July-November 2024 Semester
CS5691: Pattern recognition and Machine Learning
Programming Assignment II

Date:**13th September, 2024**
Deadline for submission of report: **5PM on Saturday, 14th September, 2024**

Dataset 1: 2-dimensional data: Linearly separable data set for 3 classes
Dataset 2: 2-dimensional data: Nonlinearly separable data set for 2 classes
Dataset 3: Image data set (Dimension of feature vector: 35) for 5 classes

Classifiers for Dataset 1:
1. K-nearest neighbours classifier, for K=1, K=5 and K=9
2. Bayes classifier with a Gaussian distribution for every class
    a. Covariance matrices for all the classes are the same
    b. Covariance matrices are different

Classifiers for Dataset 2:
1. K-nearest neighbours classifier, for K=1, K=5 and K=9
2. K-nearest representatives classifier, for K=1, K=3 and K=5: Use 10 representatives per class
3. Bayes classifier with a Gaussian distribution for every class
    a. Covariance matrices for all the classes are the same
    b. Covariance matrices are different
4. Naive-Bayes classifier with a Gaussian distribution for every class:
    a. Covariance matrices for all the classes are the same
    b. Covariance matrices are different

Classifiers for Dataset 3:
1. K-nearest neighbours classifier, for K=1, K=9 and K=15
2. K-nearest representatives classifier, for K=1, K=5 and K=9: Use 10 representatives per class
3. Bayes classifier with a Gaussian distribution for every class
4. Naive-Bayes classifier with a Gaussian distribution for every class

Use the validation method to choose the best values of hyperparameters.

Report should include the following for every classifier and for every dataset:
1. Table of classification accuracies of the model on training data, validation data and test data
2. Confusion matrix for the best configuration of the model, on training data and test data
3. Decision region plots for each of the classifiers for Datasets 1 and 2**.** Superpose the training data on the decision region plot. For the Bayes classifiers and Naïve-Bayes classifiers, superpose the plots of level curves on the training data.