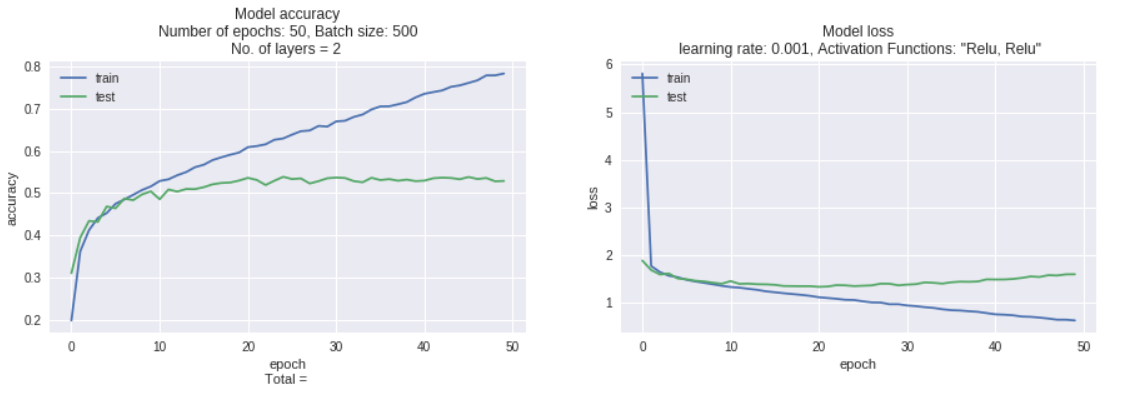
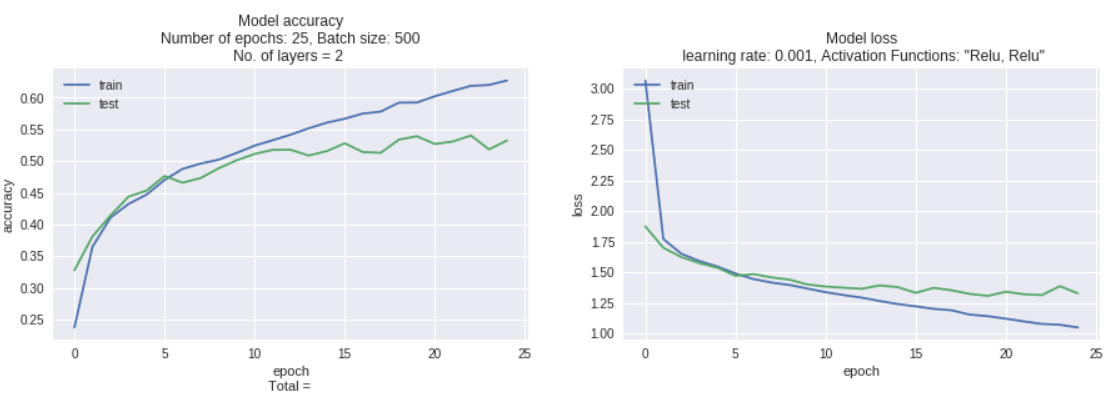
***Problem 1***

1.Effect of changing ***Number of epochs*** and keeping remaining unchanged.

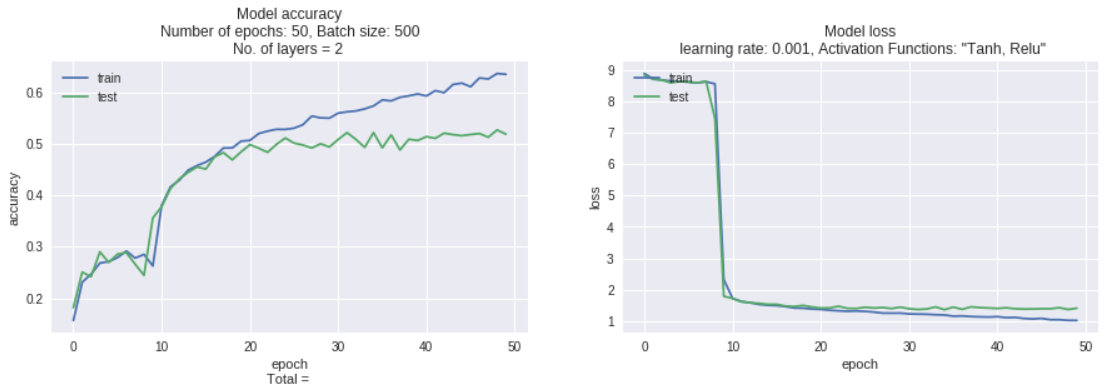
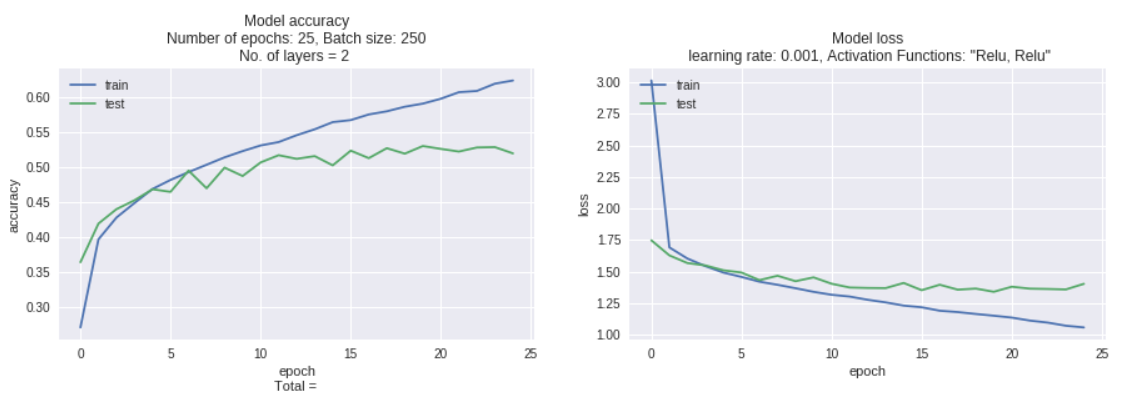
In this case as the number of epochs increased (from 25 in fig 1.1 to 50 in fig 1.2) the model began to ***overfit.***



***Figure 1.1 Figure 1.2***

**2. Effect of changing *Batch size* and keeping remaining unchanged.** (From 250 to 500)

Here as the batch size increased from 250 (fig 2.1) to 500 (fig 2.2). The model performed better with batch size = 500 till 20 epochs but then began to overfit.



***Figure 2.1 Figure 2.2***

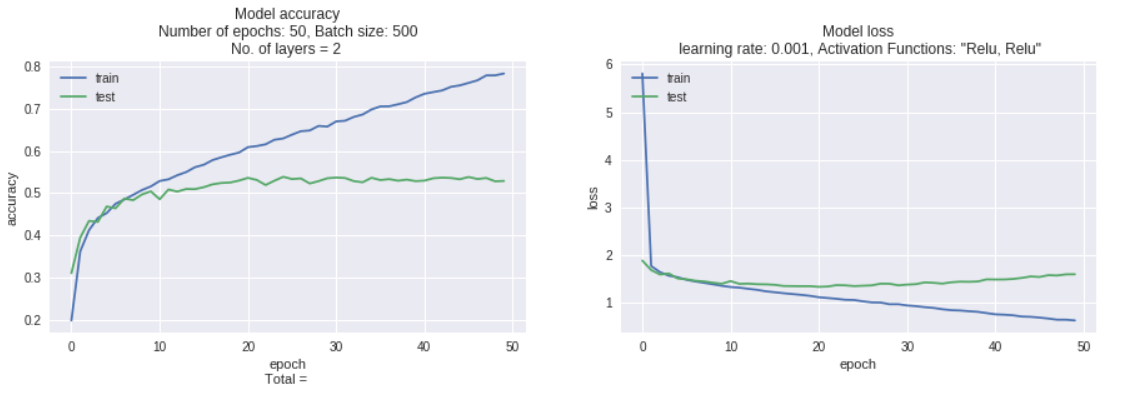
***3. Effects of changing Network Configuration***

***a) Number of neurons in a layer – Model 1 2-neurons, Model 2 3-Neurons***

***b) Number of layers – Model 1 2- layers, Model 2***

1. ***Changing Number of neurons in a layer:*** Fig 3.1 has 1024 neurons where as Fig 3.2 has 1537 neurons

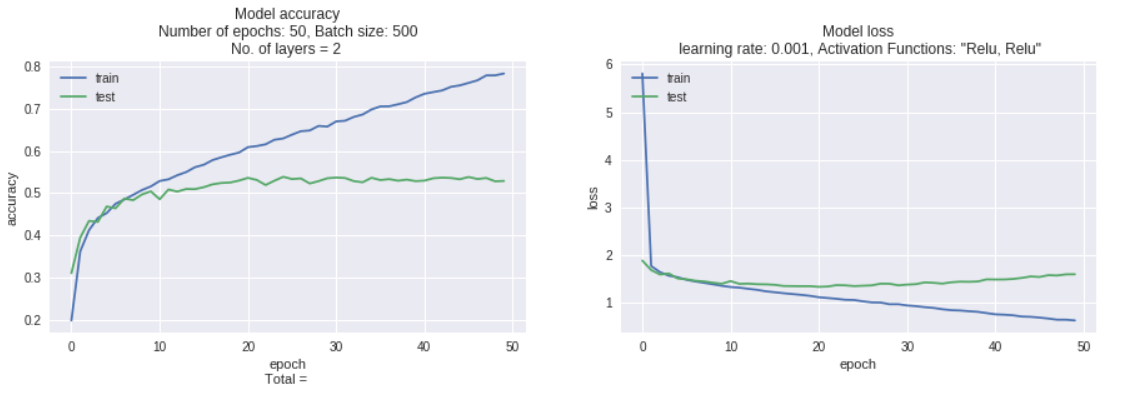
Increasing the number of neurons resulted in overfitting in this model



***Figure 3.1 Figure 3.2***

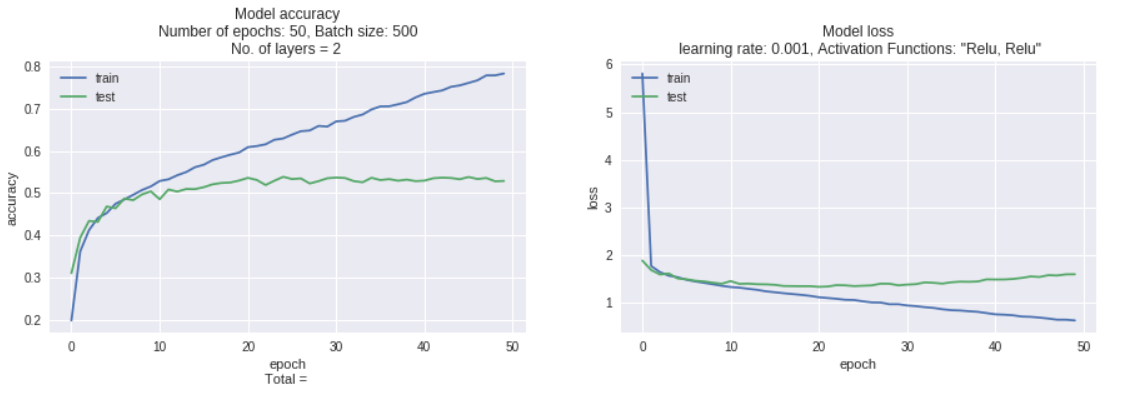
***b) Changing Number of Layers:*** Fig 3.3 had 2 layers where as Fig 3.4 had 3 layers

Form this we can see that overfitting is reduced here by inserting a new layer.



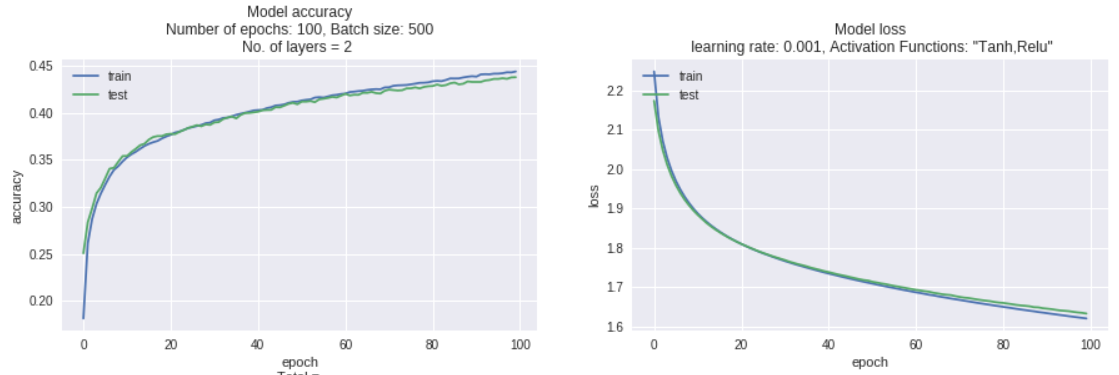
***Figure 3.3 Figure 3.4***

**5) Activation functions:** Model1 is with activation layers “Relu”, “Relu” followed by “Softmax” for probability distribution. Model 2 is with activation layers “tanh”, “Relu” followed by “softmax”.

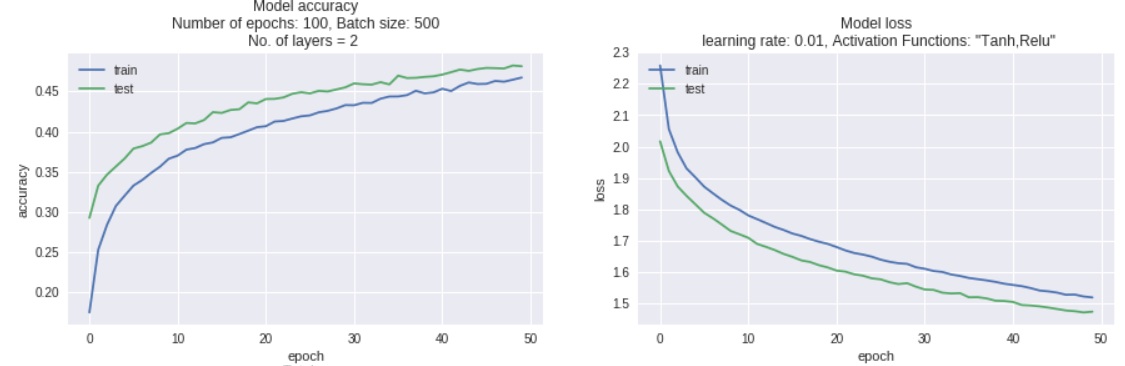
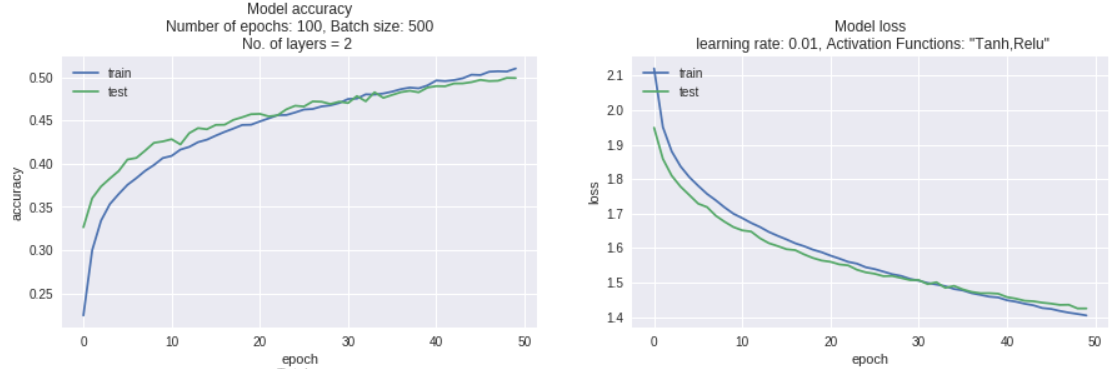


In this case we can see that all though at 50 epochs neither performed better but among these. Tanh, Relu performed better when compared to Relu, Relu

6) **Learning Rate:** Model 1 is with a learning rate of 0.01. Model 2 is with a learning rate of 0.001 with all other parameters kept unchanged. With decreasing the learning rate accuracy increased and loss decreased. But model took longer to converge.

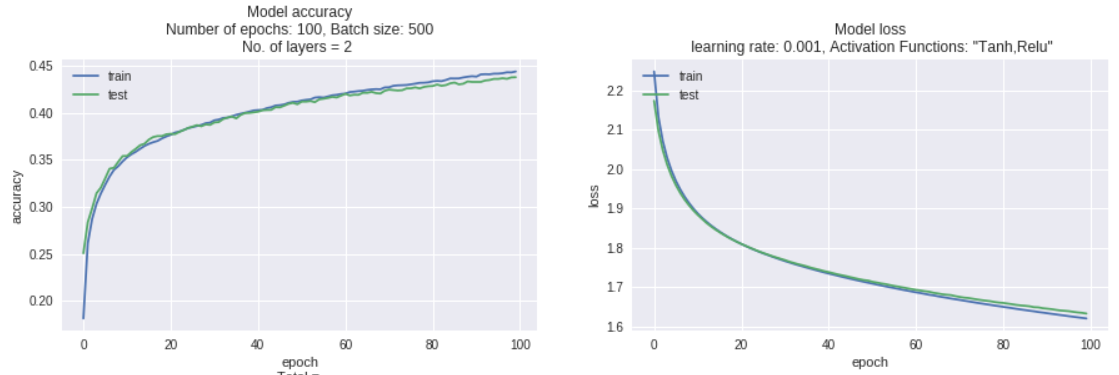


7) **Dropout Rate:** Model 1 Dropout rate 0.2, Model 2 Dropout rate 0.5. At 0.5 test accuracy began to fall at about 40 epochs where as with a dropout rate of 0.5 the test accuracy gradually increased with each epoch.



**Which model (with parameter values) would you choose and why?**

Model with the following parameters performed better although the accuracy is a bit low when compared to other models tested. This model fits the data more accurately.



**Comment on how good your model is? Does it overfit/underfit data? What could you do to improve the model?**

The model above is better when compared to other models tested in above but not the optimal model for this problem. But it is chosen as a good model because this one did not overfit/underfit the data as other models. We could use more epochs, using data augmentation or using CNN (which is proven to work better on image data).