□ Fake News Detection using Machine Learning

This project uses Natural Language Processing (NLP) and Machine Learning to detect whether a news article is real or fake. It includes a web-based interface built with **Streamlit**, allowing users to easily input news content and get predictions in real-time.

□ Project Goal

To build an intelligent system that can automatically classify news articles as real or fake using NLP techniques and a machine learning model.

✅Features

- Classifies news articles as **Real** ✅or **Fake** ❌
- Clean and simple **Streamlit UI**
- Real-time prediction based on user input
- Trained using **TF-IDF** and **Logistic Regression**
- End-to-end pipeline: data preprocessing → model training → deployment

□ Technologies Used

- **Python**
- **scikit-learn**
- **pandas, numpy**
- **NLTK**
- **TF-IDF Vectorizer**
- **Logistic Regression**
- **Streamlit** for UI

□ Folder Structure

fake-news-detection/

├── app.py # Streamlit app

├── train_model.py # Model training script

├── utils.py # Preprocessing functions

├── fake_news_model.pkl # Trained ML model

├── vectorizer.pkl # TF-IDF vectorizer

├── dataset/ │ └── news.csv # Combined dataset

├── requirements.txt # Python dependencies

└── README.md # Project overview

How to Run the Project

1. **Install dependencies** ```bash pip install -r requirements.txt
2. **Train the model**

```
bash
CopyEdit
python train_model.py
```

3. **Run the Streamlit app**

```
bash
CopyEdit
streamlit run app.py
```

Dataset

This project uses a merged dataset of real and fake news articles, labeled for binary classification:

- `label = 0` → Fake
- `label = 1` → Real

You can replace the sample dataset with full datasets from Kaggle Fake News Dataset.

Skills Used

- Machine Learning
- Text Preprocessing (NLP)
- Model Training and Evaluation
- Streamlit UI Development
- TF-IDF Feature Extraction
- Git & GitHub

Author

Neha Pinnamreddy
Project developed as part of a Machine Learning minor project.