Machine Learning Lab Assignment-2

Name: Aitik Gupta

Roll No: 2018IMT-010

Objective:

To implement a crude Naive Bayes classifier (without using any library function for naive-bayes) on MNIST data set.

The MNIST data comprises of digital images of several digits ranging from 0 to 9. Thus, the data set has 10 levels of classes.

Part-a:  
Without applying smoothing, It’s accuracy is very poor. Below are the following errors:

Accuracy: 0.1092

For 0 error % = : 0.20408163265306123

For 1 error % = : 99.8237885462555

For 2 error % = : 97.96511627906976

For 3 error % = : 99.20792079207921

For 4 error % = : 99.59266802443992

For 5 error % = : 99.55156950672645

For 6 error % = : 99.68684759916492

For 7 error % = : 93.09338521400778

For 8 error % = : 100.0

For 9 error % = : 99.90089197224975

Total Error: 89.08

Part-b:

After using smoothing function, the accuracy has raised and it’s about 84.12%. Below are the following errors:

Accuracy: 0.8412

For 0 error % = : 9.591836734693878

For 1 error % = : 4.405286343612335

For 2 error % = : 17.151162790697676

For 3 error % = : 16.33663366336634

For 4 error % = : 19.45010183299389

For 5 error % = : 29.7085201793722

For 6 error % = : 11.273486430062631

For 7 error % = : 15.369649805447471

For 8 error % = : 22.279260780287473

For 9 error % = : 16.15460852329039

Total Error: 15.880000000000006

Inference:

If one of the conditional probability is zero, then the entire expression becomes zero. So, here the concept of laplace smoothing helps by adding virtual counts. Thereby getting better posterior probabilities. Thus it increases the accuracy of the model.

Here is the link to the code for reference:

<https://github.com/aitikgupta/ITIT-4103-2021>