The D2K Learning Lab

Applied Machine Learning & Data Science Projects
(DSCI 435 / 535 & COMP 449 / 549)

Data Science & Machine Learning Capstone Laboratory (DSCI 400)

Spring 2020
Table of Contents

• D2K Lab Overview
• What is the D2K Learning Lab? (Overview of Courses)
• Who should take these courses?
• Why take these courses?
• How to register & apply for these courses?
• Other ways to get involved
Training exceptional students.

Making an **impact** through data science.
**D2K Lab Overview**

**Immersive Experience for Students**

**D2K Courses**
- D2K Learning Lab: Team-based data science projects
- D2K Consulting Clinic: Data science consulting

**D2K Programs & Events**
Focused on engagement, enrichment & excellence

**Solutions for D2K Partners and Affiliates**
- Rice Researchers
- Texas Medical Center
- Government & Non-profits
- Industry
Students work on integrated teams solving real-world data science problems.
D2K Learning Lab

Neuron Tuning in the Visual Cortex

Optimizing Fleet Allocation for Houston Fire Department

Bill.com Fraud Detection System

Oil & Gas Well Analysis & Natural Language Processing of Well Reports

Check out other past projects: d2k.rice.edu/d2k-learning-lab-projects
What are these courses?

Basic Course Info:

Time:
Monday & Wednesday 5pm – 6:30pm

Instructors:
Genevera Allen & Christopher Tunnell (DSCI 400)

Coursework:
• Final report, presentations (oral & poster) and software.
• Interim assignments to provide feedback & mentoring.

Course Description:
In this project-based course, student teams will complete semester-long data science research or analysis projects selected from a variety of disciplines and industries. Students will also learn best practices in data science.

DSCI 400 vs. DSCI 435/535 & COMP 449/549:
• DSCI 400 – non-client facing projects / student-directed projects
• DSCI 435/535 & COMP 449/549 – client facing projects – teams work with project sponsors
What will I learn in these classes?

- **Data Science Pipeline** & Pipeline Design
  - Learn how to creatively apply various data science and machine learning techniques to design and implement a Data Science Pipeline
- **Reproducibility** & Reliability in Data Science
- Technology for large-scale, collaborative data science
- Teamwork, Leadership & Project management
- How to **Communicate** data science work to both a technical and general audience
Course Schedule:

Phase I:
Lectures & Get started on projects.
• DSCI 400 & DSCI 435/535 & COMP 449/549 together
• Meets both Mondays & Wednesdays

D2K Showcase!
• End of semester showcase – oral & poster presentations
• Awards presented

Phase II:
Mentoring Phase
• DSCI 400 meets on Mondays
• DSCI 435/535 & COMP 449/549 meets on Wednesdays

End of Semester
• Final reports & software due
What is DCSI 400?

DSCI 400

- Student teams do not work with client sponsors.
- Teams get to choose from available data sets (or choose their own) and propose their own project objectives. These require instructor approval.
- 3 credit hours
  - Outside of class, students expected to meet once a week as a team
  - Work expectation outside of class: 6 hours per week
What is DSCI 435/535 & COMP 449/549?

DSCI 435/535 & COMP 449/549

- Student teams work with client sponsors.
- 3 types of projects: Industry-sponsored, public sector, or research projects.
- Students can review all available projects and apply to projects that interest them.
- 4 credit hours
  - Outside of class, teams must meet once a week with mentors and sponsors
  - Outside of class, teams must meet once a week with just students
  - Work expectation outside of class: 10 hours per week
Who can take these courses?

Who should take DSCI 400?

- DS minor students (who are not STAT, COMP, ELEC, or CAAM majors)
  - Pre-Reqs: DSCI 301, 302 & 303
  - DSCI 400 or DSCI 435 required for DS minor
- ELEC junior design students (DS track)
  - Pre-Reqs: ELEC 301 & ELEC 303
- Juniors interested in gaining experience before taking DSCI 435 / COMP 449 their senior year.
- Students who weren’t assigned to a team in DSCI 435
Who can take these courses?

Who should take DSCI 435/535 & COMP 449/549?

- STAT seniors (capstone requirement)
- COMP, ELEC, CAAM, or Math-Econ seniors (counts as elective / DS minor requirement)
- Professional Masters students in COMP, STAT, ELEC, or MCS&E
- PhD students who want to do team-based applied data science / machine learning research
- **Recommended Pre-req:** Machine Learning
Why take these courses?

These classes are great for:

- Gaining **real-world experience** in data science & machine learning
- Completing projects with a **real-world impact** that you can highlight on a resume & talk about in a job interview
- Open-ended & creative data science exploration & research
- Introductions to companies that want to hire Rice data science students
How do I apply?

Courses require an application & instructor permission

• **NOW: Register for DSCI 002 / COMP 002**
  • These are placeholders until applications are completed and instructor permission is granted to register for DSCI 400 or DSCI 435/535 & COMP 449/549.

• Applications due the beginning of spring semester.
  • We expect to take 50-75 students in DSCI 435/535 & COMP 449/549.
  • We expect to accommodate most other interested undergrads in DSCI 400.
  • Any student taking this course as part of a degree requirement will be accepted.
How do I apply?

Courses require an application & instructor permission

This Fall
Interested in DSCI 400?
Register for DSCI 002 3 credit hours

Interested in DSCI 435 / 535 / COMP 449 / 549?
Register for DSCI 002 or COMP 002 4 credit hours

January 6th
Project descriptions & Online course applications sent to all registered students

January 13th @ 5pm (class time)
Welcome pizza party! + Questions on courses answered + Intro to available projects

January 14th @ 12pm
Course Applications Due!

January 15th
Students notified of section placement & assigned teams + Instructor permission given to enroll
@5pm - Class & Meet your teams (+ sponsors)!

Before January 24th
Students must drop DSCI 002 or COMP 002 and enroll in appropriate course
How do I apply?

Courses require an application & instructor permission
How do I apply?

What is required for the application?

• Basic Info (name, department, degree program, etc.)
• Background & Experience in Data Science and Machine Learning
  • Rice courses, internships, and other experience
• Project Preferences (sponsored projects) or Project Interests (DSCI 400)
  • You can indicate which course you’d prefer
• Resume Upload & Personal Statement (optional)
Data Science Consulting (DSCI 415 / 515):

- **What?** Students learn consulting skills and gain experience advising on real data science problems in the D2K Consulting Clinic.
- **Who?** PhD or advanced undergrad / masters students who are interested in consulting & collaborative data science and machine learning research.
- **How?** Registration by application & Instructor Permission. See d2k.rice.edu/students to apply.
The D2K Lab: Ways to Get Involved

Upcoming D2K Events:

• D2K Distinguished Speaker: Hadley Wickham - November 22\textsuperscript{nd} @ 3pm McMurtry
  • Lunch interview + Book signing for students at 12pm in DH 3092
• Fall D2K Showcase – December 4\textsuperscript{th} @ 5pm in McMurtry
• Rice Datathon – January 24\textsuperscript{th} – 25\textsuperscript{th}
• Data Science Research Showcase – March
Other ways to Get Involved:

• Rice DataSci Club (datasci.rice.edu)
  • Undergraduate DS club
  • DEEP (Data Education & Exploration Projects)
• Machine Learning Seminar Series (machinelearning.rice.edu/ml-lunches)
  • Grad student organized internal seminar series + Free Lunch!
• D2K Fellows Program
  • Fellowships for advanced PhD students & Postdocs
  • D2K Fellows help mentor student teams
  • Learn more & apply: d2k.rice.edu/students