Assignment – 4

Sagar Kulkarni

20767929

Q) Report your classification accuracy results in a table with three different activation functions in the hidden layer (ReLu, tanh, sigmoid). What effects do different activation functions have on your results? What effect does addition of l2-norm regularization have on your results? What effect does dropout have on results? Explain your intuition briefly.

	Dropout		
	0.0	**0.2**	**0.4**
Sigmoid	0.76045	0.7448375	0.746
Tanh	0.7757375	0.770075	0.7661625
ReLu	0.7805125	0.7815125	0.780775

Table-1: Test accuracy with L2-norm(0.001)

	Dropout		
	0.0	**0.2**	**0.4**
ReLu	0.8025375	0.8003	0.8011375

Table-2: Test accuracy without L2-norm

It seems that ReLu activation function preforms the best with an average accuracy of ~78%, followed by Tanh activation function with an average accuracy of ~77% and sigmoid function performs the worst with an average accuracy of ~75%. Also, ReLu activation functions seems to be less computationally expensive.

Running the model without l2-normalization seems to improve the overall accuracy of the model for ReLu.

Having a dropout layer(0.0 – 0.4) does not seem to show any clear affect of the accuracy results for this dataset. Higher dropout rate with l2-norm seems to indicate worse performance however, this is not very clear.