Course Information

Course Code	9700557
Course Section	1
Course Title	STATISTICAL LEARNING AND SIMULATION
Course Credit	3
Course ECTS	8.0
Course Catalog Description	All students of the Institute are welcome to this course. Various methods from statistics, discrete mathematics, numerics and information theory are presented and combined from the view-point of modern algorithms and applications. The computational aspect is taken into account. Throughout the course, we discuss and perform the practical means of simulation and learning. The purpose of the exercises is to familiarize the student with the most usual numerical techniques and their applications.
Prerequisites	No prerequisites
Consent of Dept./Inst.	Consent of the instructor
Corequisites	Linear Algebra and Advanced Calculus
Schedule	Not available
Lab Hours & Location	This is an online course; however, we will virtually meet regularly
Course Website	https://odtuclass.metu.edu.tr
Learning Management System	https://odtuclass.metu.edu.tr

Instructor Information

Name/Title	Prof.Dr. ÖMÜR UĞUR
Office Address	Middle East Technical University, Institute of Applied Mathematics, S-227
Email	ougur@metu.edu.tr
Personal Website	http://ougur.iam.metu.edu.tr
Office Phone	210 5617
Office Hours	Fridays, 10:40 - 12:30 (tentative)

Course Objectives

At the end of the course, the student will learn:

- the fundamentals of Statistical Learning, regression and classification
- linear and nonlinear regressions including splines
- Generalised Additive Models for both regression and classification problems
- regularisation techniques including Ridge regression and the Lasso
- the tree-based methods for regression and classification
- Support Vector Machine which is highly appreciated among Data Science and Machine Learning Community
- the difference between supervised and unsupervised learning methods

Course Learning Outcomes

Student, who passed the course satisfactorily will be able to:

- present the data and its descriptive analysis
- distinguish between regression and classification problems
- apply regression or classification algorithms to solve related problems
- code their own algorithms for specific applications in Statistical and Machine Learning
- understand the fundamentals of Support Vector Machine and be able to apply to specific problems
- distinguish between supervised and unsupervised learning methods in related applications

Instructional Methods

The following instructional methods will be used to achieve the course objectives: Lecture, questioning, discussion, group work, simulation.

Tentative Weekly Outline

Week	Topic	Relevant Reading	Assignments
1	Brief introduction to Statistical Learning a) Regression versus Classification		
2	Linear Regression a) simple and multiple Linear Regression		
3	Classification a) Logistic Regression b) Discriminant Analysis (Linear and Quadratic)		
4	Resampling Methods a) Cross-Validation b) the Bootstrap		
5	Regularisation a) Subset Selection b) Ridge Regression c) the Lasso d) Principle Components Regression e) Partial Least Squares Regression		
6	Nonlinear Models a) Polynomial and Splines b) Generalised Additive Models		
7	Tree-Based Models a) Decision Trees b) Random Forest c) Boosting		
8	Support Vector Machines		
9	Unsupervised Learning a) Principle Component Analysis b) Clustering Methods		

Course Textbook(s)

Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning - with Applications in R, Springer, 2013 (Corrected at 8th printing 2017)

Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning - with Applications in R, Second Edition. Springer, 2021

Course Material(s) and Reading(s)

Material(s)

Books (Textbook):

- Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed., Springer, 2009 (Corrected at 12th printing 2017)

Reading(s)

Books (Supplementary):

- Kevin P. Murphy, Machine Learning: A Probabilistic Perspective, The MIT Press, 2012
- Peter Harrington, Machine Learning in Action, Manning Publications Co., 2012
- Charu C. Aggarwal, Neural Networks and Deep Learning: A Textbook, Springer, 2018
- G. Jay Kerns, Introduction to Probability and Statistics Using R, 1st ed., 2015
- Robert V. Hogg, Elliot A. Tanis, Dale Zimmerman, Probability and Statistical Inference, 9th ed., 2015
- Larry Wasserman, All of Statistics - A Concise Course in Statistical Inference, 2004
- W. N. Venables, D. M. Smith, and the R Core Team, An Introduction to R - Notes on R: A Programming Environment for Data Analysis and Graphics, Version 3.4.2 (2017-09-28)

Other

Resources:

- *The R Project for Statistical Computing*: https://www.r-project.org/
- *python*: https://www.python.org/
- *RStudio*: https://www.rstudio.com/
- *Anaconda*: https://www.anaconda.com/

Supplementary Readings / Resources / E-Resources

Readings

It is suggested that you should read the *documentations* of each resource below:

- *The R Project for Statistical Computing*: https://www.r-project.org/
- *python*: https://www.python.org/

Assessment of Student Learning

Assessment	Dates or deadlines
Assignments & Quizzes: The homework assignments and quizzes will be designed to help you learn specific skills covered in class. Mostly, they will be available on the course pages on ODTUClass (Moodle) system. The assignments and quizzes must be completed by the student ALONE; unless otherwise is explicitly stated/permitted, NO collaboration is allowed!	
Exams: There will be *one or two* exams that you have to write in this course. The structure of each exam may change: although it might change, the exams will be an online and possibly be an open-book one.	
Final Project: This is a take-home exam that will be given to you before the last day of classes. The final project will involve programming as well as reporting the findings of your assigned project and will be regarded as your final exam. You may be asked to **present your works** as well in class (possibly via online meeting).	

Course Grading

Deliverable	Grade Points
Assignments & Quizzes	50
Exams	30
Final Project	20

Deliverable	Grade Points
Total	100

Course Policies

Class Attendance

Students taking the course must fulfill the requirements stated in the course website. You may feel free to attend the regular meetings weekly (schedule), where we will discuss the topics of the week.

Late Submission of Assignments

Late assignments will not be graded!

Make up for Exams and Assignments

There will be no make-ups for assignments and the projects in this course.

Information for Students with Disabilities

Students who experience difficulties due to their disabilities and wish to obtain academic adjustments and/or auxiliary aids must contact ODTU Disability Support Office and/or course instructor and the advisor of students with disabilities at academic departments (for the list: http://engelsiz.metu.edu.tr/en/advisor-students-disabilities) as soon as possible. For detailed information, please visit the website of Disability Support Office: https://engelsiz.metu.edu.tr/en/

Academic Honesty

The METU Honour Code is as follows: *"Every member of METU community adopts the following honour code as one of the core principles of academic life and strives to develop an academic environment where continuous adherence to this code is promoted. The members of the METU community are reliable, responsible and honourable people who embrace only the success and recognition they deserve, and act with integrity in their use, evaluation and presentation of facts, data and documents."*