

MAJOR IN DATA SCIENCE, BIOINFORMATICS CONCENTRATION

Data Science is the discovery of knowledge and insight through the analysis of data. As such, it draws on the study of algorithms and their implementation from computer science, the power of abstraction and of geometric and topological formalism from mathematics, and the modeling and analysis of data from statistics. It has emerged as a separate field in response to the avalanche of data from web enabled sensors and instrumentation, mobile devices, web logs and transactions, and the availability of computing power for data storage and analysis. Modern data is challenging not only due to its large scale, but also because it is increasingly heterogeneous and unstructured. Information gleaned from this data none-the-less is revolutionizing diverse areas of human endeavor from health policy to high energy physics.

Learning Objectives

Upon successful completion of this program, students will be able to:

1. Select appropriate data analysis techniques for diverse biological contexts, develop and evaluate statistical models, perform analyses on biological data, and effectively communicate the results.
2. Effectively communicate analysis results and ideas through graphical, oral, and written methods. Develop the skills to collaborate and engage with professionals across diverse fields, translating complex biological concepts for varied audiences.
3. Work effectively in teams, collaborating toward the achievement of shared goals and contributing to collective success.
4. Develop algorithms and methodologies to integrate genetic, cellular, molecular, and biochemical data from large datasets, enabling them to formulate and test rigorous hypotheses about the cellular and biochemical processes that maintain normal physiological function and contribute to disease mechanisms.