MSDS Capstone Sponsored Projected Guidelines (Fall 2018)

1. Project Goals
    a. Deliverable Guidelines
    This course envisions student teams working as consulting teams for sponsors. The student team is responsible for providing professional proposal,software, and documentation.
    Examples of possible deliverables include, but are not limited to, data retrieval interfaces, classifier for data management, and predictive model creation.
    b. Process
    Steps for this project are laid out in the course syllabus. While the student teams are responsible for a number of professional quality products, they will be coached through each step with the course structure and support. Processes specific to deliverables will be described in the project proposals.
    c. Student Learning
    There is no substitute for practical experience in data science. Students entering the capstone course will be familiar with many different parts of a data science projects, and many different algorithmic solutions to data science problems. The capstone course is designed to allow students to follow one problem through all phases of the solution.
    d. Sponsor Benefits
    The sponsor benefits from this process by having a knowledgeable team work on their problem. It is the intention that the sponsor receives a usable product at the conclusion of the course, although this is not guaranteed.

2. Project Scope
    The intent of the MSDS capstone experience is for students to complete all phases of a data science solution. This implies that students will be responsible for querying and processing data, developing feature sets and applying appropriate algorithms, and a basic user interface for receiving data analysis.

    The time-line for this work is tight, however, so there are limitations on how many features may be implemented. The focus will be on getting an end-to-end working project, instead of on a wide array of features. It may be the case that diligent student work fails to provide a final solution for the specified project. In that case the student team will provide an explanation for failure as well as insights into recommended next steps.

3. Project Schedule:
    a. Sponsor schedule
    Sponsors should have a problem statement, along with desirable deliverable features at the start of the fall quarter. Sponsors should expect to spend significant time facilitating data access during the fall quarter. Sponsors should

remain available through the winter quarter for consultation, and will be able to take delivery at the end of the winter quarter.

b. Student schedule
Students will make a project proposal in mid fall. Students should have access to the data during the fall quarter. Students will refine their proposal at the beginning of the winter quarter. Students will deliver final product at the close of the winter quarter.

4. Student Commitments
   a. Solution Proposal
   Students will create a product proposal that details which data (features) are expected to be available, which algorithms are proposed for analysis, and how output calculations will be processed and presented. The proposal will also include a proposed schedule and process.
   It is anticipated that an initial proposal is completed in the fall quarter, with a revised and final proposal completed in the early winter, after the student team has had access to the data.
   b. Intermediate Deliveries
   Students will schedule intermediate deliverable dates to meet key milestones. Intermediate deliverables may consist of goals such as data-reading, feature vector creation, model creation, interface demonstration, etc. Intermediate deliverables will be described and dated in the solution proposal.
   c. Documentation
   Students will fully document their product. This will include external documentation, as well and commented code. Documentation will be sufficient to allow an outside user to use the product.
   d. Solution Delivery
   Students will provide all code and documentation to the sponsor. It is anticipated that this product will be developed and maintained through Github.

5. Sponsor Commitments
   a. Feature Specifications
   Sponsors should have a good idea of what issues are concerning to them. Before the start of the first quarter, sponsors should have a clear idea of what the final product should look like, as well and some criteria for determining success in the product. The sponsors should have a list of both necessary product features, and desirable but optional product features.
   b. Data Provisions
   Sponsors should have data available for students to work with. It the is the sponsor's job to provide the data is an organized manner (a .csv file, or database, or similar). This data must be available within the first few weeks of

the fall quarter. Data science success is often predicated on the availability of high quality data, so it is in the sponsor's best interest to provide as much data as possible, including labeled data and testing data were appropriate. It is also in the sponsor's interest to facilitate data access and interpretation. The sponsor should stop short, however, of instructing the students in how to process or use the data, as that is a fundamental part of the data science problem.

c. Communications Response
   Sponsors are required to respond student communications within one week of a contact. While the sponsor is not required to coach the students, reasonable questions or requests should be answered in a timely manner. The sponsor should expect queries about data access and interpretation which may require additional documentation or support to resolve. The sponsor should expect queries about desireable project features which may require project review to resolve.
   Successful teams have regular (weekly, ideally) meetings with the sponsor. Meetings may be held in person, or electronically.

6. Problem Resolution
   It may be expected that some problems will arise during this process. If there are concerns from either sponsor or student that can not be easily resolved, the instructor should be contacted. The instructor will moderate a solution. It is of high priority that, once a project is started, it is seen through to completion by both parties. However, if problems prove intractable after instructor moderation the project may be terminated.

7. Intellectual Property Summary
   UW and the MSDS program have developed IP agreements that may be used to protect sponsor and student interests. Copies of those letters are available upon request. Generally speaking, students should be able to share any algorithmic developments, and summary conclusions. Limitations on detailed data publication are permissible. Sponsors and student-teams are responsible for reaching an understanding regarding IP.

8. Key Dates

   Autumn Quarter - 9/27/2018 through 12/9/2018

   - **9/20** Short (1 p.) problem description. Should include a brief description of the sponsor, the big picture criteria for success, and any other notable points of interest. Please fill in the Google form. https://docs.google.com/forms/d/e/1FAIpQLSd1FXm6FT_9BjkRN1Pry76SliNR83eep0eRgOKixUTdg-6GcQ/viewform?usp=sf_link
   - **10/10-10/24** Meet with available team. Sponsors should meet with the team who has chosen their project (on site, or via skype/hangouts) to detail the project

description. At this time teams are provided with a coherent list of items needed for success, a detailed description of how data is accessed, an idea of how to communicate and work with the sponsor. Students will be prepared to ask questions in order to clarify their understanding in anticipate of project proposal.
- **11/7-11/14** Review project proposal with team
- **11/15-12/7** Help with data extraction. Sponsors should be available to facilitate data extraction and ensure that teams are able to gather necessary information for success.

Winter Quarter - 1/7/2019 through 3/9/2019

- **1/9/2019, 2/6** Review project progress. Sponsors should expect a formal check in with students to check on progress. This may be done on-site, via skype/hangouts, or via email.
- Help with project development. Sponsors should be available for project questions.
- **3/7** Attend the poster presentation, take delivery.