ISE 5103 Intelligent Data Analytics
Final Project

Instructor: Charles Nicholson

Due dates
Proposals: 11/08/15
First Draft: 12/06/15
Peer Critique: 12/13/15
Final Version & Reflection: 12/17/15

REQUIREMENTS

You will choose a data-intensive problem to explore and solve using a variety of analytics techniques. The problem should be of sufficient complexity to challenge you. You must submit a problem proposal by 11/08/15 at 11:59 pm. You will not be allowed to work on the same problem as another student – proposals will be considered on merit as well as first-come, first-serve.

Problems may be based on current research you are pursuing, problems from industry (e.g. from your place of employment), or from analytics competition websites such as,

- kaggle.com
- crowdanalytix.com
- KDD Cup (sigkdd.org/kddcup/index.php)
- tunedit.org/challenges
- www.kdnuggets.com/competitions/past-competitions.html

Due dates
Project Proposals are due on 11/08/15 at 11:59 pm. Prior to this date please share your ideas with classmates on the discussion board to ensure that you pursue a unique topic.

The *First Draft* of the report is due 12/06/2015 by 11:59 pm. Please upload it to janux.ou.edu *and* email it to the person you have been matched with for the Peer Critique.

The *Peer Critique* is due 12/13/15 by 11:59 pm. Please upload your critique to janux.ou.edu and email it to the person who wrote the report.

Both the *Final Version* of the report and the *Reflection Paper* are due on 12/17/15.

Format of the Project Report
You will submit:

- One PDF file about 12 pages + appendix
- Complete commented, and "compilable" R script text file and raw "output" file from R script – submitted via janux.ou.edu

Grading
The entire project is worth 30% of your final grade and is graded based on the quality of your effort, creativity of your approach, and your ability to communicate the problem, technique(s), and solution effectively.

DELIVERABLES

Project Proposal
The *Project Proposal* is a brief (one page max) description of the problem context, the type of problem (prediction, classification, clustering, etc.), and your initial thoughts on the techniques you will use. Please include a hyperlink to the problem description if you are using one of the competition websites.

Project Report (First Draft & Final Version)
The *Project Report* should include the following:

- Executive Summary: 1 page
  - Concise problem statement
  - List of major concerns/assumptions (if any)
  - Summary of findings
  - Recommendations
- Problem description and background, including related literature for problem, solution techniques, etc.
- Exploratory data analysis – the highlights; not the kitchen sink
- Analysis plan
  - Explanation of modeling choice – Why choose this technique? Strengths, weaknesses?
  - Feature selection, engineering, missing value, outlier plan
  - Validation plan (including how do your findings compare with others?)
- Results and validation of analysis
- Conclusion
- Appendix:
  - Data visualizations, tables, etc. which support the work, but are not of primary importance
  - List of data transformations, missing value imputations, outlier treatment, etc.
  - List of any important assumptions not otherwise included
  - Important code excerpts or algorithms used / developed if any.

The project report should not exceed 12 pages. The appendix does not have a page limit.

Peer Critique
You will write a 3-5 page *Peer Critique*. The Peer Critique gives you the opportunity to review another student's project and receive feedback on your own project before submitting the final version for a grade.

You will be asked to:
1. Evaluate the project's strengths
2. Identify weaknesses or omissions
3. Make suggestions for improvements

You will provide (1) strengths (2) weaknesses and/or (3) suggestions for each of the following areas:
- Project Understanding
- Data Understanding
- Data Preparation
- Modeling
- Evaluation
- Overall/Summary

Your Peer Critique will be evaluated on the substance of your feedback. Your *Critique* will be assessed in the following areas: communication, analysis, evaluation, synthesis and reflection.

Reflection Paper

The *Reflection Paper* is a 2-3-page paper that will be submitted when you upload the final version of your *Project Report*. In this paper you will reflect on your process, the feedback you received and the various iterations of your project. The paper should address the following:

- Briefly describe your process
- Summarize the feedback you received from your partner
- Discuss the feedback you received:
    - What was most valuable? Why?
    - What did you reject, discount or disagree with? Why?
- Reflect on iterations:
    - What did you do to improve your project?
    - What specifically led you to make those changes?