

## **Active Learning for Computer Vision Curriculum**

### **Project 4: Pool-based Active Learning – Part 2**

**Total Points Possible: 50**

**Datasets:** The datasets required for this project are included in the folder. For the feature matrix in each dataset, each row denotes a sample and each column denotes a feature. Each dataset contains an initial labeled training set, an unlabeled set and a test set. Also, each experiment needs to be run 3 times (with different initializations of training, unlabeled and test sets) and the average results should be reported.

#### **Problem 1 (50 points)**

The VidTIMIT dataset is widely used for facial image recognition and contains images of subjects reciting short sentences under unconstrained natural conditions. We will use images of 25 randomly selected subjects in this experiment.

The MindReading dataset is widely used for facial expression recognition and contains images belonging to 6 different emotion classes.

Implement the following QBC based active learning algorithms (refer to the query-by-committee paper by Liere and Tadepalli for reference):

- i) active-majority
- ii) passive-majority
- iii) active-single
- iv) passive-single

Run the process for 1000 iterations. Plot the training set size vs. accuracy graph.