

FALL 2017 - STAT 440 E100

LEARNING FROM BIG DATA (3)*Class Number: 4488 Delivery Method: In Person***COURSE TIMES + LOCATION:**

Mo 4:30 PM – 5:20 PM

BLU 10031, Burnaby

We 4:30 PM – 6:20 PM

BLU 10031, Burnaby

INSTRUCTOR:

Jack Davis

jackd@sfu.ca

PREREQUISITES:

90 units including STAT 350 and one of STAT 341 or CMPT 225, or instructor approval. STAT 240 is also recommended.

Description

CALENDAR DESCRIPTION:

A data-first discovery of advanced statistical methods. Focus will be on a series of forecasting and prediction competitions, each based on a large real-world dataset. Additionally, practical tools for statistical modeling in real-world environments will be explored.

COURSE DETAILS:

STAT 440 is suitable for senior students who have a minimum of 90 units.

Course Outline

The course will be split into 4 (possibly 3 or 5) modules of roughly equal length. Each module will focus on a particular dataset. While some will be private to the course, as many as possible will be public competitions, such as those on Kaggle.com. At the outset of each module, students will be randomly assigned to teams (of size 1+, depending on module). All teams will be given a subset of the data, and for each week of the module every team will submit their predictions on a withheld, testing portion of data. The teams will be ranked each week, and all teams will distribute a short description of their prediction method with accompanying code to all other teams. As such, the course will take a data-first approach, starting with data and having the students learn from each other. In addition to the competitive data modules, the course will cover a series of tools for statistical modeling in real-world environments. Some examples include version control, agile software development, bagging, boosting, parallelization, collaborative model development, cross-validation, model validation and verification. These tools are often found lacking by students who enter industry, and as such the goal is to fill some of these gaps.

Grading

Competition Results	50%
Competition Writeup	30%
Participation	20%

NOTES:

Assignments and Grading Procedures

Competition Results (50%): There will be weekly competitions throughout the semester. The weekly grades will be between 10/10 (the top team) and 5/10 (the bottom team) each week, with intermediate teams' grades depending on their relative scores. All competitions will be equally weighted into the 50%. Teams will be re-randomized with each new data set (roughly every 3-4 weeks).

Competition Writeups (30%): Every week teams will release their code, along with a descriptive writeup, to make their contributions accessible to the other teams. This is a strict requirement --any missing writeups/code will result in an incomplete grade for the entire course.

Participation (20%): Students are encouraged to participate in class by asking questions and engaging in discussion.

Above Grading is subject to change.

DEPARTMENT UNDERGRADUATE NOTES:

Students with Disabilities:

Students requiring accommodations as a result of disability must contact the Centre for Students with Disabilities 778-782-3112 or csdo@sfu.ca

Tutor Requests:

Students looking for a Tutor should visit <http://www.stat.sfu.ca/teaching/need-a-tutor-.html>. We accept no responsibility for the consequences of any actions taken related to tutors.

REGISTRAR NOTES:

SFU's Academic Integrity web site <http://students.sfu.ca/academicintegrity.html> is filled with information on what is meant by academic dishonesty, where you can find resources to help with your studies and the consequences of cheating. Check out the site for more information and videos that help explain the issues in plain English.

Each student is responsible for his or her conduct as it affects the University community. Academic dishonesty, in whatever form, is ultimately destructive of the values of the University. Furthermore, it is unfair and discouraging to the majority of students who pursue their studies honestly. Scholarly integrity is required of all members of the University. <http://www.sfu.ca/policies/gazette/student/s10-01.html>

ACADEMIC INTEGRITY: YOUR WORK, YOUR SUCCESS