import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.feature\_extraction.text import TfidfVectorizer

from sklearn.linear\_model import PassiveAggressiveClassifier

from sklearn.metrics import accuracy\_score, confusion\_matrix

# Load dataset

# Dataset should have two columns: 'text' and 'label' (where label is 'FAKE' or 'REAL')

df = pd.read\_csv('news.csv') # Replace with your dataset path

# Split data

X = df['text']

y = df['label']

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Vectorize text using TF-IDF

tfidf\_vectorizer = TfidfVectorizer(stop\_words='english', max\_df=0.7)

tfidf\_train = tfidf\_vectorizer.fit\_transform(X\_train)

tfidf\_test = tfidf\_vectorizer.transform(X\_test)

# Train classifier

classifier = PassiveAggressiveClassifier(max\_iter=50)

classifier.fit(tfidf\_train, y\_train)

# Predict and evaluate

y\_pred = classifier.predict(tfidf\_test)

score = accuracy\_score(y\_test, y\_pred)

cm = confusion\_matrix(y\_test, y\_pred)

print(f'Accuracy: {score \* 100:.2f}%')

print('Confusion Matrix:')

print(cm)