

## PART A

### Ques2) a

#### 1) BATCH NORMALIZATION (RELU)

- a) Adding the batch normalization increases the test accuracy while increasing the training time. It normalizes the network input weights between 0 and 1. This allows using higher learning rates when using SGD and for some datasets, eliminates the need for dropout layer.
- b) Test accuracy: 0.63690000000000002
- c) Train Accuracy is :- 0.7581 if run for 1 hrs
- d) To converge the data more number of epochs will be required and time taken would atleast more than 3-5 hr

#### 2) DROPOUT(0.25)

- a) Adding the dropout layer increases the test accuracy while increasing the training time. Dropout layer adds regularization to the network by preventing weights to converge at the same position. During forward propagation, nodes are turned off randomly while all nodes are turned on during forward propagation.
- b) Accuracy is 0.71 for 2 hrs and 20 epoches will require more number of hrs for complete convergence.

### Ques2) b

#### 1) Sigmoid

- a) Code is running fast and but giving less accuracy and not converging for 2 hrs gave around 0.54 accuracy

#### 2) TANH

- a) Code is running is fast but again accuracy is almost equal to that of sigmoid around 0.59

#### 3) RELU

- a) Giving accuracy around 69% but code is running bit slow depending upon the number of layers used

#### 4) LEAKY RELU

- a) Giving accuracy equivalent to Relu but also taking time to converge on my cpu machine.