Assignment 1 – Logistic Regression

You are required to implement logistic regression to predict whether a patient has heart disease or not based on some features.

1) **Data**

   - The attached dataset "heart.csv" contains 303 records of patients who have heart disease or don't have heart disease according to the 13 features in it.

   - The detailed description of the dataset is attached in "heart_dataset_description.pdf".

2) **Requirements:**

   Write a python program in which you:

   a- Use these 4 features **(trestbps, chol, thalach & oldpeak)** as input and the output to be predicted is the **"target".**

   b- Split the dataset into **training and testing** sets.

   c- Use **sklearn's logistic regression** to build a model that fits the training data. Then, use that model to predict the output of the test data and print the error of the model.

   d- Implement logistic regression (from scratch) using **gradient descent** to optimize the parameters of the hypothesis function.

   e- Use the optimized hypothesis function to **make predictions** on new data.

   f- **Calculate the cost (error)** in every iteration to see how the error of the hypothesis function changes with every iteration of gradient descent.

   g- **Plot the cost** against the number of iterations.

3) **Notes:**

   - You may need to try different values of the learning rate and see how this changes the accuracy of the model.

   - You may need to normalize the feature data before applying logistic regression. You can use minmax normalization where z is the normalized value and z = (x – min) / (max – min).

4) Submission Remarks:

- The **maximum** number of students in a team is **3** and the minimum is 2.

- **No late submission** is allowed.

- **Cheating students will take ZERO** and no excuses will be accepted.