BIOL/PHYS 5566: Machine Learning Methods In Natural Science Modeling

Spring 2023
Time: Monday 3:30pm
Location: Lynch 318 and DRL 4N9 on alternate weeks

Instructor

Prof. Bhuvnesh Jain and Prof. Junhyong Kim
Email: bjain@physics.upenn.edu, junhyong@sas.upenn.edu

Teaching Assistants

TBD

Course Description

This is a course for PhD students in natural sciences with interests in applying the latest machine learning and AI approaches to their problem domains. The course will consist of directed readings and tutorials with weekly discussions. The goal is to motivate mutual self-learning through guided discussions. Weekly participation and completion of readings or other assigned materials is essential. The following topics will be covered (after refinement based on student interest at the first meeting).

- Fundamentals of statistical learning. Quick introduction to 'bread and butter ML': kNN, logistic regression, decision trees, support vector machines and neural nets.
- Introduction and in-depth exploration of selected key topics in ML: reinforcement learning, generative adversarial networks, and transformers. This will occupy the bulk of the semester and students will be expected to actively participate by reading and presenting sections of papers. Tutorials and light coding will also be part of the exploration and will include applications to Biology, Physics and other natural sciences.
- Finally, we will discuss foundation models, built on the transformer architecture underlying chatGPT and related products, and how they are/will change research in the sciences.

Readings:

A longer reading list will be provided here as the course progresses. For the first two weeks, students are asked to read Artificial Intelligence, by Melanie Mitchell. Along with the introductory lectures, this book will bring everyone on the same page before we study more technical topics.