

ABSTRACT

The purpose of this Machine Learning model is to identify fake news or reports that has been circulating over the internet more accurately . The phenomenon of Fake news is experiencing a rapid and growing progress with the evolution of the means of communication and Social media. Fake news detection is an emerging research area which is gaining big interest. It faces however some challenges due to the limited resources such as datasets and processing and analyzing techniques. In this work, we propose a system for Fake news detection that uses machine learning techniques.

We have used term frequency-inverse document frequency. We also propose a dataset of fake and true news to train the proposed system. TF (Term Frequency): The number of times a word appears in a document is its Term Frequency. IDF (Inverse Document Frequency): Words that occur many times a document, but also occur many times in many others, may be irrelevant. IDF is a measure of how significant a term is in the entire corpus. CONCEPTS INVOLVED IN MACHINE LEARNING MODEL: Dataset, Importing the dataset into the machine learning model, TRAIN_TEST_SPLIT, Working of the model, TFIDF VECTORIZER, ACCURACY SCORE, Passive Aggressive Classifier, Accuracy score.

TFIDF VECTORIZER: Transforms text to feature vectors that can be used as input to estimator. vocabulary_ Is a dictionary that converts each token (word) to feature index in the matrix, each unique token gets a feature index.. Through this Machine Learning model, we are able to detect the fake news and eliminate them and reduce the chances of being circulated.

REDUCING THE PROBABILITY OF PEOPLE PRONE TO FAKE NEWS:

Since detecting fake news using Machine Learning Algorithms has 88% DOCUMENTARY REPORT: Using our model, we can get a documentary report containing a table with columns title, text, and label where the label indicates whether the particular news or report is REAL or FAKE.