Assignment 1 - report - Intro To ML

Mtech - 2nd Sem

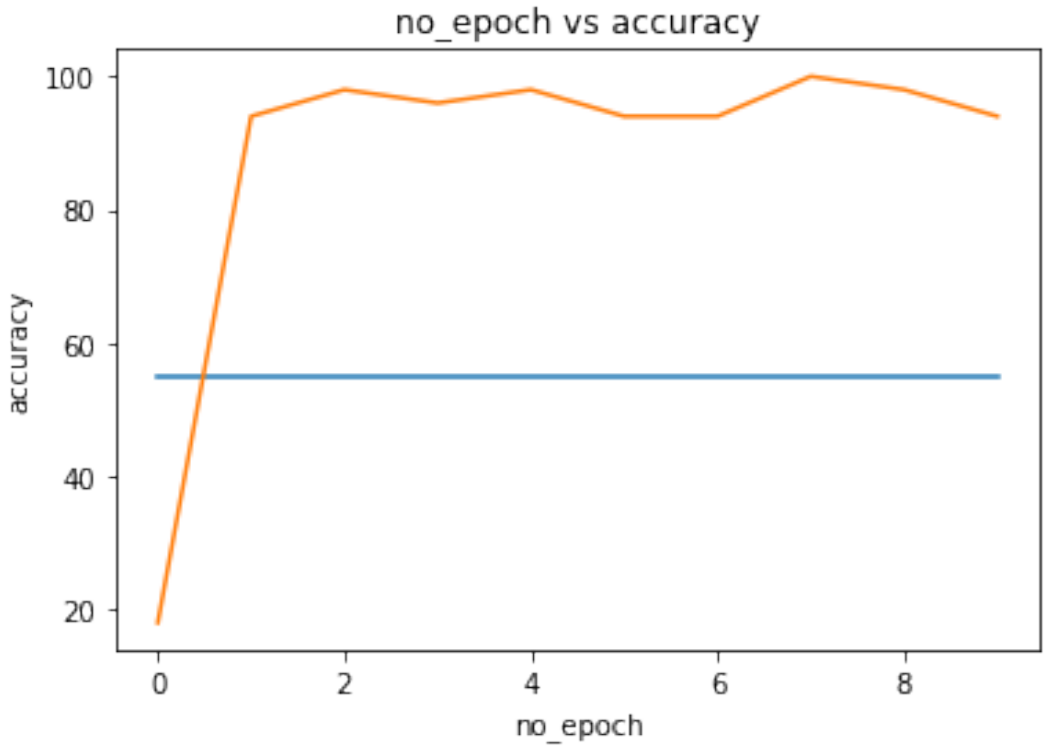
K.Ranjith - MIT2020017

21-02-2021

IIIT Allahabad

problem 1b - accuracy with respect to number of epoch - KNN vs Logistic Regression

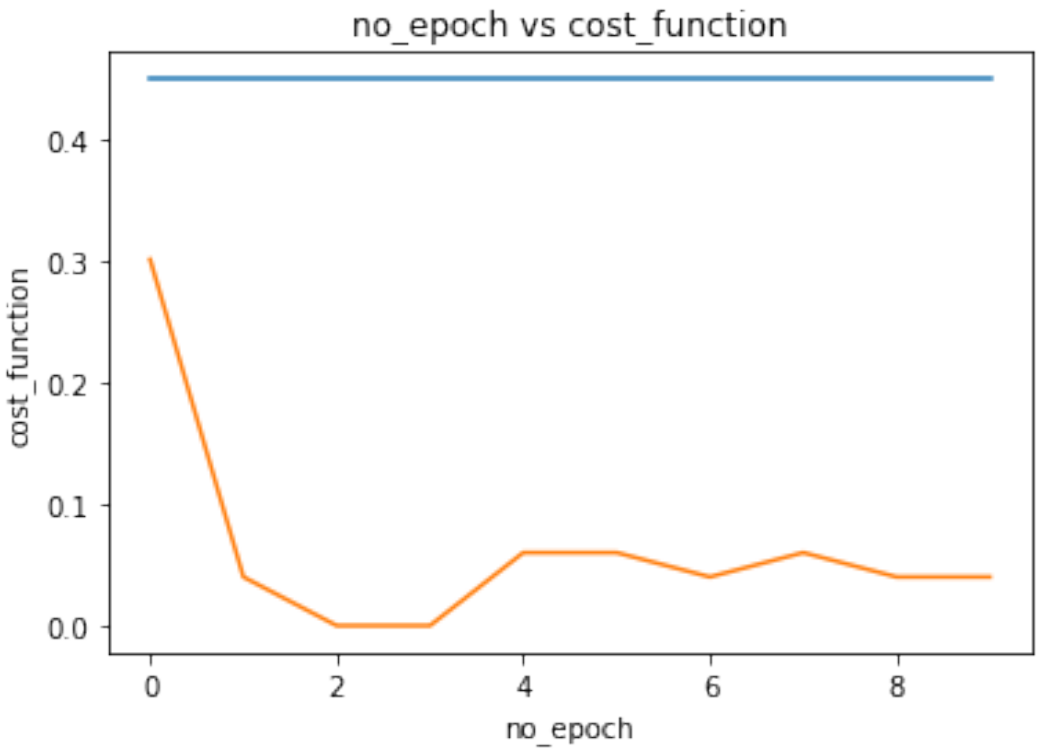
hyperparameters[no-train = 600, no-test = 50, k=5, no-epoch = [0,10], $\alpha = 0.01$]
orange - logisticRegression, blue - knn



accuracy of KNN model is 57.6 percent where as for Logistic Regression Model it is above 97 percent. so we can conclude Logistic regression works way more better than KNN for MNIST data set

problem 1b - loss with respect to number of epoch - KNN vs Logistic Regression

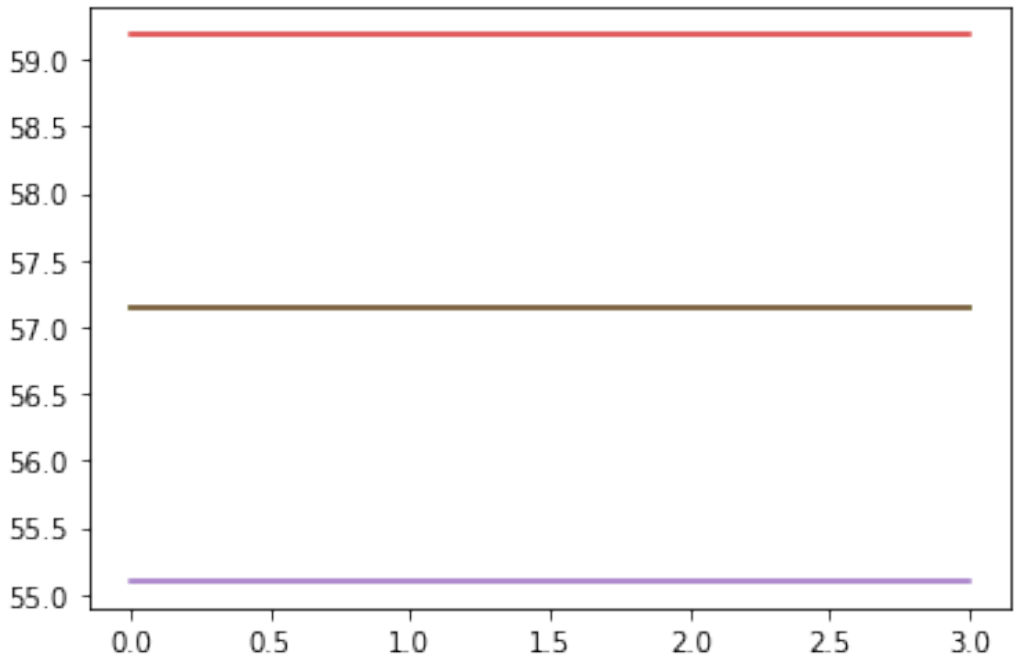
hyperparameters[no-train = 600, no-test = 50, k=5, no-epoch = 10, $\alpha = 0.01$]
orange - logisticRegression, blue - knn



error of KNN model is 0.46, where as for Logistic Regression Model it is below 0.1 .
so we can conclude Logistic regression works way more better than KNN for MNIST data set

problem 1b - accuracy with respect to number of epoch - for KNN - for varying K - for train data = 600

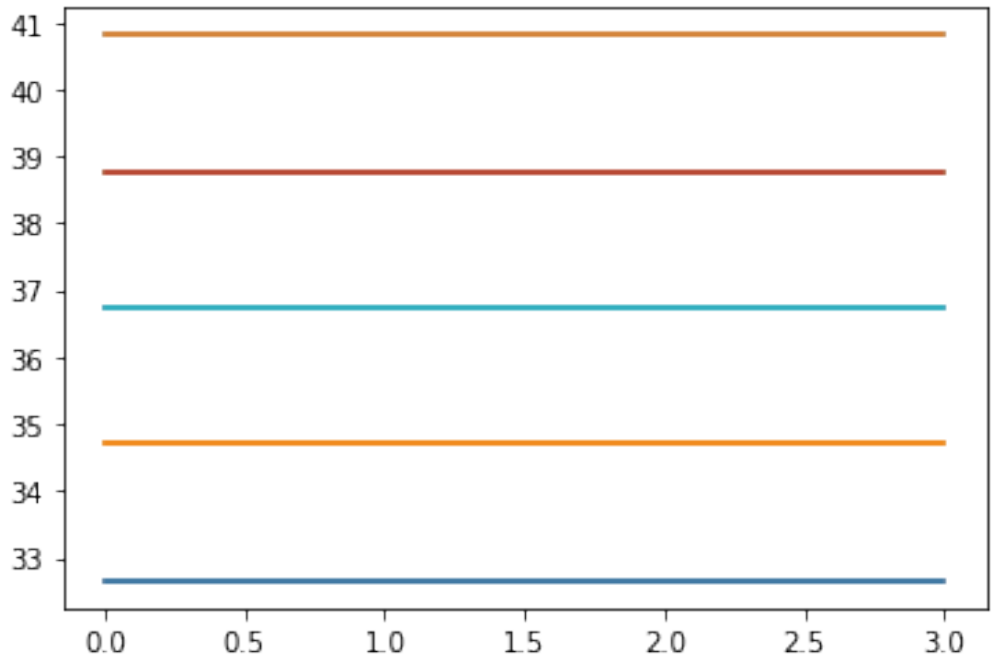
hyperparameters[no-train = 600, no-test = 50, k=[1,7], no-epoch = [1,4]]
for k= [1,2,3,4,5,6] corresponding accuracy values = [57.14285714285714, 57.14285714285714, 57.14285714285714, 59.183673469387756, 55.10204081632652, 57.14285714285714]
x-axis-epoch , y-axis-accuracy



for K = 4, we can see highest accuracy i.e 59.18. so take k=4 for this hyperparameter

problem 1b - accuracy with respect to number of epoch - for KNN- for varying K - for train data = 100

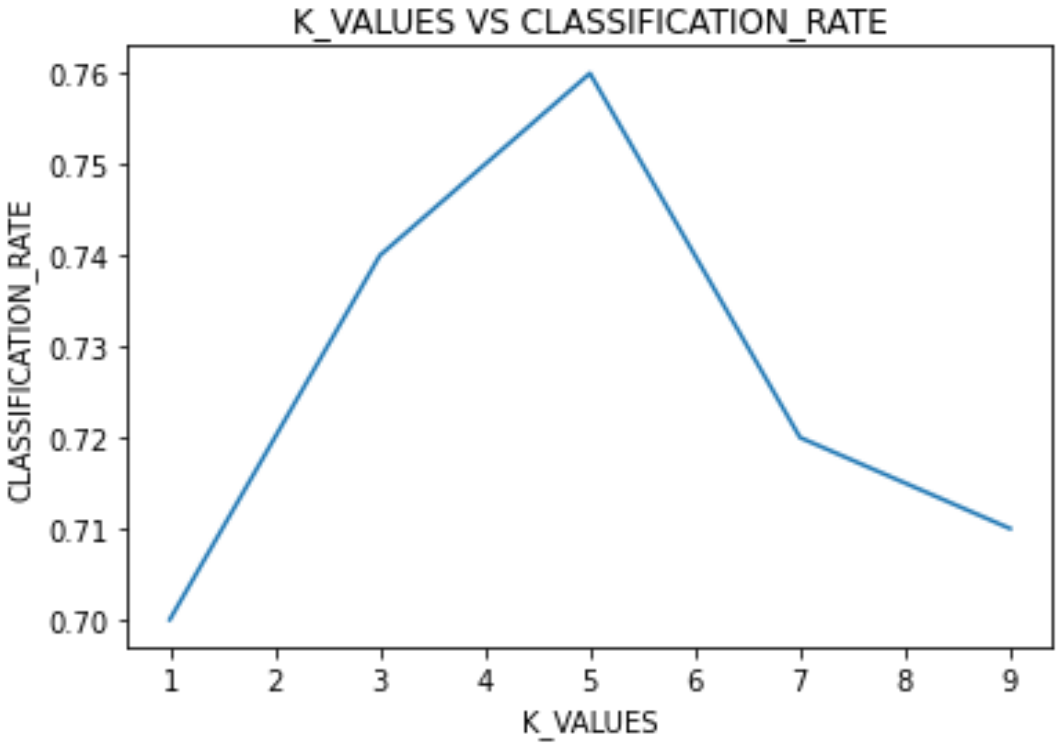
hyperparameters[no-train = 100, no-test = 50, k=[1,13], no-epoch = [1,4]]
x-axis-epoch , y-axis-accuracy



for K = 1,2, we can see highest accuracy i.e 40.816 .. then for further values of k accuracy decreases and fluctuates between [32.65,36.73]

problem 3 - Classification rate vs K

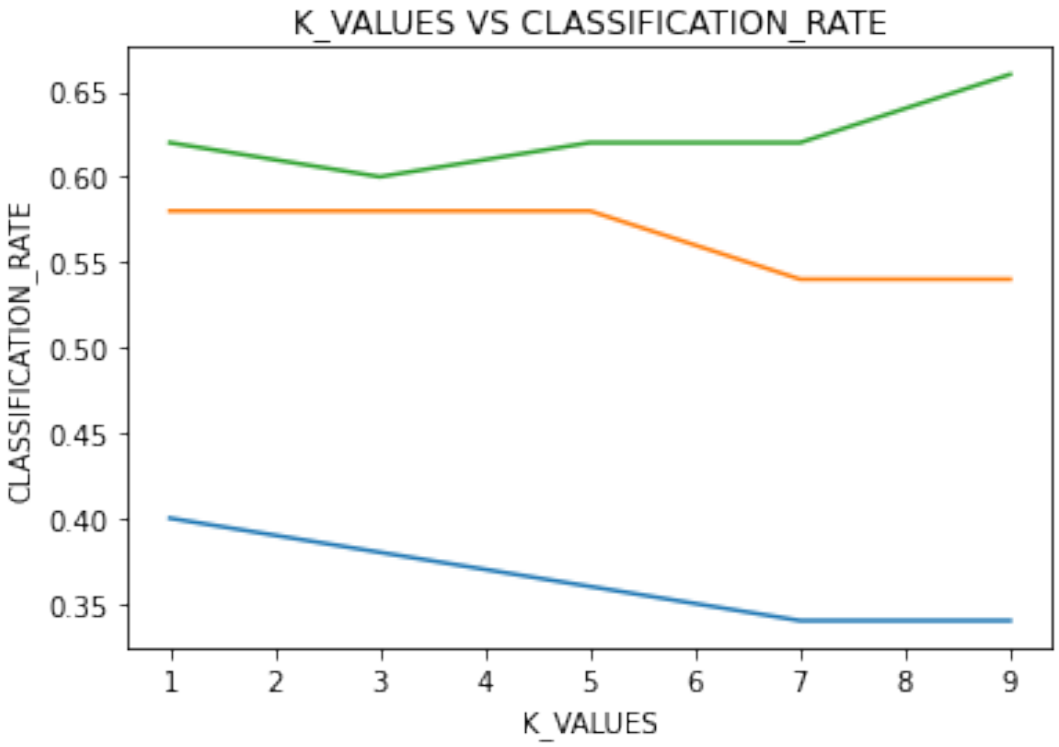
hyperparameters[no-train = 6000, no-validation = 100, no-epoch = 1]



classification rate peaks at K =5, it means for this set of parameters K=5 best suits the model

problem 3 - Classification rate vs K - for varying traindata

hyperparameters[no-train = [100, 500, 1000], no-validation = 50, no-epoch = 1]



classification rate increases when we take more train date for training model
But the value of K for which classification rate is high is varying as trian data increases and unpredictable