

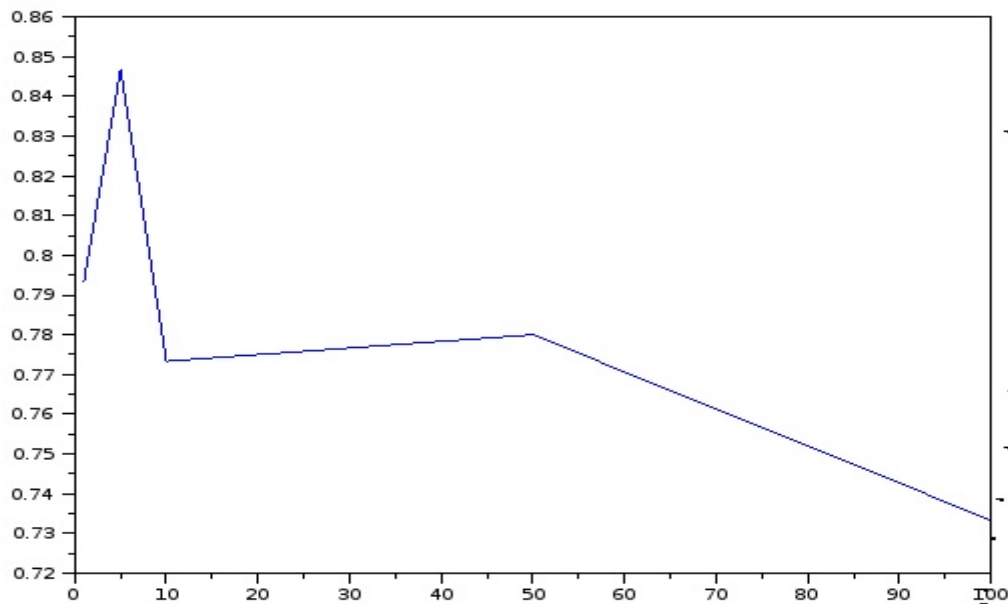
CS 419 ASSGN 2

120050009

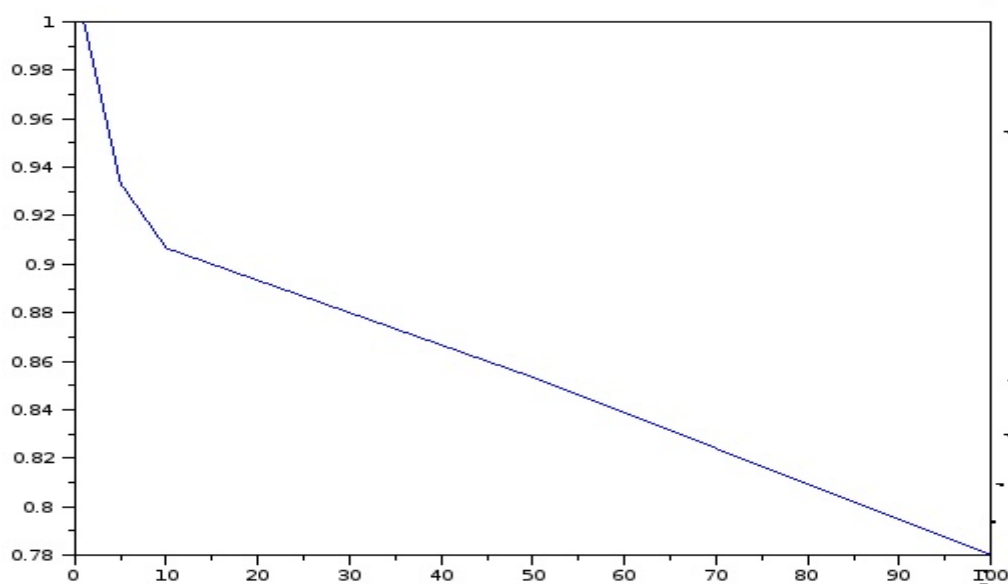
Deepanjan Kundu

\ As the data is too big and the compiler hung several times ,so i have used half the training set for first part and treated rest half as validation data as sir had mentioned in the question as there are 20000 possible attributes .

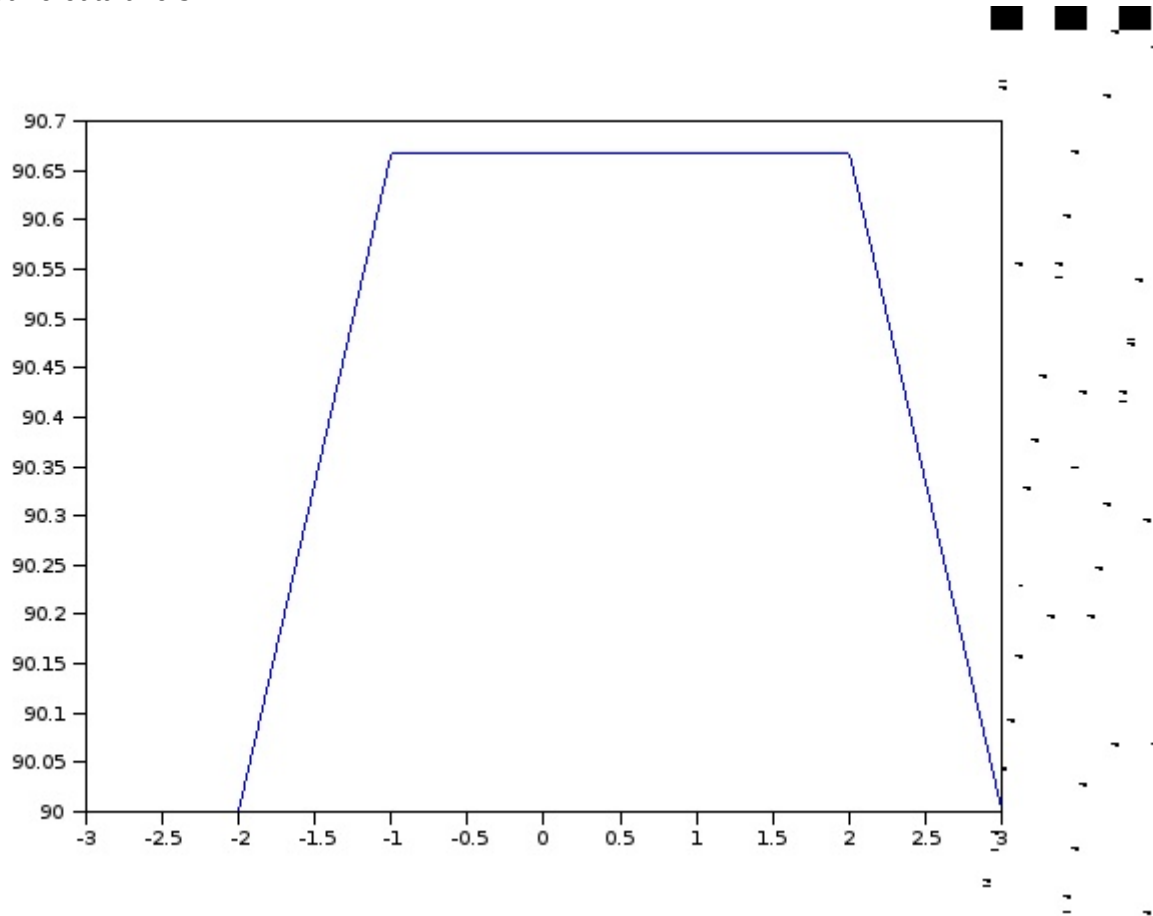
Q1)a)for validation data



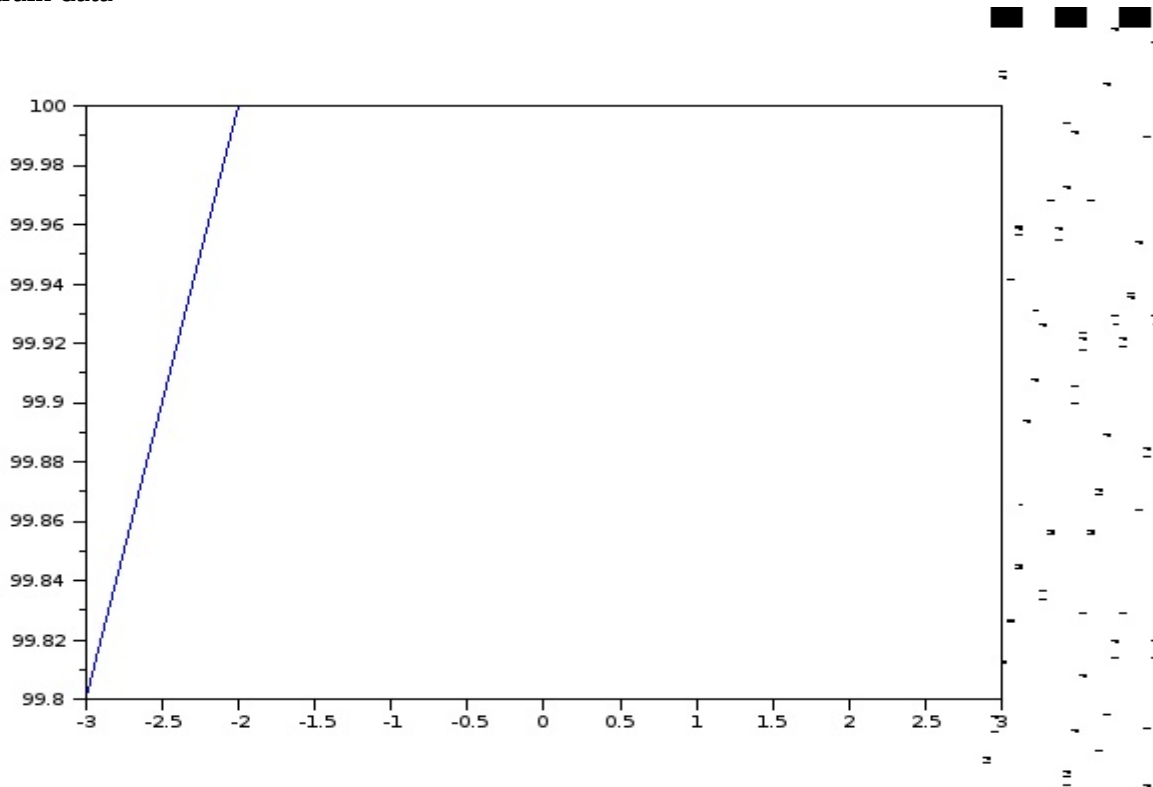
b)for training data



2)(x i.e exponent part of 10^x on x-axis)
a)with valid data and $s=7$



b)with train data



Write up:

Highest accuracy between 1 and 2 is with 2 i.e. logistic regression.

The answer is quite expected.

In k -neighbours as the k increased the precision of prediction decreased. However when we used the training set for validation the precision was larger than valid as expected.

In the second one also the precision of training set was close to 100 percent for most of the cases as expected .The precision is like 90 percent for validation data which is larger than logistic regression(70-80 percent).