ML Assignment 3
Abhishek Agarwal
2016126

Ques1)

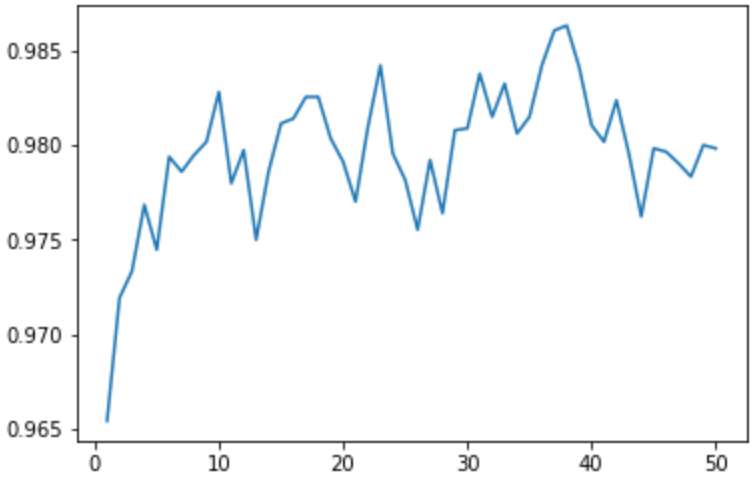
Learning Rate: **0.01**
Epochs: **50**
Batch Size: **20**
Accuracy obtained for:

- 1 hidden layer
    - Train: **0.9850877192982456**
    - Test: **0.9775517362329007**

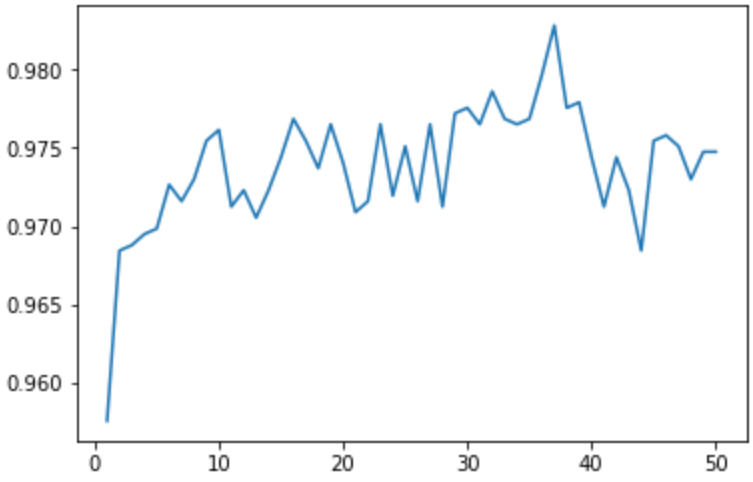
- 3 hidden layers
    - Train**: 0.883070175438596**
    - Test: **0.8877586811645037**

Best weights and biases have been saved in “weights_1.pickle” , “bias_1.pickle” and “weights_3.pickle” and “bias_3.pickle” respectively.

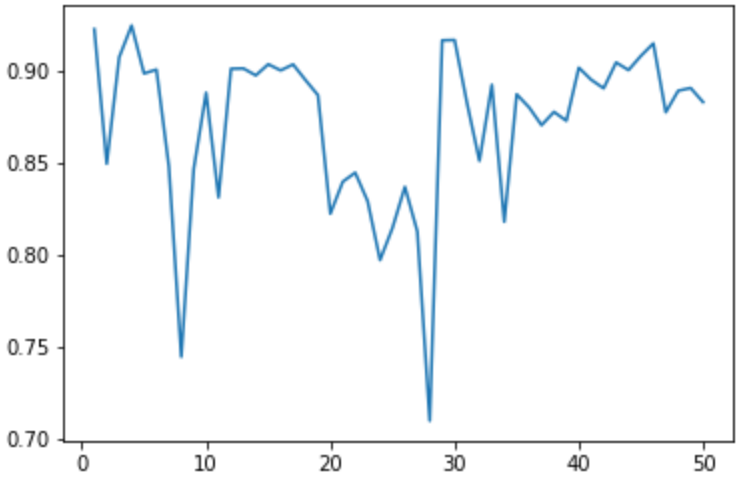
- **Train accuracy graph with Number of Epochs for 1 hidden layer**



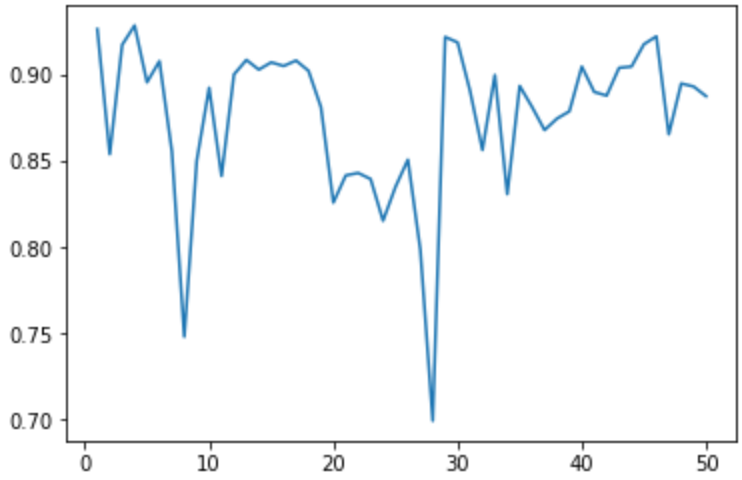
- **Test accuracy graph with Number of Epochs for 1 hidden layer**



- **Train accuracy graph with Number of Epochs for 3 hidden layers**



- **Test accuracy graph with Number of Epochs for 3 hidden layers**



Since the train and test accuracies are very close, there is no underfitting or overfitting of data.

<u>Challenges Faced along with counter measures</u>:
Correcting the dimensions of the matrices while doing back propagation was the major challenge.Deciding the learning rate was also challenging, on taking it to be 0.1, the accuracy was fluctuating between 50 and 90 again and was not giving good results. While deciding the alpha for 5 layer neural network, high alpha had to be chosen as for low alphas it was giving a very low accuracy and learning very very slowly. The training time required for 3 layer Neural Network was a lot more as compared to 1 layer, thus I couldn't test much values of alpha and obtained a graph with fluctuating accuracies.

(d) Using Sklearn MLP Classifier:
Accuracy obtained is reported as follows:
Accuracy obtained for:

- 1 hidden layer
  - Train: **0.9925438596491228**
  - Test: **0.9845668186601193**

- 3 hidden layers
  - Train**: 0.9810526315789474**
  - Test: **0.9740441950192915**

The difference in accuracies observed is as expected. In inbuilt methods, the initialization of weights and learning rate must take place in an appropriate manner contrary to our hit and trial.

Ques2)

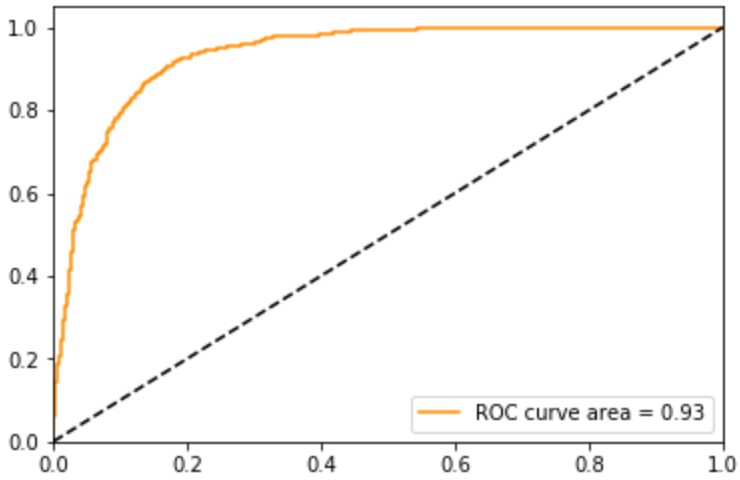
Grid Search CV was applied to obtain the best C value of 0.01

Test accuracy obtained: **86.5%**
Train accuracy is **88%**
Confusion Matrix:

	Class 0	**Class 1**
Class 0	825	175
Class 1	95	905

ROC Curve:

- Test:



- **Train**