Question 2:

Effect due to Batch normalization and dropout

The given data was split into 80-20 for train and test.

Without Batch normalization and Drop Layer:
convergence time=19 sec   accuracy=9.8 %

With Batch normalization and without Drop Layer:
convergence time=26 sec   accuracy=66 %

Without Batch normalization and with Drop Layer:
convergence time=19 sec   accuracy=10.14 %

With Batch normalization and Drop Layer:
convergence time=28 sec   accuracy=75 %

Effect due to change in the activation function:

1) RELU
convergence time=28sec   accuracy=76%

2)ELU
convergence time=28sec   accuracy=76%

3)SIGMOID
convergence time=26sec   accuracy=76%

4)TANH
convergence time=26sec   accuracy=68.8%