-IDF bi-gram features allows us to isolate the predictive value of PCFGs for our application. The results are displayed in Table V. Surprisingly, all of our models give the same result. Diving into the individual predictions, we find that all models produce the same rank order of scores. We've switched from a 0.70 threshold for classification to a top-k of 0.05 because the distribution of scores for these models have a particularly low mean with a tight range, such that determining an appropriate threshold for categorization was tedious and the 0.70 results weren't illuminating. All this goes to indicate that in the case of this classification task, PCFGs do not add a strong source of information for classification on their own.

F. SVM model performance

The dataset which has been processed by the combination of both the TF-IDF and PCGF is being taken and a feature matrix is formed from the documents. This is then passed through a classifier. The classifier consists of different machine learning algorithms. The support vector machine performs well since there is an extremely high number of features in a text classification problem but generally requires a lot of tuning and is memory intensive.

Once we have labelled training data (supervised learning), the algorithm generates the best possible hyperplane which categorizes new data automatically. In a two-dimensional space this hyperplane would be a line dividing a plane in two parts with each class lying on either side.

G. Execution

The URL/Tweet's ID of the article is the input from which text is extracted. The extracted text is then passed through a data preprocessing unit consisting of processes like tokenization and the generation of a word cloud. The data is analyzed by using 2 methods that are Stance detection and document similarity. Then it is passed through a classifier for calculation of f-score. Here the fake news can be classified from a range of 1 to 10 where 1 is completely true and 10 is completely false.

V. RESULT AND ANALYSIS

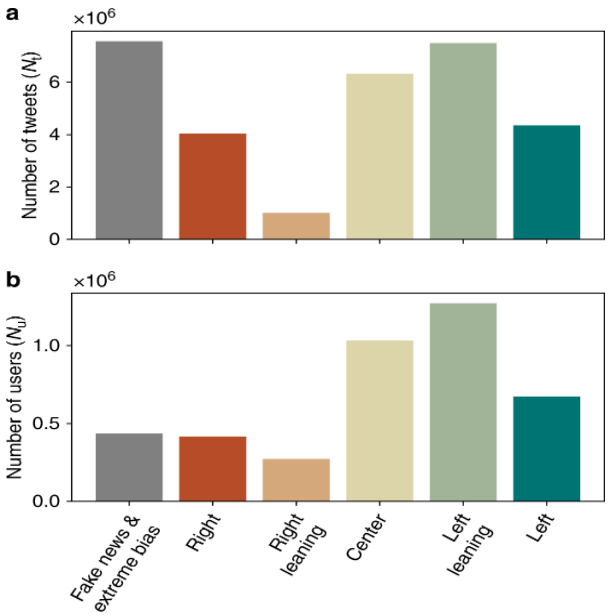

Fig 2. Classification of the news

The above Fig. shows importance of different types of news outlets in Twitter. Number of distinct tweets (**a**) and number of distinct users having sent tweets (**b**) with a URL pointing to a website belonging to one of following categories: fake or extremely biased, right, right leaning, center, left and left leaning news outlets. While the tweet volume of fake and extremely biased news is comparable to the tweet volumes of center and left volume (**a**), users posting fake and extremely biased news are around twice more active in average (see Table 1). Consequently, the share of users posting fake and extremely biased news (**b**) is smaller (12%) than the share of tweets directing toward fake and extremely biased news websites (25%)

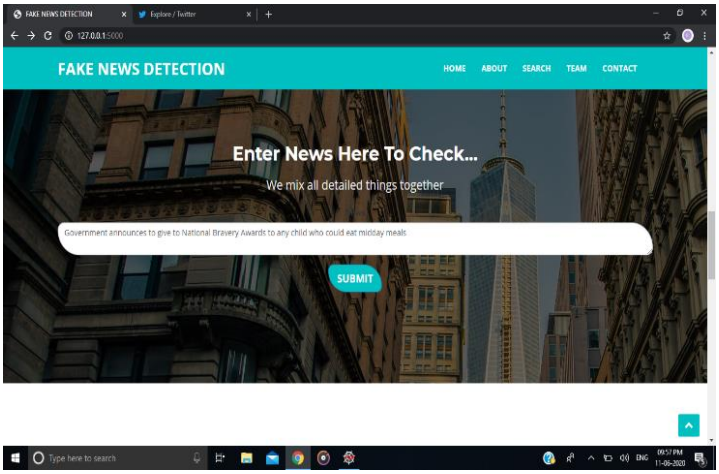

Fig 3. Insertion of the Fake News

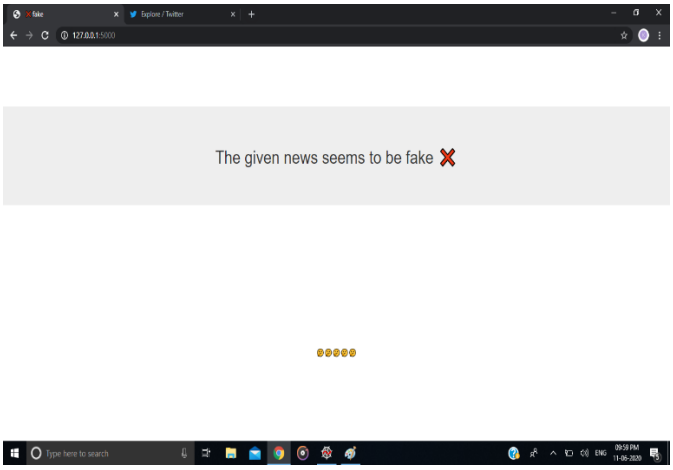

Fig 4. Result of the Fake News

VI. CONCLUSION AND FUTURE WORK

Machine Learning uses a statistical technique to give the computer the ability to learn with data hence it is

widely used in the detection of fake news. Methods used for taking parameters and for categorizing the type of news are also discussed. It was observed that the dataset is first preprocessed using preprocessing techniques such as stop word removal, tokenization and stemming. The techniques TF-IDF and probabilistic context-free grammar that are used to extract features are also identified. From the literature review it has been observed that the accuracy for predicting fake news in social media is much higher than any other online news media hence we have targeted online news media fake news detection along with website verification. In future work, our proposed model will be tested for fake news detection by using URL as an input which will not only validates headline but will also validate site behavior and other related parameters.

VII.REFERENCES

[1]. Gottfried, J. & Shearer, E. (2016, May 26). "News use across social media platforms 2016." Retrieved from http://www.journalism.org/2016/05/26/news-use-across-social-mediaplatforms-2016/

[2]. Allcott, H. and Gentzkow, M. (2018). Social Media and Fake News in the 2016 Election. [online] NBER. Available at: http://www.nber.org/papers/w23089.

[3]. Silverman, C. (2016, November 17). "This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook." Retrieved from https://www.buzzfeed.com/craigsilverman/viral fake-election-news-outperformed-real-news-on-facebook

[4]. Shu, K., Sliva, A., Wang, S., Tang, J. and Liu, H. (2018). Fake News Detection on Social Media: A Data Mining Perspective. [online]

[5]. Arxiv.org. Available at: https://arxiv.org/abs/1708.01967 [Accessed 16 Mar. 2018].

[6]. Bajaj, S. "Fake News Detection Using Deep Learning." Available at: https://web.stanford.edu/class/cs224n/reports/2710385.pdf

[7]. Singh, V., Dasgupta, R., Sonagra, D., Raman, K. & Ghosh I. Automated Fake News Detection Using Linguistic Analysis and Machine Learning.

Cite this article as :

Prof. Dipti Chaudhari, Krina Rana, Radhika Tannu, Snehal Yadav, "Fake News Detection on Social Media", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6, Issue 3, pp.1132-1137, May-June-2020. Available at doi : https://doi.org/10.32628/CSEIT2063219
Journal URL : http://ijsrcseit.com/CSEIT2063219