Syllabus for DS 6011: Capstone Work I

Last updated: August 29, 2024

Overview

Instructor Name and Contact Information:

Adam Tashman

Email: apt4c@virginia.edu

Subject Area and Catalog Number: Data Science DS 6011

Year and Term: Fall 2024

Level: Graduate

Credit Type: Graded

Meeting Time and Location
Classes meet in person each Friday (unless otherwise noted)
Time: 11:00AM – 11:50AM
Location: SDS 305

Course Description

The purpose of the capstone experience is to provide students with an understanding of data science as it may be practiced in the context of real-world problems. It allows the student to collaborate as a member of a small team to pull together knowledge from the diverse areas of data science and integrate them in an effective and professional manner. The focus is on addressing an actual client's need by building a data product that can be shared with the client. The course comprises two semesters, DS 6011 and DS 6013, each devoted to roughly half of the data science pipeline. A simplified version of this pipeline has these four phases: (1) framing the question, problem or value proposition, (2) acquiring, transforming, and exploring the data, (3) analyzing and modeling the data, and interpreting the results, and (4) packaging, communicating, and sharing the results. In general, DS 6011 covers parts 1 and 2, while DS 6013 focuses on parts 3 and 4. During this process, students will be introduced to both technical and

professional challenges and will be guided by basic project management protocols in meeting them. The course sequence culminates in the production of a publishable paper and a presentation of results to the client.

This Document

This document focuses on the first semester (DS 6011) of this experience. A separate syllabus will be provided for the second semester (DS 6013). The following is designed to provide general guidelines for each individual capstone seminar. Capstone advisors may amend as necessary.

Learning Objectives

1. Gain an understanding of data science as it may be practiced in the context of real-world problems

2. Collaborate as a member of a team to pull together knowledge from the diverse areas of data science and integrate them in an effective and professional manner

3. Address an actual client's need by building a data product that can be shared with the client

4. Demonstrate oral and written communication skills through a formal paper and presentation of project outcomes

General Process

Each capstone mentor has the discretion to oversee the students' learning experience and interaction with the client using tools and methods they find appropriate. In general, each individual capstone group will make use of a task management tool, such as Trello or Smartsheet, to plan, organize, and share activities. In addition, teams will establish a communication protocol for meeting among team members and clients.

Class meetings

Classes meet once a week for one hour. A typical one-hour course meeting might follow this pattern:

1. Begin with any announcements and course business issues. (5 minutes)

2. Have each capstone team report on progress and issues arising, using a task management tool to guide discussion. Allow other teams to respond and discuss. Teams may also demonstrate data and code at this point. Advisors may suggest that members of each team rotate the duty of reporting. (10 minutes)

3. Have teams break out into their own group and work, making sure to update their task board with new activities. Optionally reconvene and show updated tasks.

4. Steps 2 and 3 may be reversed as well.

5. 15 min discussion on a project management topic.

In general, these meetings may be run as business meetings or stand-up meetings, where the goal is to assess the current state of things and make decisions about moving forward. In

addition, time may be taken in these meetings to present information or provide guidance on technical topics as they arise.

If a team member cannot attend a weekly meeting due to a valid excuse, they should notify the instructor and the team members ahead of time. Skipping a meeting without notification may result in a score of 0 for the weekly evaluation.

Milestones

In addition to the general process for each class meeting, the structure of the semester is loosely structured as a project, with the expectation that the following sequence of milestones be met:

1. <u>Overview meeting</u>. This is the first meeting of the class after capstone projects and teams are selected.

2. **<u>Team contract</u>.** As soon as teams are formed, students should produce a document that establishes expectations for participation in the capstone team.

3. <u>Kick-off meeting with client</u>. This should happen as soon as possible. Students are expected to set up a meeting with their clients.

4. <u>Review of the literature</u>. In addition to discussing project scope and goals with the client, students are expected to review existing literature on the project problem, in order to focus attention on current methods and approaches.

5. <u>Data acquisition</u>. Parallel to the process of reviewing the literature and framing the problem, students should be actively engaging with the client to acquire the data required to complete the project.

6. **<u>Project proposal</u>**. The project proposal defines the scope of the project and its general requirements. It also outlines the process by which the project will be completed. A template is provided for the proposal.

7. **<u>Progress Presentation</u>**.

Items in bold are associated with assessments. Instructions for these are provide in separate documents.

Communication

Teams will be used as the central channel for course collaboration. Please use Teams to post questions regarding project problems, objective identification, data collection and/or compilation, data processing, data structure and management, and exploratory data analysis. We'll use this for current updates for dates/times…etc.

Grading

Participation	10%
Team Contract	15%

| Project Proposal | 40% |
| Progress Presentation | 35% |

Course GitHub Repo

The course has a GitHub repo saved here: https://github.com/UVADS/ds6011/
A list of topics can be found at that site. Some of the potential topics include:

1. Understanding business requirements
2. ML problem framing
3. Effective communication
4. Effective statistics and error analysis
5. Giving effective presentations

Assessments

Team Contract. The team contract is a document the students in each term create to codify and agree on behavioral and communicative expectations in the project. It is meant to make all team members aware of the professional expectations for participating in the capstone.

Project Proposal. The project proposal defines the scope of the project and its general requirements. It also outlines the process by which the project will be completed. It is a more or less binding document agreed to between the team and the client.

Progress Presentation. At the end of the first semester, teams will present on their progress in the first two phases of the pipeline—-framing the problem and establishing the data. This presentation should contain a summary of work to date including questions, problems and objective identification, data collection or compilation, data processing, data structure and management, and exploratory data analysis. Students are expected to cite key literature where appropriate. After each presentation, the class will offer constructive criticism and feedback to help guide future work.

University of Virginia Honor System

All work should be pledged in the spirit of the Honor System at the University of Virginia. The instructor will indicate which assignments and activities are to be done individually and which permit collaboration. The following pledge should be written out at the end of all quizzes, examinations, individual assignments and papers: "I pledge that I have neither given nor received help on this examination (quiz, assignment, etc.)." The pledge must be signed by the student. For more information, visit www.virginia.edu/honor.

Special Needs

It is my goal to create a learning experience that is as accessible as possible. If you anticipate

any issues related to the format, materials, or requirements of this course, please meet with me outside of class so we can explore potential options. Students with disabilities may also wish to work with the Student Disability Access Center to discuss a range of options to removing barriers in this course, including official accommodations. Please visit their website for information on this process and to apply for services online: sdac.studenthealth.virginia.edu. If you have already been approved for accommodations through SDAC, please send me your accommodation letter and meet with me so we can develop an implementation plan together.