

Multi-Layer Perceptron(MLP)

Analysis

(2011cs1012)

Hidden layer size	Learning rate M=3432(# of training examples/fold) (2-fold cross validation)	Training accuracy(mean over all 100 epochs and 2 folds)	Validation accuracy(mean over all 100 epochs and 2 folds)
80	15/m	91.34	89.82
	20/m	92.93	91.14
	25/m	93.84	91.96
	30/m	94.33	92.4
	35/m	94.87	92.7622
	40/m	95.2	93.1222
160	15/m	91.85	90.3458
	20/m	93.410	91.6640
	25/m	94.29	92.4203
	30/m	94.88	92.9456
	35/m	95.16	93.1685
	40/m	95.67	93.6373
240	15/m	91.86	90.345
	20/m	93.415	91.641
	25/m	94.195	92.428
	30/m	94.791	92.863
	35/m	95.170	93.295
	40/m	95.385	93.419

Performance Parameters tested:

1. Number of nodes in the hidden layer: The training accuracy increases with the increase in the number of nodes in the hidden layer but the model overfits the data set as the numbers of nodes are increased. This is because more the number of nodes in the hidden layer more complex and non-linear the hypothesis learnt. Time taken to train the MLP

increases with increase in the size of hidden layer as the sizes of Theta1 and Theta2 matrices increase.

2. The learning rate(alpha): A balance has to be kept between large and small values of the learning rate. If it is very less then it might take a large number of epochs to converge to a global minimum of the cost function. If it is too large it might overshoot the minimum or might oscillate around the global minimum of the cost function .

The best cross validation accuracy:

97.29%(alpha=0.01165=40/m,hiddenlayernodes=240)

Below is the plot for the best cross validation accuracy:

