DATA SCIENCE (PHD)

Department Website (http://cds.nyu.edu)

NYSED: 36834 **HEGIS:** 1702.00 **CIP.** 30.7001

Program Description

Advances in computational speed and data availability, and the development of novel data analysis methods, have birthed a new field: data science. This new field requires a new type of researcher and actor: the rigorously trained, cross-disciplinary, and ethically responsible data scientist. Launched in Fall 2017, the pioneering CDS PhD Data Science program seeks to produce such researchers who are fluent in the emerging field of data science, and to develop a native environment for their education and training. The CDS PhD Data Science program has rapidly received widespread recognition and is considered among the top and most selective data science doctoral programs in the world. It has been recognized by the NSF through an NRT training grant.

The CDS PhD program model rigorously trains data scientists of the future who (1) develop methodology and harness statistical tools to find answers to questions that transcend the boundaries of traditional academic disciplines; (2) clearly communicate to extract crisp questions from big, heterogeneous, uncertain data; (3) effectively translate fundamental research insights into data science practice in the sciences, medicine, industry, and government; and (4) are aware of the ethical implications of their work.

Our programmatic mission is to nurture this new generation of data scientists, by designing and building a data science environment where methodological innovations are developed and translated successfully to domain applications, both scientific and social. Our vision is that combining fundamental research on the principles of data science with translational projects involving domain experts creates a virtuous cycle: Advances in data science methodology transform the process of discovery in the sciences, and enable effective data-driven governance in the public sector. At the same time, the demands of real-world translational projects will catalyze the creation of new data science methodologies. An essential ingredient of such methodologies is that they embed ethics and responsibility by design.

These objectives will be achieved by a combination of an innovative core curriculum, a novel data assistantship mechanism that provides training of skills transfer through rotations and internships, and communication and entrepreneurship modules. Students will be exposed to a wider range of fields than in more standard PhD programs while working with our interdisciplinary faculty. In particular we are proud to offer a medical track for students eager to explore data science as applied to healthcare or to develop novel theoretical models stemming from medical questions.

In short, the CDS PhD Data Science program prepares students to become leaders in data science research and prepare them for outstanding careers in academia or industry.

The PhD. program also offers students the opportunity to pursue their study and research with Data Science faculty based at NYU Shanghai. With this opportunity, students generally complete their coursework in New York City before moving full-time to Shanghai for their research. For more information, please visit the NYU Shanghai PhD page (https://shanghai.nyu.edu/academics/graduate/data-science-phd-program/).

Admissions

All applicants to the Graduate School of Arts and Science (GSAS) are required to submit the general application requirements (https://gsas.nyu.edu/nyu-as/gsas/admissions/arc.html), which include:

- Academic Transcripts (https://gsas.nyu.edu/nyu-as/gsas/admissions/arc/academic-transcripts.html)
- Test Scores (https://gsas.nyu.edu/nyu-as/gsas/admissions/arc/test-scores.html) (if required)
- Applicant Statements (https://gsas.nyu.edu/nyu-as/gsas/admissions/arc/statements.html)
- Résumé or Curriculum Vitae
- Letters of Recommendation (https://gsas.nyu.edu/nyu-as/gsas/admissions/arc/letters-of-recommendation.html), and
- A non-refundable application fee (https://gsas.nyu.edu/admissions/arc.html#fee).

See Data Science (https://gsas.nyu.edu/admissions/arc/programs/data-science.html) for admission requirements and instructions specific to this program.

Program Requirements

Course	Title	Credits
Major Requirements		
DS-GA 2003	Introduction to Data Science	3
DS-GA 1002	Probability and Statistics for Data Science	3
DS-GA 1003	Machine Learning	3
DS-GA 1004	Big Data	3
DS-GA 1005	Inference and Representation	3
DS-GA 2001	Research Rotation (taken six times, for a total of 18 credits)	18
Electives		
Other Elective Credits		39
Total Credits		**72**

Additional Program Requirements

PhD students will be reviewed annually. Students are also required to pass a Comprehensive Exam, the Depth Qualifying Exam (DQE), a teaching requirement (at least two courses at the Center for Data Science by the end of the fourth year of study), the Dissertation Proposal presentation, Defense, and the Dissertation. All Graduate School of Arts & Science doctoral candidates must also be approved for graduation by their department for the degree to be awarded.

Sample Plan of Study

Course	Title	Credits
1st Semester/Term		
DS-GA 2003	Introduction to Data Science	3
DS-GA 1002	Probability and Statistics for Data Science	3
DS-GA 2001	Research Rotation	1-3
	Credits	**9**
2nd Semester/Term		
DS-GA 1003	Machine Learning	3
DS-GA 1004	Big Data	3
DS-GA 2001	Research Rotation	1-3
	Credits	**9**

3rd Semester/Term		
DS-GA 1005	Inference and Representation	3
DS-GA 2001	Research Rotation	1-3
General Elective		3
	Credits	**9**
4th Semester/Term		
DS-GA 2001	Research Rotation	1-3
General Elective		3
General Elective		3
	Credits	**9**
5th Semester/Term		
DS-GA 2001	Research Rotation	1-3
General Elective		3
General Elective		3
	Credits	**9**
6th Semester/Term		
DS-GA 2001	Research Rotation	1-3
General Elective		3
General Elective		3
	Credits	**9**
7th Semester/Term		
General Elective		3
General Elective		3
General Elective		3
	Credits	**9**
8th Semester/Term		
General Elective		3
General Elective		3
General Elective		3
	Credits	**9**
	Total Credits	**72**

Following completion of the required coursework for the PhD, students are expected to maintain active status at New York University by enrolling in a research/writing course or a Maintain Matriculation (MAINT-GA 4747) course. All non-course requirements must be fulfilled prior to degree conferral, although the specific timing of completion may vary from student-to-student.

Learning Outcomes

Upon successful completion of the program, graduates will have:

1. Novel data-science methodology that transcends the boundaries of traditional academic disciplines.
2. Skills to clearly communicate their research results and their relevance to the broad data-science community.
3. The ability to effectively translate fundamental research insights into data science practice in the sciences, medicine, industry, and government.
4. An awareness of the social and ethical implications of data-driven methods, alongside tools to address and mitigate associated biases.
5. An awareness for the ethical implications of their work.

Policies
NYU Policies

University-wide policies can be found on the New York University Policy pages (https://bulletins.nyu.edu/nyu/policies/).

Graduate School of Arts and Science Policies

Academic Policies for the Graduate School of Arts and Science can be found on the Academic Policies page (https://bulletins.nyu.edu/graduate/arts-science/academic-policies/).