MACHINE LEARNING
CS 412
Fall 2023 - 2024

Instructor: Onur Varol, PhD	**Email**: onur.varol@sabanciuniv.edu
TAs: - Atra Zeynep Bahçeci - Berke Odacı - Ali Najafi - Yasser Zouzou	**LAs:** - Oytun Kuday Duran - Efe Tüzün - Ahmet Furkan Ün
All course related communication will be carried out using "cs412.fens@sabanciuniv.edu" email address, so please DO NOT send emails to any other email addresses.	
Time and location - Tuesdays 8:40 – 9:30, FENS G077 - Fridays 15:40 – 17:30, UC G030 **Recitations** (Tuesday 18:40 – 19:30): About every week or two to cover basic tools and techniques for hands-on experience. **Office hours and contact information off the TAs and LAs shared on SUCourse**	
Website: SUCourse platform will be used to share course material and information.	

Main references: This is a restricted list of various interesting and useful books that will be used during the course. You may need to consult them occasionally, but none of them are require.

- Ethem Alpaydın, *Introduction to Machine Learning*, 2010
- Tom Mitchell, Machine Learning, 1997
- Christopher M. Bishop, *Pattern Recognition and Machine Learning*, 2011 (Available online here)
- Jerome H. Friedman, Robert Tibshirani, and Trevor Hastie, *The Elements of Statistical Learning*, 2001 (Available online here)
- Jure Leskovec, *Mining of Massive Datasets*, 2020 (Available online here)

Course summary: This is an introductory machine learning course that will aim a solid understanding of the fundamental issues in machine learning (overfitting, bias/variance), together with several state-of-art approaches such as decision trees, linear regression, k-nearest neighbor, Bayesian classifiers, neural networks, logistic regression, and classifier combination. In addition to supervised approaches, unsupervised approaches will be covered, and model evaluations strategies will be introduced for different tasks.

Objectives and learning outcomes:

Objectives	Outcomes
Understand the basic concepts, issues, assumptions and limitations in machine learning (e.g. overfitting, error measures, curse of dimensionality).	Have a solid understanding of the basic concepts, issues, assumptions, and limitations in machine learning and how they apply to various machine learning techniques.
Have a working knowledge of the basic background on probability and linear algebra, and algorithms behind common machine learning techniques; together with their suitability in given situations.	Have a working knowledge of the basic mathematics and algorithms behind common machine learning techniques; together with their suitability in given situations.
Given a machine learning problem, be able to implement and evaluate one of the standard machine learning algorithms.	Given a machine learning problem, select, implement and evaluate one of the appropriate machine learning algorithms using Python.

Sabancı Universitesi

Tentative Course Outline:

Weeks	Deadlines	Topics
Week 1 (3-6 Oct.)		Introduction to ML concepts
Week 2 (10-13 Oct.)	Traveling for a conference	NO CLASS ON FRIDAY
Week 3 (17-20 Oct.)		Feature extraction and selection
Week 4 (24-27 Oct.)		Decision tree learning and ensemble learning
Week 5 (31 Oct.-3 Nov.)	HW#1 release	Nearest neighbor classifier
Week 6 (7-10 Nov.)		Linear and logistic regression
Week 7 (14-17 Nov.)	HW #1 due	Bayesian approaches
Week 8 (21-24 Nov.)		Midterm exam (24 Nov., Friday in class)
Week 9 (28 Nov. – 1 Dec.)		Kernel methods and support vector machines
Week 10 (5-8 Dec.)	HW#2 release	Neural networks and Introduction to deep learning
Week 11 (12-15 Dec.)	Traveling for a project meeting	NO CLASS ON FRIDAY
Week 12 (19-22 Dec.)	HW #2 due	Practical issues
Week 13 (26-29 Dec.)		Unsupervised learning - Dimensionality reduction
Week 14 (2-5 Jan.)		Unsupervised learning - Clustering
Finals		Final exam

Grading Policy: These percentages are tentative and subject to change based on enrollment numbers and the number of TAs and LAs assigned to the course.

- **Midterm and Final exams (25% and 30%)**: Exams will be help in person
- **Homeworks (20%)**: Mix of programming and written questions will be given, and responses will be provided as a report for each assignment.
- **Project (25%)**: The project will probably be about social media analysis. Teamwork is encouraged and groups can consist of 3-5, but each student will be evaluated individually. Project involves data collection, annotation, feature extraction, model building and reporting.

Class Policies and advice:

- Regular attendance is essential and class participation is expected in paper discussions.
- Late assignments. There will be 10% late penalty for up to 3 days and 20% penalty for assignments submitted in the next 10 days.
- Maximum score you can receive from the projects and assignments cannot be more than 1.5 of the exam score. *For instance, if your exam score is 40 any HW or project score higher than 60 will be lowered to 60.*
- Students have the responsibility of backing up all their data and code. At the end of the semester, they are expected to prepare public release of their code and data with a proper documentation.

Academic honesty: All students must follow the university guidelines of academic integrity.
https://www.sabanciuniv.edu/en/academic-integrity-statement