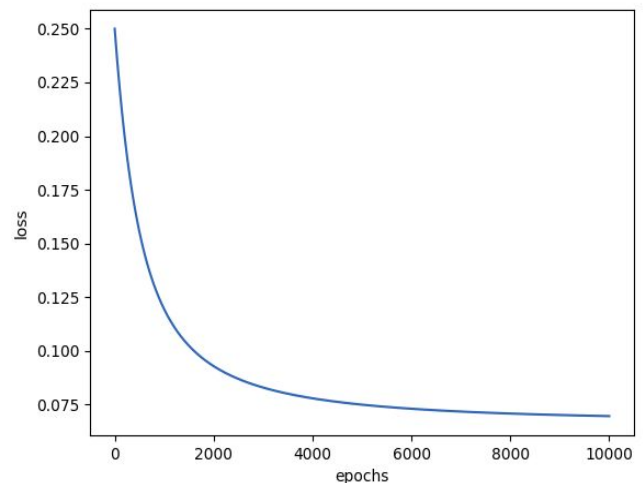
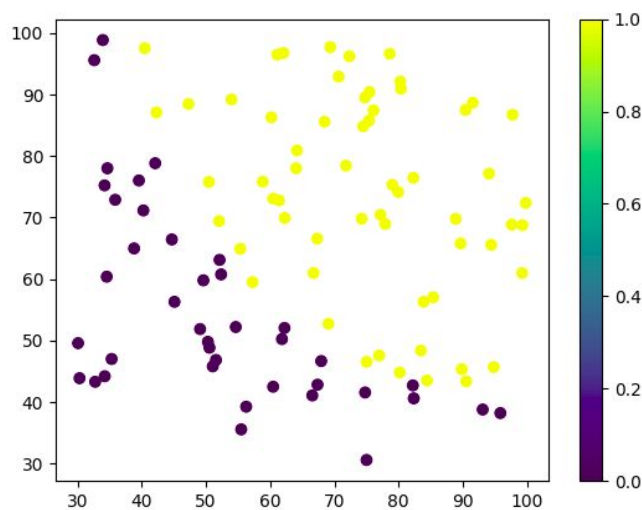


## Report on Assignment #3

**Some observations after implementing the logistic regression algorithm on both the datasets.**

**Examination dataset:**

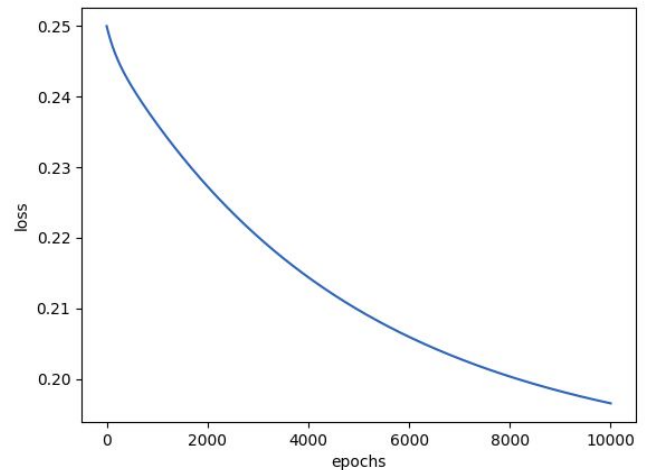
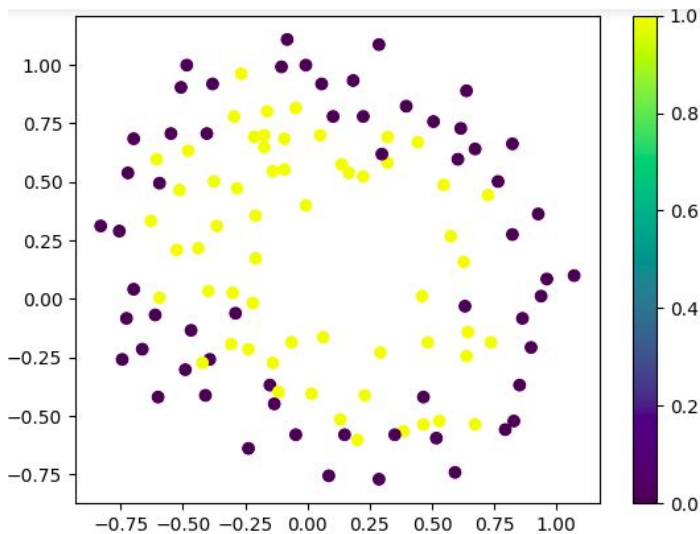
1. For the examination marks dataset the inputs were in order of 10's (like 70, 80 etc), when logistic regression was applied on this unprocessed data it was found that the loss in the descent was chattering very frequently. This behaviour can be attributed to the fact that the data as it was not very close to prediction value the results very getting very sensitive to the inputs.
2. The problem was solved by normalizing the data between 0 and 1. After normalization it was observed that the GDA was running very smoothly and showing convergence after around 5000 epochs.
3. Accuracy achieved was very good which was around 90%-95%.



Figures:- dataset and loss vs epochs

## Microchip testing dataset:

1. For this data it was observed from visually plotting the data that the data was circular it was concluded that the data cannot be linearly separated. When normal( $w_0 + w_1x_1 + w_2x_2$ ) data was used to apply logistic regression it showed poor results owing to its circular nature. Shown in figure below.
2. The problem was that we were trying to define a decision boundary that was a straight line for circular data. This was solved fitting a ellipse to the data so that better classification can be done.the equation used was  $w_0 + w_1X_1^2 + w_2X_2^2$ .
3. Accuracy achieved was very good which was around 80%-85%.



Figures: dataset and loss vs epochs