

E1213 PRNN: Assignment 1 - Basic Models

Prof. Prathosh A. P.

Submission deadline: 1st March 2022

1 Introduction

This assignment is about implementation of primitive ML algorithms for 5 different tasks. The tasks would remain fixed across all the assignments while the algorithms and experiments to be conducted would vary.

2 Problem statements

We consider **five distinct tasks** for the assignments which are described below:

2.1 Binary Classification problem

Here we consider a 2 class classification problem on image data. The dataset is PneumoniaMNIST which can be found here:

<https://zenodo.org/record/5208230#.YfqjwFhBxHQ>. It contains images from chest X-rays belonging to two classes - Normal and Pneumonia. The metrics to be computed are: classification accuracy, AUC and F1 score. Training is to be done on the training split and test on the test split.

2.2 Multi-Class Classification Problem

In this task, we look at an 8-class classification problem on Blood Cell Microscope images. The dataset is called the BloodMNIST which can be obtained from this link - <https://zenodo.org/record/5208230#.YfqjwFhBxHQ>. The metrics are same as above.

2.3 Bounding box regression Problem

In this problem, we are given images of traffic signs and the task is to find out the bounding boxes that encompass the sign in the image. The dataset is available here - <https://www.kaggle.com/andrewmvd/road-sign-detection>. There is an xml annotation file that has the co-ordinates of the boxes. For example - `<bndbox> <xmin>98</xmin> <ymin>62</ymin> <xmax>208</xmax> <ymin>232</ymin> </bndbox>`. The task is to take the input image and

regress over the co-ordinates (4) of the box. The metrics here are the mean MSE, mean MAE and mean Intersection over Union (mIoU).

2.4 Frame classification on audio data

In this problem, the task is to classify every sample of a speech/audio signal. We use the TIMIT dataset for this purpose - <https://www.kaggle.com/mfekadu/darpa-timit-acousticphonetic-continuous-speech>. Consider one of the Directories (DRx) each from the Train and Test sets for all your experiments. The task here is to classify every sample of the utterance to be belonging to vowel or not vowels. The ground truth information has to be generated from the .phn file that accompanies every .wav file. It lists the phonemes corresponding to time intervals in the utterance. Eg - 0 3050 h# 3050 4559 sh 4559 5723 ix 5723 6642 hv 6642 8772 eh 8772 9190 dcl 9190 10337 jh 10337 11517 ih 11517 12500 dcl. Define Vowels to be all phonemes that contain /a,e,i,o/ and u in them. The metrics are average true positive, average true negative, average false positive and average false negative. For this problem, use short segments of speech signals (of duration 10 to 40 ms) as data points. Either use RAW speech or features such as MFCCs or LPCs may be used as input space.

2.5 Generative Models

In this module, we build generative models on Tinyimage net (<https://www.kaggle.com/c/tiny-imagenet/data>). Use Frechet Inception Distance (FID) between 1000 generated and real data as the metric for evaluation.

3 Models for Assignment 1

- Bayes's Classifier with several Class Conditional Densities such as Gaussian, GMMs (Have to code up EM)
- Bayes's Classifier with different density estimates (ML, MAP and Parzen Window and nearest neighbor estimates)
- K-Nearest Neighbor and Naive Bayes classifiers
- Linear Models - Linear Classification/Regression, Linear Models with Polynomial Kernels, Logistic Classifier/Regression, Fisher LDM
- All the Linear Models with different Regularizers such as L1, L2 and Elastic
- GMM for the Generative model part - Fit a GMM and sample more points from it

4 Expected Outcomes

- Submit a **single** python notebook
- You are not allowed to use any ML libraries such as Sklearn
- Try as many different hyper-parameters as possible
- Submit a 4-page report with your graphs/results and interpretations
- Reports **SHOULD** be in IEEE double column format and strictly prepared in LaTeX
- Your evaluation depends on the quality of your code, experiments and **your observations of the results** and the report that you submit
- Slightest detection of any sort of plagiarism (code/report) will immediately lead to heavy penalty