SVM CA:

Libraries: ISLR

e1071

Data: *Default* (data size = 10,000, 9667=No, 333=Yes)

- Customer default records for a credit card company
- Four columns: *default* (Yes/No), *student* (Yes/No), *balance*, *income*

Task: to fit SVM model for prediction of ***default*** of customers in "No" and "Yes", using the other variables as predictors

Requirements

1. Randomly pick 80% of the data to train SVM
2. Try using two different kinds of kernels
3. For each of the models from above, use tune() function to find the best set of parameters (i.e: *gamma*, *cost*)
4. Using your trained SVM for predication on the test data (20% of the given data set), and summarize the accuracy of different models with different settings.
5. Instead of using all the three variables (student, balance, income) as predictors, use two of them to build SVM models and compare the performances of different combinations.
6. Discuss your findings and understanding (if any).

Submission (team basis)

A report including:

1. The R code
2. The model summary and the prediction table of all the models you have built (screen cut will do)
3. Your findings and understanding

- End -