

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 implement machine learning approaches appropriately and intelligently is a crucial component of data analysis. BIOF 509 provides a broad practical introduction to machine learning concepts, analysis design, and implementation. The course will give a broad and conceptual overview of the most popular machine learning algorithms, followed by examples of how and when to apply them to real data. Best practices in designing machine learning analyses will be emphasized and reviewed, along with how to avoid common pitfalls and how to interpret analysis results. Through homework and in-class assignments, students will implement machine-learning models in Python, utilizing state-of-the-art machine learning Python packages, such as scikit-learn and tensor flow. Algorithms that will be covered include, but are not limited to linear and logistic regression, random forest, K-means clustering, and deep learning. Note that the course emphasizes hands-on application of algorithms, and mathematical derivation will not be reviewed. Further, depending on the students' familiarity with Python, completing the weekly homework assignments can take one to four hours. The course will culminate in a short research project utilizing machine learning to analyze either the student's own dataset or a public dataset that the student chooses. INDIVIDUAL LAPTOP IS NEEDED FOR EACH CLASS. Photo by Markus Spiske temporausch.com from Pexels

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 Current Term

Session Session B