Spring 2-1-2018

M 567.03: Advanced Big Data Analytics Projects

Brian M. Steele
University of Montana - Missoula, brian.steele@umontana.edu

Syllabus: Math 567 Big Data Analytic Projects

Course Format: Three meetings per week MWF 11:00-11:50 a.m. Class time is split (25/75) between lecture and team solutions to data analytic problems.

Learning Outcomes: At the completion of this course, the student will

1. Understand the theory supporting predictive analytics.
2. Be able to program core algorithms of predictive analytics.
3. Be able to extend and adapt core predictive analytic algorithms to complex problems and varied data.
4. Be able to develop and plan a data analytic project.
5. Be able to effectively describe and communicate solutions and problems related to team-oriented problem solving.
6. Be able to write concise and focused reports on data analytic projects.
7. Have gained experience and understanding regarding team-based problem solving.

Course Content:

The course focuses on several core algorithms of predictive analytics that are related to the projects. Undergraduates will focus on implementing the algorithms in `Python`. The emphasis of the course is on developing practical solutions to the projects. Teams are comprised of three undergraduate students for the purpose of solving problems originating from three sources: businesses, researchers, and competitions (private and public). A central focus of the course is communicating solutions to others in written and oral form. M 567(G) co-convenes with M 467(U).

Course content differs however. In particular, theoretical and foundational aspects of predictive algorithms is a significant component of M 567 but a relatively minor component of M 467. Problems assigned to graduate student teams are more difficult. Furthermore, M 567 requires a significantly more sophisticated level of writing compared to M 467.

Textbook: Algorithms for Data Science (ISBN-10: 3319457950)

Prerequisites: M 561 or consent of the instructor.

Homework: Exercises emphasizing theory and principles will be assigned weekly during the first half of the semester.

Projects: Projects originate from three sources: businesses, researchers, and competitions (private and public) sponsored by companies and researchers. Students are organized in teams of usually three graduate students for developing solutions to a problem. Teams are necessary since the solutions are involved, sometimes requiring disparate methods from statistics, mathematics, and computer science. Solutions will require innovation—the problems will not be *textbook* problems with a methodological solution apparent from the start and no hidden difficulties. Students will work on two projects during the course of the semester.

Grading: Your course grade will be based on homework, written project reports and oral presentations. Homework assignments are worth 30% of the course grade. Home work exercises emphasizing theory and proofs will be assigned weekly during the first half of the semester. Proofs MUST be complete and concise.

Expectations and grading are commensurate with graduate standing (undergraduates generally are not asked to prove theorems).

The first project, due near the midpoint of the semester is worth 20% and the remainder is attached to the second project due at the end of the semester. Project grade is based on written reports and oral presentations (mostly short reports on progress). Written reports MUST justify the methods from the standpoint of theoretical considerations and grades are based in part (25%) on this aspect of the paper. (Generally, undergraduates are not expected to focus on theoretical considerations). Oral presentations at the project end are expository and aimed at communicating methods and results. Graduate students are encouraged to submit original solutions for publication.

The meeting time of the final is Thursday, May 10, 10:10-12:00.