Artificial Intelligence

Lab 11 Tasks

Name: Dua Amir

Sap ID: 47849

Batch: BSCS-6th semester

Lab Instructor:

Ayesha Akram

Task1.
Solution:

```python
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
from sklearn.preprocessing import LabelEncoder

# Load dataset
df = pd.read_csv("C:/Users/ACG/Downloads/public-data.csv")

# Clean and encode
df = df.replace( to_replace: " ?", pd.NA).dropna()

le = LabelEncoder()
for column in df.columns:
    if df[column].dtype == 'object':
        df[column] = le.fit_transform(df[column])

# Define features and label
X = df.drop( labels: "Salary", axis=1)
y = df["Salary"]

# Train-test split
X_train, X_test, y_train, y_test = train_test_split( *arrays: X, y, test_size=0.2, random_state=42)

# Train model
model = GaussianNB()
model.fit(X_train, y_train)

# Predict
y_pred = model.predict(X_test)

# Results
print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nPredicted labels:", y_pred)
print("Actual labels:   ", y_test.values)
print("\nClassification Report:\n", classification_report(y_test, y_pred))
```

Output:

```
Accuracy: 0.760325502840473

Predicted labels: [0 0 0 ... 1 0 0]
Actual labels:    [0 0 1 ... 1 0 0]

Classification Report:
              precision    recall  f1-score   support

           0       0.87      0.80      0.84      4942
           1       0.50      0.63      0.56      1571

    accuracy                           0.76      6513
   macro avg       0.69      0.72      0.70      6513
weighted avg       0.78      0.76      0.77      6513
```
