Probabilistic Machine Learning

Course Subject/Number BST570-1 (4 credits)

Term and Year Spring 2025 (Jan 21st to May 2nd)

Day/Time and Location T/Th 11:00 AM -12:40 PM Saunders Research Building 1.410

Instructor Seong-Hwan Jun

Office Hours Tuesday 1-2 PM at DBCB Yakovlev (SRB 4th floor).

Contact Information

- Email: seonghwan_jun AT urmc.rochester.edu
- Office: Saunders Research Building 4168
- Website: `junseonghwan.github.io/teaching/`

Academic Integrity: You may discuss homework problems with others, but you must not retain written notes from your conversations with other students, or share data via computer files to be used in completing your homework. Your written work must be completed without reference to such notes, with the exception of class and recitation notes, which may be retained in written form.

Disability resources: The University of Rochester respects and welcomes students of all backgrounds and abilities. In the event you encounter any barrier(s) to full participation in this course due to the impact of a disability, please contact the Office of Disability Resources. The access coordinators in the Office of Disability Resources can meet with you to discuss the barriers you are experiencing and explain the eligibility process for establishing academic accommodations. You can reach the Office of Disability Resources at: `disability@rochester.edu`; (585) 276-5075; Taylor Hall; `www.rochester.edu/college/disability`.

Course Description This course provides an introduction to probabilistic modeling techniques and inference methods. Topics covered include graphical models, sampling-based and optimization-based inference methods. Strong emphasis is placed on implementation using Python, utilizing libraries such as PyTorch and PyMC. The students will gain hands-on experiences developing probabilistic models and inference methods for various applications.

Course Objectives

- Understand the foundational concepts of probabilistic reasoning.
- Learn to formulate and solve inference problems using probabilistic models.
- Gain practical experience with probabilistic machine learning tools and libraries.

Assessments

- Homework assignments: 4 x 15%.

- Literature review and presentation: 10%.
- Final project: 30%.

Grading Procedures

- A: 90-100%
- A-: 85-90%
- B+: 80-85%
- B: 75-79%
- B-: 70-74%
- C: 60-69%
- D: 50-59%
- E: $<$ 50%

Textbooks

- *Machine Learning: a Probabilistic Perspective* by Kevin P. Murphy or
- *Probabilistic Machine Learning: Advanced Topics* by Kevin Murphy. Draft PDF available online: `https://probml.github.io/pml-book/book2.html`.
- *Probabilistic Machine Learning: An Introduction* by Kevin P. Murphy. `https://probml.github.io/pml-book/book1.html`.

Other resources

- *Pattern Recognition and Machine Learning* by Christopher Bishop. PDF available online: `https://www.microsoft.com/en-us/research/people/cmbishop/prml-book/`.
- *An Introduction to Probabilistic Graphical Models* by Michael I. Jordan. `https://people.eecs.berkeley.edu/~jordan/prelims/`.

Course Outline

Module 1: Background

- Probability theory and random variables.
- Probabilistic reasoning and Bayesian statistics.
- Graphical models.
- Optimization and sampling.
- Intro to Python programming and Python libraries for machine learning.

Module 2: Exact Inference

- Directed and acyclic graphical models: Bayesian networks, trees, and hidden Markov models.
- Belief propagation and variable elimination.
- Factor graphs and message passing.

Module 3: Asymptotically Exact Inference

- Directed and undirected graphical models: Markov random fields, state space models, mixture models, and latent Dirichlet Allocation.
- Markov chain Monte Carlo methods (Metropolis-Hastings, Gibbs, and gradient-based methods).
- Sequential Monte Carlo methods.

Module 4: Approximate Inference

- Kullback-Leibler divergence and evidence lower bound optimization (ELBO).
- Variational Inference.
- Variational auto-encoders.
- Topics: e.g., normalizing flows, expectation propagation, approximate Bayesian computation, transformers.