

# Biomedical Data Science: Mining and Modeling

MB&B 452, CB&B 752, CPSC 752, MB&B 752, MB&B 753,  
MB&B 754, MCDB 452, MCDB 752

Spring 2020  
MW 1:00-2:15 PM, Bass 305

Instructor-in-charge: Mark Gerstein

Guest Instructors: Corey O'Hern, Steven Kleinstein, Jesse Rinehart,  
Matthew Simon, Kei Cheung, Martin Renqiang  
Min, and Carl Zimmer

Rapid developments in bio- and information- technology and are changing the way that biomedical scientists interact with data. Data are increasingly generated much earlier in the scientific workflow and are much larger in scale. Also, before the data can be interpreted, extensive computational processing is often necessary. Thus, the data deluge in biomedicine now requires mining and modeling on a large scale - ie biomedical data science.

This course aims to equip students with some of the concepts and skills relevant to biomedical data science, with an emphasis on bioinformatics, a sub-discipline of this broader field, through examples of mining and modeling of genomic and proteomic data. Specific topics to be covered include sequence alignment, large-scale processing, next-generation sequencing data, comparative genomics, phylogenetics, biological database design, geometric analysis of protein structure, molecular-dynamics simulation, biological networks, mining of functional genomics data sets, and machine learning approaches for data integration.

For more info, email [cbb752@gersteinlab.org](mailto:cbb752@gersteinlab.org) or visit <http://cbb752b20.gersteinlab.org>