

PLEASE SUBMIT YOUR WORK ON E-Mail (Also keep a copy on sent items for just in case).

Due Date: Apr 24

The assignment is individual, so please do not make any information exchange or do not discuss about the answers or ask anybody for help except the instructor. You can use any Internet resources for reading, you can search for the answers on any search engine (like Google) or you can use any textbooks.

For all the parts below, please write your code in Lisp Programming Language and test with Dr. Racket.

Problem Definition

You are asked to find a machine learning solution for classifying Smith College students.

Data Set

You are free to use any attributes for your data set and you can collect data from any source or even you can create your own fictive data set. Prepare your data set in Excel or with your favorite spread sheet software and also submit your data set as a CSV file with your submission.

Steps

- 1) Create your own data set
- 2) Save as CSV File
- 3) Import into Rapid Miner
- 4) Create a knowledge flow and find a machine learning algorithm which can classify your sets
- 5) Try different machine learning algorithms (like ANN, K-NN, Naïve Bayes, etc)
- 6) Try different test / train set ratio.
- 7) Try to change the k value for k-fold cross validation
- 8) Report your findings after each experiment
- 9) Comment and Discuss about your findings.
- 10) Submit your CSV file and Rapid Miner Process.

Steps

A sample Rapid Miner process is provided below:

