SYNOPSIS OF THE PROJECT for SE

Tentative Title of Project: Fake News Detection

Team Details:

Team No.: B4-10

Name: Ishita Jain **Er. No.:** 201B126
Email: 201B126@juetguna.in
Ph. Number: 8817153225 **Signatures**

Name: Ishu Agarwal **Er. No.:** 201B127
Email: 201B127@juetguna.in
Ph. Number: 8595293573 **Signatures**

Name: Jaivardhan Singh **Er. No.:** 201B128
Email: 201B128@juetguna.in
Ph. Number: 8009147802 **Signatures**

Problem Formulation:

Fake news occupies large sphere of cyber space today world-wide. These days fake news is creating different issues from sarcastic articles to a fabricated news.Fake news and lack of trust in the media are growing problems in our society with huge implications. This may also affect the importance of serious news media.

Key words-

Fake News, Natural Language Processing, Text Classification, Logistic Regression, Classification, Feature Extraction, Stopwords, Data Pre-Processing, Stemming, Vectorizer, TfidfVectorizer, Accuracy Score, Prediction.

Objectives/Aim-

The main objective is to detect the fake news, which is a classic text classification problem. It is needed to build a model that can differentiate between "Real" news and "Fake" news.

Description (not more than 100 words)-

The primary objective is to identify the authenticity of the news and categories it into fake and real news. In this project, Machine Learning model i.e Logistic Regression is leveraged to distinguish real and fake news and it has achieved an accuracy of approx. 97.13%. We are using Fake News dataset in the project. After pre-processing of data where pre-processing of data include three main steps i.e filling up of missing values, stemming and vectorizer, we use Logistic Regression to train the model and then find out the model's accuracy score on testing data and even on new data.

Technical Details:

Project starts with data collection, then data pre-processing, then take a pre-processed data and split it into training and testing data, later train this data using Logistic Regression model. After this we get a trained model then check accuracy score of our model on both training as well as testing data. At last we check our model prediction on new data so we feed new data into the trained model and check the prediction.

Details of Methodology/ Approach of Development-

1. Importing Libraries and Datasets
2. Data Preprocessing
3. Preprocessing and analysis of News column
4. Converting text into Vectors
5. Model training, Evaluation, and Prediction

Tools/ Languages to be used –

For this project we use jupyter notebook as an ide, language used is Python and also we use a Fake News dataset as a data.

Synopsis Status

Faculty Remark-

Approved / Approved with Changes / Not Approved

Signature - **Date:-**