

ЛАБОРАТОРНА РОБОТА № 6

Наївний Байєс в Python

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

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.

2, 7, 12	Outlook = Overcast Humidity = High Wind = Strong	Перспектива = Похмуро Вологість = Висока Вітер = Сильний
----------	--	--

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'])

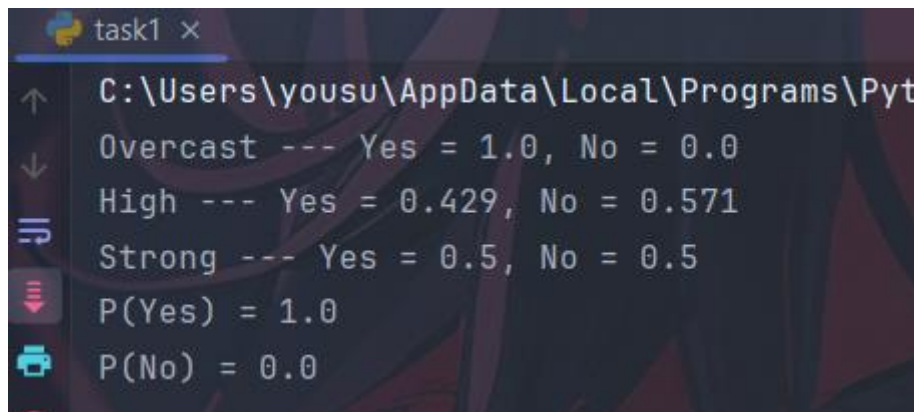
P_Y_s1, P_N_s1, PYes1, PNo1 = ChansGame(data, freq_table_outlook, "Overcast")
P_Y_s2, P_N_s2, PYes2, PNo2 = ChansGame(data, freq_table_humidity, "High")
P_Y_s3, P_N_s3, PYes3, PNo3 = ChansGame(data, freq_table_wind, "Strong")

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:



```

task1 x
C:\Users\yousu\AppData\Local\Programs\Pyt
Overcast --- Yes = 1.0, No = 0.0
High --- Yes = 0.429, No = 0.571
Strong --- Yes = 0.5, No = 0.5
P(Yes) = 1.0
P(No) = 0.0

```

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]

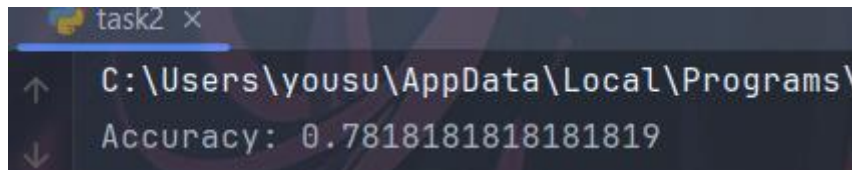
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3,
random_state=8)

```

```
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:

A screenshot of a terminal window titled 'task2'. The terminal shows the execution of a Python script. The first line of the command prompt is 'C:\Users\yousu\AppData\Local\Programs\'. The second line shows the output 'Accuracy: 0.7818181818181819'.

```
task2 x
C:\Users\yousu\AppData\Local\Programs\
Accuracy: 0.7818181818181819
```

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