CS412: Machine Learning

Syllabus, Spring 2024

Machine Learning is centered on automated methods that improve their performance through learning patterns in data. In many data-rich domains, machine learning provides solutions in business (image processing, speech recognition, recommendation, and information retrieval systems) and science (folding of protein structure, etc.). In this undergraduate-level class, students will get an introduction to the methodologies and algorithms for machine learning. Topics include supervised learning, unsupervised learning, evaluating performance and model selection and a basic introduction to deep learning. The course will also discuss recent applications of machine learning. Programming and analytical assignments include hands-on practice with various learning algorithms. Students entering the class are expected to have a pre-existing working knowledge of probability, statistics, linear algebra, programming, and algorithms.

Schedule

Lectures:	Mon 16:40 – 17:30	FENS G062
	Tue 15:40 – 17:30	FASS G062
Recitation :	Mon 17:40 – 19:30	FENS G062

Contact Information

Instructors: Oznur Tastan, `otastan@sabanciuniv.edu`

Office Hours: Oznur Tastan, by appointment via e-mail will be announced.

TAs:

Mert Pekey	`mpekey@sabanciuniv.edu`
Ali Enver Bilecen	`bilecen@sabanciuniv.edu`
Mansur Kiraz	`mansur.kiraz@sabanciuniv.edu`
Ekin Başar Gökçe	`basar.gokce@sabanciuniv.edu`
Atra Zeynep Bahçeci	`atra.bahceci@sabanciuniv.edu`

Delivery format

We will have physical lectures. Attendance will be taken. Your active participation is expected.

Course Webpage

We will be using SuCourse. Please check regularly the SuCourse of the course for lecture notes, homework assignments, discussions, and announcements.

Textbook: No required textbooks. There will be **required readings, videos** posted on SuCourse.

Some reference textbooks:

- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. An Introduction to Statistical Learning: with Applications in R (online version available).
- Christopher M. Bishop, Pattern Recognition and Machine Learning, Springer, 2011.

- Ethem Alpaydin, Introduction to Machine Learning, 2e. The MIT Press, 2010.
- Kevin P. Murphy, Machine Learning: a Probabilistic Perspective, The MIT Press, 2012.
- Tom Mitchell, Machine Learning, McGraw Hill, 1997.
- Introduction to Machine Learning with Python, Andreas C. Müller & Sarah Guido, 2016.
- Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques for Building Intelligent Systems, Aurélien Geron, 2017.

Tentative Topics

Introduction
Basic concepts
Model selection and evaluation
Performance metrics
Practical issues (Handling imbalance, missing values,..)

Supervised learning
Bayes Classifier Naive Bayes classifier
Gaussian naive Bayes classifier
Logistic regression
Linear regression
Support vector machines
Kernels
Decision trees
Ensemble methods: Bagging, boosting
Neural networks
Introduction to deep learning

Unsupervised Learning
Clustering; k-means, spectral clustering, DBScan
Principal components analysis

Additional Topics (if time permits)
Interpretability
Matrix factorization and collaborative filtering
Active learning

Grading

The final grades will be based on the following:

- One midterm exam, 25 % .
- One final exam, 30 %
- Homework (20%)
- Team project (15%)
- In class quizes (10 %)

IMPORTANT: One of the following conditions will result in an automatic NA (Not attended) or F regardless of other grades:

1. Your overall grade is **below 35.** (F)
2. Your final exam grade is **below 30.** (F)
3. Not submitting **two or more** homework assignments (empty homework does not count as a submission) (NA)
4. Average of the homework grade is **below 30.** (F)
5. Missing an exam without an execuse (NA)
6. Missing **5 or more** quizzes (NA).

** The instructor/TA can call for an oral exam to explain your homework/exam/quiz solution and answer course related questions. Students who fail to explain their work or answer related questions will get zero (0) credit from the related grade item.

** Attendance will be taken and will have an affect in the border line cases.

** Time conflict requests can be accepted for one hour only. Students who are registered to the course with time-conflict override accept the responsibility of any inconvenience that might occur due to missed content and/or (if any) quizzes and exams. No make-up will be available for such cases.

Homework assignments: Not all homework will be equal weight. Both written and programming questions.

Exams: Midterms and final exams will be paper-based.

Quizes: Quizes will be online but attendance will be taken. Participants who does not have a physical presence in the class room will get zero. The lowest two quiz will be dropped.

Projects: Project will be conducted in teams of 4-5. There will be a peer grade evaluation. Each student will be evaluated individually. We will make announcements on this.

Make up policy: There will be a single make-up exam for the exams and it will cover all topics. You can take a make-up only if you have a valid health report approved by the University Health Services. There will be no make-up for other components of the course.

Late day policy (IMPORTANT: Each student will have 4 free late days to use only for the homework assignments. 4 days is the total number of late days for each homework. Late days can only be used on the homework, and not on the project deliverables. You do not need to explain why you are submitting late and no need to notify us. <= 24 hours late counts as 1 day late, etc. Once these total of four late days are exhausted, any assignments turned in late will be penalized and will incur a reduction of 33% in the final score, for each day (or **part thereof**) it is late. For example, if an assignment is up to < 24 hours late, it incurs a penalty of 33%. Else if it is up to more than 24 hours and less than 48 hours late, it incurs a penalty of 66%. And if it is 72 or more hours late, it will receive no credit.

Regrade policy: Important: You may object to a grade only within 14 days after the grades are announced. If you feel that an error was made in grading your homework or midterm, please get an appointment to discuss.

Honor code: Students are expected to comply with the Sabanci University Academic Integrity Statement. Any form of academic dishonesty will be penalized with a failing grade and disciplinary actions will be taken. Students may discuss and work on homework problems in groups in an

abstract way. However, each student must write down the solutions independently, and without referring to written notes from the joint session. In other words, each student must understand the solution well enough to reconstruct it by him/herself. In addition, each student must write on the problem set the names of the people with whom s/he collaborated or discussed. Otherwise, it will be considered as an incidence of academic honesty violation.