

MSCI-641 Assignment 4

Nauman Ahmed

20743448

	Sigmoid	ReLu	Tanh
	Test Accuracy		
No regularization + No Dropout	0.744	0.810	0.793
L2-regularization- =0.001	0.766	0.790	0.786
L2-regularization- =0.01	0.500	0.775	0.775
Dropout=0.4	0.788	0.811	0.794
Dropout=0.6	0.787	0.807	0.794
L2-regularization- =0.001 + Dropout=0.4	0.733	0.790	0.787
L2-regularization- =0.1 + Dropout=0.4	0.500	0.754	0.755

The best accuracy (0.811) I got was with the Relu activation function when no regularization was used but a dropout layer with a value of '0.4' was used. Relu is better than other activation functions because its gradient does not saturate. The dropout layer randomly drops out hidden units and therefore helps to prevent overfitting. This makes the model to perform better on the testing data.