Large amounts of heterogeneous data have become available in various healthcare organizations (payers, providers, and pharmaceuticals). Driven by mandatory requirements and the potential to improve the quality of healthcare delivery and reduce costs, these massive quantities of data (known as ‘big data’) hold the promise of supporting a wide range of medical and healthcare decisions, including clinical decision support, disease surveillance, and population health management, among others.

In this course, we will (1) introduce the process and methods of big data analytics, (2) examine the characteristics and related mining challenges on dealing with big healthcare data, and (3) review exemplar analytic methods for automated analysis of healthcare/medical data with different applications.

Topics to be included:
Fundamentals of Big Data Analysis (process, methods, challenge and applications); Healthcare Data Models; EHR Analytics; Patient Analytics; Text Analytics in Healthcare; and Predictive Analytics. Potential applications include, but do not limit to, outpatient appointments, prediction of inpatient length of stay, clinical decision supports, reducing readmissions, detecting errors, misuse and fraud, identifying high-risk patient, and infection control, etc.

Target Audience:
Junior or senior undergraduate students.

Evaluation Methods:
The course will include a class project on a topic proposed by a team of two students. Evaluation of the course will be based on mid-term report and final reports of the project, an in-class 20 minute presentation of a relevant topic of a student’s choice, homework, and in-class activities.