**ЛАБОРАТОРНА РОБОТА № 6**

Наївний Байєс в Python

**Мета роботи:** набути навичок працювати з даними і опонувати роботу у Python з використанням теореми Байэса.

**Варіант 7**

**Хід роботи:**

Посилання на GitHub:

<https://github.com/Dubnitskyi/AI_all_labs/tree/master/Lab6>

**Завдання 1**

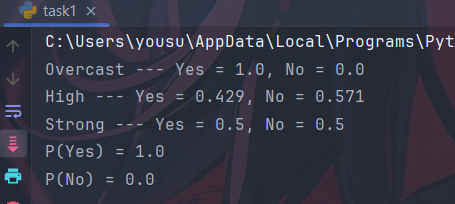
Визначити відбудеться матч при наступних погодних умовах чи ні.



**Програмний код:**

import pandas as pd  
  
  
def ChansGame(data, freq\_table, elem):  
 all\_Yes = sum(freq\_table["Yes"])  
 all\_No = sum(freq\_table["No"])  
 S\_YaN = freq\_table['Yes'][elem] + freq\_table['No'][elem]  
  
 all\_Elem = len(data)  
 PElem = round(S\_YaN / all\_Elem, 8)  
 PYesAElem = round(freq\_table['Yes'][elem] / all\_Yes, 8)  
 PYes = round(all\_Yes / all\_Elem, 8)  
 P\_Y\_s = round((PYesAElem \* PYes) / PElem, 8)  
  
 PNoAElem = round(freq\_table['No'][elem] / all\_No, 8)  
 PNo = round(all\_No / all\_Elem, 8)  
 P\_N\_s = round((PNoAElem \* PNo) / PElem, 8)  
  
 print(f"{elem} --- Yes = {round(P\_Y\_s, 3)}, No = {round(P\_N\_s, 3)}")  
 return P\_Y\_s, P\_N\_s, PYes, PNo  
  
  
data = pd.read\_csv('data.csv')  
  
freq\_table\_outlook = pd.crosstab(data['Outlook'], data['Play'])  
freq\_table\_humidity = pd.crosstab(data['Humidity'], data['Play'])  
freq\_table\_wind = pd.crosstab(data['Wind'], data['Play'])  
  
P\_Y\_s1, P\_N\_s1, PYes1, PNo1 = ChansGame(data, freq\_table\_outlook, "Overcast")  
P\_Y\_s2, P\_N\_s2, PYes2, PNo2 = ChansGame(data, freq\_table\_humidity, "High")  
P\_Y\_s3, P\_N\_s3, PYes3, PNo3 = ChansGame(data, freq\_table\_wind, "Strong")  
  
Probability\_yes = round(P\_Y\_s1 \* P\_Y\_s2 \* P\_Y\_s3 \* PYes1 \* PYes2 \* PYes3, 8)  
Probability\_no = round(P\_N\_s1 \* P\_N\_s2 \* P\_N\_s3 \* PNo1 \* PNo2 \* PNo3, 8)  
  
print(f"P(Yes) = {round(Probability\_yes / (Probability\_yes + Probability\_no), 3)}")  
print(f"P(No) = {round(Probability\_no / (Probability\_yes + Probability\_no), 3)}")

Результат виконання:



**Завдання 2.**

Застосуєте методи байєсівського аналізу до набору даних про ціни на квитки на іспанські високошвидкісні залізниці.

**Програмний код:**

import pandas as pd  
import matplotlib.pyplot as plt  
from sklearn.naive\_bayes import GaussianNB  
from sklearn.model\_selection import train\_test\_split  
from sklearn.naive\_bayes import GaussianNB  
from sklearn.metrics import accuracy\_score  
  
data = pd.read\_csv('renfe\_small.csv')  
freq\_table = pd.crosstab(data['price'], data['destination'])  
array = freq\_table.values  
X, y = array[:, :-1], array[:, -1]  
  
X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.3, random\_state=8)  
gnb = GaussianNB()  
gnb.fit(X\_train, y\_train)  
y\_pred = gnb.predict(X\_test)  
  
accuracy = accuracy\_score(y\_test, y\_pred)  
print(f'Accuracy: {accuracy}')

Результат виконання:

Изображение выглядит как текст, Шрифт, снимок экрана, Графика

Автоматически созданное описание

**Висновок:** Під час лабораторної роботи я набув навичок працювати з даними і опонував роботу у Python з використанням теореми Байэса.