# Project1

1. Write KNN classifier
2. Write Centroid Classifier
3. Write Linear Regression classifer
4. Download C.J.Lin 's libsvm and learn to use it on the matlab

Use 4 classifiers to run on AT&T50 and AT&T200 files to learn the model paramters and predict the class labels for test images

# Project 2

Do 5-fold cross-validation on AT&T400 and Hand-written-letters image data
Repeat this for all 4 classifiers

# Project 3

Data Cluster using K-means algorith provided by matlab.
Run k-means on AT&T 400 images, set K=10. Obtain confusion matrix. Re-order the confusion matrix and obtain accuracy.
Run k-means on Hand-written-letters data, set K=26, as above.

# Project 4

Computer f-statistic and select top 10, 20, 30, 50, 100, 200 features (image pixels) from AT&T400 and Hand-written-letters image data. Run 5-fold Cross-validation on selected data.

For each data set, for each classifier, plot the CV accuracy as a curse using whatever curve-plotting software (matlab has it ! )

# Project 5

Dimension Reduction and Laplacian Embedding

On AT&T400 and Hand-written-letters image data,
(a) Do PCA to reduce to 10,20,30,50,100 PCA dimensions.
(b) Do Linear Discriminant Analysis to reduce to K-1 dimension.
(c) Do Laplacian embedding of the data in 5, 10, 20, 30, 50 dimensions
(d) Do cross-validation on the transformed (reduced) data using SVM and Centroid, Linear Regression