DERS TANITIM BİLGİLERİ (İNGİLİZCE)

Course Information

Course Name	Code	Semester	Theory (Saat/Hafta)	Application (Saat/hafta)	Laboratuary (hours/week)	National Credit	ECTS
Machine Learning	**ECON 484**	**Spring 2022**	**3**			**3**	**5**
Prequisites	Students should understand basic probability and statistics, and college-level algebra and calculus. Knowledge of linear algebra is also expected, and some knowledge of mathematics underlying probability models will be useful. Programming assignments will be expected to be submitted as Python, Java or R code.						
Course Language	English						
Couse Type	Elective						
Mode of Delivery (face to face, distance learning)	Face-to-Face with Distance Learning Components If for any reason, courses need to be conducted remotely, a permanent Zoom link for the lectures will be provided through Moodle and e-mail.						
Learning and Teaching Strategies	**Flipped Classroom:** Example code and workbooks in Knime, Python, Java and R will be shared through Github. Students ae expected to review these before coming to class. https://github.com/boragungoren-portakalteknoloji/ATILIM-ECON484-Spring2022 **Active Homeworks and Projects:** Students will form teams early in the semester and pursue practical aspects of **Use of Github:** Students are required to use Github actively.						
Instructor(s)	Bora GÜNGÖREN						
Course Objective	Machine learning is used to design and implement software systems that learn from experience. Such systems are not directly programmed to solve a problem based on a fixed algorithm, but instead they further develop their own algorithm based on examples of how they should behave. This requires some form of trial and error experience, trying to solve the problem. Therefore machine learning is very promising for problems that are solved many times, which includes many decision problems in business and economics, as well as many problems in engineering and sciences. Machine learning has proven itself through many applications in economics and life sciences, and its use has spread towards almost all fields where a class of problems I solved over and over. Learning performance of any machine learning approach is very dependent on the type of trial and error experience, therefore researchers not only develop new algorithms but also focus on evaluating which algorithms are best used in which circumstances. This course will focus on “a core set” of machine learning methods that have proven valuable and successful in many practical applications. This course will contrast these methods, with the aim of explaining the circumstances under which each is most appropriate.						
Learning Outcomes	Upon the completion of this course, the student will be able to: 1. Define machine learning. 2. Evaluate a given model.						

	3. Apply typical techniques such as dimension reduction, using decision trees, naive bayes 4. Work with features 5. Work with real world data
Course Content	1. Concepts of Machine Learning (Standard Examples, Relations with Other Fields, Approaches); Building an initial estimation model; Model evaluation criteria 2. Building a first kNN Model 3. Model Selection Criteria; Linear models (Linear regression, Logistic regression, Ridge, Lasso); Comparison of advantages and disadvantages of techniques 4. Dimension reduction; Feature selection strategies; Feature projection strategies (PCA, NMF, Kernel PCA, LDA, GDA); Comparison of advantages and disadvantages of techniques 5. Decision trees; Basic terminology; Basic methods for splitting trees (Variance reduction, Knowledge Gain, Gini, Chi-square); Alternative methods to decision trees (Random forest, XGBoost) 6. Working with Features (Date-Time data, Categorical data, Reframing Numerical Quantities, Scaling Numerical Quantities) 7. Working with Text Files (Using Regex, Text Cleanup) 8. Naive Bayes 9. Factor Analysis
References	1. Bishop. CM. 2006. Pattern Recognition and Machine Learning. Springer. 2. Murphy, KP. 2012. Machine Learning: A Probabilistic Perspective. MIT Press. 3. Kelleher, A, Kelleher, A. 2019. Machine Learning in Production: Developing and Optimizing Data Science Workflows and Applications. Addison Wesley. 4. Burkov, A. 2019. The Hundred-Page Machine Learning Book. (Free E-Book)

Course Schedule: Wednesdays 12:30 – 15:20

Office Hours: Available upon request, through Zoom (evenings).

Weekly Course outline

Week	Topics	Pre-study

1	Concepts of Machine Learning
	• Standard Examples
	• Relations with Other Fields
	• Approaches
	• Building an initial estimation model
	• Model evaluation criteria
	  • Confusion Matrix
	  • Precision
	  • Accuracy
	  • Recall
	  • R-Square
		Install necessary software (Knime, Python, R, Java, MySQL or SQLite)
		First assignment: Establish Github account and follow course Github page

2	Building a first kNN Model Model Selection Criteria • Over-fitting • Under-fitting • Verification • Bias	Form Project Teams
3	Linear models • Linear regression • Logistic regression	Select Team Project Topics
4	Linear models (cnt'd) • Ridge • Lasso • Comparison of advantages and disadvantages of techniques	

5	Dimension reduction
	• Feature selection strategies
	• Feature projection strategies
	  ○ PCA
	  ○ NMF
	  ○ Kernel PCA	

6	Dimension reduction (cnt'd)
	• Feature projection strategies
	  ○ LDA
	  ○ GDA
	  ○ Comparison of advantages and disadvantages of techniques	

7	Mid-Term Examination	
8	Decision trees • Basic terminology • Basic methods for splitting trees	

	○ Variance reduction ○ Knowledge Gain ○ Gini ○ Chi-square	
9	Decision trees (cnt'd) • Alternative methods to decision trees ○ Random forest ○ XGBoost	
10	Working with Features • Date-Time data • Categorical data • Re-framing Numerical Quantities • Scaling Numerical Quantities	
11	Working with Text Files • Using Regex • Text Cleanup	
12	Naive Bayes	
13	Factor Analysis	
14	Project Presentations	

Assesment methods

Course Activities	Number	Percentage %
Attendance **(mandatory after add-drop week)**	12	%10
Assignments (**programs and/or reports, submissions through Github**)	4	%40
Midterm Exam **(in-class or online, short duration, with discussion questions)**	1	%20
Final Project **(a program, with a report, submissions through Github)**	1	%30
Total	18	100
Percentage of semester activities contributing grade success	18	100
Percentage of final exam contributing grade success		
Total	18	100

Attendance: Attendance after add-drop period is mandatory. Students are required to attend at least %70 of classes. Typical excuses such as health reports should be communicated with the instructor.

Use of Programming Languages: Course discussions are language-independent. Some examples will be visualized through Knime.

Assignments will be primarily based on Java but students can freely submit Python or R code when there is no explicit Java requirement. Students are expected to have a fundamental understanding of programming language. Some limited guidance is available.

Course Category: Major Area Courses

Workload and ECTS Calculation

Activities	Number	Duration (Hours)	Total Work Load
Course Duration (Including Exam Week: 16 x Total Hours)	14	3	42
Laboratory			
Application			
Specific practical training (if any)			
Field Activities			
Study Hours Out of Class (Preliminary work, reinforcement, etc)	13	3	39
Presentation / Seminar Preparation			
Projects	1	30	30
Homework assignment	4	6	24
Midterms (Study duration)	1	25	25
Final (Study duration)			
Total Workload			160

Matrix of the Course Learning Outcomes Versus Program Outcomes

Program Outcomes	Contribution Level*				
	1	**2**	**3**	**4**	**5**
1 Deepen and develop advanced level and current knowledge in economics to an expert level based on the competencies of the economics education and acquire the skills make original thoughts, views, and contributions to science.					x
2 To comprehend the interaction between economics and related fields; to achieve original results by using expert knowledge in analysis,and evaluation of new and complex ideas.					x
3 To acquire high level skills in applied methods in economics and to be able to systematically follow new theoretical developments in applied economics.					x
4 To be able to apply existing techniques to an original research idea.					x
5 To find a new scientific method.					x
6 To broaden the knowledge of economics by publishing scientific articles related to economics in national or international refereed journals and to publish these articles in a scientific ethical way.					x
7 To be able to identify, participate in, organize and lead activities, such as panels, workshops and conferences, that aim to develop solutions for novel and interdisciplinary problems.			x		
8 To follow pioneering and innovative theories and methods in economics, to be able to use them in analysis and research and to be able to contribute to the process of becoming an information society.					x
9 To be able to analyze the relationships and processes in economics and to develop functional and effective communication networks.			x		
10 To make some scientific contributions to projects aiming to increase the productivity in business.					x
11 To be able to share his/her work and to criticize works of others by using academic networks in economics efficiently.					x

1: Lowest, 2: Low, 3: Average, 4: High, 5: Highest