Machine Learning and Data Mining (SOFE 4620)

Winter 2024
Course Outline

Instructor

Dr. Mennatullah Siam
Home page: https://engineering.ontariotechu.ca/people/ecse/mennatullah.siam.php
Office: SIRC-3388
Office Hours: XX-XX Mondays
Email: Only through Canvas

Teaching Assistants

TBD

Course Description

This course covers machine learning basics of both supervised and unsupervised learning algorithms. It covers classical methods such as support vector machines while including some of the more recent methods in deep learning. It also goes through the probabilistic framework by covering the basics of maximum likelihood, maximum posteriori and information theory.

Course Content:

08/01: Introduction to Machine Learning (**Tutorial:** Jupyter Notebooks Intro)
10/01: Parametric Methods - Linear Regression (**Tutorial:** Linear Regression Closed Form - Code)
15/1: Optimization for Beginners (**Tutorial:** Linear Regression with Gradient Descent - Code)
17/1: Parametric Methods - Linear Classification
22/1: Linear Classification cont. (**Tutorial:** Logistic Regression - Code)
24/1: Overfitting and Regularization (**Tutorial:** Regularized Linear Regression - Code)
29/1: Parameter Estimation I - Maximum Likelihood and Maximum Posteriori
31/1: Parameter Estimation II - Probabilistic View and Information Theory Basics (**Tutorial:** Linear Regression and Logistic Regression within a ML framework).
5/2: Naive Bayes Classification (**Tutorial:** Numerical Example - Code)
7/2: SVM I (**Tutorial:** Lagrange Multipliers Introduction)
12/2: SVM II (**Tutorial:** SVM - Code)

14/2: Revision Lecture
19/2, 21/2: Reading Week
26/2: Midterm Exam

28/2: Non-parametric Methods - K-Nearest Neighbour (**Tutorial:** KNN - Code)
4/3: Image Processing Basics - How to read Images, Convolution, More on Image Filtering (**Tutorial:** OpenCV Intro - Code)—> Necessary to understand Convolution to work w/ Convolutional Neural Networks.
6/3: Neural Networks I - Multi Layer Perceptron (**Tutorial:** MLP construction - Code**)**
11/3: Neural Networks II. - Backpropagation (**Tutorial:** MLP learning - Code)
13/3: Deep Learning I - CNNs & Advanced Regularization (**Tutorial:** CNNs - Code)
18/3: Deep Learning II - CNNs cont. & Transformers (**Tutorial:** Self Attention - Code)

20/3: Deep Learning III - Advanced Optimization (**Tutorial:** Adaptive Learning Rate Methods - Code)
25/3: Deep Learning IV - LLMs Intro. & Prompting, In-Context Learning
27/3: Clustering (**Tutorial:** Kmeans - Code)
1/4: Dimensionality Reduction (**Tutorial:** PCA)

1/4 Week: Projects Discussion Week - Refer to Canvas for the Schedule
8/4 Week: Study Break
15/4 Week: Final Exams

Textbook

- Murphy, Kevin P. Probabilistic machine learning: an introduction. MIT press, 2022.
- Machine Learning Handbook, Predrag Radivojac and Martha White.

Mark Distribution

- Attendance and Participation: 10%
- Assignments: 30%
- Mini-Research Project: 15%
- Midterm Exam: 15%
- Final Exam: 30%

Course Organization

Two 80-minute lectures that include in-class tutorials.

Lectures

- It is strongly recommended that all students attend all lectures and pay full attention at all times.
- Mondays 2:10-3:30 pm (XX) and Wednesday 2:10-3:30 pm (XX).
- Lectures will start on Monday 8th January, 2024.

Office Hours

- Office hours start on Monday January 15th.
- Run by the instructor:
    - Dr. Mennatullah Siam: XX-XX (SIRC-3388)

In-class Tutorials:

- Each lecture will cover an in-class tutorial that is mainly focused on introducting the practical aspects through programming or written exercises (e.g., derivations and numerical examples).
- The tutorial will include 1 or 2 questions that are left for the students to work on and deliver on Canvas during the lecture.
- Tutorials are graded for attendance not correctnes due to the limited lecture time. The later is evaluated through the assignments.

Assignments

- There are two programming exercises as assignments.
- These should be implemented in Python programming language and without the use of third party libraries unless stated in the assignment.

Mini-Research Project:

- Project groups of 3-4 students maximum are allowed.
- You can either choose from the suggested list of topics or choose your own.
- Pre-approval of your own topic from the instructor is required.
- Project presentations will happen before the study break in the April 1st week.
- Deliverables: Report and code. A more detailed rubric and deliverables guide will be released later.

Mid-term Exam

- There will be a mid-term exam on the 26th February during the lecture time
    - Monday: XX-XX (XX)
- There will be a revision lecture prior to the midterm exam to help the students prepare for it.

Final Exam

- Students are responsible to know the entire material taught in the course for the final examination.
- Final exam will be scheduled during the exam period **TBD**.
- Final exam is for the evaluation purposes only and will not be returned to students.