GLS University
Faculty of Computer Applications & IT
IMCA SEM VI
222301605 Practicals on Machine Learning
Practical Practice

1.	Write a machine learning code to create a histogram using 5 bins with edge color set to black. Given the data [10, 20, 20, 25, 30, 35, 40, 50], write
2.	Create a box plot for the dataset [1, 3, 5, 7, 9, 15, 20, 25] and find the outliers.
3.	Write a machine learning code to create a scatter plot for the following data points: • x = [2, 4, 6, 8, 10] • y = [3, 6, 9, 12, 15]
4.	Write a machine learning code to visualize the frequency of letters in the dataset {'Letters': ['A', 'B', 'A', 'C', 'B', 'C', 'A']}.
5.	Write a machine learning code to visualize the frequency of each fruit category in a bar chart. dataset {'Fruits': ['Apple', 'Banana', 'Orange', 'Apple', 'Banana', 'Banana', 'Orange', 'Apple']},
6.	Write machine learning code to generate a box plot for the dataset: • data = [5, 7, 8, 12, 15, 18, 19, 20, 35, 37] Identify outliers using the IQR method and explain your findings.
7.	Generate a bar chart for a dataset where the categories are ['Dog', 'Cat', 'Bird'] and their respective counts are [3, 5, 2].
8.	Write machine learning code to Compute the correlation matrix and visualize it as a heatmap. Dataset: • Feature_1 = [10, 20, 30, 40, 50] • Feature_2 = [15, 25, 35, 45, 55] • Feature_3 = [5, 10, 15, 20, 25]
9.	Create a heatmap using a random dataset with 4 features and 50 samples.
10.	You're working with a dataset containing the age of customers for a marketing campaign with 10 customers. Make a dataset of it, take into dataframe and display the histogram of it with the following

	conditions: • Bins should be 5 • Give title “Age distribution • Edge color should be blac • Color should be blue • Give x and y axis labels accordingly
11.	2. A house price dataset contains a column with property prices. 100 Prices should be generated randomly. Display the histogram of it with the following conditions: • Bins should be 10 • Give title “Price distribution • Edge color should be gra • Color should be pink • Give x and y axis labels accordingly
12.	3. {'Age': [25, 30, 35, None, 28, 32, 45, None, 30, 28]} Create a histogram after deleting or impute missing values.
13.	models = ['Logistic Regression', 'SVM', 'Random Forest'] training_times = [2.1, 12.5, 8.4] Create a bar graph with following conditions: • Graph color should be gree • It should be horizontal graph
14.	Create a group bar graph and Stacked bar graph for the following data: Shape = ['Square1', 'Rect', 'Square2'] Length = [10, 15, 20] Height = [12, 18, 25]
15.	Create a scatter plot for the following data 1. house_size = [1000, 1500, 2000, 2500, 3000] house_price = [200000, 250000, 300000, 350000, 400000] 2. actual_prices = [200000, 250000, 300000, 350000, 400000] predicted_prices = [210000, 245000, 290000, 360000, 390000]

16.	Display box plot and show outliers for the following data 1. Transaction_amounts = [10, 15, 20, 25, 30, 35, 1000, 2000] 2. cv_scores_model_1 = [0.82, 0.84, 0.83, 0.81, 0.95] cv_scores_model_2 = [0.88, 0.87, 0.89, 0.86, 0.88] 3. daily_returns = [0.01, 0.02, 0.015, -0.03, 0.02, -0.1, 0.03, 0.08]
17.	Display heatmap for the following data 1. data = { 'Feature1': np.random.rand(100), 'Feature2': np.random.rand(100) * 2, 'Feature3': np.random.rand(100) * 3, 'Target': np.random.rand(100) * 4, } 2. true_values = np.random.rand(10) predicted_values = true_values + np.random.normal(0, 0.1, 10)