

Data Analytics (DA)

Search DA Courses using FocusSearch (<http://catalog.northeastern.edu/class-search/?subject=DA>)

DA 5020. Collecting, Storing, and Retrieving Data. (4 Hours)

Studies how to build large-scale information repositories of different types of information objects so that they can be selected, retrieved, and transformed for analytics and discovery, including statistical analysis. Analyzes how traditional approaches to data storage can be applied alongside modern approaches that use nonrelational data structures. Through case studies, readings on background theory, and hands-on experimentation, offers students an opportunity to learn how to select, plan, and implement storage, search, and retrieval components of large-scale structured and unstructured information repositories. Emphasizes how to assess and recommend efficient and effective large-scale information storage and retrieval components that provide data scientists with properly structured, accurate, and reliable access to information needed for investigation.

DA 5030. Introduction to Data Mining/Machine Learning. (4 Hours)

Introduces the fundamental techniques for data mining, combining elements from CS 6140 and CS 6220. Discusses several basic learning algorithms, such as regression and decision trees, along with popular data types, implementation and execution, and analysis of results. Lays the data analytics program foundation of how learning models from data work, both algorithmically and practically. The coding can be done in R, Matlab or Python. Students must demonstrate ability to set up data for learning, training, testing, and evaluating.