

**Завдання 3. Використовую данні з пункту 2 визначити відбудеться матч при наступних погодних умовах чи ні: Розрахунки провести з використанням Python.**

Варіант	Умова	
1, 6, 11	Outlook = Overcast Humidity = High Wind = Weak	Перспектива = Похмуро Вологість = Висока Вітер = Слабкий
2, 7, 12	Outlook = Overcast Humidity = High Wind = Strong	Перспектива = Похмуро Вологість = Висока Вітер = Сильний
3, 8, 13	Outlook = Sunny Humidity = High Wind = Weak	Перспектива = Сонячно Вологість = Висока Вітер = Слабкий
4, 9, 14	Outlook = Sunny Humidity = Normal Wind = Strong	Перспектива = Сонячно Вологість = Нормальна Вітер = Сильний
5, 10, 15	Outlook = Rain Humidity = High Wind = Strong	Outlook = Дощ Вологість = Висока Вітер = Сильний

Посилання на Git: <https://github.com/PanchukPetro/SShILabsPanchuk/tree/main/Lab6>

Humidity	Play Yes	Play No	Total
High	3/9	4/5	7/14
Normal	6/9	1/5	7/14
Total	9/14	5/14	

Wind	Play Yes	Play No	Total
High	6/9	2/5	8/14
Normal	3/9	3/5	6/14
Total	9/14	5/14	

Outlook	Play Yes	Play No	Total
Sunny	2/9	3/5	5/14
Overcast	4/9	0/5	4/14
Rainy	3/9	2/5	5/14
Total	9/14	5/14	

Yes =  $P(\text{Outlook} = \text{Overcast} \mid \text{Yes}) * P(\text{Humidity} = \text{High} \mid \text{Yes}) * P(\text{Wind} = \text{Strong} \mid \text{Yes}) * P(\text{Yes})$

No =  $P(\text{Outlook} = \text{Overcast} \mid \text{No}) * P(\text{Humidity} = \text{High} \mid \text{No}) * P(\text{Wind} = \text{Strong} \mid \text{No}) * P(\text{No})$

Yes =  $4/9 * 3/9 * 3/9 * 9/14 = 0.0317$

No =  $0/5 * 4/5 * 2/5 * 5/14 = 0$

Yes =  $0.0317 / (0.0317 + 0) = 1$

Модель передбачає ймовірність гри при заданих умовах в 100%

					ДУ «Житомирська політехніка».24.121.02.000 – ІПЗ		
Змн.	Арк.	№ докум.	Підпис	Дата			
Розроб.		Панчук П.С			СШІ Лабораторна робота №6	Лім.	Арк.
Перевір.		Голденко М.Ю.					1
Керівник		Використання на Python				Аркуші	4
Н. контр.						ФІКТ Гр. ІПЗ-21-3	
Зав. каф.							

```

#Variant 2
# Outlook = Overcast
# Humidity = High
# Wind = Strong

#Probability tables:
WeatherTable = [
    [2/9,3/5,5/14], #Sunny
    [4/5,0/5,4/14], #Overcast
    [3/9,2/5,5/14], #Rainy
    [9/14],[5/14]   #Total
]
HumidityTable = [
    [3/9,4/5,7/14], #High
    [6/9,1/5,7/14], #Low
    [9/14,5/14]     #Total
]
WindTable = [
    [6/9,2/5,8/14], #High
    [3/9,3/5,7/14], #Low
    [9/14,5/14]     #Total
]
YesTotal = 9/14
NoTotal = 5/14

CurrentWeather = 1
CurrentHumidity = 0
CurrentWind = 0

YesProbability = WeatherTable[CurrentWeather][0] *
HumidityTable[CurrentHumidity][0] * WindTable[CurrentWind][0] * YesTotal
NoProbability = WeatherTable[CurrentWeather][1] *
HumidityTable[CurrentHumidity][1] * WindTable[CurrentWind][1] * NoTotal

NormalizedYesProbability = YesProbability / (YesProbability + NoProbability)
NormalizedNoProbability = 1-NormalizedYesProbability

print("Вірогідність проведення гри при даних умовах:")
print(round(NormalizedYesProbability*100,2),"%")
print("Вірогідність що гру не проведуть:")
print(round(NormalizedNoProbability*100,2),"%")

```

```

Вірогідність проведення гри при даних умовах:
100.0 %
Вірогідність що гру не проведуть:
0.0 %

```

Рис. 1. Виведення в консоль

		Панчук П.С			ДУ «Житомирська політехніка».24.121.02.000 – ІПЗ	Арк.
		Голенко М.Ю				2
Змн.	Арк.	№ докум.	Підпис	Дата		

**Завдання 4.** Застосуєте методи байєсівського аналізу до набору даних про ціни на квитки на іспанські високошвидкісні залізниці.

Програмний код:

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import classification_report, accuracy_score
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import confusion_matrix

url = "https://raw.githubusercontent.com/susanli2016/Machine-Learning-with-Python/master/data/renfe_small.csv"
data = pd.read_csv(url)

#Choosing necessary fields in table
data = data[['price', 'train_type', 'fare', 'origin', 'destination']]
data = data.dropna()

#Turning price into binary value
threshold = data['price'].median()
data['price_category'] = (data['price'] > threshold).astype(int)
categorical_columns = ['train_type', 'fare', 'origin', 'destination']
data = pd.get_dummies(data, columns=categorical_columns, drop_first=True)

X = data.drop(columns=['price', 'price_category'])
y = data['price_category']

#TrainTestSplit and predict
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3,
                                                    random_state=42)
model = GaussianNB()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)

#Model evaluation
print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nClassification Report:\n", classification_report(y_test, y_pred))
cm = confusion_matrix(y_test, y_pred)

#Confusion matrix visualisation
plt.figure(figsize=(6, 6))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=['Cheap', 'Expensive'],
            yticklabels=['Cheap', 'Expensive'])
plt.title('Confusion Matrix')
plt.xlabel('Predicted')
plt.ylabel('True')
plt.show()
```

		Панчук П.С			ДУ «Житомирська політехніка».24.121.02.000 – ІПЗ	Арк.
		Голенко М.Ю				3
Змн.	Арк.	№ докум.	Підпис	Дата		

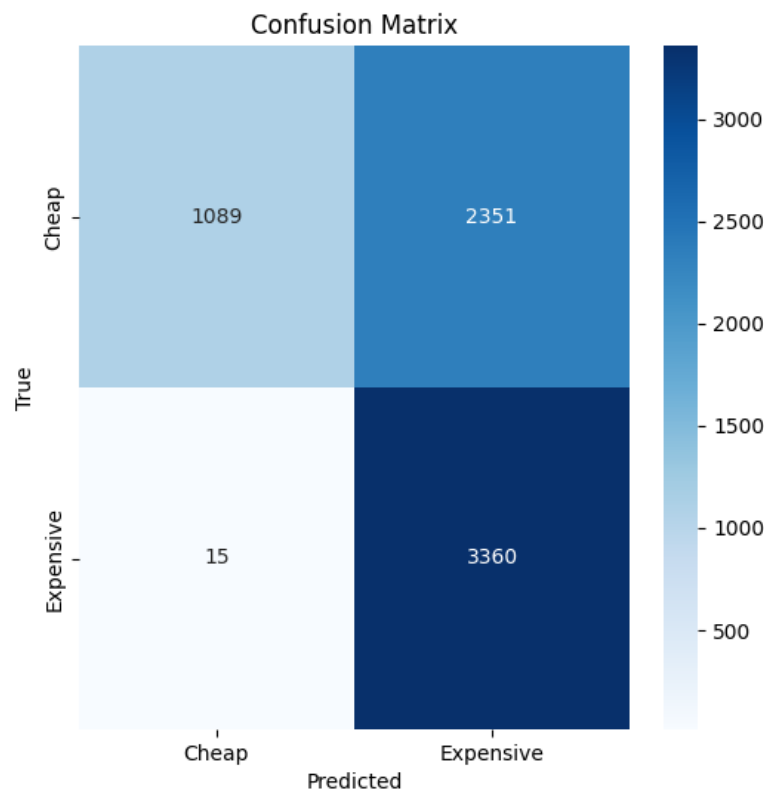
```
Accuracy: 0.6528246515040352

Classification Report:
              precision    recall  f1-score   support

    0       0.99         0.32         0.48         3440
    1       0.59         1.00         0.74         3375

 accuracy          0.65          0.65          0.65          6815
 macro avg         0.79         0.66         0.61          6815
weighted avg         0.79         0.65         0.61          6815
```

Виведення в консоль



Візуалізації Confusion Matrix