**MLA0201-Fundamentals of Machine Learning**

**Name: Kaviya Kumara (192425055)**

Experiment 5:

Write a program for Implementation of K-Nearest Neighbours (K-NN) in Python

**Code:**

import pandas as pd

import matplotlib.pyplot as plt

from sklearn.neighbors import KNeighborsClassifier

from sklearn.preprocessing import LabelEncoder, StandardScaler

from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import accuracy\_score

data = pd.read\_csv("car\_data.csv")

X = data.iloc[:, :-1].copy()

y = data.iloc[:, -1]

encoders = {}

for column in X.columns:

if X[column].dtype == 'object':

encoders[column] = LabelEncoder()

X[column] = encoders[column].fit\_transform(X[column])

target\_encoder = LabelEncoder()

y = target\_encoder.fit\_transform(y)

scaler = StandardScaler()

X = scaler.fit\_transform(X)

X\_train, X\_test, y\_train, y\_test = train\_test\_split(

X, y, test\_size=0.3, random\_state=1

)

k = 5

model = KNeighborsClassifier(n\_neighbors=k)

model.fit(X\_train, y\_train)

print("K-NN model trained successfully")

new\_sample = pd.DataFrame([{

'Brand': 'Toyota',

'Fuel': 'Petrol',

'Transmission': 'Automatic',

'EngineCC': 2000,

'Mileage': 15,

'Seats': 5

}])

for column in new\_sample.columns:

if column in encoders:

new\_sample[column] = encoders[column].transform(new\_sample[column])

new\_sample\_scaled = scaler.transform(new\_sample)

prediction = model.predict(new\_sample\_scaled)

predicted\_class = target\_encoder.inverse\_transform(prediction)

print("Predicted Class for new sample:", predicted\_class[0])

y\_pred = model.predict(X\_test)

accuracy = accuracy\_score(y\_test, y\_pred)

print("Model Accuracy:", accuracy)

feature\_x = data['EngineCC']

feature\_y = data['Mileage']

plt.figure(figsize=(8, 6))

plt.scatter(feature\_x, feature\_y, c=y, cmap='viridis', label='Training Data')

plt.scatter(

new\_sample['EngineCC'],

new\_sample['Mileage'],

color='red',

marker='X',

s=200,

label='New Sample'

)

plt.xlabel("EngineCC")

plt.ylabel("Mileage")

plt.title("K-NN Visualization (EngineCC vs Mileage)")

plt.legend()

plt.show()

plt.figure()

plt.bar(['K-NN Accuracy'], [accuracy])

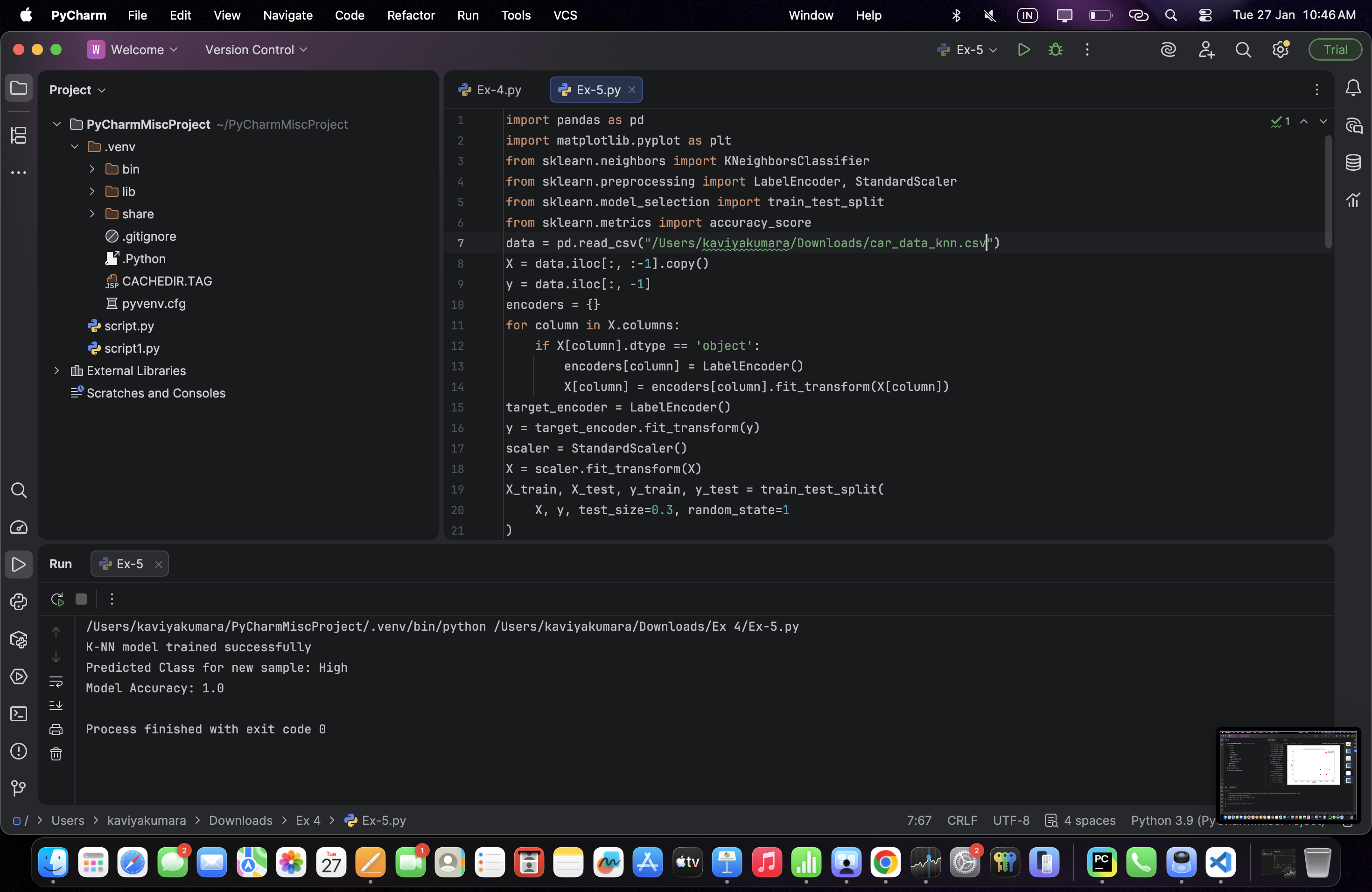
plt.ylim(0, 1)

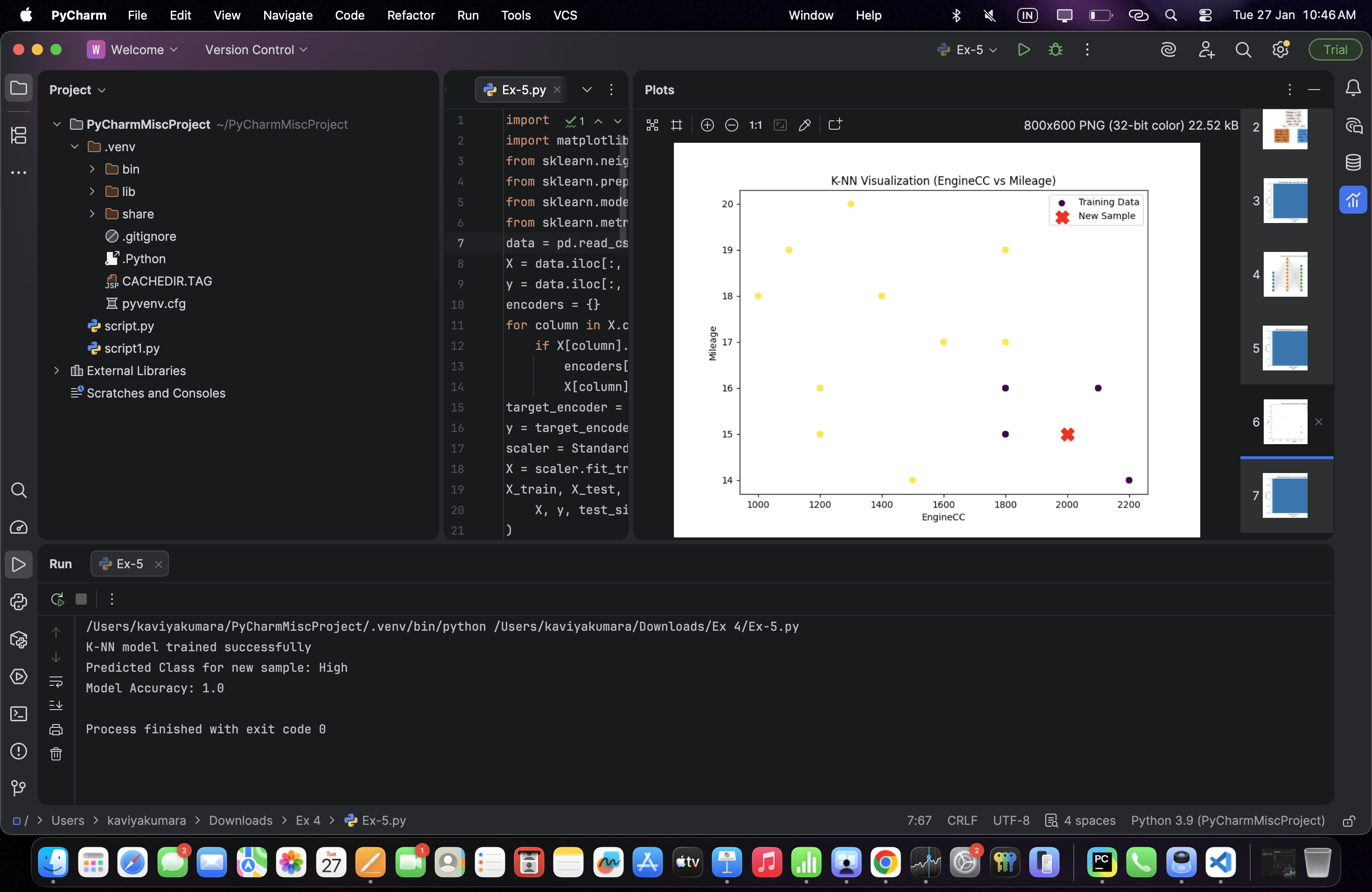
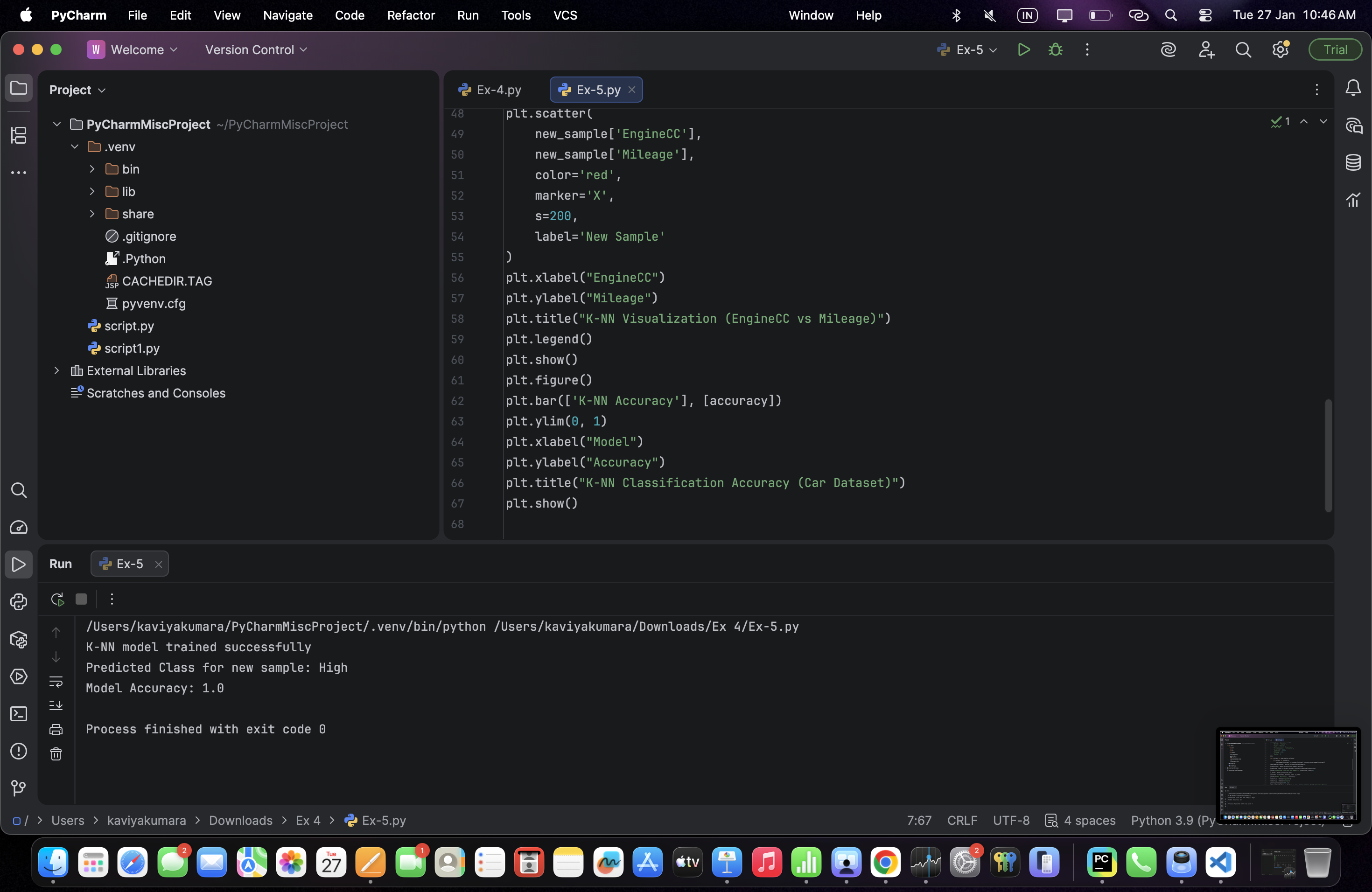
plt.xlabel("Model")

plt.ylabel("Accuracy")

plt.title("K-NN Classification Accuracy (Car Dataset)")

plt.show()

**Output:**

****