Parshvanath Charitable Trust's
A. P. SHAH INSTITUTE OF TECHNOLOGY, THANE
(All Programs Accredited by NBA)
Department of Information Technology

A Web Framework to Predict Fake News Using ML

Group No: 13

Jaagrut Shah 19204002
Jigar Desai 19204003
Yash Jain 19204013

Guided by:
Prof. Apeksha Mohite
Prof. Geetanjali Kalme

Contents

- Introduction
- Objectives
- Problem Definition
- Technological Stack
- Review Suggestions (Given in Last meeting)
- Proposed System Architecture/Working
- Prototype Design Demonstration
- Implementation Status
- Status of Paper Draft & Targeted Conference

1. Introduction

- The term "fake news" was relatively unknown and unpopular a few decades ago, but it has emerged as a massive monster in the digital era of social media.
- Fake news, information bubbles, news manipulation, and a lack of faith in the media are all issues that are becoming increasingly prevalent in our culture.
- The proposed approach will assist in determining if the news provided is true or fraudulent based on prediction.

2. Objectives

- To show the news relevancy and analysis to attain accuracy in anticipating real and dependable news.
- To work on this issue, a layered model is proposed, which fine-tunes the informational insight received from the data at each phase before attempting a prediction.
- To use a variety of Machine Learning approaches, achieve demonstrable success in the prediction of fake news and posts.
- To eliminate the propagation of false information on social media that may mislead users.
- To be able to give more and more accurate news on the screen.

3. Problem Definition

- Because of the simple availability and exponential growth of information available on social media networks, distinguishing between false and real information has become difficult.
- The ease with which information may be shared has aided in the exponential expansion of information deception.

4. Technological Stack

- Software Requirements:
  - Python libraries
  - ReactJS
  - Bootstrap
  - Framework: Flask
  - PostgreSQL and for monitoring: PG Admin
  - Machine-Learning Classification models
  - OS Requirements: Windows XP or above
  - JavaScript supported browsers

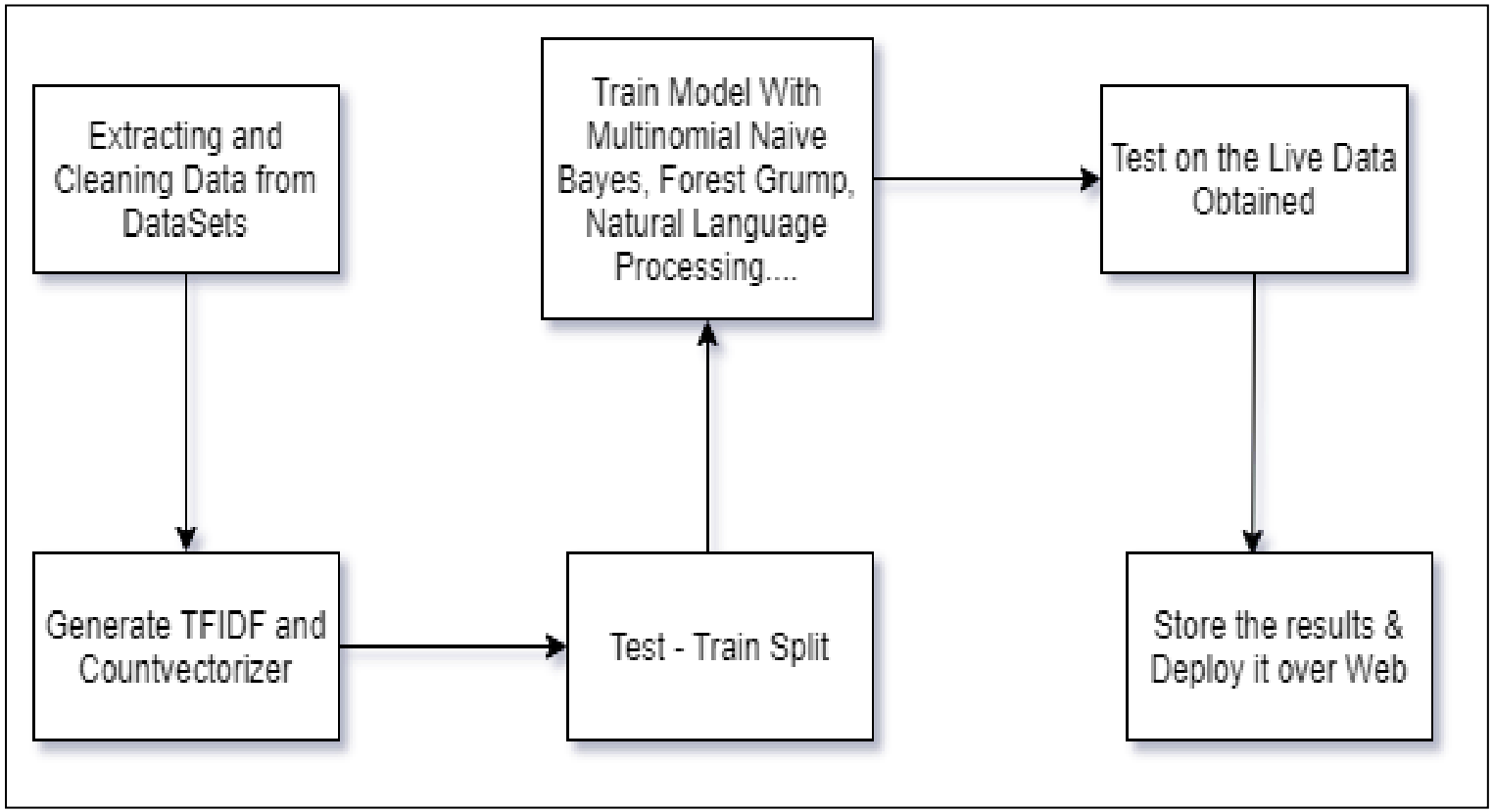
- Hardware Requirements:
  - Minimum 2GB Ram
  - Pentium / Intel i3 Processors or above

5. Review Suggestions

- Our product received positive feedback, although there were a few additions that are noted below:
  1. Including Sentiment Analysis in our project
  2. Adding a feedback mechanism for our website's visitors

- We intend to utilize natural language processing for sentiment analysis, and for the feedback system, we would record user information in a database together with the votes he offered to be fake or true. (This will be combined with machine prediction and sentiments with the news as well.)

6. Proposed System/Architecture/Working



7. Prototype Design Demonstration

- Ongoing

8. Implementation Status

- Our first goal was to obtain a standard data set from Kaggle. Which we began and finished cleaning the data gathered with operations such as:
  - NA values are Dropped.
  - Using Regular Expressions to Clean Text
  - Tokenization is the process of converting a text into words.
  - Stopwords are being investigated.
  - Lammetizer is applied to the words.
  - CountVectorizer, and TF-IDF are used to convert text into machine intelligible form.
- For our project, the data pre-processing processes were as follows. Furthermore, we are splitting the data into train and test sets in order to apply it into Machine Learning Algorithms.
- So far, we've practically completed our first objective, and after using ML models such as Logistic Regression, Decision Tree, SVM, Naive Bayes, and NLP, we'll come up with the model that provides the highest accuracy

9. Status of Paper Draft & Targeted Conference

- We've already begun preparing our paper by combining research from previously published IEEE papers with our project's objectives and concerns.
- Because our study involves data manipulation and machine learning, we chose the following conference:
  - International Conference On Big Data, Machine Learning and Applications (NATIONAL INSTITUTE OF TECHNOLOGY SILCHAR) Deadline: September 25th 2021
- The task at hand is to
  - Add some implementation and scenarios that came up while working on our first objective.
  - Diagrams and charts that could be used
  - Conclusion
  - And necessary changes by our Guides.

- We are ready to publish our paper after completing these tasks..

Thank You...!!