Course Syllabus for SIADS 699

Course Overview and Prerequisites

The Capstone is a project-based course in which students propose and build end-to-end data science projects in their domains of interest. Students are asked to demonstrate mastery of data science concepts and methods from their MADS training and produce a creative, original, and technically rigorous portfolio piece. Projects will be supervised by instructors with regular peer review.

The core of this class is a group data science project which will be independently designed and executed by students. The intention is to create a resume-worthy portfolio project in a format that is ready to share and present with the broader community of data science professionals.

The following are prerequisites for Capstone 697, 698 & 699:

- SIADS 694 & 695: Milestone II A and B
- 1 data science application course (currently, these are SIADS 682: Social Media Analytics, SIADS 680: Learning Analytics, SIADS 685: Search and Recommender Systems, and SIADS 688: Data Science for Social Good)

Instructor and Course Assistants

- Instructors: Winston Featherly-Bean, Elle O'Brien, Michelle LeBlanc, Ceren Budak, and Rachael Tatman
- Course manager: Kirtana Choragudi

Course Communication Expectations

- Course-wide chat via Slack and Zoom.
- Live Office Hours (see Live Events section in Coursera)

How to Get Help

If you have questions concerning the degree program, encounter a technical issue with Coursera, or issues using Slack, please submit a report to the ticketing system at <u>umsimadshelp@umich.edu</u>.

If you have an issue specific to the Coursera environment, you can also begin a <u>live chat session</u> with Coursera Technical Support (24/7) or view <u>Coursera troubleshooting guide</u>s. (you may be asked to log in to your Coursera account).

For questions regarding course content, refer to the **Communications Expectations** section.

Weekly Readings or Textbook Information

- Week 1: Syllabus and course resources
- Week 2 suggested reading: None

- Week 3 suggested reading: None
- Week 4 suggested reading: *How to Avoid Machine Learning Pitfalls,* by Michael A. Lones <u>https://arxiv.org/pdf/2108.02497.pdf</u>
- Week 5 suggested reading: Get rid of AI saviorism, by Shreya Shankar https://www.shreya-shankar.com/ai-saviorism/
- Week 6 suggested viewing: Data engineering principles: build frameworks, not pipelines. By Gatis Seja <u>https://www.youtube.com/watch?v=pzfgbSfzhXg</u>
- Week 7 suggested reading: *How to review code like a human,* by Mike Lynch <u>https://mtlynch.io/human-code-reviews-1/</u>
- Week 8 suggested reading: *The ghosts in the data*, by Vicki Boykis <u>http://veekaybee.github.io/2021/03/26/data-ghosts/</u>
- Week 9 suggested reading: Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing, by Deborah Raji et al. https://dl.acm.org/doi/pdf/10.1145/3351095.3372873
- Week 10 suggested reading: So your data science project isn't working, by Jacqueline Nolis <u>https://jnolis.com/blog/project_isnt_working/</u>
- Weeks 11 & 12: None

Learning Outcomes

The core of this class is a group data science project which will be independently designed and executed by students. The intention is to create a resume-worthy portfolio project in a format that is ready to share and present with the broader community of data science professionals. Student groups will be submitting the following at the end of the course:

- A blog post that tells the story of your project. Blogging is a popular way working data scientists share their ideas, network, and find jobs. While you don't have to actually publish your blog, you'll need to turn in something that's both interesting to a broad audience of data professionals and technically accurate. Note: if you're aiming to write a scientific paper, a draft of that paper can count. Just make sure it's readable to non-experts!
- A GitHub repository full of the code required to reproduce your analysis and figures. Your repository should be documented with a nice, informative README so a data scientist from outside MADS could understand what you did!

And one of the following high-level overviews of your work designed to be shared:

- A 3-5 minute video from your team about what you made.
- A poster appropriate for a data science conference, such as the MIDAS Data Science Symposium.

Student groups will participate in a midway check-in meeting with the teaching team and submit three informal video stand-ups throughout the course.

Course Schedule

- This course begins on February 1, 2023, and ends on April 25, 2023
- Group video standups are **due in Slack by 11:59 pm** (Ann Arbor, Michigan time Eastern Time) on the following days:.
 - Wednesday, March 1 by 11:59 pm (Ann Arbor, Michigan time Eastern Time
 - Wednesday, March 15 by 11:59 pm (Ann Arbor, Michigan time Eastern Time)
 - Wednesday, March 29 by 11:59 pm (Ann Arbor, Michigan time Eastern Time)
- Students are required to respond to 2 other teams about their standups by Saturday of the standup week.
- Teams must schedule a check-in meeting with their designated contact on the teaching team between March 1 and March 15.
- Projects must be submitted to the instructional team by **April 19**, **2023 at 11:59 pm** (Ann Arbor, Michigan time Eastern Time). *Note this is one week before the official last day of class!*

Office Hours via Zoom

After your project has been proposed, you will be assigned a team mentor from the faculty. One-on-one team meetings with your mentor are available by request at any time during the course, so don't hesitate to ask. Think of us as coaches or consultants who can help you when you're up against a wall or need another set of eyes.

Grading

Course Item	Number of Points	Percentage of Final Grade
Team matching survey	2	2%
Project proposal	4	4%
Course policy quiz	1	1%
Midway check in with teaching team	4	4%
Weekly mini-deliverables	3 pts each, 8 assignments	24%
Standups & responses	4 pts each, 3 assignments	12%

Team project	50	50%
First & Final Week Reflections	1.5 pts each, 2 assignments	3%
Total	100	100%

Note: All assignments are required to earn credit for this course.

Late Submission Policy

Several assignments cannot be turned in late: the team-matching survey, your three standups, and your midway-check-in. Credit cannot be awarded for these assignments if they are incomplete.

All other assignments have a 25% penalty for each day they are late. This penalty is automatically assigned by Coursera.

Letter Grades and Course Grades

Refer to the <u>MADS Assignment Submission and Grading Policies</u> section of the UMSI Student Handbook (access to Student Orientation course required).

The grading scale for this course is as follows- note that an A+ can only be awarded at the discretion of the instructors. An A+ is a mark of a truly distinguished final project, and they are rare.

А	90%
B+	87%
В	80%
C+	77%
С	70%
D+	67%
D	60%
F	0%

Academic Integrity/Code of Conduct

Refer to the <u>Academic and Professional Integrity</u> section of the UMSI Student Handbook. (access to Student Orientation course required).

Accommodations

Refer to the <u>Accommodations for Students with Disabilities</u> section of the UMSI Student Handbook (access to the Student Orientation course required). Use the <u>Student Intake Form</u> to begin the process of working with the University's Office of Services for Students with Disabilities.

Accessibility

Refer to the <u>Screen reader configuration for Jupyter Notebook Content</u> document to learn accessibility tips for Jupyter Notebooks.

Library Access

Refer to the <u>U-M Library's information sheet</u> on accessing library resources from off-campus. For more information regarding library support services, please refer to the <u>U-M Library Resources</u> section of the UMSI Student Handbook (access to the Student Orientation course required).

Student Mental Health

Refer to the University's <u>Resources for Stress and Mental Health website</u> for a listing of resources for students.

Student Services

Refer to the <u>Introduction to UMSI Student Life</u> section of the UMSI Student Handbook (access to the Student Orientation course required).

Technology Tips

- Recommended Technology
 - This program requires Jupyter Notebook for completion of problem sets and Adobe or other PDF viewer for reading articles.
- Working Offline
 - While the Coursera platform has an integrated Jupyter Notebook system, you can work offline on your own computer by installing Python 3.5+ and the Jupyter software packages, including pyspark. For more details, consult the Jupyter Notebook FAQ.