DATASET	METHOD	DATA PREPROCESSING	REGULARISATION	ACCURACY	COMPUTATION COMPLEXITY
EXAM DATASET	NEWTON'S METHOD	Normalization	None	~93%	HIGH
MICROCHIP DATASET	NEWTON'S METHOD	None	None	~72%	High

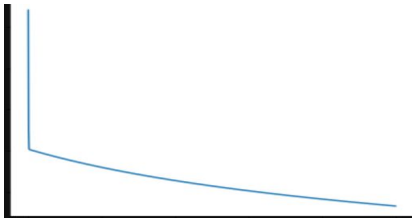
ANALYSIS (EXAM DATASET) :

Exam data set has two feature column. The data is linearly separable.

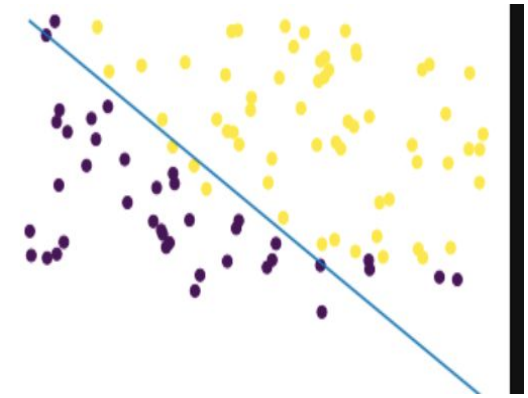
Dataset Scatterplot

Used same features as given in the dataset with normalizations.

The model converged with very few epochs, but as we have to calculate the inverse for hessian matrix, the computation complexity was very high and training time was high even for 100 epochs.



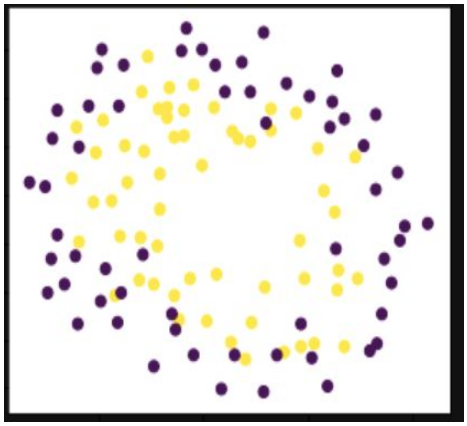
Loss function plot



Decision Boundary Plot

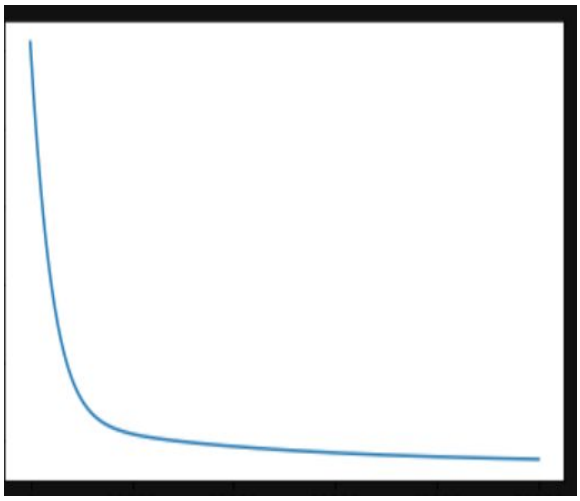
ANALYSIS (MICROCHIP DATASET) :

Microchip Dataset is non linear with two feature column.

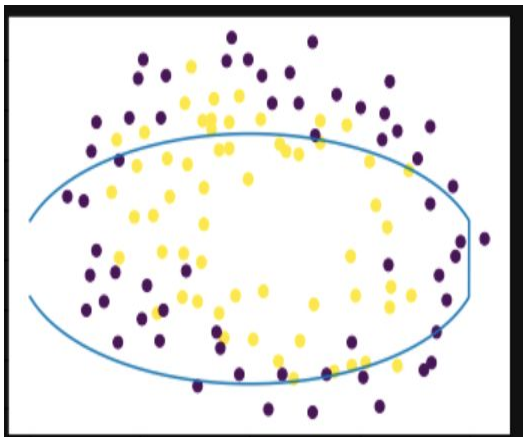


Dataset Scatterplot

As the dataset is not linearly separable, used the square of columns as features.
The model convergence was slow (~1000 epochs) and computational complexity was very high.



Loss function plot



Decision Boundary plot