CS 224/EE242A: FUNDAMENTALS OF MACHINE LEARNING

CLASS: Sproul Hall 1102, MWF 9.00-9.50AM

OFFICE HOURs: Monday 1000-1050am in-person MRB 3128
If you cannot make it, email me for a one-time meeting in-person or virtual.

TEXT BOOK: Probabilistic Machine Learning: An Introduction, K. Murphy
https://probml.github.io/pml-book/book1.html
This will be the book for the majority of the course. However, we will be occasionally referring to some other books and papers, which I will indicate as appropriate.

Course purpose: Provide the mathematical foundations for machine learning from both discriminative and generative models points-of-view.

Prerequisites: Probability and random variables (equivalent to STAT155/EE114), Linear Algebra (equivalent to MATH 31), Programming (CS100).

Instruction Mode: The course will be taught in-person.

Grading: The following will be the grading rubric used.
Programming Assignments: 45% (exact percentage for each will be in the assignment)
Quizzes: 20%
Final: 35%

Teaching Assistant: There will be a TA and a grader for the course. For questions on homework practice problems and programming assignments, it is best that you contact the TA first. For questions about the material covered in class, please contact me first.

Homeworks, Assignments, and Exams: Homework practice problems will be assigned every week and solutions provided. These will not be graded and you do not need to submit them. Please try to solve the problems without looking at the solutions.

There will be 2 quizzes during class on dates that will be announced at least a week ahead. The quiz questions will be similar to the homework practice problems.

Three programming assignments will need to be completed and submitted by the deadlines provided in the assignment. The mode of submission will be provided by the TA or grader.

The final will be comprehensive and cover material similar to the homework practice problems.

Late Work: Each student has 6 late days over the duration of the course to handle individual disruptions. I will not accept work later than allowed by the late days, except for medical and family emergencies that cannot be reasonably covered within the allocated 6 days. The last date

for submission of all assignments, including late days, will be Dec 09, 2024. This is necessary in order for us to submit grades by the campus deadline.

Student Code of Conduct: Problem sets and exams are to be completed on your own. You may discuss with others, but what you submit must be your work. This applies to the use of AI tools like ChatGPT and others. You may not submit answers (written or programming) to problem sets that contain material you did not produce yourself for the express purpose of this offering of this course. If I find that you have submitted work that is not your own or is work you submitted in a different course, I will assign you a zero on that assignment, and I will forward the case to Student Conduct & Academic Integrity Programs for campus-level consideration. Please read the policies on Student Conduct.

Copyright: As instructor, I retain all copyrights on all lecture videos, slides, problem sets, other course materials, and the syllabus, as provided and protected by U.S. copyright law and University policy.

Students in the course may take notes and make copies of the course materials for their own use. You may also share those materials with other students who are registered and enrolled in the course. You may not reproduce, distribute or display (post or upload) lecture notes or recordings or course materials in any other way – whether or not a fee is charged - without my express written consent. You may not allow others to do so. If you do, you may be subject to student conduct proceedings.

Similarly, you own the copyright on your original problem set solutions, code, and exams. If I wish to post such answers, I will ask for your written permission.

CLASS SCHEDULE

The following is to give you an idea of the topics covered and the pace of the course. The exact dates are likely to change somewhat, especially as we get later into the quarter.

(S27) Class 01 – Course Overview; Introduction to ML

(S30) Class 02 – Overview of ML; Basics of Probability

(O02) Class 03 – Discrete Random Variables; Sigmoid Function

(O04) Class 04 – Multiple Random Variables, Marginals

(O07) Class 05- Gaussian Random Variables

(O09) Class 06 – Estimation Basics; Maximum Likelihood Estimation

(O11) Class 07 – Bayesian Estimation; Maximum A Posteriori Estimation

(O14) Class 08 – Least Squares

(O16) Class 09 – Eigenvalue Decomposition, PCA, Singular Value Decomposition

(O18) Class 10 – Eigenvalue Decomposition, PCA, Singular Value Decomposition

PROGRAMMING ASSIGNMENT 1 – PCA/SVD

(O21) Class 11 – Clustering

(O23) Class 12 – Bias-Variance Tradeoff, Overfitting, Cross-validation, Evaluation Measures

(025) Class 13 – Quiz 1

(O28) Class 14 – Information Theory Basics

(O30) Class 15 – Optimization: Gradient Descent, Stochastic Gradient Descent

(N01) Class 16 – Linear Discriminant Analysis

PROGRAMMING ASSIGNMENT 2 - Linear and Logistic Regression

(N04) Class 17 – Naive Bayes

(N06) Class 18 – Logistic Regression

(N08) Class 19 – Linear Regression

(N13) Class 20 – Deep Neural Networks

PROFRAMMING ASSIGNMENT 3 – Deep Neural Networks

(N15) Class 21 – Deep Neural Networks

(N18) Class 22 – Attention Mechanisms; Large Multimodal Models

(N20) Class 23 – Decision Trees and Random Forests

(N22) Class 24 – Quiz 2

(N25) Class 25 – KNN, Kernel Density Estimation

(N27) Class 26 – Gaussian Mixture Models; Expectation Maximization

(D02) Class 27 – Generative Models

(D04) Class 28 – Generative Models

(D06) Class 29 – Summary of the course

FINAL: Finals week, exact date TBD