Assignment 4
Logistic Regression

In this assignment, you need to classify the given dataset using logistic regression.

Database
Independent/Predictor Variable -

https://drive.google.com/file/d/1eEBjqnIMlViYOwoOBqZ11tMbn9cZLLEo/view?usp=drive_link

Dependent/Response Variable -

https://drive.google.com/file/d/1hnowuwHw_qgRl1a0LLa6y0eItyMU_lZ0/view?usp=drive_link

Use the same cost function that we had discussed in class.

Answer the following questions based on your observations

1. Use logistic regression to find decision boundary For the given database. Set your learning rate to 0.1. What is the cost function value and learning parameter value after convergence?

2. Plot cost function v/s iteration graph for the model trained in question 1. Plot the line as shown here - https://pythonguides.com/matplotlib-plot-a-line/#Matplotlib_plot_a_line_chart
Do not use scatter plots for this.

3. Plot the given dataset on a graph, use different colours for different classes and also show the decision boundary you obtained in question 1. Do not use scatter plot.

4. Train your model with a learning rate of 0.1 and 5. Plot the cost-function v/s iteration curve for both learning rates on

the same graph. For this task, only train your model for 100 iterations.

5. Find the confusion matrix for your training dataset. Using the confusion matrix to calculate the accuracy, precision, recall, F1-score.

Note
1. Use batch gradient descent unless mentioned in the question.
2. Your code should be in Python.
3. You are required to implement the model yourself. You are not allowed to use libraries like sci-kit learn, Tensorflow, Pytorch, etc, and use their built-in models.
4. You may use libraries like numpy, pandas, matplotlib, etc to read and manipulate the dataset and plot the graphs.
5. Your plots should be labelled properly.
6. Normalize your independent variables befores training.

How to Submit?
1. You need to submit a single PDF report file and the name of that file should be your roll numbers.
2. Your report should have answers to all the questions with the plots. Upload your code on github and mention the link in the report. Do not copy paste code in the report.