Name : Dion D Rodrigues Assignment.no: 05

Roll.no : 53 Class : FYMCA B Subject : ML Lab Batch : B3

-----------------------------------------------------------------------------------------------------------------------------

# Implementation of Naive Bayes for Fake Job Posting

import pandas as pd

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

from sklearn.feature\_extraction.text import CountVectorizer

from sklearn.naive\_bayes import MultinomialNB

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

# Step 1: Load Dataset

df = pd.read\_csv('C:/Users/Admin/Desktop/ML/Assignment\_5/Gender\_Detection.csv', encoding='ISO-8859-1')

print(df)

# Map gender to binary values

df['Gender'] = df['Gender'].map({'F': 0, 'M': 1})

print(df)

# Split the dataset into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(df['Name'], df['Gender'], test\_size=0.2, random\_state=42)

# Convert the data into a bag-of-words model using CountVectorizer

vectorizer = CountVectorizer()

X\_train\_transformed = vectorizer.fit\_transform(X\_train)

X\_test\_transformed = vectorizer.transform(X\_test)

# Step 3: Train the Naive Bayes Model

model = MultinomialNB()

model.fit(X\_train\_transformed, y\_train)

# Step 4: Evaluate the Model

y\_pred = model.predict(X\_test\_transformed)

# Print the accuracy

print(f"Accuracy: {accuracy\_score(y\_test, y\_pred)}")

# Print the classification report

print("\nClassification Report:\n", classification\_report(y\_test, y\_pred))

# Print the confusion matrix

print("\nConfusion Matrix:\n", confusion\_matrix(y\_test, y\_pred))

