CS 184A/284A: Artificial Intelligence in Biology and Medicine Fall 2023

Instructor	Xiaohui Xie
Lectures	MW **8:00-9:30am**
Office Hours	TBD
Other Links	Ed Discussion

Course description

Introduction to artificial intelligence and machine learning with a focus on how to apply them to solve problems in biology and medicine. This class will familiarize you with a broad cross-section of models and algorithms from AI and machine learning. Applications will focus on problems from bioinformatics, genomics, medicine and healthcare.

Topical Outlines for Lectures (expect to vary in time):

- Introduction. Course Framework
- Nearest neighbor methods, Linear regression
- Perceptron, Logistic regression, Support vector machines, Decision trees
- Application 1: Gene expression analysis, Biomarker discovery, Precision medicine
- Unsupervised learning, Principal Component Analysis, Clustering
- Application 2: Single cell RNA-seq analysis, other genomic applications
- Probabilistic models, Markov models, EM algorithm
- Application 3: Gene discovery, Regulatory motif discovery, CpG islands
- Neural networks, Deep learning
- Application 4: Biomedical image analysis

Lecturess

- Course Introduction, Logisitcs

- Intro to Machine Learning, Data Exploration, Visualization
- Linear Regression, Cost Function, Gradient Descent, Cross Validation, Regularization
- Perceptron, Logistic Regression, Multi-Class, Cross-Entropy
- Neural Nets, Back-propagation
- Decision trees, Emsemble model
- VC-dim, Model Complexity
- SVM, Margin maximization, Lagrangian, Dual, Kernel
- Application: AI in digital pathology
- Clustering, K-means, Hierarchical clustering, Mixture of Gaussians
- PCA, Dimension reduction
- Visualization, TSNE
- Application: interpreting the human genome
- Jupyter notebooks for some example problems

Prerequisites

Intro to AI, Calculas, Linear Algebra, Python Programming

Grading Policy

- Course Project: 40%
  - Project proposal: 5%
  - Presentation: 5%
  - Project report: 30%
- Final: 35%
- HWs: 20%
- Participation: 5%
  - Class participation, Polls, Discussions on Ed Discussion, etc

Programming Assignments

Assignment 1

Description

PDF

Code/Data

HW1-code.zip

Assignment 2

Description

PDF

Code/Data

HW2-code.zip

Final Project

- The general scope of the project is to develop new or use existing machine learning appoaches to solve an interesting problem in biology or medicine.
- Project proposal
    - Write two-page proposal on the final project. In the proposal, specify:
        - Team members.
        - The problem you plan to work on.
        - Background and significance of the problem.
        - Proposed method
    - Please use the following template to write your project proposal. Submit the proposal in PDF format.
- Please use the following final project report template to write your final project report. Submit the reprot in PDF format.
- Some resources for project ideas
    - Kaggle competitions
    - Grant Challenge
    - Dream Challenge

Homework Policy

- Homework: You may discuss each assignment with others, but are required to code and write up each assignment independently.
- Late homework policy: If you get a note from the Student's Office (personal problems) or infirmary (medical problems) requesting a postponement, it will be honored. Otherwise, late homework will not be accepted.

Textbooks (not required)

- Machine learning by Kevin P. Murphy
- Deep Learning by Goodfellow, Bengio and Courville

Ed Discussion

We will be using Ed Discussion for class discussion. The system is highly catered to getting you help fast and efficiently from classmates and myself. Rather than emailing questions to me, I encourage you to post your questions on Ed Discussion.

Academic Honesty

- For assignments you are allowed to discuss the assignments verbally with other class members, but you are not allowed to look at or to copy anyone else's written solutions or code. All problem solutions and code submitted must be material you have personally written during this quarter, except for any standard library or utility functions.
- For class projects all reports submitted must be written by you or members of your project team. Code generated for class projects can be a combination of code written by team members and publicly-available code. You should clearly indicate in your reports and in your code documentation which parts of your code was written by you or your team and which parts of your code was written by others.
- Academic honesty is a requirement for passing this class. Any student who compromises the academic integrity of this course is subject to a failing grade. The work you submit must be your

own. Academic dishonesty includes, but is not limited to copying answers from another student, allowing another student to copy your answers, communicating exam answers to other students during an exam, attempting to use notes or other aids during an exam, or tampering with an exam after it has been corrected and then returning it for more credit. If you do so, you will be in violation of the UCI Policies on Academic Honesty (see link). It is your responsibility to read and understand these policies. Note that any instance of academic dishonesty will be reported to the Academic Integrity Administrative Office for disciplinary action and may be cause for a failing grade in the course.