Foundations of Machine Learning

Generative and Discriminative Modelling Project

Moustapha Cisse

November 20, 2019

Instructions

1. Write a python class to implement Gaussian Discriminant Analysis(GDA) and another class to implement Naive Bayes (NB) algorithm from scratch. Test your models with the dataset.

2. Using the dataset given, compare Gaussian Discriminant Analysis and Naive Bayes with Logistic regression using a different size of the dataset. Use the following sizes of the data for comparison;

- 10% of the data
- 30% of the data
- 60% of the data
- 100% of the data

For each of the sizes, write your observation and show the reports of the comparison on the algorithms.

Dataset You would be working with a text data from the UCL repository data here. Download the data and considering you would be splitting your data into different sizes, you can combine both the train and test data into one large dataset then do your splitting and experiment.

Submission

The deadline for submission is September 30th, before midnight. You are to Submit a python notebook with submission file as *firstname_lastname_project.ipynb*. Send with subject **Foundations of Machine Learning project** to ghtutors2019@aimsammi.org