Assignment 3

Machine Learning

- Mayank Chauhan, MT18008

1.   Neural Networks

1.1. Implementation of Backpropagation and Forward pass

Details about the dataset: Images of digits (7&9). There are total 14251 images of size 28*28. The dataset was split into the following training, validation and test set.

Set	#samples
Train	11414
Validation	1411
Test	1426

Preprocessing: Flatten the images to vector of length 784 (features). Then, apply MinMaxScaler using sklearn. The range is features is [0, 1].

Parameters :

No. of hidden layers: 3 [100, 50, 50].

No. of units in output layer: 2

No. of units in input layer: 784

Learning rate: 0.01

No. of epochs: 10

Optimization: Stochastic Gradient Descent.
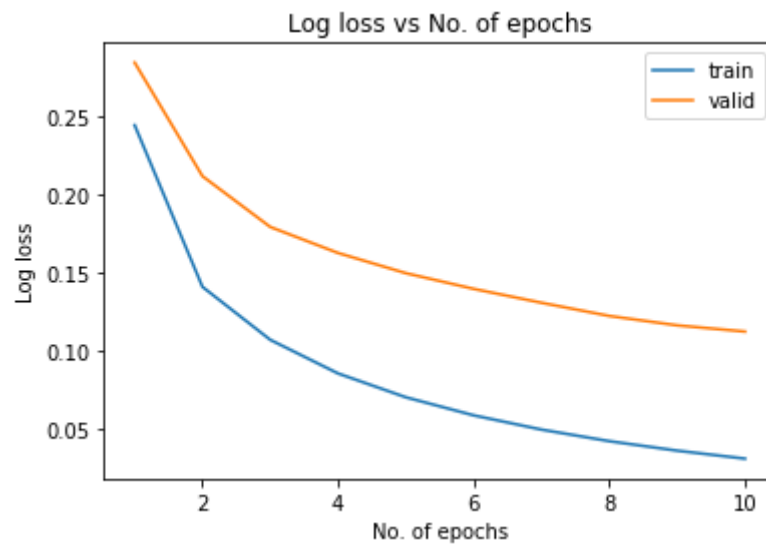
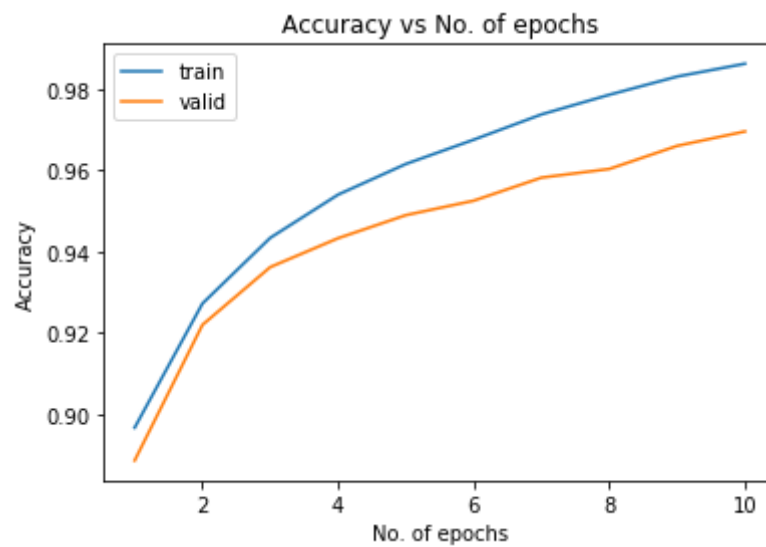## 1.2. Loss and Accuracy Graphs

Figure 1: Loss vs No. of epochs

Figure 2: Accuracy vs No. of epochs

Final results on Training, Validation and Test set

Table 1: Results obtained after using model trained on 10 epochs

Set	Accuracy	Log loss
Training	0.9861573506220431	0.035344691167555155
Validation	0.969525159461375	0.11212213213713688
Test	0.9719495091164095	0.07762462637156711

Challenges

- Numerical overflow problem in using standard softmax function. To avoid this, exp-normalize trick is trick is used from the below link.
  https://timvieira.github.io/blog/post/2014/02/11/exp-normalize-trick/
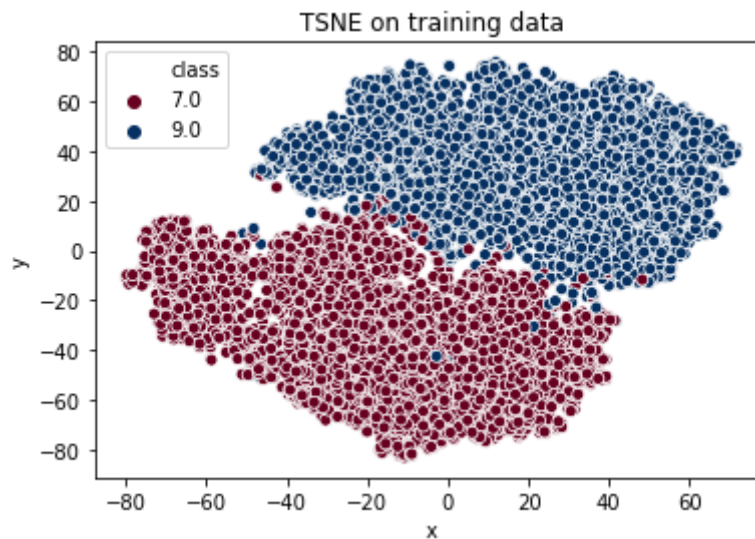
1.3. Visualizing features

Figure 3: Visualizing features obtained from last hidden layers using tSNE

1.4. Implementation using sklearn

Table 2: Results using MLPClassifier

Set	Accuracy	Log loss
Training	0.9874715261958997	0.04570688745988210
Validation	0.9744861800141743	0.07954908733625361
Test	0.9831697054698457	0.05498110216309294

The learning rate was set to constant value 0.1. The no. of epochs is same, as used in part 1.1. Solver was set to 'sgd' and activation = 'logistic'. The sklearn implementation used other various things like batch size and regularization, so we cann't compare our results with the model as it is.

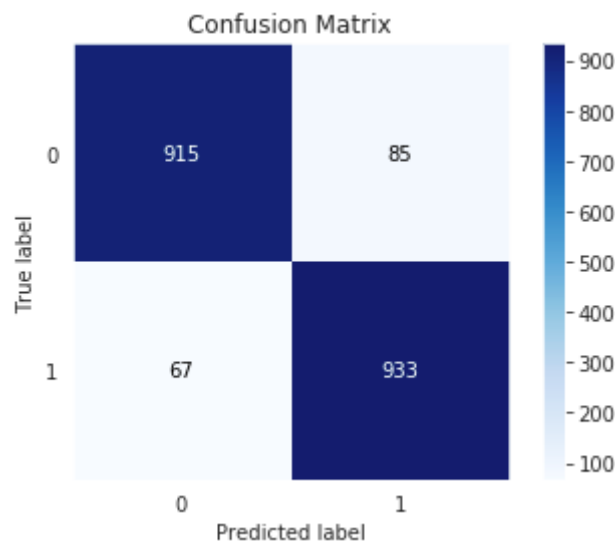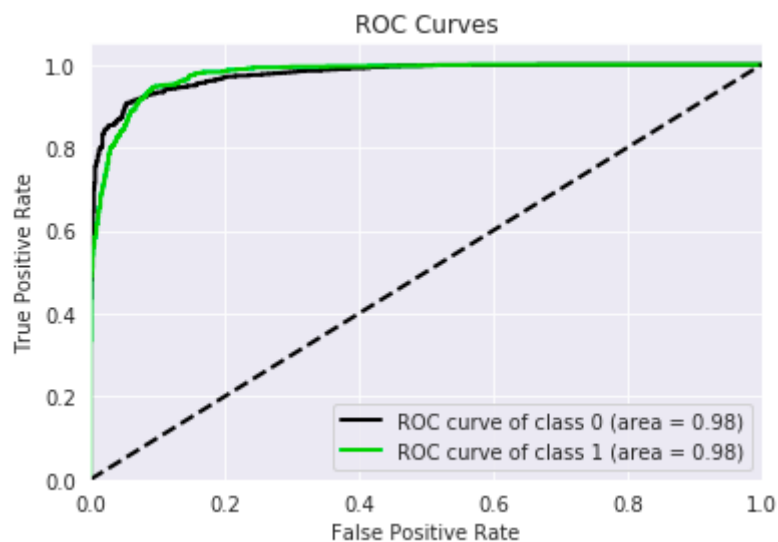## 2. AlexNet

Test accuracy: 0.924

Figure 4: Confusion matrix

Figure 5: ROC Curve along with AUROC values