Data Science Projects [Stat 435 / 535] (or the D2K Learning Lab)

In this project-based course, student teams will complete semester-long, real-world data science research or analysis projects sponsored by Rice or Texas Medical Center researchers, community partners, or companies. Students will also learn best practices in data science.

More Information

- Lecture: Wednesdays 4 - 5:30pm.
- In this course, students will work on diverse teams (undergrad + grad AND from at least two disciplines) of 4-6 students to solve a real data science challenge. Each team will have a sponsor mentor who knows the discipline well and a data science mentor for technical help. Students will be assigned to a team based on their project preferences as well as the expertise required for each project. Teams will work on the data science project throughout the semester, give an oral and poster presentation of their findings in the D2K Learning Lab Showcase, and prepare a report along with all code used to produce results (adhering to reproducible research standards) that will be sent to the project sponsor.
- This is a lecture / lab course. In the lecture component, students will learn best practices in large-scale data science projects including: designing a data science analysis pipeline, team-based project management & software development, and computational and scientific reproducible research for data science.
- Spring 2019 will include several projects sponsored by industry (tech, finance, energy), by government and non-profits, and by researchers. These projects represent a wide array of disciplines from social sciences and policy to science and engineering.
- There is limited enrollment for this course and students must apply (see instructions below). We anticipate enrolling 40 - 50 students in the course with about ½ undergraduates and ½ graduate students.
● Students from ANY background or major can apply (not just Stat students). We are especially looking for students with expertise in machine learning, signal processing, and GIS / spatial statistics.

● This course is appropriate for undergraduate juniors or seniors AND professional masters or beginning PhD students who have a strong technical background in data science. Students who are discipline experts and would like to join a team may be able to do so by enrolling in another course through their home department.

● Note that any STAT or COMP senior taking this course as part of a capstone requirement will automatically be granted a spot in the course, but you still must submit an application.

Apply by completing this form!

● Applications due: Friday, November 2, 2019. Decisions will be communicated to applicants by Wednesday, November 7, 2019.

Data Science Consulting [Stat 415 / 515] (affiliated with the D2K Consulting Clinic)

Students in this course will advise clients from Rice and beyond in a data science consulting clinic, while learning best practices in consulting and gaining exposure to a wide variety of real data science problems.

More Information

● Clinic: Mondays 1-4pm. Lecture: Wednesdays 3 - 4pm.

● Students in this course will learn best practices in data science consulting through real experiences by serving as consultants in the D2K Consulting Clinic. This is a free, drop-by data science consulting clinic open to the community (Rice as well as beyond the hedges). Students will work on short-term teams to help real clients with their data science problems in the clinic and then then follow-up with a written report. Most consultations involve framing the data science problem (e.g. setting up the problem, providing background research, and recommending technical approaches) and will not often involve direct data analysis.

● This is a lab (clinic) / lecture course. During the lecture component, we will review cases brought in to the clinic and learn more about common data science problems and techniques that arise in the clinic.
• This course is a great way to gain exposure to a wide variety of data science problems, learn about data science consulting through real experience, and (especially for PhD students) find opportunities for fruitful collaborative research.
• This course is appropriate for PhD students AND advanced undergraduate seniors or masters students who have technical expertise in one or more areas of data science.
• Enrollment is limited to 12 students (6 grad + 6 undergrad), and students must apply for this course (see instructions below).
• Students from ANY major (not just Stat) can apply. We especially welcome advanced students from COMP or ELEC who have expertise in signal processing, machine learning, and databases.

Apply by completing this form!

• Applications due: Friday, November 2, 2019. Decisions will be communicated to applicants by Wednesday, November 7, 2019.