Assignments on GDA #5 (Ref Video Lectures 21-22)

(Timely submission of assignments is essential. Copying/plagiarised submission from others will fetch fail (F) grade on this subject)

You have been provided with Microchip data with two different quality assurance test results. The third digit tells you whether the microchip has passed the quality assurance test (1 means pass, 0 means fail) or not.

1. Using raw data set as given, create three more features, and from there develop a GDA model. Thereafter, utilize the same to predict whether a Microchip component will be accepted or rejected. May use 70% data for training and 30 % data for testing.
2. Using the same data set and features, and same 70% of the data for training and 30% for testing, now use Box-Muller transformation to create a new data set having Gaussian distribution within the range of the given data set and create a new Gaussian Discriminant Analysis (GDA) model. Thereafter, utilize the model to predict where a component will be accepted or rejected using the testing data.
3. Compare the performance of GDA in both the above cases and write a comparative analysis report with results.

Deadline for submission: October 2nd, 2020, mid night.

Full marks: 50+50=100