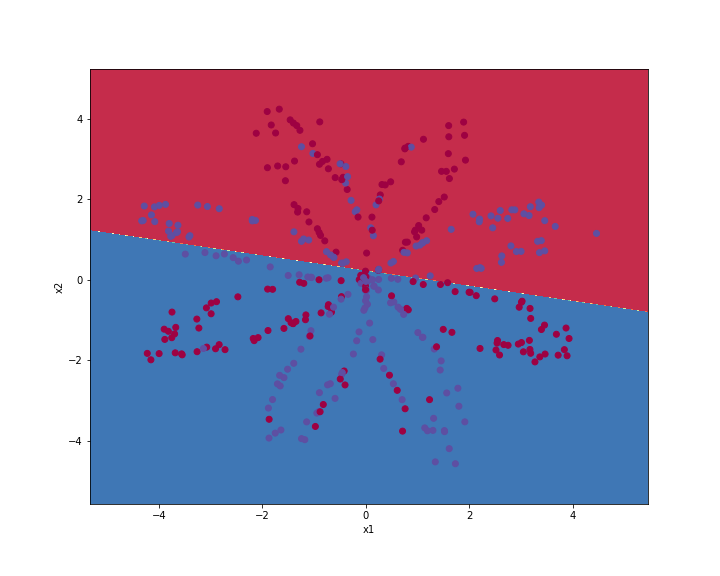
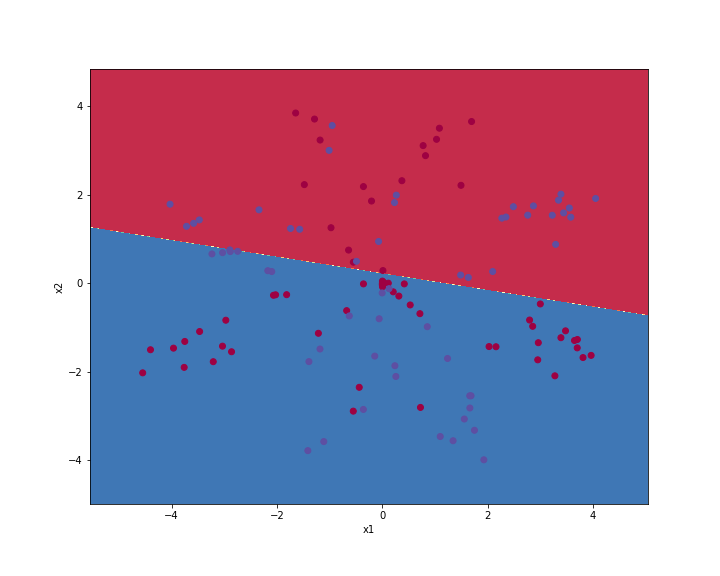
Neural Networks

a) implemented neural network

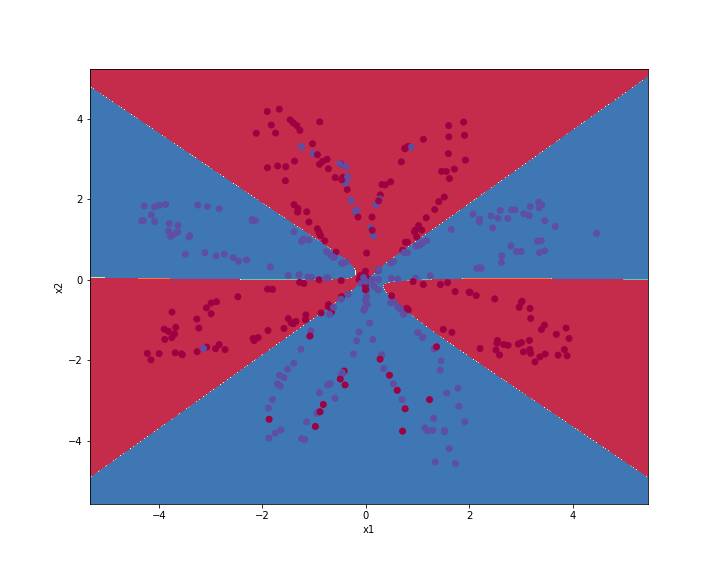
b) i) Train accuracy is: 0.457894736842

Test accuracy is : 0.383333333333



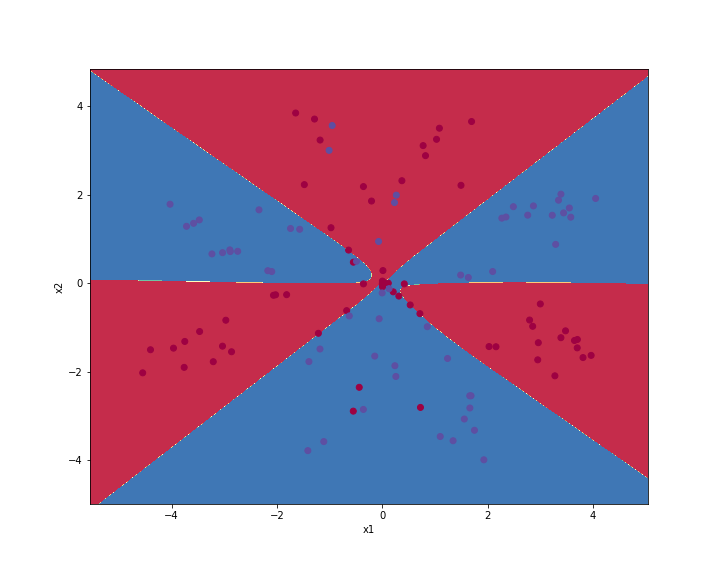


ii) Train accuracy is 0.9



Test accuracy is :- 0.875

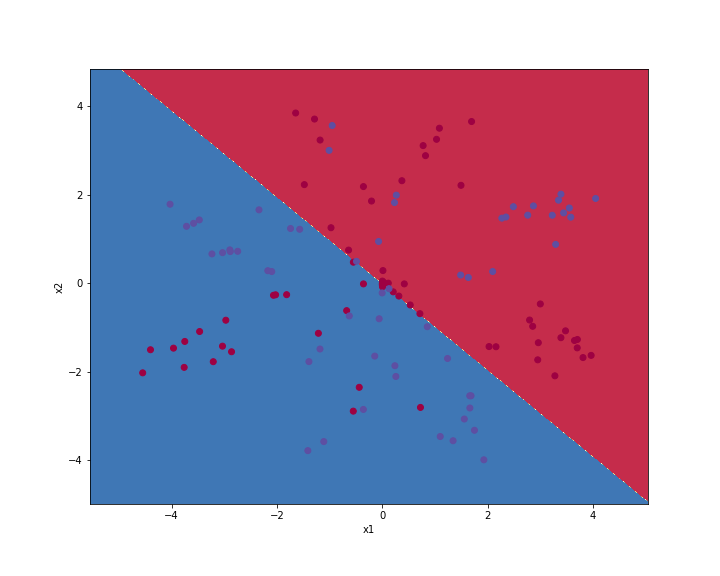
The accuracies increased significantly in comparison to logistic regression.



iii) hidden layer nodes 1:

train accuracy : 0.6342105263157894

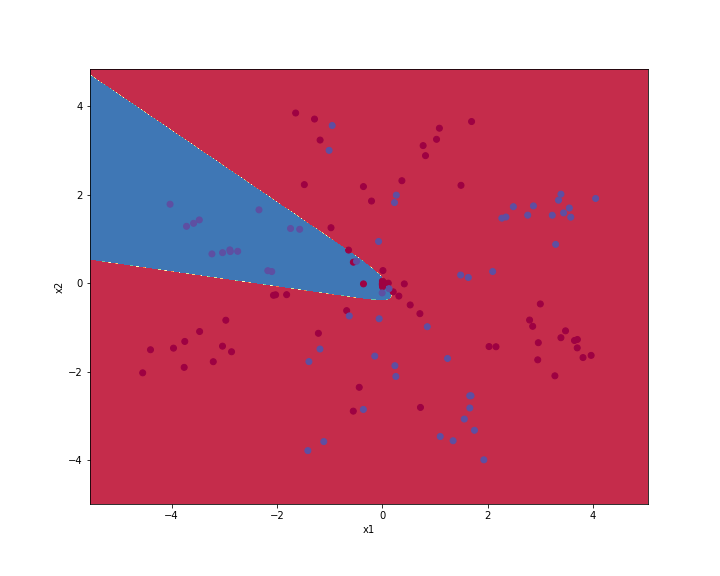
train accuracy : 0.625



hidden layer nodes 2:

train accuracy : 0.618421052631579

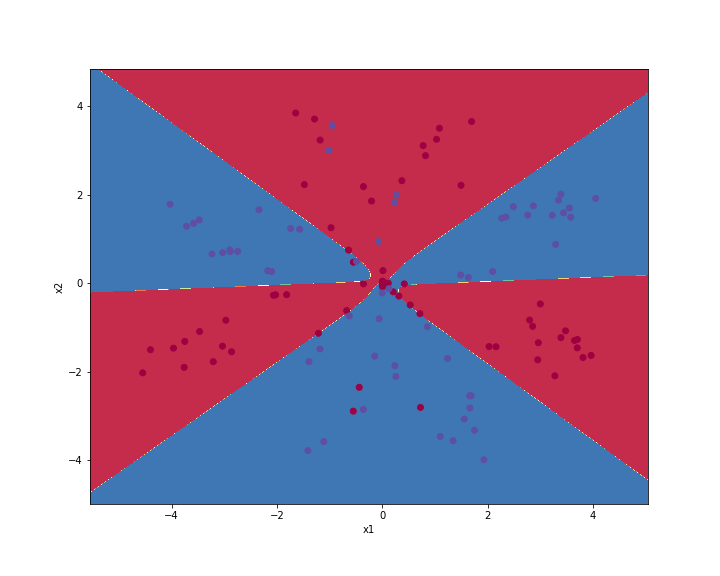
test accuracy : 0.6



hidden layer nodes3:

train accuracy: 0.8921052631578947

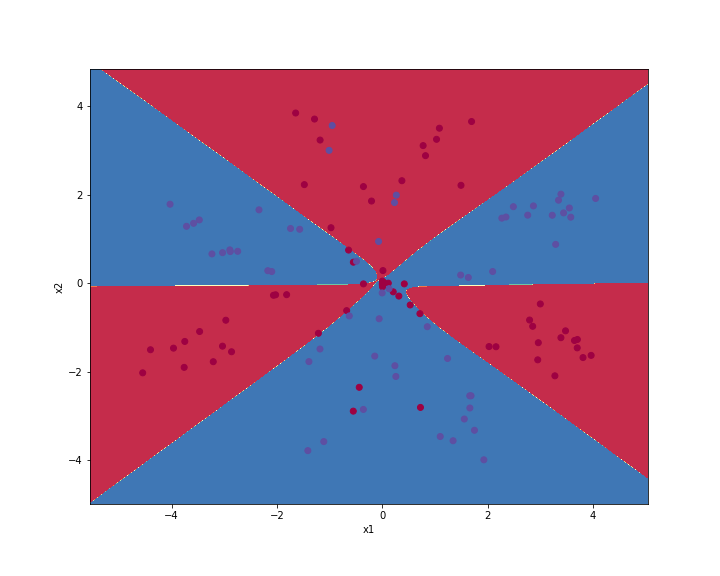
test accuracy: 0.8583333333333333



hidden layer nodes: 10

train accuracy : 0.8868421052631579

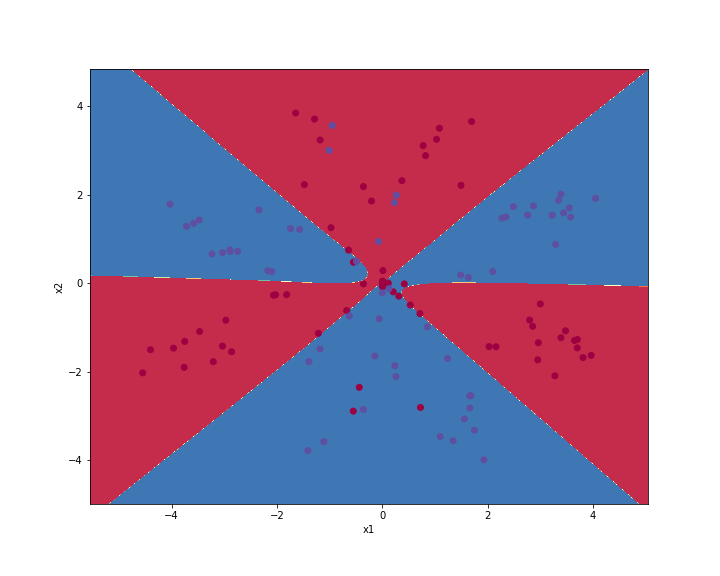
test accuracy : 0.825



hidden layer nodes: 20

train accuracy : 0.9

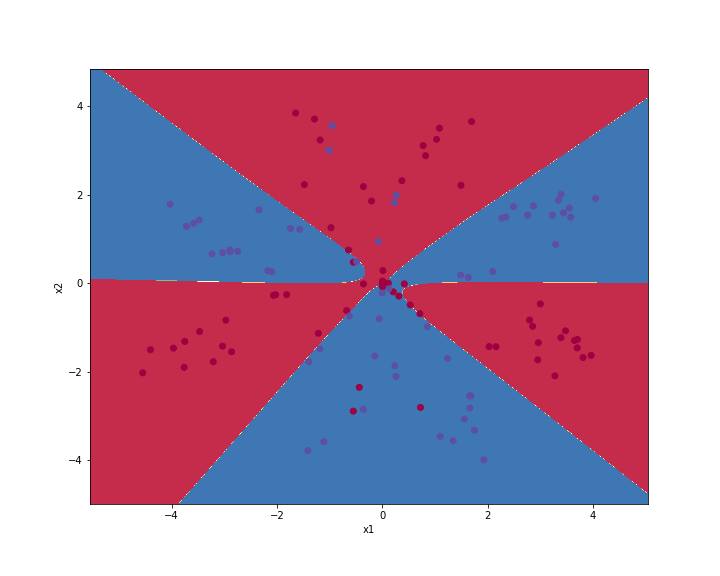
test accuracy : 0.875



hidden layer nodes: 30

train accuracy : 0.8973684210526316

test accuracy : 0.875



As we increase the number of units the accuracies increased gradually and then became constant.also the decision boundary changed from linear to nonlinear , classifying data properly.

Optimal number of units I found were either 5 or 10 having same accuracies.

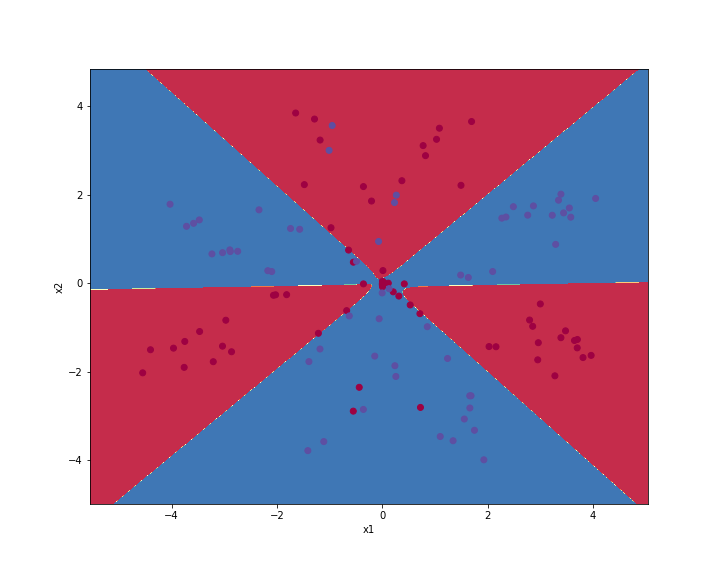
iv) 2 hidden layer having 5 nodes each

train accuracy : 0.8921052631578947

test accuracy : 0.8333333333333334

results remain same in comparison to single hidden layer with 5 nodes.

Double hidden layer with 5 nodes is equivalent to single hidden layer with 5 nodes



C)

i)libsvm:-

train accuracy :99.8

test accuracy:98.8

single perceptron:-

train accuracy : 0.5

test accuracy : 0.5041407867494824

ii) stopping criteria:-

difference in loss values at new thetas should be less than eita, where eita = math.exp(-9)

loss function :- negative log likelihood

learning rate: 5

train accuracy : 0.86825

test accuracy : 0.8405797101449275

iii)

loss function negative log likelihood

learning rate: 0.006

train accuracy : 0.9265

test accuracy : 0.8659420289855072

accuracy with reLU increased by 6 percent in training data and 2.5 percent in test data.