Pattern Recognition

ECE 4363 / ECE 5363

Project 1

- Load the dataset "fisheriris" into the workspace.
  - Study the dataset in terms of (a) Number of classes, (b) Number of features, and (c) What the data represents, i.e., gain some intuition about the problem domain. Based on your study, would you expect the features to perform well in this problem?
- Compute the following quantities for each feature. Do you observe anything of interest from these statistics?

	Sepal Length	**Sepal Width**	**Pedal Length**	**Pedal Width**
Minimum				
Maximum				
Mean				
Variance				
Within-Class Variance	$sw(i) = \sum_{j=1}^{M} P_j \sigma_{ji}$, where σ_{ji} is variance of i-th feature in class j, and P_j is a-prior probability of class j			
Between-Class Variance	$sb(i) = \sum_{j=1}^{M} P_j (\mu_{ji} - \mu_i)^2$, where μ_{ji} is mean of i-th feature in class j, and μ_i is the mean of the i-th feature			

- Compute and display the correlation coefficients exactly as shown below (left figure). Do you observe anything interesting from this display?
- Display each of the four features versus the class label, exactly as shown below (right figure). What can you state about how well the features may perform in classification?
- Perform the following classification tasks.

Setosa Vs. Versi+Virigi	All Features	Batch_Perceptron and LS
Setosa Vs. Versi+Virigi	Features 3 and 4 Only	Batch_Perceptron and LS
Virgi Vs. Versi+Setosa	All Features	Batch_Perceptron and LS
Virgi Vs. Versi+Setosa	Features 3 and 4 Only	Batch_Perceptron and LS
Setosa Vs. Versi Vs. Virigi	Features 3 and 4 Only	Multiclass LS

- For each case, (a) report whether the method converged, (b) No. of epochs, (c) Computed weight vector, (d) No. of training misclassifications, and whenever appropriate, (e) plot of feature vectors, as well as the computed decision boundary.

Upload your .m or .py file to Blackboard prior to the deadline.