**ECE 8527: Introduction to  
Machine Learning and Pattern Recognition**

# HW No. 1: Gausssian Distributions and MAXIMUM LIKELIHOOD DECODING

For this assignment, you will use the data set located here:

*https://www.isip.piconepress.com/courses/temple/ece\_8527/resources/data/set\_04/*

We will focus on the files *train.txt* and *dev.txt*, which containing training and development test set data.

The tasks to be accomplished in this homework assignment are:

1. Select the two dimensions of the data that have the greatest variance and do a scatter plot of each class. Estimate the mean and covariance for each class. Write an expression for a Gaussian distribution that models the data in each class. Superimpose this model on your scatter plot (hint: using support regions).
2. Repeat this for the two dimensions that have the greatest overlap.

Consider classification of the data using:

1. The feature with the largest covariance.
2. The two features with the largest covariance.
3. The feature with the smallest covariance.
4. Two features: the feature with the largest covariance and the feature with the smallest covariance.

Comment on your ability to classify the data using these features.