CSE455/CSE552 – Machine Learning (Spring 2015)

Homework #1

Handed out: 9am Tuesday March 3, 2015.

Due: 11:55pm Sunday March 22, 2015.

Hand-in Policy: Via Moodle. No late submissions will be accepted.
Collaboration Policy: No collaboration is permitted.
Grading: This homework will be graded on the scale 100.

Description: Experiments with KNN and SVM on IRIS classification data.

We will explore supervised learning techniques on the well known IRIS data (for classification of flowers based on a few dimensional measurements). The data is available in R releases. It has also been made available here.

In essence this project is expecting you to write four different functions to test your solutions to the problem. You are expected to use the R language. You will prepare a report including your code and results. The report format is attached.

- Part 1: Build a classifier based on KNN (K=5 for testing) using Euclidean distance.
  - o You are expected to write the classifier by yourself.
  - o Report performance using an appropriate k-fold cross validation using ROC curves and confusion matrices.
- Part 2: Build a classifier based on KNN using Manhattan distance.
  - o You are expected to write the classifier by yourself.
  - o Report performance using an appropriate k-fold cross validation using ROC curves and confusion matrices.
- Part 3: Build a classifier based on linear SVM.
  - o You may use an available implementation of SVM in R.
  - o Report performance using an appropriate k-fold cross validation using ROC curves and confusion matrices. Find the best algorithm based on SVM thresholding as described in the note by Fawcett.
- Part 4: Build a classifier based on polynomial SVM.
  - o You may use an available implementation of SVM in R.
  - o Report performance using an appropriate k-fold cross validation using ROC curves and confusion matrices. Find the best algorithm based on SVM thresholding as described in the note by Fawcett.

What to hand in:

- **Homework1_lastname_firstname_.doc** (template provided).
- **Homework1_lastname_firstname_.pdf** (pdf version of the above).