COMP 4745/6745 Machine Learning

Fall 2024

Administrative Information

Instructor: Xiajun Jiang
Email: Xiajun.Jiang@memphis.edu
Time: Tuesday/Thursday 9:40 am - 11:05 am
Classroom: Dunn Hall 233
Office Hours: Scheduled by appointment
TA: Md Abdul Mazid Adan

Textbooks

There is no specific textbook since machine learning is a fast-growing area. A lot of reference books are mentioned below, most of which are available online. Reference chapters will be mentioned in the lecture slides.

References:

- Pattern Recognition and Machine Learning, Chris Bishop, Springer-Verlag, 2006 [1]
- Machine Learning: A Probabilistic Perspective, Kevin Murphy, MIT Press [2]
- Machine Learning, Tom Mitchell, McGraw Hill, 1997 [3]
- Deep Learning, Goodfellow et. al., MIT Press [4]

Prerequites

- Reasonable mathematics maturity such as multivariable calculus, linear algebra, probability, and statistics
- Foundation of algorithm analysis and data structure
- Programming skill in python/R/Matlab

Learning Objectives

- Understand basic principles and techniques of machine learning and its connection to other fields
- Build suitable machine learning models for any given problem
- Derive the algorithm (equations etc.) needed to learn and apply the model
- Implement the algorithm in reasonably efficient programming languages

[1] https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf
[2] https://probml.github.io/pml-book/book0.html
[3] http://www.cs.cmu.edu/~tom/files/MachineLearningTomMitchell.pdf
[4] https://www.deeplearningbook.org/

Topics

1. Foundations
   (a) Bayesian learning: Bayes rule, MAP and MSE estimation, marginal likelihood
   (b) Frequentist learning: maximum likelihood (ML) estimation, asymptotic
   (c) Model selection: validation and cross-validation, regularization and prior, bias-variation trade-off
2. Supervised learning
   (a) Classification problems: ROC curves, decision theory, loss functions
   (b) Classification model: naive Bayes, nearest neighbors, decision trees, neural network
   (c) Linear models: linear regression, logistic regression
   (d) Learning via optimization: gradient descent, Newton's method, boosting
   (e) Kernel methods: support vector machines (SVMs)
3. Unsupervised learning
   (a) Clustering: K-means, mixture models, expectation maximization (EM) algorithm
   (b) Dimensionality reduction: factor analysis, PCA
4. Reinforcement learning
5. Advanced topics in machine learning

Course Materials

- Will be hosted on Canvas, including lecture notes, project descriptions, data sets, etc.
- Assignments will be submitted to Canvas.

Evaluation

1. Midterm: 15% (TBD)
2. Final: 20% (TBD)
3. Problem Sets: 20%
4. Programming Tasks: 15%
5. Project: 25%
6. Class Participation: 5%
7. 6745 students only: Some extra work may be assigned in assignments/projects

Policies

1. No late homework will be accepted unless well-documented reasons are presented.
2. All homework must be individual work. Plagiarizing assignments or code sharing is not permitted. If plagiarism or cheating occurs, the student will receive a failing grade on the assignment and (at the instructor's discretion) a failing grade in the course. The course instructor may also decide to forward the incident to the office of student accountability office for further disciplinary action. For further information on the U of M code of student conduct and academic discipline procedures, please refer to: `https://www.memphis.edu/osa/`.

3. **Use of Generative AI in Course Assignments:** You cannot use generative AI tools such as ChatGPT in any manner to write your project reports or codes. When learning fundamental skills, you need to ensure that you master the basics. If the authorship is doubted, you may be asked to explain the code or re-create aspects of the code – you must show that you have mastered the fundamentals. You may use generative AI to polish your writing or debug your code, but beware that generative AI tools make errors and it is your responsibility to ensure the correctness of your work. You might consider generative AI tools to be an error-prone partner whose work you need to check.

4. Regular class attendance is mandatory. There is a strong correlation between regular attendance and obtaining a good grade. The instructor reserves the right to lower grades for lack of attendance. Students are responsible for any material and contents of missed lectures.

5. No early or late exams will be given unless under extreme situations.

6. Any grading errors in assignments should be noticed **within a week** to the TA/instructor.