**Fake NEWS Detection**

**Phase-1 Document Submission**

**Project:** Fake NEWS Detection

**Team Members:**

Logeshwaran M - 110521106012 [Team leader]

Karthick D - 110521106313

Gunasegkaran S V - 110521106311

Shalini R - 110521106024

**OVERVIEW:**

A fake news detection project is an endeavor aimed at developing a system or model to identify and classify news articles, stories, or information that is intentionally misleading, deceptive, or fabricated. The goal of such a project is to promote information accuracy, combat misinformation, and enhance media literacy.

**DESIGN THINKING:**

**Empathize :**

* Understand the needs and concerns of the target users, which may include journalists, fact-checkers, and the general public.
* Conduct interviews, surveys, and observations to gain insights into how people consume and interact with news
* Throughout the design thinking process, it's essential to remain flexible and open to new insights and ideas

**Define :**

* Clearly articulate the problem you are addressing. For example, "How might we effectively identify and combat fake news in online media?"
* Create user personas to represent the different types of users who will interact with your fake news detection system.

**Ideate :**

* Brainstorm creative solutions to the problem, considering various aspects such as data sources, algorithms, user interfaces, and educational components. ● Encourage a diverse team to generate a wide range of ideas.

**Prototype :**

* Create low-fidelity prototypes of your fake news detection system. These can be paper sketches, wireframes, or simple mock-ups.
* Develop a proof-of-concept model for the machine learning component to test its feasibility.

**Test :**

* Gather feedback from potential users by presenting them with your prototypes.
* Adjust your design based on user feedback and iterate on your ideas.
* Test the performance of your machine learning model using validation data to assess its accuracy and effectiveness.

**Develop :**

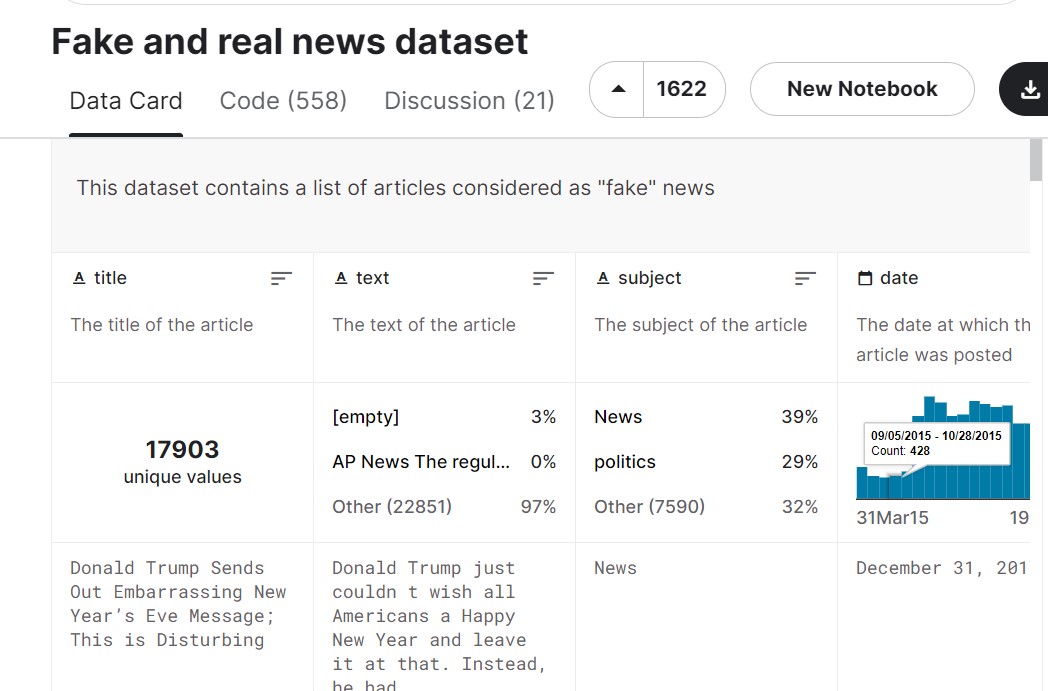
* Build the full-fledged fake news detection system, incorporating the refined design and insights from the testing phase.
* Implement a user-friendly interface and any educational components that help users understand the system's results.

●

**DataSet Link :**

[**https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset**](https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset) **Sample**

**Images :**



**Steps :**

**1.Data Collection :**

* Gather a diverse dataset of news articles, including both real and fake news examples. You may find such datasets through academic institutions, research organizations, or online repositories.
* Ensure that the dataset is labeled, indicating whether each article is real or fake.

**2.Data Preprocessing :**

* Clean and preprocess the text data by removing HTML tags, special characters, and irrelevant information.
* Tokenize the text into words or subword tokens.
* Perform tasks like stemming or lemmatization to reduce words to their base form.
* Split the dataset into training, validation, and test sets.

**3.Model Selection :**

* Choose appropriate machine learning or deep learning algorithms for classification, such as **Logistic Regression, Random Forest**, **Support Vector Machines**, or **Neural Networks**. ● Depending on the dataset size and complexity, select the most suitable algorithm.

**4.Model Training :**

* Train the selected model using the training data.
* Fine-tune hyperparameters through techniques like cross-validation.
* Evaluate the model on the validation set to monitor its performance during training

**5.Model Evaluation :**

* Assess the model's performance using various metrics such as accuracy, precision, recall, F1-score, and confusion matrix.
* Use the test dataset (not used during training) to get an unbiased estimate of model performance.

**Program :**

**Importing Necessary Modules :**

|  |
| --- |
| **import pandas as pdimport pandas as pd** |
| **import numpy as np** |
| **from sklearn.model\_selection import train\_test\_split** |
| **from sklearn.feature\_extraction.text import TfidfVectorizer** |
| **from sklearn.linear\_model import LogisticRegression from sklearn.metrics import** |
| **accuracy\_score,precision\_score,recall\_score,f1\_score,roc\_auc\_score** |
| **import pickle as pk** |

**Testing the Trained Models :**

|  |
| --- |
| **n=input("enter:")** |
| **a=[]** |
| **a.append(n)** |
| **mod=pk.load(open("fakenews.pkl","rb"))** |
| **news=vect.transform(a)** |
| **prediction=mod.predict(news)** |
| **if prediction==[0]:** |
| **print("it is fake news")** |
| **else:** |
| **print("it is real news")** |

**Accuracy :**

**Accuracy score: 0.988641425389755 precision score: 0.9856807511737089 recall score: 0.9903301886792453 f1\_score : 0.988**

**rou\_auc\_score: 0.9887304951835044**

**Conclusion :**

In conclusion, a fake news detection project is a multifaceted endeavor aimed at

combating the spread of misinformation and disinformation in the digital age. Such projects are critical for safeguarding the accuracy of information, promoting media literacy, and preserving the integrity of news sources.