**Classification using Logistic Regression: Many v/s One**

We know MNIST dataset contains features and response variable as hand written digits to identify digits ranging from 0 to 9 resulting in 10 classes.

**Model 1: Many binary logistic classifiers to build multi class logistic classification**

We build 10 binary classifiers and learn the weights of the features wrt to each class separately. Then we use the learnt weights of each class to predict the class of the data.

*Training set Accuracy:92.716%*

*Validation set Accuracy:91.42%*

*Testing set Accuracy:91.9%*

**Model 2: One classifier for multi class logistic classification**

We build just one multiclass classifier and learn the weights of all classes together. Then we use the learnt weights of each class to predict the class of the data.

**Support Vector Machines:**

**SVM kernel as linear:** sampled train data pf 10000 observations

*Training set Accuracy:93.002%*

*Validation set Accuracy:91.41%*

*Testing set Accuracy:91.59%*

**SVM kernel as radial basis function:**

Choosing hyper parameters – gamma and cost to tune the svm for randomly sampled 10000 from train data.

**Gamma: Default**

*Training set Accuracy:91.902%*

*Validation set Accuracy:91.95%*

*Testing set Accuracy:92.45%*

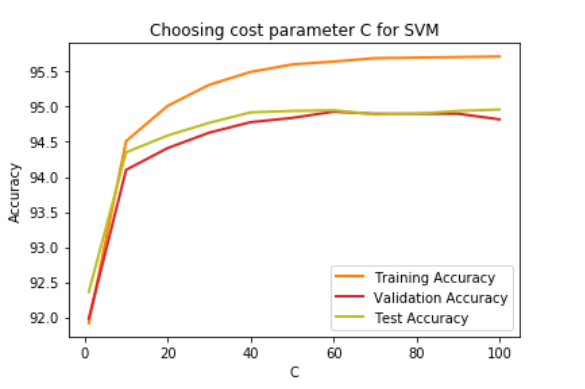
**Gamma:1**

*Training set Accuracy:32.884%*

*Validation set Accuracy:14.26%*

*Testing set Accuracy:15.8%*

**C- Cost Parameter: Gamma default and C = [1,10,20,30,40,50,60,70,80,90,100]**



Testing accuracy was high when C=10 and it does not improve much and gets flat thereafter. But test accuracy when C=10 and C=40 has improved by just 1% over sampled 10000 observations of the training data.

Having higher C increases complexity of the model so we choose C as 10.

From the results on sampled data for SVM with kernel as radial basis function we consider gamma as default and C=10 and fit the model.

**Final Model:** Gamma default and C=10 trained on full train data.

*Training set Accuracy:97.13199999999999%*

*Validation set Accuracy:96.17999999999999%*

***Testing set Accuracy:96.1%***

**Conclusion:** For given dataset with linear SVM the test accuracy is over 90% and for radial basis function with tuning hyper parameters – C and gamma we achieved 96% of accuracy. If model complexity and time a constraint its better to choose linear model over radial basis SVM.