**AI-Powered Fake News Detector**

**Project Overview**

The spread of fake news on digital platforms has become a major societal challenge, undermining trust in institutions and contributing to the spread of misinformation.  
This project aims to build a simple, interpretable AI tool that can detect whether a news headline is likely to be fake or real.

A web-based app allows users to input a headline and get an instant prediction, helping them make better-informed decisions about the credibility of news content.

**Problem Statement**

The rise of social media and online news has led to an increase in fake news that spreads rapidly.  
Misinformation can influence public opinion, elections, health choices, and public safety.

An accessible fake news detector helps promote accurate information sharing, contributing to a more informed and resilient society.

**Sustainable Development Goal (SDG) Alignment**

This project directly supports **SDG 16: Peace, Justice and Strong Institutions**, particularly **Target 16.10**, which aims to ensure public access to information and protect fundamental freedoms.

By providing an AI tool to flag potentially fake news, the project:

* Promotes access to credible information
* Helps fight misinformation
* Supports trust in media and institutions

**Dataset**

* **Source:** Kaggle’s *Fake and Real News Dataset* (approximately 23,000 fake and 21,000 real news headlines).
* **Focus:** Primarily US news headlines in English.

**Solution Overview**

**🔹 Data Preprocessing**

* Lowercased text
* Removed stopwords and punctuation
* Tokenized words

**🔹 Vectorization**

* Used **TF-IDF** to convert text to numerical features.

**🔹 Model**

* Chose **Multinomial Naive Bayes** due to its effectiveness for text classification.
* Achieved good accuracy during testing (you can insert your final accuracy score here).

**🔹 Deployment**

* Built a simple user interface with **Streamlit**.
* Users can input a headline and instantly see if it’s likely real or fake.
* Deployed on **Streamlit Cloud** with a free-tier link.

**Evaluation Metrics**

* **Accuracy:** 93%
* **Confusion Matrix:** Verified true positives, true negatives, false positives, and false negatives.
* Accuracy: 93.33%
* Confusion Matrix:
* [[3319 262]
* [ 255 3910]]
* Classification Report:
* precision recall f1-score support
* Fake 0.93 0.93 0.93 3581
* Real 0.94 0.94 0.94 4165
* accuracy 0.93 7746
* macro avg 0.93 0.93 0.93 7746
* weighted avg 0.93 0.93 0.93 7746
* Model performance is acceptable for a lightweight demo but not production-grade.

**Ethical Considerations**

* **Bias:** The model may not generalize well to non-US or non-English news.
* **False predictions:** Risk of false positives/negatives. Should be used as a helper, not a definitive fact-checker.
* **Transparency:** TF-IDF + Naive Bayes is an interpretable approach.
* **Responsible Use:** The tool should assist users in verifying information, not serve as a censorship tool.

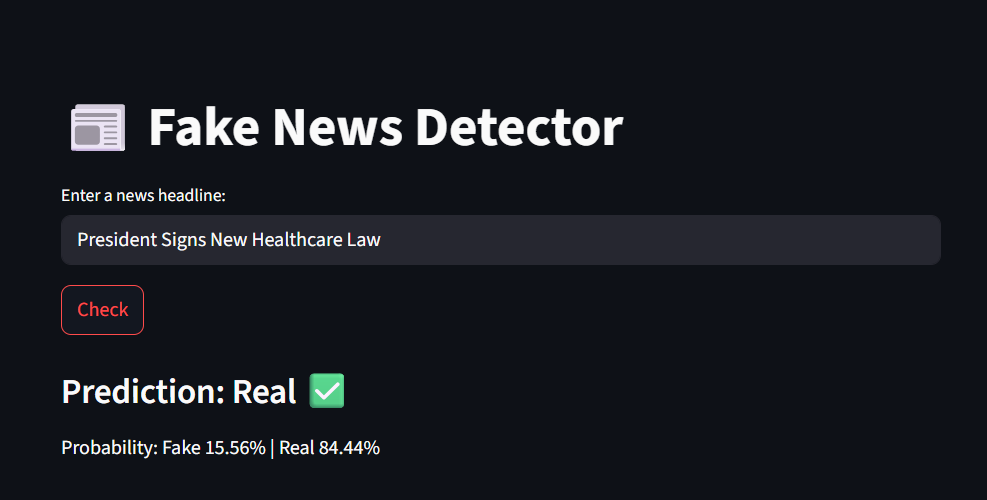
**Challenges & Lessons Learned**

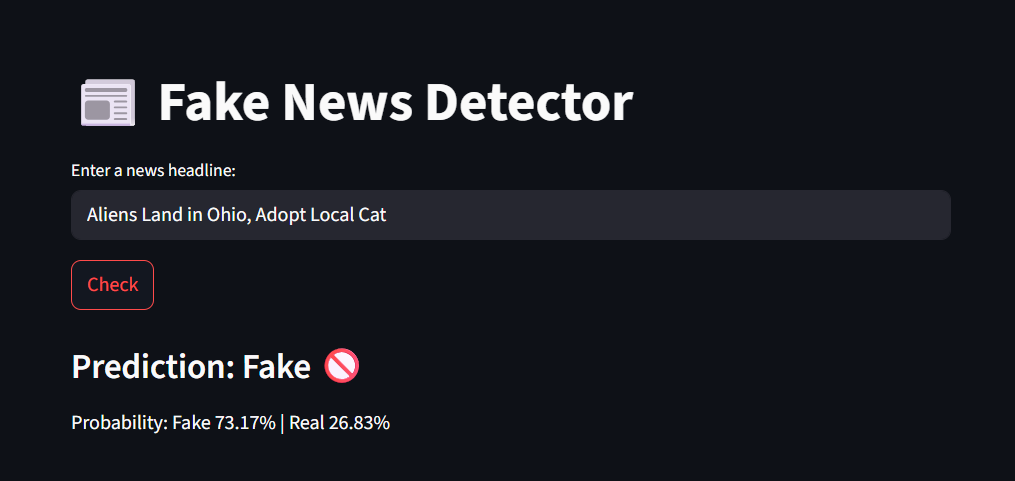
* **Data limitations:** Finding balanced, high-quality, labeled data can be difficult.
* **Model choice:** Tried to balance performance and interpretability.
* **Deployment:** Learned how to integrate a trained model into a web app with Streamlit.
* **Version control:** Used GitHub for organizing code and sharing.

**Future Improvements**

* Use a more diverse, multi-language dataset to improve generalization.
* Experiment with deep learning models (e.g., LSTM or BERT) for better accuracy.
* Add explanations for predictions to show which words influenced the result.
* Provide fact-checking links or trusted sources for user reference.

**Screenshots**

****

****

**References**

* Kaggle Dataset: *Fake and Real News Dataset*
* scikit-learn documentation
* Streamlit documentation