FAKE NEWS DETECTION USING MACHINE LEARNING

PRESENTED BY

STUDENT NAME:KUTHADI MANASA VENKATA NAGA SRUTHI

COLLEGE NAME:DMSSVH COLLEGE OF ENGINEERING

DEPARTMENT: CSE

EMAIL ID: KM1842376@GMAIL.COM

AICTE STUDENT ID:STU652E56F366D251697535731

AINSI_115449

OUTLINE

- **Problem Statement**
- **Proposed System/Solution**
- **System Development Approach**
- **Algorithm & Deployment**
- **Result**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

- In the digital age, the spread of misinformation and fake news has become a significant concern. With the rapid growth of social media and online platforms, it is easier than ever to circulate false information. This fake news can influence public opinion, create panic, and even affect elections and social harmony. Therefore, identifying and mitigating fake news is a pressing need in modern society.

PROPOSED SOLUTION

- We propose an AI-based Fake News Detection system that utilizes Natural Language Processing (NLP) and Machine Learning algorithms to analyze the textual content of news articles and classify them as "Real" or "Fake". The system is trained on a labeled dataset of fake and real news articles to learn patterns and linguistic features that typically indicate fake news.

SYSTEM APPROACH

Technologies and Tools Used:

Programming Language: PythonLibraries: scikit-learn, pandas, numpy, matplotlib, seaborn, nltk, reML

Algorithm: Logistic Regression / PassiveAggressiveClassifierNLP

Techniques: TF-IDF vectorization, text preprocessing

IDE: Jupyter Notebook

Dataset: Kaggle - Fake News Dataset or similar

ALGORITHM & DEPLOYMENT

Steps:

1. Data Preprocessing:

Cleaning text (removing punctuation, symbols, lowercasing)

Removing stop words and applying stemming/lemmatization

2. Vectorization:

Using TF-IDF to convert text into numerical format

3. Model Training:

Splitting data into training and testingApplying

Logistic Regression or PassiveAggressiveClassifier

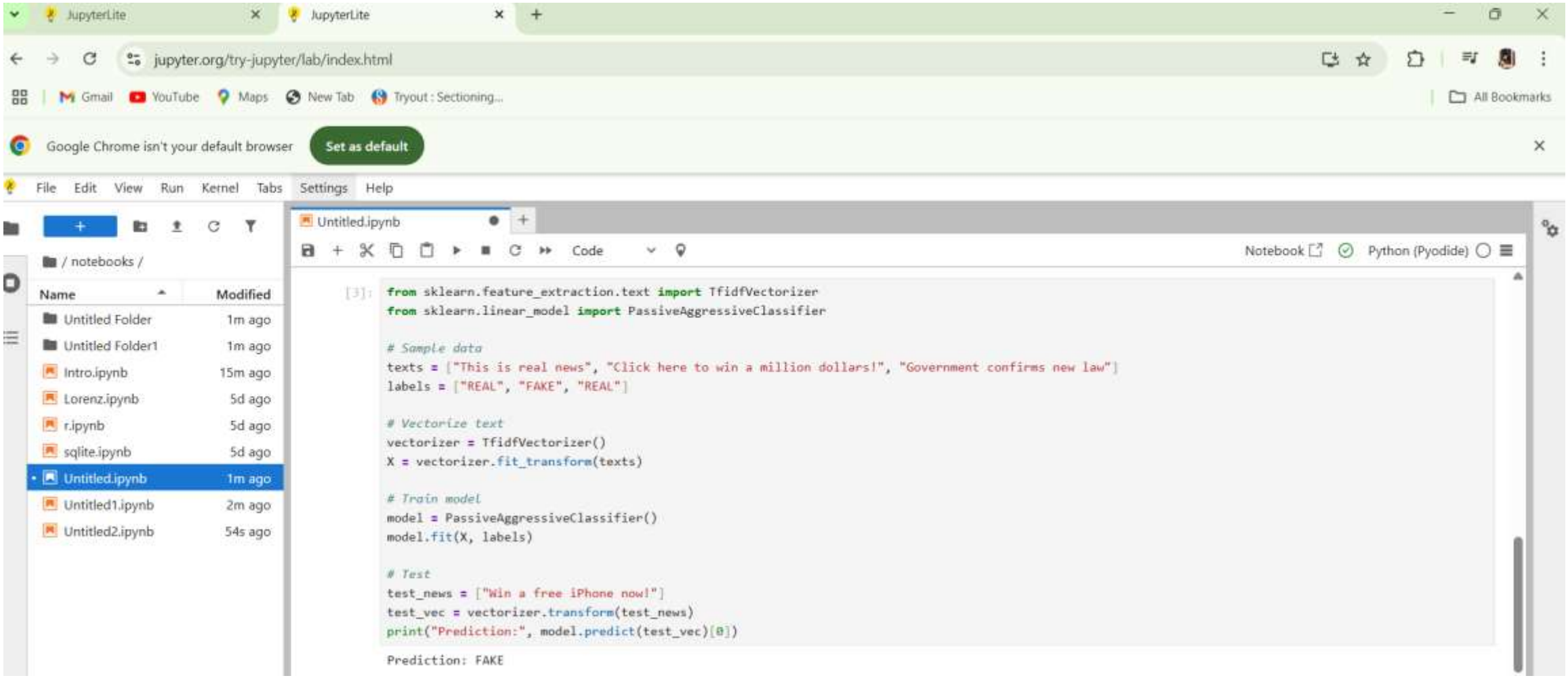
4. Model Evaluation:

Accuracy Score, Confusion Matrix, Classification Report

5. Deployment (optional):

Web interface using Flask for user interaction (submit news text → get prediction)

RESULT


JupyterLite
JupyterLite
jupyter.org/try-jupyter/lab/index.html
Gmail
YouTube
Maps
New Tab
Tryout : Sectioning...
All Bookmarks
Google Chrome isn't your default browser
Set as default
File
Edit
View
Run
Kernel
Tabs
Settings
Help
/ notebooks /
Name
Modified
Untitled Folder 1m ago
Untitled Folder1 1m ago
Intro.ipynb 15m ago
Lorenz.ipynb 5d ago
r.ipynb 5d ago
sqlite.ipynb 5d ago
Untitled.ipynb 1m ago
Untitled1.ipynb 2m ago
Untitled2.ipynb 54s ago
Untitled.ipynb
Code
Notebook
Python (Pyodide)
[3]:
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import PassiveAggressiveClassifier
# Sample data
texts = ["This is real news", "Click here to win a million dollars!", "Government confirms new law"]
labels = ["REAL", "FAKE", "REAL"]
# Vectorize text
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(texts)
# Train model
model = PassiveAggressiveClassifier()
model.fit(X, labels)
# Test
test_news = ["Win a free iPhone now!"]
test_vec = vectorizer.transform(test_news)
print("Prediction:", model.predict(test_vec)[0])
Prediction: FAKE

CONCLUSION

- Fake news detection is a vital application of AI that can contribute significantly to maintaining social harmony and spreading awareness. Our system demonstrates that machine learning models, when trained with proper data and NLP techniques, can effectively differentiate between fake and real news articles. This system can be integrated into larger platforms for real-time detection and flagging of misinformation.

FUTURE SCOPE

Integrate multilingual support for detecting fake news in various languages.

Use deep learning models like BERT or LSTM for higher accuracy.

Create a browser extension or mobile app for real-time article verification.

Enhance dataset with newer articles and social media posts for better model generalization.

REFERENCES

1. Scikit-learn Documentation
2. Natural Language Toolkit (NLTK)
3. Kaggle: Fake News Dataset
4. Research Paper: "Detecting Fake News on Social Media: A Data Mining Perspective", ACM SIGKDD

GitHub link: https://github.com/Manasa272/Manasa_Ai-project.git

Thank you