IST407/707 Data Mining

HW7: SVMs, kNN, and Random Forest for Customer Churn

Now that we have learned about the final three classification algorithms we will study in this class (SVMs, kNN, and Random Forest) let’s see which of the 5 classification algorithms we have studied is the best model for the Telecommunications Churn Data Set.

In this homework, you will use SVMs, kNN, and Random Forest algorithms on the Telecommunications Customer Churn data and compare their performance with the naïve Bayes and decision tree models you built for HW6.

Deliverables:

1. Write a report to describe what you did: including the data preparation, transformation, algorithm tuning, and model generation for each model. Describe which model you feel is the best model for this application.
   1. Please review and report on all 4 models evaluation measures (Correctly Classified Instances, Precision, Recall, F-Measure) for each of the 5 models.
   2. Please indicate which evaluation measure you think is best the measure of model accuracy, and why.