SPAM DETECTION USING AI

USER QUESTIONS

1. Name at least 3 most prominent words which made you question the sender’s intent. Out of 3
2. Frequency
3. Once 0
4. Less than 5 1
5. Less than 10 2
6. 10+ 3
7. Why ?
8. Never accessed that site / product 1
9. Not required any longer 1
10. Financial scams 4
11. Fake lottery / price notifications 2
12. Offensive content 3
13. Others 1

IMPLEMENTATION

(CODE)

from tkinter import \*

def demo1():

values = []

window = Tk()

window.title("REPORT SPAM")

# StringVar to store the value entered in the Entry widget

entry\_value1 = StringVar()

entry\_value2 = StringVar()

entry\_value3 = StringVar()

entry\_value4 = StringVar()

entry\_value5 = StringVar()

Label(window, text="Q1. What is the reason? \n1. Offensive content \n2. Financial scams \n3. Not required any longer \n4. Never accessed that site/product \n5. Fake lottery/ price notifications\n6.Others").grid(row=1, column=0)

entry1 = Entry(window, textvariable=entry\_value1, bg="LightBlue")

entry1.grid(row=1, column=1)

Label(window, text="Q2. How frequent have you received this sender's email? \n1.Once \n2.Less than 5 \n3.Less than 10 \n4.10+").grid(row = 2, column = 0)

entry2 = Entry(window, textvariable = entry\_value2, bg = "LightBlue")

entry2.grid(row=2, column=1)

Label(window,text = "Q3. What were the 3 most prominent words which made you question the sender's intent?" ).grid(row = 3, column = 0)

entry3 = Entry(window, textvariable = entry\_value3, bg = "LightBlue")

entry3.grid(row=4, column=1)

entry4 = Entry(window, textvariable = entry\_value4, bg = "LightBlue")

entry4.grid(row=5, column=1)

entry5 = Entry(window, textvariable = entry\_value5, bg = "LightBlue")

entry5.grid(row=6, column=1)

def on\_button\_click():

window.destroy()

button = Button(window, text="Submit", command=on\_button\_click)

button.grid(row=7, column=7, columnspan=2)

window.mainloop()

# Retrieve the value from the StringVar after mainloop has finished

values = [entry\_value1.get(), entry\_value2.get(), entry\_value3.get(), entry\_value4.get(), entry\_value5.get()]

return values

demo1\_result = demo1()

print(demo1\_result)

a = demo1\_result[0].lower()

b = demo1\_result[1].lower()

c = demo1\_result[2].lower()

d = demo1\_result[3].lower()

e = demo1\_result[4].lower()

lst = ["free", "money", "click", "win", "prize", "congratulations", "offer", "sale", "discount", "viagra", "urgent", "act now", "limited time", "cash", "guaranteed", "million", "lottery", "unsubscribe", "opportunity", "credit card"]

score = 0

if a=='1':

score += 3

elif a=='2':

score += 4

elif a=='3':

score += 2

elif a=='4':

score += 2

elif a=='5':

score += 3

else:

score += 1

if b=='1':

score += 0

elif b=='2':

score += 1

elif b=='3':

score += 2

elif b=='4':

score +=3

for i in lst:

if c==i or d==i or e==i:

score +=1

print(score)

x = False

if score >= 6:

print("CHECK FOR SPAM")

x = True

if x == True:

import pandas as pd

df = pd.read\_csv("emails.csv")

print("df shape:", df.shape)

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

Y = df['Prediction']

X = df.drop(columns = ['Prediction', 'Email No.'])

print("(X Shape, Y Shape):", X.shape, Y.shape)

from sklearn.naive\_bayes import MultinomialNB

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

print("(X\_train Shape, Y\_train Shape):", X\_train.shape, Y\_train.shape)

print("(X\_test Shape, Y\_test Shape):", X\_test.shape, Y\_test.shape)

model = MultinomialNB()

model.fit(X\_train, Y\_train)

print("Model prediction:", model.predict(X\_test))

print("Model score on test data", model.score(X\_test, Y\_test))

print("Model score on training data", model.score(X\_train, Y\_train))

else:

print("Not applicable")

lst2 = df.columns

lst3 = []

for i in lst2:

for j in lst:

if i==j:

lst3.append(i)

print("List 3: ", lst3)