

BIOF 509: Applied Machine Learning

Machine learning is a computational field that consists of techniques allowing computers to learn from data and make data-driven predictions or decisions. The ability to effectively implement machine learning approaches is a crucial component of data analysis. BIOF 509 provides a comprehensive overview of machine learning concepts, project design, and implementation. The course will give a conceptual overview of the most popular machine learning algorithms with examples of how/when to apply them to datasets. Algorithms that will be covered include: support vector machines, decision trees, random forests, multiple clustering approaches, and deep learning. Best practices in designing machine learning projects will also be emphasized, and this course will introduce strategies to avoid common pitfalls and to accurately interpret results. To reinforce key concepts, this course contains 4 written homework assignments and a research project. Through the homework assignments, students will (i) study theory behind common machine learning algorithms and (ii) explore examples of successful machine learning projects in biomedical research. For the research project, students will use python machine learning packages (Scikit-Learn, Tensorflow, Pytorch) to design a multistep pipeline to analyze a dataset of their choice. Students will also be expected to use Github to demonstrate proper documentation and version control practices when completing the project.

Learning Objectives

- Choose appropriate machine learning techniques for data analyses and interpret their results
- Design properly machine learning analysis pipelines and avoid common pitfalls
- Complete a short research project using machine learning

Credits: 2

Class Type: Graduate Course

Prerequisites:

Students should have previously completed BIOF 309 Introduction to Python or have equivalent experience. While the course will include a brief Python refresher, the emphasis of the course will be on applying machine learning.

Program: Bioinformatics and Data Science

Availability Available in Upcoming Term

Session Both Sessions