CSCI 3900C Project VII (50 points)

Write an R Markdown document to perform and report the analysis described below. Your report should explain (in paragraph form using full sentences) the process you followed, the results you obtained, and some interpretation of those results. Embedded in the report should be the code that was used to obtain the results, as well as any supporting figures and plots. Each result should have some accompanying verbal explanation (do NOT just show code and results!)

Publish your final compiled html file in RPubs, and submit your work as follows:

- Upload your .Rmd source file to the Project VII D2L drop box.
- With your dropbox submission, include the URL for your published file on RPubs

Download the following file from D2L:

weightlift.csv (This is the data file for the assignment)

Background

Using devices such as Jawbone Up, Nike FuelBand, and Fitbit, it is now possible to collect a large amount of data about personal activity. Such devices allow enthusiasts to take measurements about themselves regularly to improve their health, to find patterns in their behavior. One thing that people regularly do is quantify how much of a particular activity they do. However, they rarely quantify how well they do it. In this project, your goal will be to use data from accelerometers on the belt, forearm, arm, and dumbell of 6 participants. They were asked to perform barbell lifts correctly and in 5 different incorrect ways. More information is available here: http://groupware.les.inf.puc-rio.br/har (see the section on the Weight Lifting Exercise Dataset).

Analysis Task

- Use machine learning to create a model that will predict the manner in which they did the exercise. This is the "classe" variable in the training set.
- Use cross-validation.
- Narrow down the variables you use for prediction to no more than 30.
- Remember to explain your process throughout your write-up.
- Report your final in-sample accuracy and estimate out-of-sample accuracy.