Assignment 5 Analysis Report

ADABOOSTING CLASSFIER

In adaboost classifier we have used ensemble of decision stump, and 4 classifiers, which acts as one vs all classifier, for each orientation we have each classifier, for each decision stump of classifier we will randomly select 100 pair of features, and split the data point. We will find accuracy of each classifier and we will select the best split based on low error. And finally, classifier with high voting value of (alpha value) will get selected for labelling the test data.

In this approach, we have number of stump as a parameter, based on the change of the stump count we could see the changes in the accuracy we obtained in predicting the test data class.

Number of decision stump	Accuracy of prediction	Overall running time in minutes
5	53.9%	3.12
10	56.5%	6.08
15	63.2%	9.08
20	64.5%	9.25

From the above table, we see it clearly that if the stump is increased the accuracy level of the model is also increasing, further this model can be made still more accurate in predicting the test class if the number of pair selecting randomly is increased more or if the split is happening based on some entropy and gin index calculation.

If the we use subset of the train data we will able to achieve more than 80% of accuracy and the running time will be much lower. This model is prone to overfitting if the train data is of small size.

Sample Images correctly classified by adaboost classifier are




Sample images incorrectly classified by adaboost classifier are



Neural Network

In the neural network there are three layer one input one output and one hidden layer and there are 192 neurons in input layer and 4 neurons in output layer, the train data is feeded into the feed forward network and after that the output value are handled by softmax classifier which converts the value to the probability of one. Traditional Gradient descent is applied to update the weights and the errors are backpropagated using traditional gradient descent.

There are number of features which plays important role in neural network, they are number of hidden neuron, number of epoch(number of times the network has to run). Below are some of my observation based on changes in the parameter

Number of neurons and Number of Epochs	Accuracy	Overall time in minutes
200 and 100	71%	1.18
200 and 500	74%	5
200 and 1000	75%	9

Number of Neurons and Epochs	Accuracy	Overall time in minute
100 and 100	70%	1
100 and 500	73%	3
100 and 1000	74%	8

Images classified correctly by neural network correctly are,




Images misclassified by the neural network



KNN CLASSIFIER

Team mates didn’t do anything