Principles of Machine Learning – Spring'24
CSCI-B455
Tentative Syllabus

Oğuzhan Külekci
okulekci@iu.edu

Welcome to the "Principles of Machine Learning" course in the Spring 2024 semester. Throughout this course, our focus will be on understanding the theoretical underpinnings, methodologies, and essential principles that govern machine learning algorithms. By delving deep into mathematical frameworks, statistical principles, and algorithmic paradigms, we will gain comprehensive insights into various facets of machine learning, including supervised and unsupervised learning, regression analysis, classification techniques, clustering algorithms, neural networks, predictive modeling, and model evaluation. Through a balanced approach of theoretical exploration and hands-on applications, we will develop a solid foundation enabling us to critically assess, design, and innovate machine learning solutions across diverse domains.

Our primary textbook for this course will be **"Introduction to Machine Learning, 4th Edition, The MIT Press, 2020"** by **Ethem Alpaydın**, available online via the library. Additionally, we will refer to *"Python Machine Learning, 3rd Edition, Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2, PACKT books, 2019"* by *Sebastian Raschka* and *Vahid Mirjalili* to gain practical experience in applying machine learning concepts using the Python language.

This course will be divided into two modules. Following the completion of the first module, the midterm exam is scheduled for March 4th, 2024. The final exam is set for April 29th, 2024, and it will predominantly emphasize topics covered after the midterm assessment. Both exams will be in a traditional format, conducted with pen and paper only, and will be closed-book and closed-note.

A *tentative* schedule outlining the chapters we aim to cover is provided below. Please note that although our intention is to adhere to this schedule, adjustments, inclusions, or exclusions of planned subjects might occur based on our progress. Additionally, certain lectures will primarily focus on practical sessions introducing Python packages for machine learning through examples. Timely announcements will be made for these practice sessions.

- **WEEK 1-8, January 8th to February 28th:** Module–1, chapters 1, 2, 3, 4, 5, 6, 7
- **MIDTERM EXAM, March 4th**
- **WEEK 9-16, March 6th to April 24th:** Module–2, chapters 8, 9, 10, 11, 20
- **FINAL EXAM: April 29th**

Prerequisites for this course include Math-211 (Calculus) and CSCI-C200 (Introduction to Computers and Programming). A solid understanding of probability, statistics, and linear algebra will be crucial as these concepts will be frequently used throughout the course. If you feel less confident in these foundational mathematical areas, it is strongly recommended to review the related topics in the appendices of the textbook to reinforce your understanding.

Grading:

Assessment in this course will comprise a midterm exam scheduled for March 4th, 2024, and a final exam on April 29th, 2024, in addition to six homework assignments. These assignments are crafted to reinforce both your theoretical understanding and practical application of machine learning concepts, incorporating programming exercises. Python language and its related libraries will be utilized for these practice questions.

Furthermore, a seventh assignment will be offered as a BONUS opportunity. Completion of the BONUS assignment is optional. In the event of submission, the average homework score will be calculated based on the best six scores among the seven assignments.

- Exams: 45%. Midterm 25% and final 20%.
- Homework (HW): 55% HWs may contain some practice questions that will need coding in Python. the schedule for the assignments is given below.

Assignment ID	Announcement	Due
#1	1/18/2024	1/28/2024
#2	1/29/2024	2/11/2024
#3	2/12/2024	2/25/2024
#4	2/26/2024	3/9/2024
#5	3/18/2024	3/31/2024
#6	4/1/2024	4/14/2024
BONUS	4/8/2024	4/29/2024

Important notes about the execution of the course

- There are 6 assignments and a seventh one as optional! Please note that you will be doing the coding exercises in Python.
- Discussion of homework questions and programming exercises is certainly permitted. **However, all students must submit their individual work.** Cheating and plagiarism are strictly prohibited, and university policies regarding academic ethics (https://policies.iu.edu/policies/aca-33-code-academic-ethics/index.html) will be enforced. **Automated tools will be used to detect similarities between submissions.**
- **Late submission** of homework will incur penalties. There will be a 25% penalty for submissions on the first day after the due date, escalating to a 50% penalty if submitted more than 24 hours up to 48 hours late. Submissions will not be accepted after 48 hours past the due date. **These penalties will be strictly enforced.**
- **Re-taking** the midterm exam is permitted in this course. You will have the option to solve the midterm exam questions at home within two days after the midterm and submit your revised solutions. Your final score will be calculated as the average between your initial submission in the regular exam and the later revised submission. Notice that re-taking will NOT be provided for the final.
- Instructor and teaching assistants will be available during specified office hours without requiring prior appointments. Office hours will be announced in the first week of the course.
- The head teaching assistant of the course is Pavan Sai Vemulapalli (pavemu@iu.edu). You may reach him for any questions or concerns during the semester.
- Compliance with the policies of Indiana University Bloomington is mandatory. Refer to university policies at https://policies.iu.edu/index.html, specifically reviewing the code of ethics, code of conduct, and updates related to COVID-19 at https://www.iu.edu/covid/index.html.

Lec.#	Date	Topic
1	January 8, Mon.	Introduction
2	January 10, Wed.	Supervised Learning – I
-	January 15, Mon.	Martin Luther King Day-No Classes
3	January 17, Wed.	Supervised Learning – II
4	January 22, Mon.	Bayesian Decision Theory – I
5	January 24, Wed.	*Practice Session*
6	January 29, Mon.	Bayesian Decision Theory – II
7	January 31, Wed.	Parametric Methods – I
8	February 5, Mon.	Parametric Methods – II
9	February 7, Wed.	Multivariate Methods –I
10	February 12, Mon.	Multivariate Methods –II
11	February 14, Wed.	Dimensionality Reduction –I
12	February 19, Mon.	Dimensionality Reduction –II
13	February 21, Wed.	Clustering – I
14	February 26, Mon.	Clustering – II
15	February 28, Wed.	*Practice Session*
16	March 4, Mon.	MIDTERM
17	March 6, Wed.	Nonparametric Methods – I
–	March 11, Mon.	Spring Break
–	March 13, Wed.	Spring Break
18	March 18, Mon.	Nonparametric Methods – II
19	March 20, Wed.	Decision Trees – I
20	March 25, Mon.	Decision Trees – II
21	March 27, Wed.	Linear Discrimination – I
22	April 1, Mon.	Linear Discrimination – II
23	April 3, Wed.	Multilayer Perceptron – I
–	April 8, Mon.	Solar Eclipse – No Classes
24	April 10, Wed.	Multilayer Perceptron – II
25	April 15, Mon.	*Practice Session*
26	April 17, Wed.	Design and Analysis of ML Experiments - I
27	April 22, Mon.	Design and Analysis of ML Experiments - II
28	April 24, Wed.	Review and Miscellaneous Topics