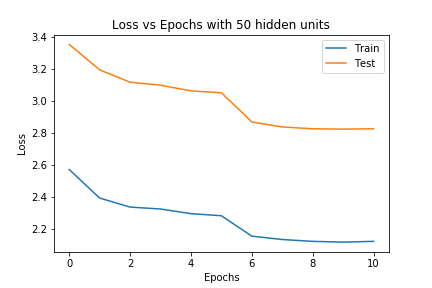
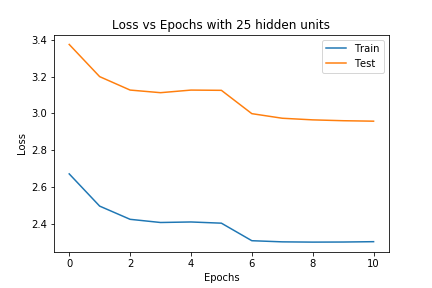
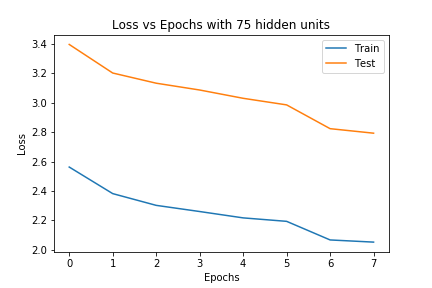
Part c

In this part, we experimented with different number of units in the hidden layer. We used 25, 50 and 75 units. The plots of loss vs epochs are shown below. Three experiments give final test errors of 2.96, 2.82 and 2.79 respectively. We observe that as number of units increases, both train and test losses decrease, which indicates that more hidden units can learn the internal features better.





Part d

In this part, we experimented with different dropout ratio – 0.1, 0.2 and 0.3. The plots of loss vs epochs are shown below. Their final train losses are very similar. For test loss, dropout with 0.3 gives the lowest value of 2.62, which indicates dropout with 0.3 helps the model achieve best generalization.

