Class Project

Robert M. Haralick

Computer Science, Graduate Center
City University of New York

Class Project

- The class project
- Is for each student
- To do a few experiments
- With variants of the N-Tuple Subspace Classifier
- On real data sets
- That can be found in repositories
- And are associated with published papers

- Although the experiments are all different
- Students wanting to work in a group to develop the code
- Are encouraged to do so
- You must first determine the subroutines needed
- The input and output arguments
- And the data structures involved

- Those who want to program alone can do so
- Those that want to divide up the programming can do so
- Do what has to be done to give confidence that the code is bugfree

- The Experimental Protocol
- Specifies how the experiment will be done
    - The training set size
    - The testing set size
    - Cross Validation
    - The efforts taken to make sure that for the given training set size
    - The requisite variety criterion is maintained
- What the fixed parameter settings will be
- What variable parameter values will be tried
- How the analysis of the experimental results will be presented
- What kind of graphs and figures will be used

The Experimental Protocol

- Must tell the reader exactly how you did your experiments
- Must be complete enough that the reader could replicate your experiment
- And get the same results modulo sampling variation

Report

- Must tell the reader all the technical detail so that what you do can be programmed by somebody else
- Use the abstract mathematical notation along with natural language for the technical description
- Describe exactly the training and test aspects – cross validation
- Training set sizes per class must be 10 times the amount of storage required for the largest subspace class conditional probabilities
- The report should be about 10 pages long
- The report does not need to provide a review of related published papers
- Experiments must be done on 3 different data sets obtained from one of the depositories
- The number of classes must be 3 or larger
- The number of dimensions must be larger than 25

- Introduction
- Technical Description
- Experimental Protocol
- Results