PHD PROGRAM IN DATA SCIENCE

Why a PhD in Data Science?

A report by International Data Corporation in 2015 observed the following potential for big data analytics and the need to analytics professionals:

- Shortage of skilled staff will persist. In the U.S. alone, there will be 181,000 deep analytics roles in 2018 and five times that many positions requiring related skills in data management and interpretation.

- Over the next five years, spending on cloud-based Big Data and analytics solutions will grow three times faster than spending for on-premise solutions.

- Adoption of technology to continuously analyze streams of events will accelerate as it is applied to Internet of Things (IoT) analytics.

Why a PhD in Data Science at BGSU?

This program was approved in Fall 2018 and the first formal cohort will start in Fall 2019. This is one of fewer than 25 data science related PhDs offered in the United States. Many of these are tracks, minors, or concentrations, while BGSU's degree is 100% data science.

Learning outcomes

- Demonstrate competency in the core concepts and techniques of data science, which come from both computer science and statistics.

- Use or develop appropriate techniques to analyze structured, unstructured, or dynamic datasets.

- Understand the principles that underlie analytical methods, articulate the strengths and limitations of analytical methods, and defend choices to use some methods over others.

- Communicate effectively to technical and non-technical audiences orally, in writing, and with effective visualization.

- Identify and respond to ethical concerns with the provenance and use of data.

- Develop new techniques for the analysis of complex datasets or real-time modeling and decision-making, or extend existing techniques to novel and challenging datasets.

- Organize data using tools appropriate to the problem, code new techniques in the appropriate computer language, optimize for performance and scalability, and distribute new tools to the data science community in a usable form.

FOR MORE INFORMATION

Contact Dr. Rob Green, Assistant Professor and Graduate Coordinator College of Arts and Sciences, at greenr@bgsu.edu or 419-372-8782

Program strengths and uniqueness

- One of the few PhD Data Science programs currently offered

- A truly multi-disciplinary program with expert faculty from three departments (Computer Science and Mathematics & Statistics from the College of Arts and Sciences and Applied Statistics & Operations Research from the College of Business) involved.

- Unique among technical degrees, courses in communication and ethics are required. Students will be aware of ethical issues related to the gathering and analysis of big data and will be able to communicate their findings to a wide variety of audiences.

Professional opportunities

The graduates of the PhD program will have developed the level of knowledge and skills necessary to take them on one of two different paths. First, some graduates will seek academic positions where they lead the development of data science and similar programs. The proliferation of master's degree programs in the US necessitates faculty with these higher-level credentials.

Second, both private industry and the public sector require high level data scientists as the leaders in applied innovation. The scale at which data science models are applied and their many hidden impacts on everyday life make ethical decision making and communications skills vital for practitioners in the field. While master's students are able to understand and utilize data science, doctoral students are able to develop and lead the field.

A search for PhD data scientist positions in January of 2018 resulted in 4,646 positions. On the same day, a search for PhD statistics positions only resulted in 3,857 positions.

Admission requirements

Prerequisite coursework includes differential, integral, and multivariate calculus, linear algebra, senior-level introduction to probability and statistics, programming skills in high level languages such as C, C++, Java, Python, and understanding of data structures and computer algorithms.

All applicants will apply directly to the PhD program whether or not they have earned a master's degree. Upon review by the admissions committee, admitted students who will not have earned a master's degree in data science, statistics, applied statistics, computer science, mathematics, or a closely related field, will be enrolled in the 90 credit pathway and will first complete the 30 credit MS in Data Science program before entering the 60 credit post-master PhD program. Others needing additional prerequisite coursework may be admitted to the 60 credit pathway with additional credit requirements.

Applicants must have a minimum GPA of 3.0 on a 4.0 scale (or equivalent). Applicants are required to submit scanned copies of official or unofficial transcripts from all institutions attended. Upon admission, final official or notarized copies of transcripts from all institutions where degrees were earned and diplomas from international institutions must be submitted. They are also required to submit official scores from the Graduate Record Examination (GRE) or the Graduate Management Admission Test (GMAT).

All applicants must submit three letters of recommendation from faculty or professionals in the field, a statement of purpose, and a current resume.

International applicants are also required to submit scores from the International English Language Testing System (IELTS), the Pearson Test of English Academic (PTEA), or the Test of English as a Foreign Language (TOEFL). Successful completion of ELS 112 will also be accepted for this requirement.

Cost of tuition

Please refer to www.bgsu.edu/offices/bursar for current information on tuition and fees.

Financial assistance

Scholarships and stipends are available on a competitive basis. For more information, please contact the coordinator.

Domestic students enrolled in four (4) or more credit hours are eligible to apply for financial aid using the Free Application for Federal Student Aid (FAFSA) to calculate student contribution and financial need. You may apply online at www.fafsa.ed.gov.

How to apply

Visit www.bgsu.edu/graduate/admissions.
Preferred application deadline, fall term: January 31

Curriculum

Curriculum for the Master of Science in Data Science (30 credit hours)

• CS 5200: Artificial Intelligence Methods	3 hours
• CS 5620: Database Management Systems	3 hours
• CS 6010: Data Science Programming	3 hours
• MATH 6410: Probability Theory I	3 hours
• MATH 6420: Mathematical Statistics II	3 hours
• STAT 5020: Regression Analysis	3 hours
• STAT 5160: Time Series Analysis	3 hours
• STAT 6440/CS 6440: Data Mining	3 hours
• OR 6610: Linear and Integer Programming	3 hours
• DATA 6910: Data Science Project	3 hours

Curriculum for the PhD in Data Science (60 credit hours)

Required Courses (23 credit hours):

Choose one of the following sequences in Computer Science (6 credit hours)

- CS 6260 Visualization (3) and CS 7200 Machine Learning (3)
- CS 6500 Big Data Analytics (3) and CS 7300 Unsupervised Feature Learning (3)

Choose one of the following sequences in Statistics (6 credit hours)

- MATH 7550 Statistical Learning I (3) and MATH 7560 Statistical Learning II (3)
- MATH 7570 Linear Stat Inference (3) and MATH 7590 Gen Linear Models and Ext (3)

Take the following courses (8 credit hours)

- DATA 7770 Data Science Exploration (1)
- DATA 7780 Data Science Communication (1)
- PHIL 6400 Ethical Issues in Data Science (3)
- STAT 7440 Advanced Data Mining (3)

Applied Data Science Experience (3 credit hours)

- DATA 7890 Internship/DATA 7930 Directed Reading (3)

Elective Courses (21 credit hours):

Choose 7 additional courses with at least 2 from CS, at least 1 from MATH, and at least 1 from OR/STAT. If not counted toward a required course sequence, CS 6260, 7200, 6500, 7300 and/or MATH 7550, 7560, 7570, 7590 may be used as electives. At most 10 credit hours of 5000-level courses may be counted toward the degree. Courses counted toward a master's degree at BGSU cannot be counted as electives.

Qualifying Examination:

Students will take qualifying examinations preferably after the first academic year of the PhD program. A student must pass qualifying examinations in both Statistics and Computer Science. A student must pass the Qualifying Examination to qualify to take the Preliminary Examination.

Preliminary Examination:

This is a research-oriented examination intended to help prepare students to begin their dissertation research. The Preliminary Examination includes a written and an oral component.

Dissertation (16 credit hours):

A doctoral candidate must enroll in at least 16 credit hours of dissertation work (DATA 7990), maximum of 30 credit hours.