MSDS Capstone Syllabus (2017-2018)

1. Introduction

The Capstone class is offered two consecutive quarters (DATA 590A in Autumn, and DATA 591A in the Winter). Successful students must complete both portions of this course. Quarter 1 will encompass team development, and project proposals. Quarter 2 will culminate in a deliverable project.

The goal of this class is for each graduating student to stand out in the field of data science. Upon completion of this course students should have extensive experience along with code and writing samples to demonstrate competence.

Successful data scientists must be familiar with many different aspects of engineering including, but not necessarily limited to
- Data file parsing
- Database query and storage
- Feature set design
- Data cleaning
- Model training or application
- Testing mechanisms
- User Interface design
- Code documentation
- Current literature
- Project presentation
- Source control software

At the same time, the primary concern of engineering managers or co-workers surround the communication and interpretation skills of the data scientist. Successful data scientists need to be able to communicate clearly about their proposed solutions, including both advantages and limitations. It is necessary to evaluate algorithmic output for reasonability of solution and impact - numbers should never exist in a vacuum, but be evaluated for meaning in context. Stand out data scientists are able to describe potential downfalls or unintended consequences of a given approach or chosen feature set.

2. Instructors

Dr. Megan Hazen (mh75@uw.edu) has more than 15 years experience using data science in both research and industrial settings (https://www.linkedin.com/in/megan-hazen-3037564/).

Teaching assistant, Huazeng Deng (hd6@uw.edu) has more than 7 years working with data driven signal processing (https://www.linkedin.com/in/huazeng-deng-b94703a4)

3. **Meeting times and locations**

Fall quarter - Wednesday, 6:00-7:50 pm, Savery Hall 164
Spring quarter - Wednesday, 6:00-8:50pm, Savery Hall 164

Office Hours - Office hours are available upon request. Conversation and discussion on the *Slack* channel is also encouraged.

4. **Tools and methods**

One motivation for this course is to allow graduating students to hit the ground running in their new careers. As such, the tools and methods for this class are designed to look like those found in industry.

Project proposals and reports will be written as professional presentations. Documents must be carefully edited for correctness as well as clarity. Details must be included to support the project funding and scheduling.

Software management tools will include *Github* and *Slack*. Slack should be viewed as one of the primary remote communication tools, while the Github repository should be maintained to professional standards, including documentation. Students are required to use the issue tracker provided with Github to structure and manage agile development techniques.

5. **Attendance**

While attendance is not graded directly, each class session will include a graded assignment. It is the responsibility of each student to refer to the CANVAS schedule for each assignment and its grading rubric.

6. **Late policy**

If a student must miss a class session or a deadline it is the responsibility of the student to contact the professor before that meeting in order to make alternative arrangements. If alternative arrangements are in place when a student misses an assignment credit will be given, if they are not the assignment will receive a grade of '0'.

7. **Student commitments**

In order to succeed in this class students must complete a long-term team project. These projects are done in conjunction with an industrial partner. This should be viewed as a client / consultant relationship. Teams will develop a contract with their project sponsors that defines work expectations. By enrolling in this course students are committing to their best effort to fulfill that contract.

The beginning of this course will require that students participate in some foundational exercises (introductions, self-organization of teams, paper presentation, tool set=up). Success at the end of the first quarter will require student-teams to meet with project sponsors (potentially off-site), formulate a proposal, and demonstrate that they are able to use the data sources provided by sponsors.

Success in the second quarter requires that student-teams complete the deliverable according to their proposal in quarter one. Communication between team-members, with project sponsors, and with teaching staff is essential.

8. **Group work and group policy**

While not every assignment in this course is a group assignment, the preponderance of the points will depend on a long-term team project. It is essential for success that this teamwork be successful. Each individual is responsible for their own contribution to the team.

Part of team development will be the determination and documentation of team communication and work standards. These will be similar to the guidelines most work-places have for attendance and professionalism.

In order to maintain accountability, one portion of your team-project grade will be determined by your team-mates.

9. **Grading policy**

Grades will be available via Canvas. Students are responsible for reviewing their grades, and will have access to rubrics which will illuminate how assignments are graded. Discrepancies or concerns should be raised with the instructor within one week of the posted grades.

Each assignment will be worth a specific number of points. A student's final percentile grade will be equal to the percentage of all the possible points awarded that quarter. If necessary, grades will be adjusted before the final grade is assigned to ensure proper scaling on the 4.0 range.

There will be an assignment due for every class session, which will be worth around ten points. The bulk of the work for this course is focused on one single project, which will be represented by approximately half of the first quarter, and the majority of the second quarter assignments. As a result, the performance on that project will weigh heavily in a student's final grade.

10. Week-by-week schedule

Week #	Date	Topic
1.1	9/27	Introductions / tool set-up
1.2	10/4	Mock interview (student elevator pitches)
1.3	10/11	Current paper presentation / project prospectuses out
1.4	10/18	Teams formed / projects chosen / career fair discussion
1.5	10/25	Teams agreements due / proposal over-view due
1.6	11/1	Meet with sponsors, iterate proposal
1.7	11/8	Work session on proposal; technology research
1.8	11/15	Detailed proposal due
1.9	11/22	No meeting; Thanksgiving Holiday
1.10	11/29	Stand-up session on data pipe development
1.11	12/6	Data pipe demonstration
2.2	1/3	Proposal revisions due; week-by-week schedule
2.2	1/10	Stand-up
2.3	1/17	Stand-up
2.4	1/24	Stand-up
2.5	1/31	Stand-up / interim delivery
2.6	2/7	Stand-up
2.7	2/14	Stand-up
2.8	2/21	Stand-up

2.9	2/28	Poster review , final standup.
2.10	3/7	Final project fair

https://join.slack.com/t/data590a/shared_invite/enQtMjQzNjcxNjg3MjY2LWExM2U5ZTY1YTIxM2U4ODgzMzY3Y2M3YjZkNzJmOWEyNWEwOWJmMzc2ODk1Njc2Yjg1ZTAwZTUwYWY3MzdkMGM