Fake vs Real News Classification - README

Fake vs Real News Classifier - README

Fake vs Real News Classification using NLP

This project classifies news articles as either Fake or Real using Natural Language Processing and Machine Learning.

Project Features

- NLP-based preprocessing and feature extraction
- Logistic Regression model for classification
- TF-IDF vectorization
- Streamlit Web App for live testing
- Clean, colorful, and user-friendly UI

Tech Stack

- Python, Pandas, NumPy, Scikit-learn
- NLTK for text preprocessing
- TF-IDF Vectorizer
- Streamlit for front-end UI

Dataset

Dataset used: Combined Fake and Real News datasets from Kaggle.

- Fake.csv -> Label 0

Fake vs Real News Classification - README

- True.csv -> Label 1

Steps to Run

1. Preprocess the data using NLTK (remove stopwords, punctuation, etc.)
2. Vectorize using TF-IDF
3. Train the model (Logistic Regression)
4. Save the model as news_model.pkl and vectorizer.pkl
5. Build a Streamlit UI (app.py)
6. Deploy on Streamlit Cloud

How to Use

- Enter or paste a news article into the input box.
- Click on "Predict".
- The app will show whether the news is real or fake with an explanation.

Deployment

This app can be deployed using Streamlit Cloud. Upload the project to a public GitHub repo and deploy using the Streamlit sharing platform.