MACHINE LEARNING
CS 412 - Fall 2023

Instructor: Onur Varol, PhD	**Email**: onur.varol@sabanciuniv.edu
TAs: - Ali Najafi - Veysel Oğulcan Kaya **LAs:** - Omay Ece Ayyıldız - Ege Aktemur	**Emails**: ali.najafi@sabanciuniv.edu vogulcan@sabanciuniv.edu ayyildizomay@sabanciuniv.edu egeaktemur@sabanciuniv.edu
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, FENS G077 **Recitations** (Tuesday 18:40 – 19:30): About every week or two to cover basic tools and techniques for hands-on experience. Please use the following link for online event https://sabanciuniv.zoom.us/j/92887595036	
Website: SUCourse platform will be used to share course material and information.	
Office hours: - Onur Varol: By email appointment - Ege Aktemur: Wednesday 9:40 – 10:30 https://sabanciuniv.zoom.us/j/3996774312 - Omay Ece Ayyıldız: Mondays 15:00 – 16:00 https://sabanciuniv.zoom.us/j/98173040182 - Veysel Oğulcan Kaya: Wednesdays 10:00 – 11:00 https://sabanciuniv.zoom.us/j/3393853446 - Ali Najafi: Wednesday 16:00 – 17:00 https://sabanciuniv.zoom.us/j/7133283748	

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 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. Given a machine learning problem, be able to implement and evaluate one of the standard machine learning algorithms.	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 mathematics and algorithms behind common machine learning techniques; together with their suitability in given situations. Given a machine learning problem, select, implement and evaluate one of the appropriate machine learning algorithms using Python.

Tentative Course Outline:

Weeks	Deadlines	Topics	Recitations
W 1		Introduction to ML concepts	
W 2		Feature extraction and selection	
W 3		Decision tree learning and ensemble learning	Onur
W 4		Nearest neighbor classifier	Omay
W 5	HW#1 rel.	Linear and logistic regression	Ege
W 6		Conference – No class	Ege
W 7	HW #1 due	Bayesian approaches	Ali
W 8		Kernel methods and support vector machines - Midterm exam	
W 9	Annotation	Neural networks	Veysel
W 10	HW#2 rel.	Introduction to deep learning	Ali
W 11	Project rel.	Deep Learning on text, image, graphs	Veysel
W 12	HW #2 due	Practical issues	
W 13	Proj #1	Unsupervised learning - Dimensionality reduction	
W 14	Proj #2	Unsupervised learning - Clustering	Omay
Finals	Proj #3	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 (or following the universities guidelines).
- **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 be about social media analysis. Last semester we build bot detection system at the account level and sentiment classifier for the tweet level. 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.

Late submissions will cause 5 points reduction for each day and no assignment will be accepted after 7 days.
To pass the course your grade as calculated following the above criteria **must be at least 40** and this is not subject to change or negotiation and **final exam score also must be above 39/100**. Attendance and class participation will be considered in border-line grades.

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