DATA SCIENCE

Overview

Data Science is the scientific study of how to extract useful knowledge from large quantities of data. The job of data scientists is crucial in today's world of many, highly-connected devices that generate incredible amounts of data. By leveraging the knowledge this data provides, data scientists can help to make new scientific discoveries, increase efficiencies in industries, and solve many problems in our society. The work of data scientists involves gathering and aggregating data from various sources, transforming it into a useful form, building models that predict and explain, and communicate the results to decision makers. This work is typically done in conjunction with domain experts from a specific area. A good data scientist is someone who can easily adapt the analysis techniques to different domains by trying to understand the target domain. This means data scientists have to work in teams and spend much time communicating with experts and decision makers. Some typical data scientist tasks include:

- Gathering and aggregating data from various sources, such as databases, websites and embedded devices. Data comes in various forms, such as text, images or videos.
- Exploring data through visualization and statistical analysis to identify potential knowledge that can be gained and problems involved in extracting this knowledge.
- Building models that help to make predictions or explain the data.
- Using large-scale computing infrastructure to process Big Data.
- Developing software that can be used to deploy learned models in applications
- Creating visualizations, websites, and reports that can communicate the knowledge gained from data. The data science programs in the Department of Engineering, Computing, and Mathematical Sciences will prepare students for performing these tasks and pursuing data scientist careers. Due to the broad societal needs for analysis of data, students with degrees in this field can apply their skills and knowledge to many areas, such as business, medicine, cybersecurity, transportation, criminal justice, and many others.

Programs Offered

The Department of Engineering, Computing, and Mathematical Sciences offers undergraduates several opportunities for studying Data Science:

- The majority of students pursue a Bachelor of Science in Data Science (https://catalog.lewisu.edu/undergraduate/aviation-science-technology/data-science/data-science-bachelor/). This program provides students with a data science education that is supported by an extensive background in mathematics and computer science. Students with this degree will be well prepared to apply data science to a variety of fields and to pursue graduate coursework in the field.
- Students who prefer to pursue a slightly less rigorous program in Data Science may pursue a Bachelor of Arts in Data Science (https://catalog.lewisu.edu/undergraduate/aviation-science-technology/data-science/data-science-bachelor-arts/) instead. This is a good choice for students who would like to focus on a specific application area for data science. It has a lower credit hour requirement that can allow a student to take many courses outside of the major or even double-major in another field of study.
- Students interested in earning a Master of Science degree in Data Science may enter the Fast Track Program (https://catalog.lewisu.edu/undergraduate/aviation-science-technology/data-science/bachelors-data-science-masters-4-1-program/) earning a Bachelor of Science in Data Science and leading to a Master of Science degree in Data Science. Both B.S. and M.S. degrees can be completed in just five years.
- Students majoring in other disciplines who are interested in interpreting and visualizing large data sources may earn a minor in Data Science (https://catalog.lewisu.edu/undergraduate/aviation-science-technology/data-science/data-science-minor/). In addition to a very interesting set of curriculum options, students also have excellent opportunities to obtain valuable work experience while attending Lewis University. Lewis provides numerous work-study jobs for its undergraduates, in which students can gain hands-on experience maintaining computer systems and assisting other students. Local businesses frequently contact the department concerning both part-time and full-time employment opportunities. Many students from the Department of Engineering, Computing, and Mathematical Sciences work at the Argonne National Laboratory as part of a cooperative education program.

Recommendations for Other Students

As data scientists play important roles in virtually every other field, students who major in Data Science are encouraged to minor, double-major, or otherwise pursue coursework in other areas that interest them. For example, courses in Biology may help data scientists understand the problems in processing genomic data. Courses in political science may help to prepare the student to develop models for predicting elections or identifying classes of voters. Marketing courses may allow the student to help leverage transactional data for increasing marketing effectiveness for companies. A student who took courses in finance will be better prepared to build software that can predict the stock market. Criminal justice courses can aid in discovering predictive rules for identifying criminal hotspots in cities. These are just some among the many potential options for data science students that allow them to apply their knowledge to particular disciplines. A student's academic advisor will help in making the appropriate course decisions.

Transfer Students

Most transfer students who enroll at Lewis starting their junior year complete the major in two years. It is important, however, that students carefully evaluate which of their prior coursework will transfer by examining course descriptions instead of course titles, as the meaning of course titles do vary across different institutions.

Programs

Bachelor

- Data Science / Bachelor of Arts (https://catalog.lewisu.edu/undergraduate/aviation-science-technology/data-science/data-science-bachelor-arts/)
- Data Science / Bachelor of Science (https://catalog.lewisu.edu/undergraduate/aviation-science-technology/data-science/data-science-bachelor/)

Minor

- Data Science / Minor (https://catalog.lewisu.edu/undergraduate/aviation-science-technology/data-science/data-science-minor/)

Courses

DATA 20000 - Introduction to Data Science (3)
This course provides an introduction to the field of data science. It examines problems that can be solved by data scientists. The course presents the data science methodology of solving problems, including data gathering, preprocessing, modeling, evaluation, and visualization. Students will use standard data science tools and be introduced to programming concepts.

DATA 23500 - Programming for Data Analysis (3)
Disciplines and industries are collecting increasing amounts of data to help guide their work. This course presents programming techniques for working with large data sets. It teaches computer programming from the perspective of developing tools to analyze data.
Prerequisite: (DATA 20000 (may be taken concurrently) or CPSC 20000 (may be taken concurrently) or ECEN 10000 (may be taken concurrently)) and (MATH 21500 (may be taken concurrently) or MATH 22000 (may be taken concurrently) or MATH 31500 (may be taken concurrently) or PSYC 30300 (may be taken concurrently) or BSAN 34900 (may be taken concurrently))

DATA 30000 - Visualizing and Communicating Data Knowledge (3)
In this course, students will study effective communication of knowledge derived from data. The course also covers visualization of data for purposes of analysis and communication. Students will use standard software tools and programming libraries for visualization. The course will require writing technical reports that present the data science process and results. It also includes a discussion of ethical issues involved in data science.
Prerequisite: DATA 23500 (may be taken concurrently) or CPSC 21000 (may be taken concurrently)

DATA 36000 - Neural Networks and Deep Learning (3)
This course will survey leading algorithms and methods for neural networks and deep learning. The first part of the course will cover the basics of neural networks through the backpropagation algorithm. Then, the fundamentals of neural networks will be discussed through radial-basis function (RBF) networks and restricted Boltzmann machines. The second part of the course will cover deep learning methods through recurrent neural networks, convolutional neural networks (CNNs), and generative adversarial networks (GANs).
Prerequisite: CPSC 21000

DATA 40000 - Big Data Systems (3)
This course covers the study of systems for storing and processing large datasets. Covered concepts include standard architectures for Big Data, use of common software frameworks, and applications to batch and real-type systems. Students will work on projects using Big Data technologies such as Hadoop, MapReduce, Hive, Spark or NoSQL databases.
Prerequisite: CPSC 21000 (may be taken concurrently) and CPSC 33000 (may be taken concurrently)

DATA 47100 - Machine Learning (3)
This course studies programs that use experience for improving their performance at solving a variety of tasks such as classification, regression, or clustering. Topics include supervised and unsupervised learning, reinforcement learning, parametric and non-parametric methods, ensemble learning and introduction to computational learning theory. Students will learn how to evaluate the performance of machine learning methods and how to utilize the techniques in various applications.
Prerequisite: CPSC 21000 (may be taken concurrently) and (MATH 31000 (may be taken concurrently) or MATH 21000 (may be taken concurrently))

DATA 47200 - Introduction to Data Mining (3)
An introduction to the concepts, techniques, and systems of data warehousing and data mining, including (1) design and implementation of data warehouse and on-line analytical processing (OLAP) systems, and (2) data mining concepts, methods, systems, implementations, and applications.
Prerequisite: MATH 21000 (may be taken concurrently) and CPSC 21000 (may be taken concurrently)

DATA 49000 - Data Science Undergraduate Capstone Project (3)
In this course, students will work in teams to develop a data-driven solution for a real-world problem using data science methods, will document their work in a scholarly report, and present their methodology and results to faculty and peers. Students will identify appropriate project topics with help of the faculty, research appropriate current methods and technologies, then apply them to find a solution. The results will be presented in a form of a technical report and an oral presentation. Additionally, this course will cover topics in professional ethics, intellectual properties, privacy and professional communication.
Prerequisite: DATA 30000 (may be taken concurrently) and DATA 40000 (may be taken concurrently) and (DATA 47100 (may be taken concurrently) or DATA 47200 (may be taken concurrently))
Program Restrictions: Must be enrolled in the following Program: Data Science.
Class Restrictions: Must be in the following Class: Senior.
Attributes: Advanced Writing, Experiential Learning Gen Ed