Department of Computational and Data Science
4400 University Drive, MS 6C3, Fairfax, Virginia, 22030
Phone: 703-993-3807, Fax: 703-993-9300
Web: cos.gmu.edu/cds

CDS 492 – Capstone Course in Data Science
Syllabus and Schedule – Spring 2022

1. Introduction

- **Instructor**: Dr. Ron Mahabir
- **Credits**: 3.0
- **Course Forum:** Blackboard. Allow 24 hours for a reply.
- **Office:** Research Hall, 225
- **Email:** rmahabir@gmu.edu
- **Phone**: 703-993-6377
- **Preferred Contact Method:** Email

2. Course Objectives

The capstone project aims to provide students an opportunity to integrate and apply core knowledge and skill components in data science, computational science, or computational social science that were acquired during the program in a real-world project driven setting. The problems that will addressed in this course will include extensive use of various data sources and questions from diverse scientific domains including Medical and Life Science, Social Science, Business and Policy Science, and other domains including urban and transportation. In the context of specific data and questions, students have to develop analysis methods and tools. In this environment students will be required to define possible scenarios, identify key challenges, explore possible solutions and deliver an effective solution. Given its nature, a capstone project often goes beyond a single discipline and require the application of varied disciplines to the solution of a single large-scale problem. In addition, a capstone project may require analysis at different scales, from local to regional or national. In light of this, the particular goals of the course are to:

- Allow students to implement and demonstrate their core skill set in their major.
- Develop integrative multi-disciplinary problem-solving skills.
- Promote critical thinking, including the ability to critically examine existing works and established methods, and develop innovative approaches.
- Enhance and develop rigorous writing and presentation skills.

- Enhance students' ability to work and manage a project individually or as a team.

3. Mason Core

This course addresses all criteria of a Capstone course:

- Minimum of 3 Credits: The course will be 3 credits.
- Later in the curriculum, after a student has taken at least 85 credits, and at the 400 course level.
- No more than 10 students in the course or equivalent instructional/mentored support.
- Emphasis on experiential/applied/integrative learning.
- Allow students to apply critical thinking skills.
- Learning outcomes defined by the degree program.

4. Course Schedule

The course (3 hours per week) will be taught as a semester-long class project, combining lectures, topic/problem oriented discussion, and directed reading assignments with corresponding class discussion.

5. Textbook

No textbook is required. Handouts distributed in the class will contain all essential material. Additional readings (selected readings from research journals, technical reports, and other sources) will be distributed via the course website.

6. Course Outline and Learning Objectives

We follow a problem-driven approach, and proceed in steps by identifying issues, solutions and their integrative analysis. Most class sessions will start will a short lecture that lasts no more than an hour, and will outline the various steps involved in the data science process. Following this students will be expected to take part in various class discussions and activities.

WEEK	TOPIC	ACTIVITY	LEARNING OBJECTIVES
1	Course introduction	Class • Formal introductions between Instructor and students, and between students. • Discuss the concepts surrounding data science and what data scientists do. For next class • Students will come up with initial project idea. **READING:** *The Roots of Data Science. How it all began*	• Students know the Instructor and other students. • Students understand class policies and procedures. • Student understand what is data science and what data scientists do.

WEEK	TOPIC	ACTIVITY	LEARNING OBJECTIVES
2	Coming up with a good project idea.	Class • Students discuss their project ideas. • Students and the Instructor review projects ideas. For next class • Students revisit project ideas based on class feedback. • Students will identify at least two relevant sources from the literature to support the need for doing such a project. • Student will identify sources and specific data for their project. • **READING:** (1) *Trends and future perspective challenges in big data* and (2) *The world's most valuable resource is no longer oil, but data*	• Students understand strategies for coming up with good projects.
3	Sources of data, access to data and ethical use of data.	Class • Discuss sources of data in general terms and ethical issues with its use. • Student present project ideas, including identifying motivation for and benefit of the project. For next class • **READING:** *Data wrangling*	• Students can identify sources of data for their project. • Student understand ethical issues surrounding the use of data.
4	Working with data/ Exploratory data analysis.	Class • Students present their project ideas and data. For next class • Student will clean and transform data for their project.	• Students understand techniques to clean and transform data. • Students can determine if data is suitable for their project.
5	Project planning.	Class • Student discuss issues with cleaning and transforming data and approaches to overcome them. For next class • Students will prepare a workflow diagram of their project plan.	• Students understand the project planning phase.

WEEK	TOPIC	ACTIVITY	LEARNING OBJECTIVES
		• **PROJECT DELIVERABLE I:** One page report • **READING:** sample annotated article *Sustainable development: Meaning, history, principles, pillars, and implications for human action: Literature review*.	
6	Reading and writing a scientific paper.	Class • Students present their **Project Deliverable I** and workflow plan. • **Project Deliverable I** submitted to Blackboard • For next class Students review and summarize 3 recent/seminal articles related to project. See Weekly Tasks for more details.	• Students understand techniques for reading and writing scientific articles. • Student will understand how to use a synthesis matrix to keep track of important information gleaned from scientific papers. • Students will understand how to use the CARS model to help write the Introduction section.
7	Statistics and modelling for data science I.	Class • Students present their articles' summaries. • Discussion on project challenges and ways to overcome them. For next class • Students will read 4 more peer review articles related to their research and add it to their synthesis matrix. • Students will develop the Methodology section for their report.	• Students understand the role of statistical modelling in data science projects.
8	**SPRING RECESS – NO CLASSES**		
9	Guest speaker from Industry/Academia. (TBA)	Class • Discussion with guest speaker • Discussion of articles • Discussion on rhetorical architecture technique. For next class	• Students understand qualities of a good scientific article. • Students will develop a new skill, or built on an existing skill pertinent to data science.

WEEK	TOPIC	ACTIVITY	LEARNING OBJECTIVES
		• Student will develop their Results section for their report. • Students will update their project report and presentation. • Students will read 5 more peer review articles related to their research and add it to their synthesis matrix.	
10	Statistics and modelling for data science II	Class • Students present their updated project. For next class • Students will begin developing their paper/report outline into a full report/paper, including a title. • **READING**: (1) *Doing better data visualization*, (2) *The good, bad, and the biased*, and (3) *How to choose the right data visualization*.	• Students understand advanced data science techniques and when to use them.
11	Visualizing and communicating results	Class • Discuss articles on data visualization. For next class • Student will revise the Methodology and Results sections of their report. • Students will add a Discussion section to their report based on their results thus far.	• Students understand different approaches for communicating their results and when to use them.
12	Student – Instructor roundtable	Class • Student discuss project progress in detail For next class • Students will prepare their final class presentation. This should be no more than 10 to 15 minutes. • Student will develop their Introduction section for their final report.	
13	Final project presentations Cohort 1	Class • **Project deliverable 2:** Student present their project	

WEEK	TOPIC	ACTIVITY	LEARNING OBJECTIVES
		• Students review presented projects For next class • Students will prepare their final paper/report	
14	Final project presentations Cohort 2	Class • **Project deliverable 2:** Student present their project • Students review presented projects • **Project deliverable 3:** Final report is submitted to the Instructor	
15	Project feedback and class summary	Class • Discussion on project feedback with the Instructor • Class and project takeaways	• The Instructor understands ways that the course can be improved. • Students understand the value of the data science process in developing data science projects

7. Grades

The grade will reflect the student performance in the classroom and on the assignments. At the end of the term all the marks will be totaled as a weighted average according to the following weights:

- Class participation - 10%
    - This score is computed based on the class attendance, in-class participation, and project submissions/updates, with the exceptions of those below.
- Project deliverable 1 - 20%
    - Project proposal
- Project deliverable 2 - 30%
    - Project presentation
- Project deliverable 3 - 40%
    - Project report

Final grades at the end of the course will be assigned using a combination of absolute achievements and relative standing in the class.

The following undergraduate grading scheme will be used:

GRADE	PERCENTAGE SCORE	GRADE	PERCENTAGE SCORE
A+	> 97.00	**C+**	77.00 - 79.99
A	93.00 - 96.99	**C**	73.00 - 76.99
A-	90.00 - 92.99	**C-**	70.00 - 72.99
B+	87.00 - 89.99	**D**	60.00 - 69.99
B	83.00 - 86.99	**F**	< 60.00
B-	80.00 - 82.99		

8. Exams

This course does not include a midterm or a final exam.

9. Assignments

The course will include a number of intermediate assignments that build on knowledge acquired in class during the course of the semester and lead towards a final report and presentation. The time allocated for every assignment will be announced in class.

All assignments are mandatory. You will be working as an individual or in groups. If working within a group, each member is expected and required to contribute equally towards that group's workload. Assignment work is typically delivered through a presentation and discussion mode.

10. Late submission policy

The acceptance of assessments beyond the stipulated course deadline is left at the discretion of the Instructor. Any request for late submissions must be accompanied by a justifiable reason and email request for permission, which must be submitted to the Instructor within 24hrs prior to the assessment deadline.

11. Attendance

You are required to attend all class meetings. Your active participation in the class is essential to the success of this course and to students grade. Attendance may be verified during each session.

12. Course Website

The course has a Blackboard website. This website will provide you a single portal through which you may obtain lecture notes, retrieve assignment data and, review links to additional materials, and receive special announcements. Please notify ITU (and, if necessary, the Instructor) if you encounter any problems accessing this website.

13. Electronic Communication

All course related email correspondence, including submission of assignments, should be made through the course Blackboard website. Please DO NOT send emails to the instructors' @gmu.edu address.

14. Student Expectations

- Academic Integrity: Students must be responsible for their own work, and students and faculty must take on the responsibility of dealing explicitly with violations. The tenet must be a foundation of our university culture. [See http://academicintegrity.gmu.edu/distance].
- Honor Code: Students must adhere to the guidelines of the George Mason University Honor Code [See http://oai.gmu.edu/the-mason-honor-code/].
- GMU Email: Students are responsible for the content of university communications sent to their George Mason University email account and are required to activate their account and check it regularly. All communication from the university, college, school, and program will be sent to students solely through their Mason email account. [See https://masonlivelogin.gmu.edu].
- Patriot Pass: Once you sign up for your Patriot Pass, your passwords will be synchronized, and you will use your Patriot Pass username and password to log in to the following systems: Blackboard, University Libraries, GMU email, myMason, Patriot Web, Virtual Computing Lab, and WEMS. [See https://password.gmu.edu/index.jsp].
- University Policies: Students must follow the university policies. [See http://universitypolicy.gmu.edu].Responsible Use of Computing - Students must follow the university policy for Responsible Use of Computing. [See http://universitypolicy.gmu.edu/policies/responsible-use-of-computing].
- University Calendar: Details regarding the current Academic Calendar. [See http://registrar.gmu.edu/calendars/index.html].
- Students with Disabilities: Students with disabilities who seek accommodations in a course must be registered with the George Mason
- University Office of Disability Services (ODS) and inform their instructor, in writing, at the beginning of the semester [See http://ods.gmu.edu].
- Students are expected to follow courteous Internet etiquette at all times; see http://www.albion.com/netiquette/corerules.html for more information regarding these expectations.

15. Student Services

- University Libraries: University Libraries provides resources for distance students. [See http://library.gmu.edu/distance and http://infoguides.gmu.edu/distance_students].
- Writing Center: The George Mason University Writing Center staff provides a variety of resources and services (e.g., tutoring, workshops, writing guides, handbooks) intended to support students as they work to construct and share knowledge through writing. [See http://writingcenter.gmu.edu]. You can now sign up for an Online Writing Lab (OWL) session just like you sign up for a face-to-face session in the Writing Center, which means YOU set the date and time of the appointment! Learn more about the Online Writing Lab (OWL).
- Counseling and Psychological Services: The George Mason University Counseling and Psychological Services (CAPS) staff consists of professional counseling and clinical psychologists, social workers, and counselors who offer a wide range of services (e.g.,

individual and group counseling, workshops and outreach programs) to enhance students' personal experience and academic performance [See http://caps.gmu.edu].

- Family Educational Rights and Privacy Act (FERPA): The Family Educational Rights and Privacy Act of 1974 (FERPA), also known as the "Buckley Amendment," is a federal law that gives protection to student educational records and provides students with certain rights. [See http://registrar.gmu.edu/privacy].

16. Additional information

Disclaimer: Any typographical errors in this Course Syllabus are subject to change as deemed necessary by the Instructor. All corrections will be announced in class or through Blackboard announcements.

Class recording: Recording is permitted in the class. Only with the prior written consent of the Instructor or if the recording is part of an approved accommodation plan will an exception be allowed.