

## SMAI Assignment Report

**Name:** Rishabh Arora

**Roll Number:** 201501010

Q1)

Used a weight vector of size 785. '1' added initially to the data to account for bias. If the label is 0 then the entire row is negated to make the update rule more general.

While training if  $W.X \leq \text{margin}$ , then the update is done in which  $X$  is added to  $W$ . If it's batch then these additions are stored in a temporary variable and added after batch size is full.

1. Single sample perceptron without margin ()

This converges on 12 epochs with validation accuracy 99.9

2. Single sample perceptron with margin-

This converges on 21 epochs with validation accuracy 99.9

3. Batch perceptron with and without margin-

Both approximately converge ( if epochs == 1000, break) on 1000 epochs with validation accuracy 99.9

Q2)

**1. Relaxation algorithm + margin :-**

Used the gradient descent algorithm with margin and put a learning rate of 0.01. This was run for 10000 epochs. It gave validation accuracy of 99.11

The gradient used was -

$$\nabla J_r = \sum_{y \in \mathcal{Y}} \frac{a^t y - b}{\|y\|^2} y,$$

2. The modified perceptron algorithm- Used a decaying learning rate which is used in the update rule while adding  $X$  to  $W$ . The learning rate is a function of the accuracy till now in this epoch and is given by  $(1 - \text{correct}/\text{total})$ .

This sort of becomes stable after 1000 epochs and gives a validation accuracy of 99.00 %.

Q3)

**Data Processing steps-**

For all the continuous values(attr 0-6), normalized them so that all of them are in the range(min(value), max(value)).

For discrete values, used each value to split.

Built the tree recursively till quality = 0(All homogenous values). For continuous values used a step size of 0.01 and picked the one which gave highest information.

This gave a validation accuracy of 97.91 %

Q4)

First added all the words to a list and created a bag of words by using `makefeaturevector`.

**KNN:**

Then for each test calculate it's manhattan distance of from each data-point in the training data and sort the the data and use the first k minium distances to classify the current node.

Validation accuracy of 95 % using  $K = 7$ .