Eckovation AIML: Programming Exam

Supervised Learning

Consider the problem of classification of Shouted and Normal Speech. A common approach is to extract the MFCC features from the speech signal. These features are extracted from overlapping frames of tagged speech intervals. Two data files containing 2000 MFCC features of normal and shouted speech are provided for training the classifier. Similarly, two files containing 500 MFCC features of normal and shouted speech are provided for classifier performance evaluation.

Training Data Set File Name	
Train_MFCC_N_2000	Train_MFCC_S_2000
Testing Data Set File Name	
Test_MFCC_N_500	Test_MFCC_S_500
File Format : npy	
Shouted (Positive)	Normal (Negative)

1. Load the two training datasets Train_MFCC_N_2000.npy and Train_MFCC_S_2000.npy into arrays ***trn_N*** and ***trn_S*** respectively using *np.load(* filename *).*
2. Append these two arrays to form Dataset ***trn_D = [trn_N , trn_S]***
3. Compute the mean and standard deviation values along each dimension of input data. Store the dimension-wise mean and standard deviation values in respective vectors **meanVec** and ***stdVec***.
4. Normalize input dataset ***trn_D*** using ***meanVec*** and ***stdVec*** to form the normalized dataset ***norm_trn_D***.

5. A simple Perceptron ***P*** having sigmoid activation function is learned to classify normal and shouted speech (binary classifier). This perceptron **P** is trained using dataset ***norm_trn_D***.
6. Load the two test data files for performance analysis. These are Test_MFCC_N_500.npy and Test_MFCC_S_500.npy. Load them into arrays ***tst_N*** and ***tst_S*** respectively using *np.load(* filename *).*
7. Append these two arrays to form Dataset ***tst_D = [tst_N , tst_S]***
8. Normalize dataset ***tst_D*** using ***meanVec*** and ***stdVec*** to form the normalized dataset ***norm_tst_D***.
9. Conduct performance analysis using learned perceptron (***P***) based classification label prediction on data from ***norm_tst_D*** and their corresponding ground truth label information. Report the binary confusion matrix and the following performance parameters.

TPR	TNR	PPV	NPV	FPR	ACC	BA	$F1^{(P)}$	$F1^{(N)}$