

Machine Learning using Python

List of experiments (up to 20.09.23)

Q1 To study the following python libraries:

- i. numpy
- ii. matplotlib
- iii. scikit-learn

Q2. To work with numpy arrays:

- a) create 1-D, 2-D and 3-D array using lists
- b) Access elements of 1D, 2D, 3D array
- c) Print the shape of the array
- d) Perform array slicing
- e) Perform boolean indexing
- f) Reshape a 1D array to 2D and 3D array
- g) Perform array arithmetic - addition, subtraction, multiplication and division, dot product
- h) Matrix multiplication using ndarray
- i) Sort numpy arrays
- j) Perform assignment, shallow copy and deep copy
- k) matrix multiplication using matrix class

Q3 To work with pandas : Series and Dataframe

- a) Create a python Series with index as date range of length 12. Access its elements using loc and iloc.
- b) Create a DataFrame reading from a csv file and perform
 - i) Slicing using loc and iloc.
 - ii) Sort the DataFrame in descending order.
 - iii) Describe the DataFrame.

Q4 Working with matplotlib : create plots using matplotlib pyplot

- a) Create a line graph for semester wise grades.(label both x & y-axis,also give a title)
- b) Create a pie Chart depicting your core competency in each of the programming languages you have learned.
 - i) Explode the wedges from the center for each language.
 - ii) Add labels to the wedges and also the legend to chart.
- c) Create a Scatter plot depicting your percentages till Semester 5(use different colors for different semesters).
- d) Create a Bar chart both horizontally and vertically depicting your percentages till semester 5,Use width and height attributes. Different colors for horizontal and vertical charts.
- e) Create subplots of 3 x 2 dimensions.

Q5 Working with Seaborn:

- a) Create a distribution plot with histogram.
- b) Create a distribution plot without histogram.
- c) Create a normal distribution plot and add labels to it.
- d) Create a binomial distribution plot and add labels to it.
- e) Use a csv file to obtain a line plot between two parameters.
- f) Draw a scatter plot between two parameters and also use hue property.
- g) Draw a boxplot between two parameters.
- h) Draw a violin plot between two parameters.
- i) Draw a swarm plot
- j) Draw a bar chart.
- k) Draw a pie chart.
- l) Draw a cat plot.

Q6 To import (any) dataset from sklearn, describe the dataset and plot scatter plot.

Q7 To load datasets (csv format) using numpy, pandas and csv modules.

Q8 To perform the outlier analysis using z-score and IQR i.e. statistical technique and visualization technique (box plot and scatter plot).

Q9 To handle the missing numerical data and perform one hot encoding on categorical data.

Q10 To implement Linear Regression.

Q11 To implement Logistic Regression.

Q12 To implement Naive Bayes classification.

Q13 To implement a Support Vector Machine.

Q14 To implement KNN(k-nearest neighbor).