Active Learning for Computer Vision Curriculum

Project 7: Multi-label Active Learning

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.

Problem 1 (25 points)

The Scene dataset consist of 2407 images of an outdoor scene, where each image is represented by a feature vector of dimesnion 294. Also, there are 6 classes in the problem and an image can belong to one or more of the 6 classes. The dataset has been divided into a training set (with 1500 samples) and a test set (with 907 samples). Each row of X_train and X_test denotes a sample and each column denotes a feature. Each row of y_train (y_test) denotes the labels of the corresponding training (testing) sample, where 1 means the class is present and 0 means the class is absent. For instance, in the training set, sample 462 belongs to classes 4 and 5.

One strategy to solve a multi-label learning problem is to train an SVM separately for each class. To predict a test sample, each SVM is applied separately on it. A positive output indicates that the corresponding class is present and a negative output indicates that it is absent. The accuracy is computed using the following expression:

$$A = \frac{|T \cap P|}{|T \cup P|}$$

where T is the true class label vector of a test sample and P is the predicted class label vector. Train an SVM classification model on the training set and test on the test set. Report the percentage accuracy on the test set using the following classification models: (i) SVM with polynomial kernel with parameter 2 and (ii) SVM with Gaussian kernel with parameter 2.

Problem 2 (25 points)

Split the training samples of the Scene dataset into 2 groups – use 500 for initial training and 1000 as an unlabeled pool. Implement the LMAL Multi-label active learning algorithm and test its performance on the Scene dataset. Use batch size 10 and 100 iterations. Use Random Sampling as a comparison baseline.