* In this project, you are required to train a model and then test on a test set.
* [traindata.csv](https://home.cse.ust.hk/~yuzhangcse/hkust_only/MSBD6000B/data/project1/traindata.csv), [trainlabel.csv](https://home.cse.ust.hk/~yuzhangcse/hkust_only/MSBD6000B/data/project1/trainlabel.csv), [testdata.csv](https://home.cse.ust.hk/~yuzhangcse/hkust_only/MSBD6000B/data/project1/testdata.csv)
* The training set includes a set of training data points stored in traindata.csv and the corresponding class labels stored in trainlabel.csv.
  + Each line in traindata.csv corresponds to a training point. There are in total 57 continuous features separated by commas.
  + Each line in trainlabel.csv contains the class label of the corresponding data point.
* You need to train a model on the training set and additional pre- or post-processing on the training set is allowed.
* You can use any classifer such as linear regression or SVM.
* You need to predict class labels of the test data points stored in testdata.csv, which has the same format as traindata.csv.
* The prediction of your model on the test dataset should be stored in a csv file where each line stores a predicted class label for one test data point.
  + The file name of the csv file should be project1\_student id.
  + For example, if your student id is 10001, then the csv file should be named 'project1\_10001.csv'.
* You have to submit a 1-2 page report in a pdf file on what model you used and pre-/post-processing (if any).
  + The file name of the pdf file should be project1\_student id\_report.
  + For example, if your student id is 10001, then the pdf file should be named 'project1\_10001\_report.pdf'.
* Grading will be based on the testing accuracy.
* Due date: 11:59pm, Nov 16, 2017.
* Please email your prediction (in csv file) and report (in pdf file) to WANG Weiyan, the TA for this course, via wwangbc@cse.ust.hk before the deadline.
* For enquries, please contact the TA