Different classifications techniques were used to classify test data of MNIST test dataset. Overall accuracy achieved was different for each classifier.

**1.** **Neural Networks:**

* Sigmoid Loss Function-Batch Learning( Probability of Error: 5.88%)
* Sigmoid Loss Function with Regularization-Batch Learning( Probability of Error: 4.99%)
* Relu Loss Function-Batch Learning( Probability of Error: 4.11%)
* Relu Loss Function with Regularization-Batch Learning( Probability of Error: 3.98%)
* Sigmoid-Stochastic Gradient

1. Sigmoid( Probability of Error: 3.87%)
2. Relu( Probability of Error: 4.89%)

Details of different parameters used:

Learning rate: 2e-5

Regularization Constant: 1e-4

RELU found to be better non-linearity because it converges faster and the difference between training and test error is significantly low. Stochastic gradient is must faster than batch learning because same accuracy is achieved in around 15 epochs.

**2.Boosting**

Adaboost algorithm: Probability of Error: 9.37%

10 classifiers were used to classify each of the digit. Classification efficiency of each of them was observed using CDF vs margin plot along with pictorial representation of each of the weak learner. Decision stumps with both positive and negative polarity were used as weak learners.