

Project : Fake News Detection using NLP

ABSTRACT :-

Social media is one of the very powerful media in spreading information. People are interested in sharing without any proper checking of any sort of false information. Unstructured text data may be classified into meaningful categorical classifications using text classification, which is a typical study area in the discipline of Natural Language Processing (NLP). The main contribution of this article is to identify a finest framework to tackle the fake news problem with the NLP and Machine Learning techniques. In this empirical research, the fake news data is analysed with the different combinations of Vectorizers and Machine Learning Classifiers. From the experimental results on five benchmark datasets namely fake_real_news dataset extracted from Kaggle, COVID-19 Constrain, Politifact, ISOT and Gossipcop, it is observed that the fake news detection with the combination of TF-IDF Vectorizer and Passive-Aggressive Classifier outperforms the other existing methods.

Keywords: Machine Learning, Pre-processing, Vectorizer, Classifiers, Natural Language Processing

INTRODUCTION

Various difficulties are being created by fake news nowadays, from satirical pieces to fake stories and deliberate government propaganda in some publications. False news and public distrust in the media are developing issues in our society that have serious consequences. The term “fake news” refers to a tale that intentionally misleads the public, but in recent months, social media has begun to redefine the term.

The importance of disinformation during the time of COVID19 pandemic is a subject of weighty attention, particularly following the emergence of multiple variants of Coronavirus. Factually false and deceptive stories produced primarily for the goal of gaining money through page views became known as “fake news.” With the use of this research, we want to develop a model that can properly identify whether or not a piece of content is fake news.

WhatsApp in India has become one of the Major fake news providers. False news is also

suspected to propagate on WhatsApp because it is circulated between small groups of friends and families who commonly believe.

The consequences of fake news and the use of the word are introduced in the 21st century. The Web was introduced to the public in the 1990s and was to provide them with the ability to access information.

RELATED WORKS

Various researches have been undergone by many researchers to find out the best classifiers which gives a maximum accuracy in identifying the fake and real news. The excerpts from some of such researches were mentioned here in order to get a clear picture of the objective of this research article.

Classification approach was utilised by Gee et.al. [8] to recognise spam profiles based on spam reports from users. They've used the Twitter API to gather ordinary user profiles and spam profiles from the “@spam” Twitter handle, and they've expressed the data in JSON. Naive Bayes method with a 27 percent error rate and SVM algorithm with a 10 percent error rate were used to classify it in CSV format. Wang et.al. [9], studied the suspicious behaviours of spam accounts on Twitter. They applied machine learning methods to automatically differentiate spam accounts from 1normal ones. Data generated by spam or fake users is substantial, according to Dutse et al. [10]. One out of every 200 social media postings and one out of every 21 tweets are spam, according to estimates. They've come up with a new way to tell spam from non-spam social media messages, and

they've revealed more about the Twitter behaviour of the spammers.

According to Jia et al. [11], SVM beat both the rule-based classifier and the Decision Tree classifier when it came to spam identification on the Internet. Aldwairi et.al [12] suggested an easy but efficient strategy to allow users to install an easy tool in their private browser to detect and filter prospective Clickbaits. Depending on Precision, Recall, F-Measure and ROC, the classifiers are compared.

METHODOLOGY

Feature Extraction Vectorizers like TF-IDF and Count Vectorizer are used to do the word embeddings and dataset cleaning in this approach. It's done by removing stop words like "the," "when," and "there," as well as utilising an n-number of the most often used words, phrases, lowercased or not, as well as using keywords that appear at least a certain number of times in a certain text corpus. The dataset is then divided into two parts: data for testing and data for training. Classifiers like Passive Aggressive Classifier, Multinomial Nave Bayes, Random Forest, and Support Vector Machine are fed the train data after the dataset has been divided. The classifiers train the data and check the relativity of the features in the test data and gives the result as 'Fake' or 'Real' along with the accuracy of the Classifiers.

Vectorization

It is an NLP approach known as "word embeddings" or "word vectorization" that uses real numbers to map words and phrases from the dictionary to the corresponding vectors of real numbers, which may subsequently be used to create word predictions and meanings. It's a method for turning text into a graphical representation. Text can be vectorized using Count Vectorizer, TF-IDF Vectorizer, or Hashing Vectorizer, among others.

EXPERIMENTAL RESULTS AND ANALYSIS

Experimental Setup

Using Jupyter Lab and five benchmark datasets from Kaggle, COVID-19 Constrain, Politifact, ISOT, and Gossipcop, the empirical evaluation is carried out in Python. A Windows 10 PC with an Intel Core i5 - 9300H CPU running at 2.40GHz, 8GB of RAM, and an NVIDIA GEFORCE GTX 1650 graphics card was used for all testing. Table 4.1 provides a thorough breakdown of the five real-world datasets. The training and testing examples were chosen at random, with 67 and 33 percent of the total being taken from each group.