# List of Practicals:

## Lab Project 1

| **Normalizing the Data** Write a code which works exactly like normalize(data). |
| --- |

## Lab Project 2

| **Cost Function** **Description**  Given below is the cost function.    Based on the given cost function, write a code for implementing it. |
| --- |

## Lab Project 3

| **Mapping Variables** **Description**  You’re given two lists, the first of which contains the name of some people and the second contains their corresponding ‘response’. These lists have been converted to a dataframe.  Now, the values that the ‘response’ variable can take are ‘Yes’, ‘No’, and ‘Maybe’. Write a code to map these variables to the values ‘1.0’, ‘0.0’, and ‘0.5’.  **Note:** It also might happen that the first letter of the three responses is not in uppercase, i.e., you might also have the values ‘yes’, ‘no’, and ‘maybe’ in the dataframe. So make sure you handle that in your code.  **Example:**  **Input 1:**  ['Reetesh', 'Shruti', 'Kaustubh', 'Vikas', 'Mahima', 'Akshay']  ['No', 'Maybe', 'yes', 'Yes', 'maybe', 'Yes']  **Output 1:** |
| --- |

## Lab Project 4

| Sigmoid Function **Description** ​The equation given below represents a model used to predict if a child is likely to choose commerce as their major in high school. The features of the model are x1 = marks in 10th class (use percentage value directly, e.g., for 60% marks, use 60), x2 = a boolean value representing whether some family member is a CA.  ​p=1/1+e-(0.005x1+0.5x2)​  For a particular child scoring 80% marks in class ​10th and their elder brother being a CA, find the probability of the child opting for commerce (Solve this by writing a code to find ‘p’). |
| --- |

## Lab Project 5

| **Gradient Descent with Two Variables** ​  Consider the minimisation of the function L(w1,w2)=w21+w22. We have already performed one iteration starting from w0=(5,−4) which moves to w1=(4,−3.2). Continue one more iteration with the same learning rate α=0.1. What is the value of W2? (Write a code to calculate W2) |
| --- |

## Lab Project 6

| Weights Calculation Description  A logistic regression model is trained on the following data set   | x  1 | x  2 | x  3 | **y** | | --- | --- | --- | --- | | 10 | -0.5 | 1 | 0 | | 55 | 0.3 | 0 | 1 |   During the gradient descent, the weights calculated in the 20th iteration are w1 = 0.04, w2 = -1, w3 = 0.4. Calculate the weights of the 21st iteration (assume the learning rate of 0.2, and use the batch-gradient algorithm). (Write a code to solve the problem)  **Hint:** Calculate the sigmoids of all the given data points, then calculate the loss for each data point and find the average. Lastly, find the updated weights. |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |

## Lab Project 7

| Predicting Class Label Description  Using the weights: w1=−0.025;w2=−0.957;w2=0.322​, predict the class of a data point with the following features:  ​x1=40; x2=0.6; x3=1​  Assume the threshold to be 0.6. Write the code to find ​σ and then based on the threshold classify the data point. |
| --- |

## Lab Project 8

| Evaluation Metrics Which among accuracy, sensitivity, and specificity is the highest for the model below?   | **Actual/Predicted** | **Not Churn** | **Churn** | | --- | --- | --- | | Not Churn | 80 | 40 | | Churn | 30 | 50 |   Write a code to find accuracy, sensitivity, and specificity and then input the given values. |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |

## Lab Project 9

| F1-Score There is a measure known as F1-score, which essentially combines both precision and recall. It is basically the [harmonic mean](https://en.wikipedia.org/wiki/Harmonic_mean) of precision and recall, and its formula is given by:  ​F=2×(precision×recall)/(precision+recall​)  The F1-score is useful when you want to look at the performance of precision and recall together.  Write a code to calculate the F1-score for the model below.   | **Actual/Predicted** | **Not Churn** | **Churn** | | --- | --- | --- | | Not Churn | 400 | 100 | | Churn | 50 | 150 | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |

## Lab Project 10

| Bayes Theorem Bag A contains 3 Red and 4 Green balls, and bag B contains 4 Red and 6 Green balls. One bag is selected randomly, and a ball is drawn from it. If the ball drawn is found Green, find the probability that the bag was chosen was A. (Write a code for Bayes Theorem and input the given values) |
| --- |

## Lab Project 11

| Distance Measures Consider points, A(7,50) and B(23,34). Write a code to compute the Euclidean and Manhattan distance between the two. [Round off the answer to two decimal places] |
| --- |

## Lab Project 12

| Hopkins Statistic Write a code defining Hopkins statistic function which checks whether the dataset is suitable for clustering or not when a dataframe is passed through it. |
| --- |

## Lab Project 13

| **Cricket Clustering** **Description**  Given a dataset [here](https://drive.google.com/file/d/1-aCZn6XqDasbPLKbvSiibTD8YpAio7j4/view?usp=share_link) about the batting figures of batsmen in ODI matches.  Choose the number of clusters as four. Does ​SR Tendulkar fall in the same cluster as Virat Kohli? |
| --- |

## Lab Project 14

| **Cricket Groups** **Description**  Given a dataset [here](https://drive.google.com/file/d/1-aCZn6XqDasbPLKbvSiibTD8YpAio7j4/view?usp=share_link) about the batting figures of batsmen in ODI matches.  Based on the clustering, given that the clusters formed are (high SR, high Ave) - A, (low SR, low Ave) - B, (High SR, Low Ave) - C, (Low SR, High Ave) - D. Who all fall in group A? |
| --- |

## 