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

# HW No. 3: Nonlinear DEcision Surfaces

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

*https://www.isip.piconepress.com/courses/temple/ece\_8527/resources/data/set\_05/yinyang.py*

This tool generates data that follows the shape of a yin yang (Taijitu) symbol. The interface is simple:

*python yinyang.py N0 N1 <overlap>*

where N0 is the number of points in class “0”, N1 is the number of points in class “1”, and <overlap> is a parameter that controls the overlap between the two classes. Its range is [-1,1].

The tasks to be accomplished in this homework assignment are:

1. Using a maximum likelihood (ML) classifier (assume the priors are equal), generate a training set with 10,000 points in each class using an overlap of -1. Train an ML classifier on this data. Generate an independent evaluation set of 10,000 points per class. Evaluate it using your classifier trained only on the training data. Measure performance on both the training set and the evaluation set. Draw the decision surface that corresponds to your classifier. Superimpose this over a scatter plot of the data.

Repeat this for a value of -0.10 for the overlap parameter. Discuss the implications of these plots and the impact the training data had on the decision surface.

1. Repeat the data generation and error computation steps of part 1 for the overlap parameter values shown in the table below. Use this template for your table:

|  |  |  |
| --- | --- | --- |
| Overlap | Train | Eval |
| -1.00 | xx.xx% | xx.xx% |
| -0.25 | xx.xx% | xx.xx% |
| -0.10 | xx.xx% | xx.xx% |
| 0.00 | xx.xx% | xx.xx% |
| 0.10 | xx.xx% | xx.xx% |
| 0.25 | xx.xx% | xx.xx% |
| 1.00 | xx.xx% | xx.xx% |

and provide error rates for each setting of the overlap parameter.

Finally, include a discussion of the performance of this classifier based on what you have observed in these experiments. What could you do to improve performance?