

CAPSTONE: Senior Research Project

School of Professional Studies

MS in Data Science

Course #: DATA 698

Spring 2019

INSTRUCTOR INFORMATION

Professor: Nasrin Khansari

E-mail: Nasrin.Khansari@sps.cuny.edu

Office hours are conducted by appointment only via GoToMeeting, Skype or phone. You are encouraged to ask questions on the “Ask Your Instructor” forum on the course discussion board where other students will be able to benefit from your inquiries. For the most part, you can expect your Instructor to respond to questions by email within 24 to 48 hours.

COURSE DESCRIPTION

The MS in Data Science program culminates with a Capstone project that represents highly sophisticated, but practical, solutions to address real world problems. As an applied program, the Capstone project focuses on applications of data science to industry. From an academic perspective, the goal of the project is to synthesize the knowledge acquired in the program and to apply it to some novel problem. There are up to five students allowed per section.

COURSE LEARNING OUTCOMES

At the end of this course, students will be able to:

- Identify a research topic and develop a research plan.
- Conduct a literature review and perform research
- Collect and analyze data
- Write-up and disseminate research results

PROGRAM LEARNING OUTCOMES (addressed by this course)

Learning outcomes of the Capstone course align with the program learning outcomes, as follows:

- Business Understanding. Build data-analytic solutions from an understanding of business goals

Students will apply what they've learned in terms of methods to collect and analyze data and apply critical thinking to draw conclusions to solve/address practical, real-world issues/topics/problems/business goals

- Data Programming. Use industry standard data science and analytics packages

Students will apply what they've learned using the libraries of the R analytics package as well as MySQL database programming package

- Foundational Math and Statistics. Emphasis on probability, statistics, and computational methods

Students will apply what they've learned in their probability, statistics, and computational math courses.

- Data Understanding. Collect, describe, model, explore and verify data

Students will apply what they have learned in terms of internal and external validity of data and data-sampling techniques

- Modeling. Select modeling techniques, generate test designs, build and assess models.

Students will apply what they have learned in terms of statistical data models and algorithmic/machine learning models to design, build, implement, and measure their predictive or descriptive power.

- Model Implementation and Deployment. Implement models for the various descriptive, predictive and prescriptive modeling techniques covered in the MSDA program

Students will apply what they have learned in terms of model implementation.

- Dissemination. Students will deliver presentations and report of their project results.

Based on what they have learned in the program, students will be expected to express articulate their research findings clearly, concisely and succinctly, and explain complex information in plain, direct, nonprofessional terms/language.

COURSE TOPIC

Topics can be wide-ranging, as the goal of the project is for students to synthesize the knowledge acquired in the program and to apply it to some novel problem.

Unless stipulated by the Instructor, students are expected to choose a topic by developing and submitting a short (1-2 page) proposal of their project (see “Proposal Approval Process” below), which must be approved by the Instructor.

Proposal should include a description of the problem, why it's interesting, and what other approaches have been tried. Then there needs to be a discussion on what the student's hypothesis is and how the student's specific solution will improve or solve the problem.

The Instructor reviews and provides feedback on the proposal, ultimately accepting it, rejecting it, or conditionally accepting it based on a change to the plan/proposal. In particular, the Instructor will ensure that the project is properly scoped for the expected time commitment, and that the problem is.

Group vs Individual Projects

As are most employees in workplaces, students are expected and encouraged to work in teams. However, a student may be allowed to work on an individual project, if the Instructor so deems.

Also given the diversity of time zones, backgrounds, professional interest of students, a group project may not be practical, and an Instructor may decide for each student to produce his/her own paper/project. An Instructor may also make this determination if the group dynamics of the class seem problematic.

Size of Groups

At most, two projects should be chosen among the group. Alternatively, everyone can work on the same project. The first week or two should focus on idea

generation and deciding on research topics. The paper will be jointly written by the members of the project, one per topic.

Proposal Approval Process

1. Students write a short 1-2 page proposal, including a description of the problem, why it is interesting, and what other approaches have been tried. Then there needs to be a discussion on what the student's hypothesis is and how the student's specific solution will improve or solve the problem.
2. The Instructor reviews and provides feedback on the proposal, ultimately accepting it, rejecting it, or conditionally accepting it based on a change to the plan/proposal. In particular, the Instructor will ensure that the project is properly scoped for the expected time commitment, and that the problem is sufficiently challenging that it serves as a good demonstration of what the student has mastered in the program. Finally, the approach and solution to the problem must also be novel and pushing the ability of the student.
3. Students conduct research, collect data, build models, produce output artifacts (i.e., data visualizations, shiny app, etc.). During this time students should attend checkpoint meetings with the Instructor weekly or every two weeks – see **KEY DATES**.
4. Students produce a medium length (10-12 pages) paper describing the project, along with a runnable demo of any code. Code listings, interpreter output, and exploratory data visualization should not be included in the paper itself. This is an expanded version of the initial proposal that includes more discussion of prior approaches, results of the student's approach, and a final assessment of whether the project effectively solved the goal stated in the initial proposal. This needs to be submitted 3 weeks before the course ends.
5. The Instructor reviews the proposal based on four factors:
 - a. Problem novelty and challenge of problem. At this point the novelty should have been established;
 - b. Depth of the analytics component of project - The project should make use of quantitative methods covered in the course, following best practices for the given method;

- c. Depth of the coding component of the project - The project should make use of programming methods covered in the course, following best practices for the given method;
 - d. Synthesis and overall ability to demonstrate that the student's solution satisfied the goal of the paper. The advisor will provide feedback within a week so that the student has time to make corrections as needed.
6. Student submits final draft of paper and makes working code publicly available (as necessary)

READING/SOURCES

[Detailed here are reading materials that may help students developing their idea and refine their Senior Research Project. This is not an exhaustive lists and students are encouraged to seek additional materials support their project.]

SUGGESTED STRUCTURE

The following is a suggested outline for your Senior Research Project. Please provide amended outline to your instructor for approval.

- Cover Page and Table of Contents
- Abstract/Executive Summary
- Literature Review/Research Conducted
- Theory and Hypotheses (if applicable)
- Data and Methods
- Results
- Discussion
- Conclusion

ASSIGNMENTS AND GRADING SCALE

Assignments and grade weightings

- | | |
|----------------------------------|-----|
| • Proposal | 10% |
| • Discussion Board Participation | 15% |
| • Mid Term Checkpoint | 15% |
| • Group Contribution | 30% |
| • Final Project & Presentation | 30% |

Grading Scale:

90-100	A
80-89	B
70-79	C
60-69	D
Below 60	F

When grading students, Instructors will use the “Capstone Student Performance Eval Spring 2018” template.

KEY DATES (Subject to change)

Week Of	Description
30-Jan-18	Introductions
6-Feb-18	Proposal development
13-Feb-18	Proposal development
20-Feb-18	Final review/approval of proposals
27-Feb-18	Weekly check-in on research strategy, data selection, methodology, etc.
6-Mar-18	Weekly check-in on research strategy, data selection, methodology, etc.
13-Mar-18	Weekly check-in on research strategy, data selection, methodology, etc.
20-Mar-18	Weekly check-in on research strategy, data selection, methodology, etc.
27-Mar-18	Mid-term submission of draft project
3-Apr-18	Weekly discussion/progress report
10-Apr-18	Weekly discussion/progress report
17-Apr-18	Weekly discussion/progress report
24-Apr-18	Spring Break (19 Apr– 28 Apr)
1-May-18	Weekly discussion/progress report
8-May-18	Submission and presentation of draft project
15-May-18	Submission and presentation of final project

SUBMISSION REQUIREMENTS

1. Final project papers are required to be medium length 10-12 pages (20-24 pages if printed single-sided) and double spaced, along with a runnable demo of any code. Code listings, interpreter output, and exploratory data visualization should not be included in the paper itself. In keeping the amount

of required pages relatively low, we expect the writing to be concise, succinct, and to-the-point. Key points should be expressed in clear, simple and easy-to-understand language. The final project paper, along with supporting commentary -- discussion of prior approaches, results of the student's approach, and a final assessment of whether the project effectively solved the goal – must be submitted 3 weeks before the course ends

2. Given that most Capstones are group projects, it is preferred that final papers are prepared in LaTeX, using the collaborative features of online LaTeX editors such as overleaf.com or sharelatex.com. Alternately, papers may be prepared in Microsoft Word format (Microsoft Word 2010) and delivered in Adobe Acrobat PDF format (Adobe Acrobat XI Pro). Written papers shall be prepared on standard 8.5 by 11 inch pages.
3. Student name and course code must be included in the heading section for all pages
4. The proposal pages shall be bound along the left side and numbered on the bottom, right side.
5. Each page shall have a one-inch margin at the top and the bottom and on each side.
6. Notations of proprietary material and any other identifying information printed on each page may be included in the margin.
7. Text shall be in Times New Roman, 12 point font. However, text included in figures or graphics in the written proposal may be reduced to 9 point font, times new roman. Table text must maintain 12 point font.
8. Should the proposal require foldout pages for graphics, one foldout page shall not exceed 18 by 11 inches. Foldouts will be counted as two (2) pages, regardless of size, and subsequent page numbering must be adjusted appropriately.
9. During the first week of the semester, you will receive an email at your SPS email address, providing credentials to the ePortfolio platform in Digication. You will be required to upload your final paper to an ePortfolio (created by you). All questions related to the use of the platform should be directed to eportfolios@sps.cuny.edu.

NOTE: Students who do not post their final project to eportfolios@sps.cuny.edu will not receive a final grade and, thus, not receive credit for completing the course.

Students may also create a website site or use of another multimedia platform if they wish to highlight specific case studies and examples (YouTube, Facebook, etc.).

ACCESSIBILITY AND ACCOMMODATIONS

The CUNY School of Professional Studies is firmly committed to making higher education accessible to students with disabilities by removing architectural barriers and providing programs and support services necessary for them to benefit from the instruction and resources of the University. Early planning is essential for many of the resources and accommodations provided. Please see: http://sps.cuny.edu/student_services/disabilityservices.html

ONLINE ETIQUETTE AND ANTI-HARASSMENT POLICY

The University strictly prohibits the use of University online resources or facilities, including Blackboard, for the purpose of harassment of any individual or for the posting of any material that is scandalous, libelous, offensive or otherwise against the University's policies. Please see:

http://media.sps.cuny.edu/filestore/8/4/9_d018dae29d76f89/849_3c7d075b32c268e.pdf

ACADEMIC INTEGRITY

Academic dishonesty is unacceptable and will not be tolerated. Cheating, forgery, plagiarism and collusion in dishonest acts undermine the educational mission of the City University of New York and the students' personal and intellectual growth. Please see:

http://media.sps.cuny.edu/filestore/8/3/9_dea303d5822ab91/839_1853cee9c9d90e9.pdf

STUDENT SUPPORT SERVICES

If you need any additional help, please visit Student Support Services:

http://sps.cuny.edu/student_resources/