



Artificial Intelligence

Lab Tasks # 11

Name: Samreen Bibi

Sap Id: 46484

Batch: BSCS-6th semester

Lab Instructor: Mam Ayesha Akram

Task # 01

Solution:

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import accuracy_score, classification_report
from sklearn.preprocessing import LabelEncoder

# Read and clean data
data = pd.read_csv("C:/Users/Lenovo/Downloads/public-data.csv").replace(to_replace="?", pd.NA).dropna()

# Encode all categorical features
for col in data.select_dtypes(include='object'):
    data[col] = LabelEncoder().fit_transform(data[col])

# Features and label
features = data.drop(columns=['Salary'])
labels = data['Salary']

# Splitting the dataset
X_tr, X_te, y_tr, y_te = train_test_split(*arrays: features, labels, test_size=0.2, random_state=1)

# Model training
nb = GaussianNB()

nb.fit(X_tr, y_tr)

# Predictions
predictions = nb.predict(X_te)

# Evaluation
print(f"Model Accuracy: {accuracy_score(y_te, predictions)}")
print("\nTrue Labels:      ", y_te.values)
print("Predicted Labels:", predictions)
print("\nReport:\n", classification_report(y_te, predictions))
```

Output

Model Accuracy: 0.7524950099800399

True Labels: [0 0 1 ... 0 0 1]

Predicted Labels: [0 0 0 ... 0 0 1]

Report:

	precision	recall	f1-score	support
0	0.88	0.79	0.83	5026
1	0.47	0.64	0.54	1487
accuracy			0.75	6513
macro avg	0.67	0.71	0.69	6513
weighted avg	0.79	0.75	0.76	6513