Assignment 5: Classification and Neural Networks

Abdul Maroof (16030010)

December 6, 2017

1 Problem 0

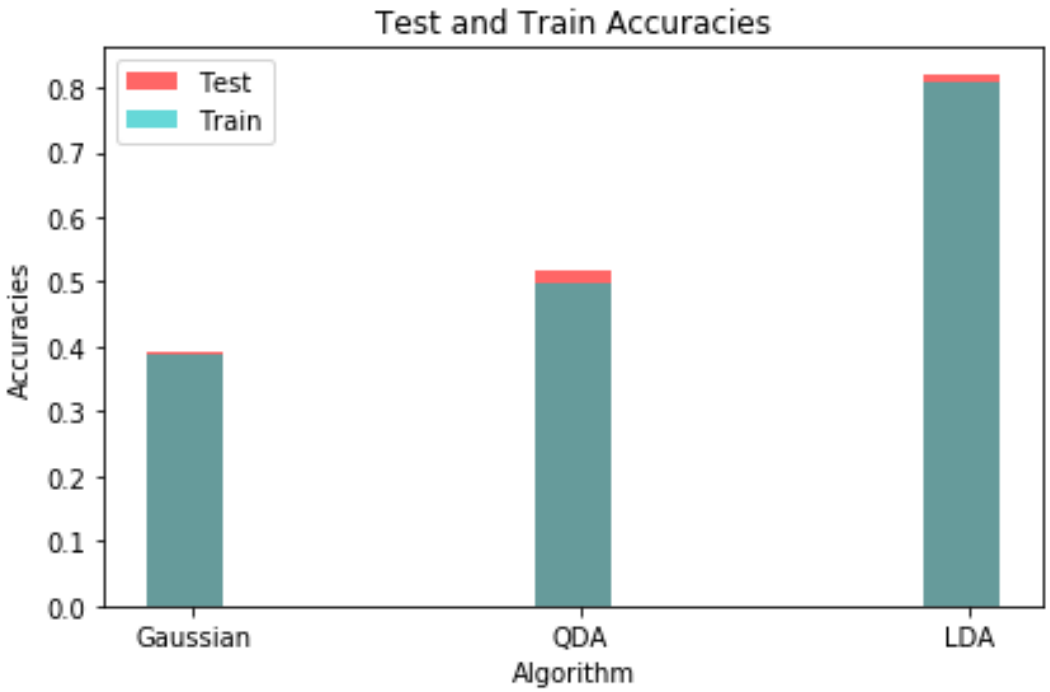

Figure 1.1: Gaussian vs QDA s LDA

1.1 Comment

Fisher Linear Discriminant is best as we know that this is a 2 class problem & we used Gaussian Naive Bayes classifier & not Multinomial Naive Bayes because the features are continuous. The best results were shown by Fisher Linear Discriminant as compared to Gaussian Naive Bayes & Quadratic Discriminant. GNB gave the worst results as Naive Bayes expects discrete values & refers to conditional independence of each of the features in the model.

Question 0 ends here.

2 Problem 1

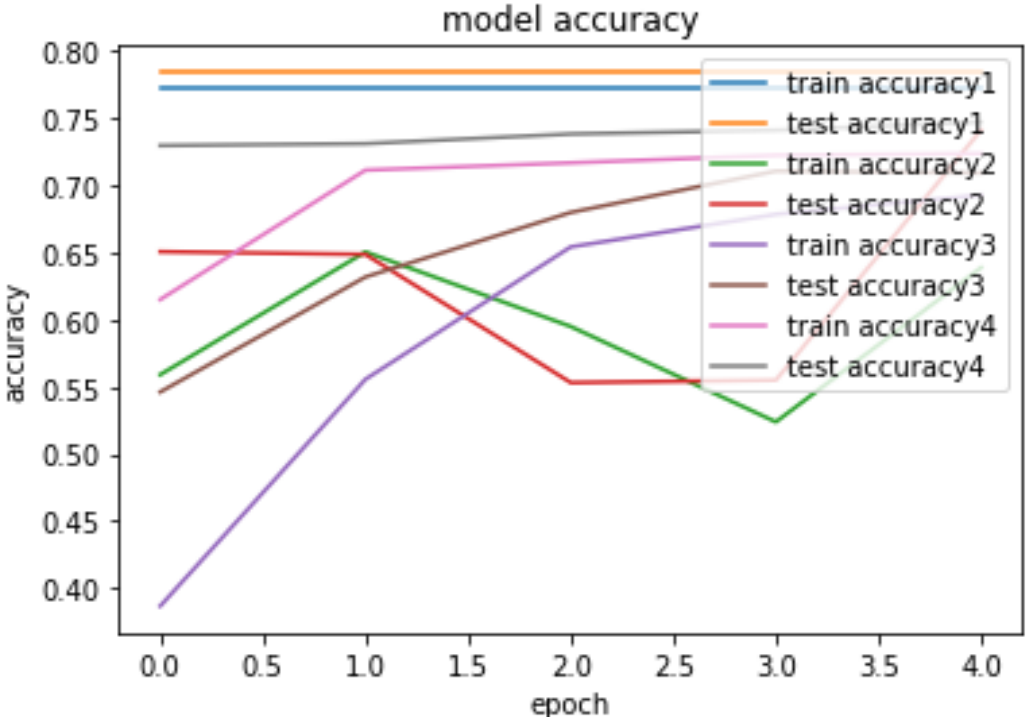

Figure 2.1: Accuracy Graph of our experiments

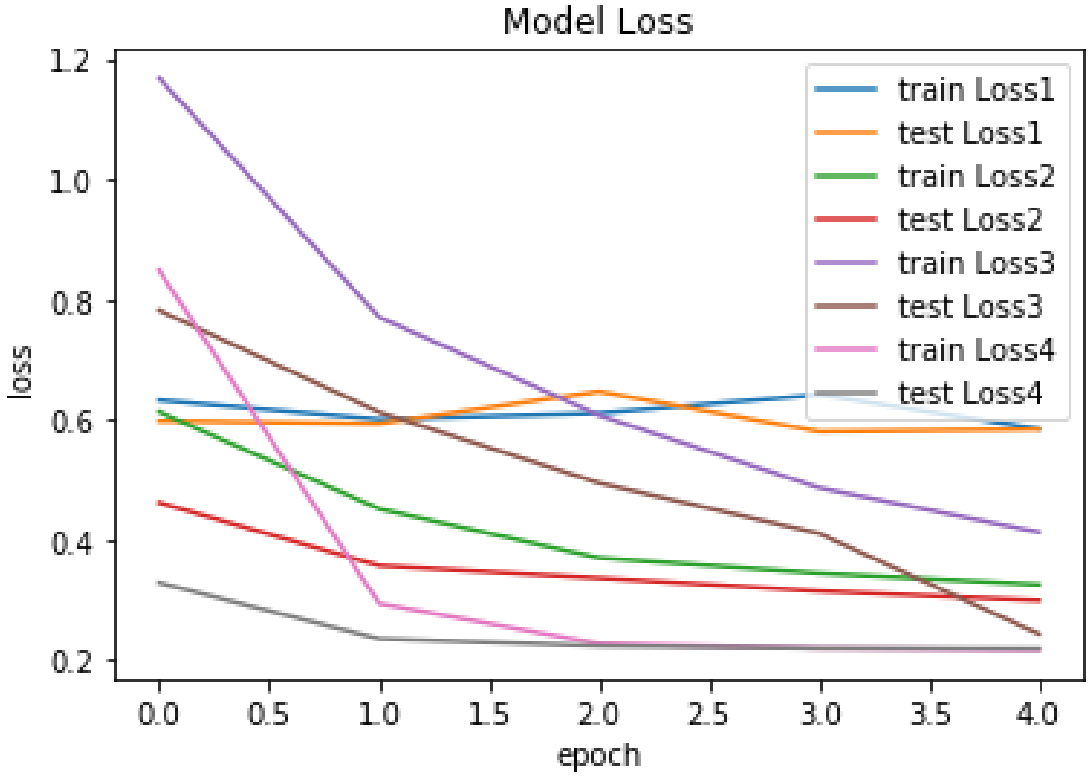

Figure 2.2: Loss Graph of our experiments

2.1 Comment

In most of the cases we see that testing accuracy is slightly better than the training accuracy. Adding hidden layers decreases our accuracy but increasing the output units with hidden layers gives better results as compare to adding layer with less output units.

Question 1 ends here.

---

3 Problem 2

All the functions are implemented and could be found in $functions$ folder. We have used the fields of downloaded data that were most likely to help us detecting duplicate entries. Two fields that are selected for this purpose are $titles$ & $description$. However other fields can be added but it did'nt help much.
We have set the random seed generation to 100 for reproducibility of the results. Following table summarizes the results obtained with accuracy more than 80%.

Table 3.1: Accuracy Graph of our experiments

-	FeedForward Network	Convolutional Network
Train Accuracy	0.8	0.9
Validation Accuracy	0.9	0.95
Test Accuracy	0.25	0.55

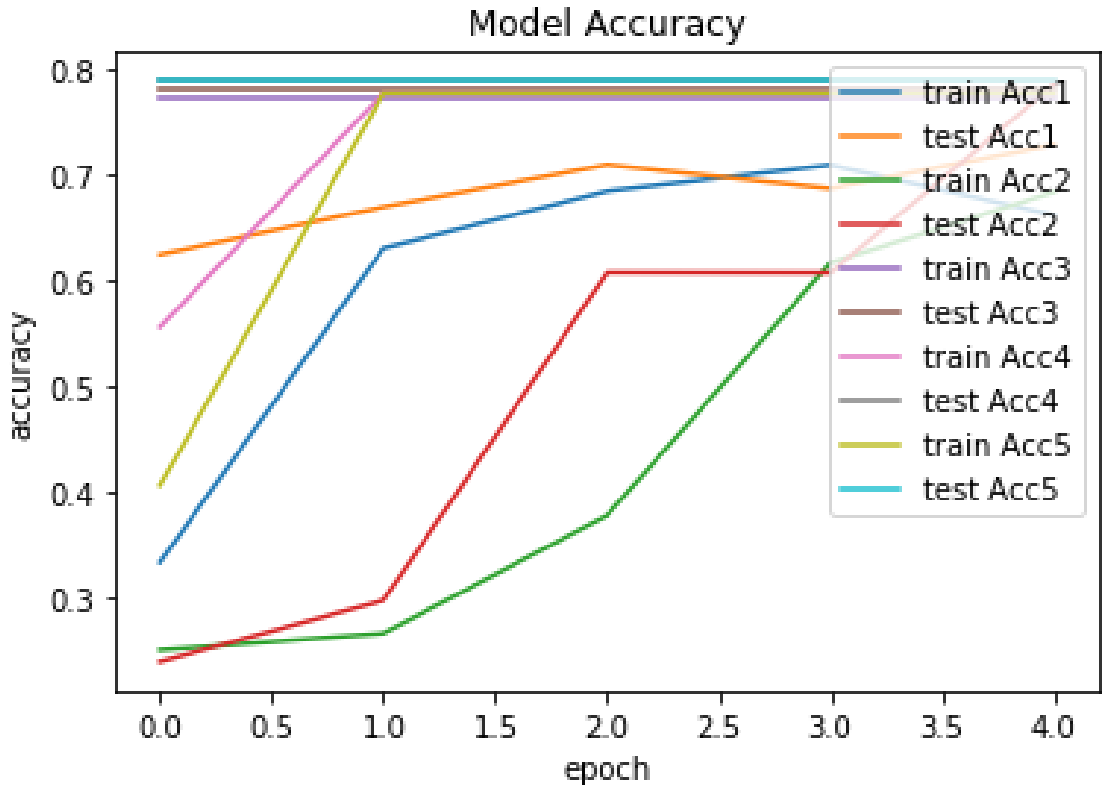

Figure 3.1: Accuracy Graph of our experiments

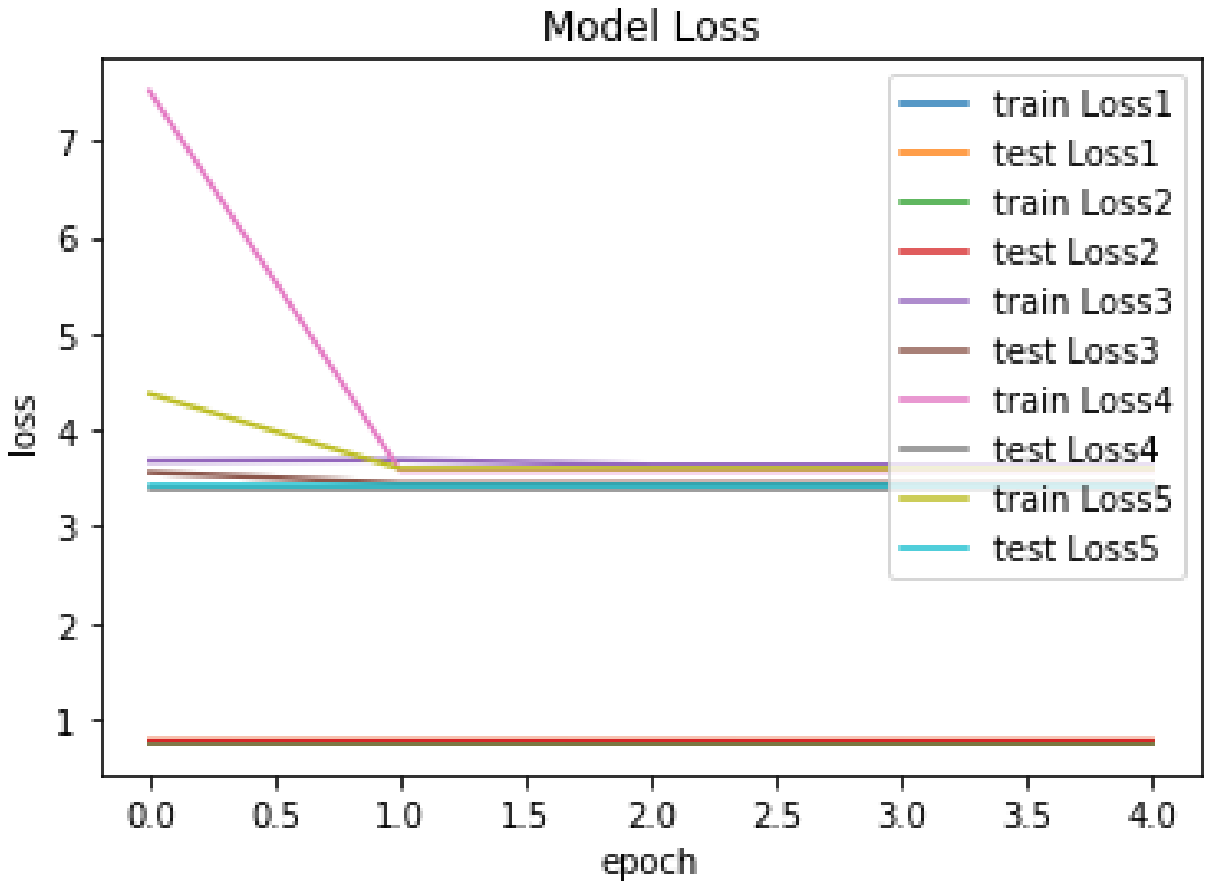

Figure 3.2: Loss Graph of our experiments

3.1 Comment

Combination of Relu(hidden) and Softmax(output) with Mean Square Error gives better results. as we have also done one hot encoding in softmax as it is used for continuous values. and also in this case we have used categorical cross entropy as it's for more then 2 classes we have use softmaax

---

Question 2 ends here.

4 Problem 3

Everything Regrading Problem 3 goes here.

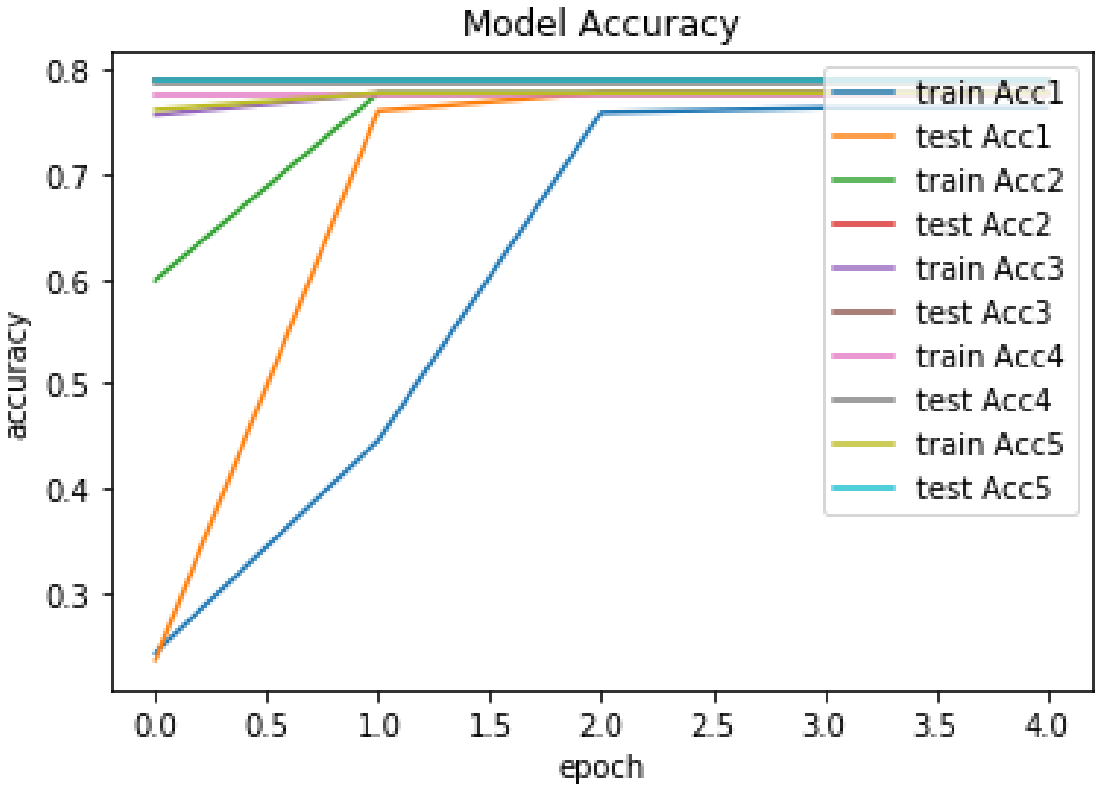

Figure 4.1: Accuracy Graph of our experiments

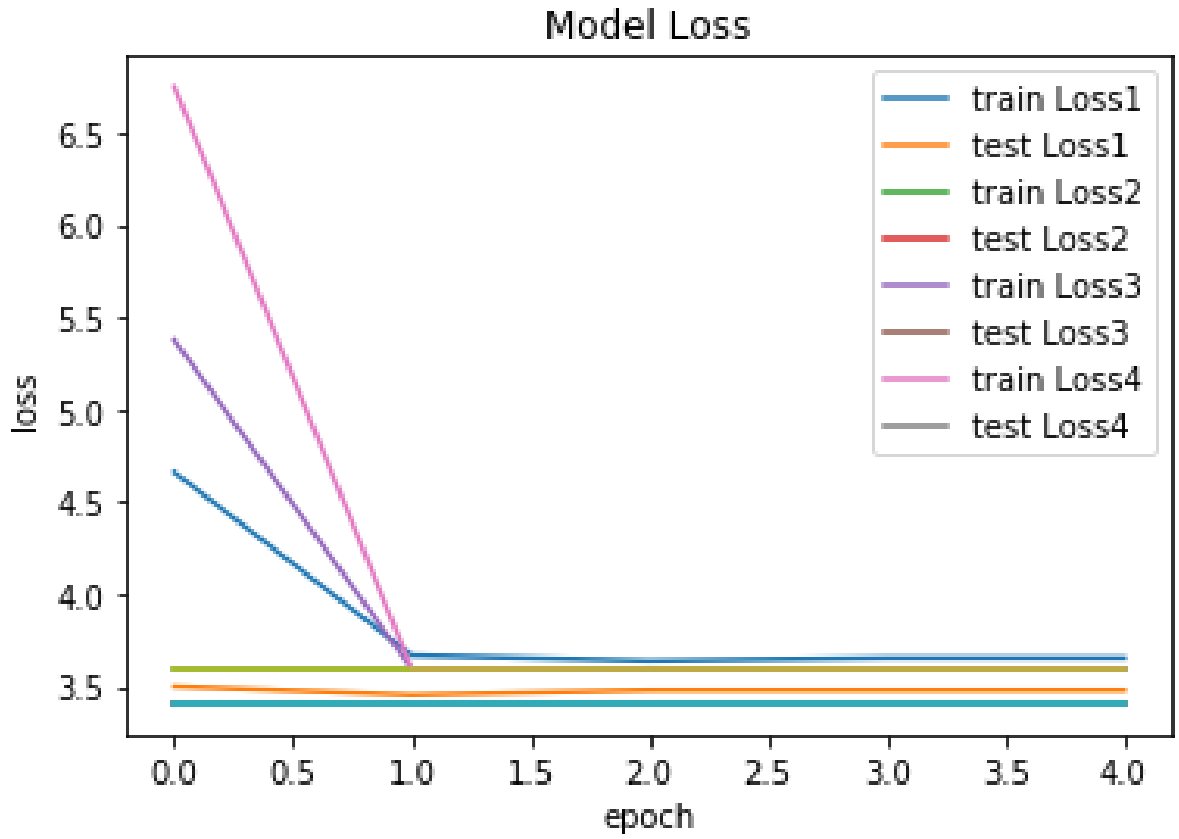

Figure 4.2: Loss Graph of our experiments

4.1 Comment

The accuracy & loss coverged eventually for all variaites of momentun & learning rate so this isn't has much affect on our model

Question 3 ends here.

5 Problem 4

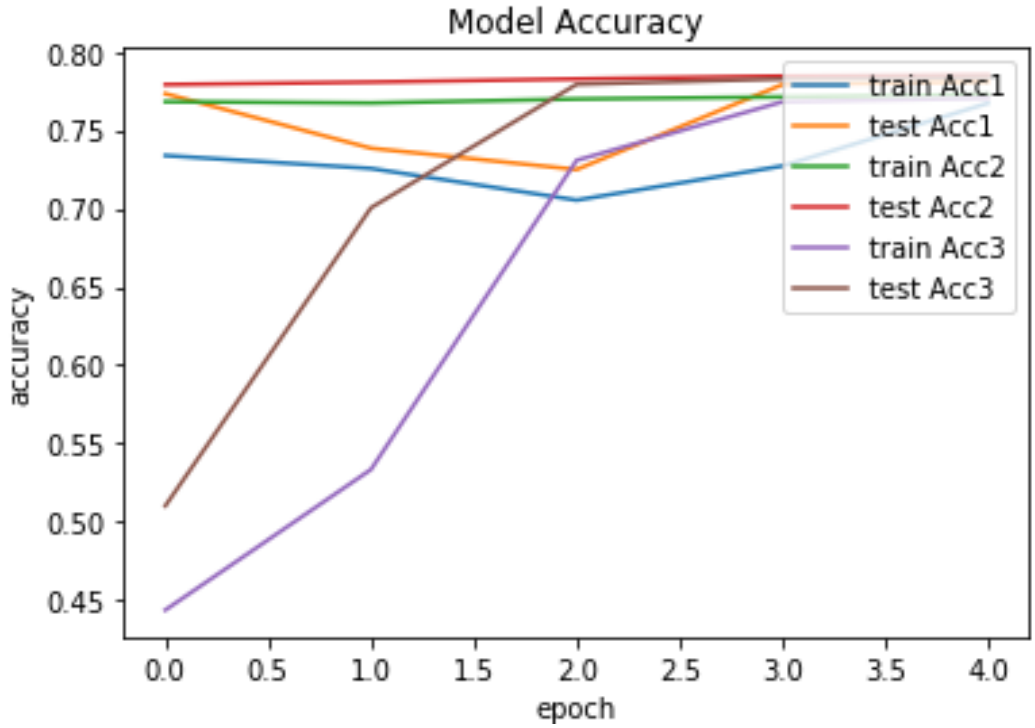

Figure 5.1: Accuracy Graph of our experiments

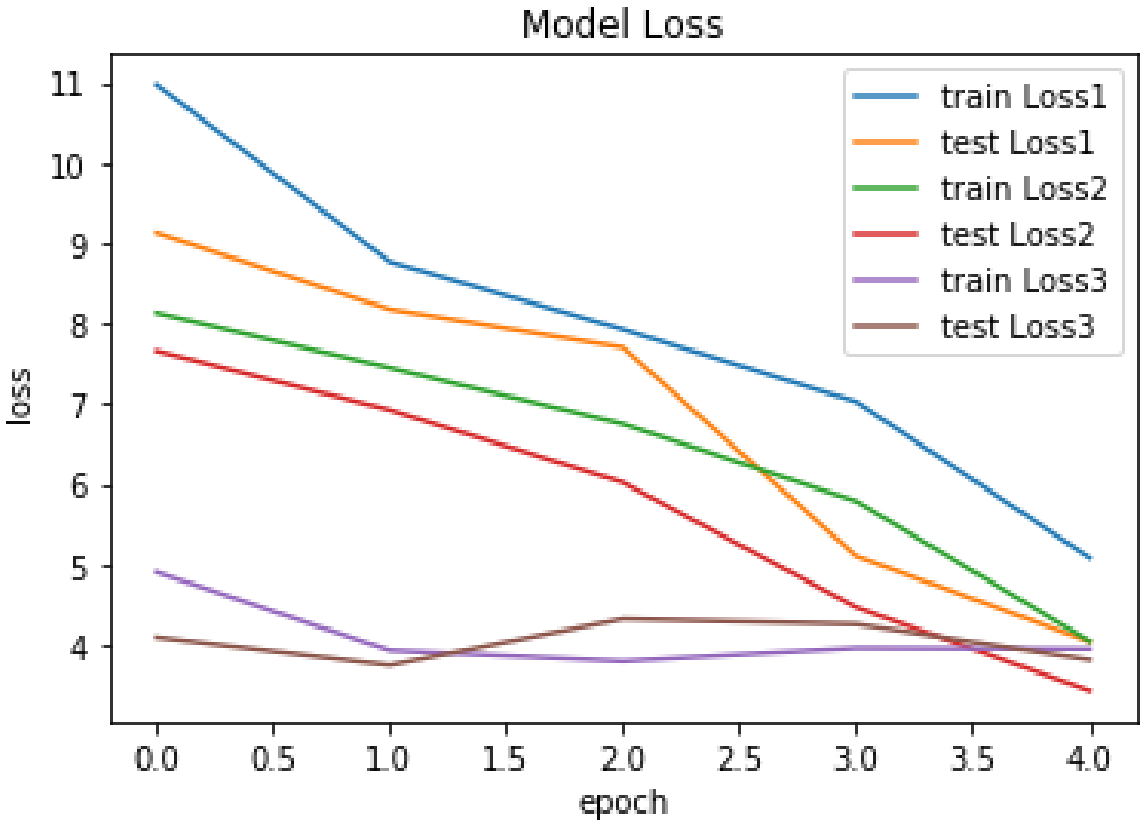

Figure 5.2: Loss Graph of our experiments

5.1 Comment

Results are better in case of SGD& rmsprop. Results are also good for adam & nadam as we can see in figure 5.1 that they are approx. converging after some time.

Question 4 ends here.

6 Problem 5

6.1 Comments

Convolution layer learns about the features.
Neural Networks has to learn only the extracted features when we add convolutional layer & when we dont then it would have to learn the whole model. Max pooling is necessary to add after convolution layer so that we can have a bit of distortion to aviod linearity.
Adding more layers in case when you have a lot of data & diversified data then then it's a smart and good choice and adding layers when you have very limited data may not perform that well.
Deep learning is better to get classifications of a diversified data. It is basically combination of many nerual network layers like convolution and fully connected layers and some other layers like max pool etc. Deep Learning refers to Neural Network models with generally more than 2 or 3 hidden layers. It's a large multi layered Neural Network model that requires lots of data.

Question 5 ends here.