

Fake News Prediction project

Submitted by:

Sai Chander Reddy I

**INTRODUCTION**

* Business Problem Framing
* The proliferation of misleading information in everyday access media outlets such as social media feeds, news blogs, and online newspapers have made it challenging to identify trustworthy news sources, thus increasing the need for computational tools able to provide insights into the reliability of online content.
* Motivation for the Problem Undertaken

Our Primary motive of the problem is to predict the fake news and real news of each and every news so that the online news portals can be free of fake news by using our project.

**Analytical Problem Framing**

* Data Preprocessing Done

The Data Pre-processing steps are done on the object data[news] using Natural language processing . The Pre-process steps involve removing stop words, html tags, unnecessary tags and making the data clean and ready for vectorization . We had used TFIDF and Hashing Vectorization techniques for the news.

* Data Inputs- Logic- Output Relationships

The Data Input is only news column and the output is label which has two outputs ‘0’ and ‘1’. If the output is zero then the news is real and viceversa. The input news column is cleaned, vectorized Describe the relationship behind the data input, its format, the logic in between and the output. Describe how the input affects the output.

* Hardware and Software Requirements and Tools Used
* Software requirements: python programming language, pandas,nltk,seaborn,numpy packages are required
* Hardware requirements: ram:8gb,Rom:512gb,intel i5 or more processor.

**Model/s Development and Evaluation**

* Identification of possible problem-solving approaches (methods)

The Problem is solved using Natural Language processing, where the news are cleaned , transformed into vectors and used for training the model.

* Testing of Identified Approaches (Algorithms)

We used MultinomialNB and Passive Agressive algorithms.

* Key Metrics for success in solving problem under consideration

The Key Metrics used for the model are confusion matrix. Accuracy\_score and classification\_report .

**CONCLUSION**

* Key Findings and Conclusions of the Study

This project FakeNewsPrediction is used to detect the fake news developed using TFIDF Vectorization technique and trained using Passive Aggressive Classifier .