BA in Management Program
Fall 2020
OPIM407 – Advanced Business Analytics

Instructor: M. Erdal Balaban, PhD.
Office: YBF 1034
Phone: (216) 483-<xxxx>
Fax: (216) 483-9699
E-mail: erdal.balaban@sabanciuniv.edu
mebalaban@gmail.com
Web: SuCourse
Office Hours: <by appointment, or exact time specified>

Type	Time	Days	Where
Class	17:40 - 19:30	W	Online
Class	9:40 am - 10:30 am	R	Online
Recitation			

Course Objective:
The main objective of this course is to present predictive and prescriptive analytics tools in the context of business cases, with an emphasis on implementing analytical approaches within an organization. Throughout this course, students will learn: use tools and ideas from Analytics and R's environment to solve interesting and exciting business problems; learn how to formulate relevant business questions that can be answered using data; understand the various steps of data preparation like data cleaning and feature extraction; develop and test hypotheses to inform managerial decisions and to collaborate meaningfully with multiple stakeholders; evaluate the effectiveness of a solution through statistical learning methods (supervised & unsupervised methods); communicate results to both technical and non-technical audiences.

Learning Outcomes:
Upon successful completion of the course, the student should be able to:
1. Understand the business analytics, data mining and related terms, big data, data science.
2. Use R language commands as a business analytic tool.
3. Use visual techniques to describe data and make better business decisions using R language.
4. Use data mining (supervised and unsupervised learning algorithms) in solution real life business problems.
5. Explain the assumptions of various techniques such as K-Means Clustering, k-Nearest Neighbor, Naïve Bayes Clasifier, Decision Trees, Multiple Regression, Logistic Regression, and Artificial Neural Networks.

6. Write an insightful and well-organized report for a real-world case study, including thoughtful and convincing details.

Course Material:
Main Textbook:

R for Everyone: Advanced Analytics and Graphics, 2nd Edition, Pearson Global Edition.
https://www.homerbooks.com/urun/r-for-everyone-advanced-analytics-and-graphics
Author: Jared P. Lander (https://www.jaredlander.com/about/)

Optional Reading Material:

Data Mining for Business Analytics: Concepts, Techniques, and Applications in R
Author: Galit Shmueli, Peter C. Bruce, Inbal Yahav, Nitin R. Patel, Kenneth C. Lichtendahl, Jr. Publisher: John Wiley & Sons, R Edition (2017)

Additional List of References in Turkish:

- Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları (Beyoğlu, İstanbul: Çağlayan Kitabevi, 2.Baskı 2018)
  Yazarlar: M.Erdal Balaban, Elif Kartal
- R ile Veri Madenciliği Uygulamaları, (Beyoğlu, İstanbul: Çağlayan Kitabevi, 1.Baskı 2016)
  Editörler: M.Erdal Balaban, Elif Kartal
- Veri Madenciliği ve Makine Öğrenmesi Temel Kavramlar, Algoritmalar, Uygulamalar (Beyoğlu, İstanbul: Çağlayan Kitabevi, 1.Baskı 2019)
  Editörler: M.Erdal Balaban, Elif Kartal

Software:

- We require the R Statistical Software, which is powerful and free. R can be downloaded at the link below: http://www.cran.r-project.org/
- RStudio is a free platform for both writing and running R, available at www.rstudio.org. Some students find it friendlier than basic R (especially in Windows OS).
- The learning curve is very steep. Students can become proficient in a few weeks. Some manuals are very helpful to learn R, e.g., http://cran.r-project.org/manuals.html

List of Cases

Case 1	Date:	**December 17**
	Case:	Teaching Assistant Evaluation https://archive.ics.uci.edu/ml/datasets/Teaching+Assistant+Evaluation
	Type:	Chapter Case
	Subject:	**k-Nearest Neighbor Algorithm**
	Teamwork?:	No
	Grading:	Individual
Case 2	Date:	**December 24**

	Case:	Credit risk prediction https://archive.ics.uci.edu/ml/datasets/South+German+Credit	
	Type:	Chapter Case	
	Subject:	**Naïve Bayes Classifier**	
	Teamwork:	No	
	Grading:	Individual	
Case 3	**Date:**	**December 31**	
	Case:	Prediction of potential bank customers https://archive.ics.uci.edu/ml/datasets/Bank+Marketing	
	Type:	Chapter Case	
	Subject:	**Logistic Regression**	
	Teamwork:	No	
	Grading:	Individual	
Case 4	**Date:**	**January 7**	
	Case:	Prediction of tennis match results https://archive.ics.uci.edu/ml/datasets/Tennis+Major+Tournament+Match+Statistics	
	Type:	Chapter Case	
	Subject:	**Decision Tree**	
	Teamwork:	No	
	Grading:	Individual	

Course Web:
In this course, students will actively use the SUCourse online system. Lecture notes, slides, and additional material will be available on SUCourse. Students will be expected to visit the course web site at least two or three times a week. SUCourse will also be used for any in-class exercises, short assignments, and uploading/downloading all course-related files. Students will submit all assignments via SUCourse. Sabanci University uses a very powerful web-based tool called Turnitin. Turnitin is the worldwide standard in online plagiarism prevention. It allows instructors to compare student papers against a database composed of millions of articles. Every paper submitted by students will be scanned and cross-checked by Turnitin, and results will be reflected in students' grades.

Instructional Design:
This course is intended to be highly interactive, engaging students in active learning through hands-on exercises, guest speakers, etc. in addition to standard lecture material. To facilitate this process, students are expected to come to class prepared by reading the assigned material, and to actively and meaningfully participate in class discussions.
Grading:

Participation	: 10%
Homework and Case Assignments	: 40%
Midterm Exam	: 25%
Term Project	: 25%

Requirements:

General requirements regarding the grading items listed above are as follows: a) Participation score is awarded for in-class participation only (not just attendance). Participation means joining in class discussions and intellectually contributing to the learning in the classroom by voicing one's ideas, comments, feedback, etc. regarding the subject matter. b) In-class exercises are meant to be completed during the lecture hour. Students will submit the result of their work at the end of the lecture. Students who are not physically present in class will not receive marks from such exercises. Make sure you sign the attendance sheet on the day of the exercise. c) Homeworks and reports are to be individually submitted via SUCourse no later than the posted due date and time. SUCourse will be closed for submission and late homework will not be accepted. There will be no deadline extensions for any homework or report. d) Midterm Exam will include multiple-choice, short essay, and problem-solving type of questions. The exam will be open book and open notes and laptops, with the exception of multiple-choice and short-essay sections. A make-up exam will be offered at the end of the semester to those who miss the Midterm Exam and bring an official doctor's report for it. No other excuses will be accepted for make-up purposes. e) The Term Project will be assigned in the second half of the semester and the students will be expected to turn in a Final Report during the first week of university-wide final exams. The project can be done Individually. The project will be about a real business case with business data where students will be expected to apply the concepts and techniques learned in class to data-driven decision making. f) If you miss a particular assignment (including class attendance) due to sickness, accident, etc., you must bring in an official doctor's report describing the situation before you can request a make-up for the missed grade. No other excuses (e.g. grandmas or grandpas passing away) will be accepted for make-up purposes.

Academic Honesty:

Learning is enhanced through cooperation and as such you are encouraged to work in groups, ask for and give help freely in all appropriate settings. At the same time, as a matter of personal integrity, you should only represent your own work as yours. Any work that is submitted to be evaluated in this class should be an original piece of writing, presenting your ideas in your own words. Everything you borrow from books, articles, or web sites (including those in the syllabus) should be properly cited. Although you are encouraged to discuss your ideas with others (including your friends in the class), it is important that you do not share your writing (slides, MS Excel files, reports, etc.) with anyone. Using ideas, text, and other intellectual property developed by someone else while claiming it is your original work is *plagiarism*. Copying from others or providing answers or information, written or oral, to others is *cheating*. Unauthorized help from another person or having someone else write one's paper or assignment is *collusion*. Cheating, plagiarism, and collusion are serious offenses that could result in an F grade and disciplinary action. Please pay the utmost attention to avoid such accusations.

Classroom policies and conduct

Sabancı BA in Management Program values participatory learning. Establishing the necessary social order for a participatory learning environment requires that we all:

- Come prepared to make helpful comments and ask questions that facilitate your own understanding and that of your classmates. This requires that you complete the assigned readings for each session before class starts.
- Listen to the person who has the floor.
- Come to class on time.
- If you use your laptop during class, it is only to be used for class activities such as taking notes or referring to a spreadsheet. You are not to connect the laptop to the network and should not be doing any non-class activities during class time.

Course Schedule:

Week 1	**Date:**	**October 7-8, 2020**
	Topic:	*Course Introduction:* What Is Business Analytics and Business Intelligence?, What Is Data Mining?, Data Mining and Related Terms, Big Data, Data Science.
	Requirements:	Reading (before class) –Chapter 1, Optional textbook (Galit Shmueli, …)
Week 2	**Date:**	**October 14-15, 2020**
	Topic:	*R Language for Business Analytics I* Getting R: Downloading R, R Version, 32-bit vs. 64-bit, Installing R, The R Environment: Command Line Interface, RStudio, R Packages: Installing Packages, Loading Packages, Basics of R: Basic Math, Variables, Data Types, Vectors, Calling Functions, Missing Data
	Requirements:	Reading (before class) – Chapter 1-2-3, R for Everone OR Chapter 3, Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları, (Balaban & Kartal, 2018) **Assign Homework #1**
Week 3	**Date:**	**October 21-22, 2020**
	Topic:	*R Language for Business Analytics II* Advanced Data Structures: data.frames, Lists, Matrices, Arrays, Reading Data into R: Reading CSVs, Excel Data, R Binary Files, Extract Data from Web Sites, Writing R functions: Hello, World!, Function Arguments, Return Values, Control Statements: if and else, switch, ifelse, Loops (for, while), Controlling Loops
	Requirements:	Reading (before class) – Chapter 4-5-6, R for Everone **Homework #1 is due** **Assigned Homework #2**
Week 4	**Date:**	**October 28-29, 2020**

	Topic:	**Overview of the Data Mining Process** Data Mining, Supervised Learning, Unsupervised Learning, , Process the term "Knowledge Discovery in Databases" (KDD), Sampling, Types of Variables, Outliers, Missing Data, Normalizing Data, Overfitting, Partitioning the Data.
	Requirements:	Reading (before class) –Chapter 1, Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları, (Balaban & Kartal, 2018) **Homework #2 is due**
Week 5	**Date:** Topic:	**November 4-5, 2020** *Visualization:* Base Graphics with ggplot2, Line Graphs, Bar Charts, Scatterplots, Boxplots, Histograms.
	Requirements:	Reading (before class) –Chapter 7, R for Everone OR Chapter 3- Optional textbook (Galit Shmueli,…) **Assigned Homework #3**
Week 6	**Date:** Topic:	**November 11-12, 2020** *Evaluating Classification & Predictive Performance:* Mean Absolute Error, Average Error, Mean Absolute Percentage Error, Root-Mean-Squared-Error, Total Sum of Squared Errors, Confusion Matrix, Accuracy, Error Rate, Sensitivity, Specificity, Precision, ROC Curve.
	Requirements:	Reading (before class) – Chapter 1.6.5, Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları, (Balaban & Kartal, 2018) OR Chapter 5- Optional textbook (Galit Shmueli, …) **Homework #3 is due**
Week 7	**Date:** Topic:	**November 18-19, 2020** *Descriptive Analytics and Clustering* (Unsupervised Learning) Correlation and Covariance, Exploring the data (Average, Median, Minimum, Maximum, Standard deviation, Counts & percentages), Correlation Matrix. K-means
	Requirements:	Reading (before class) –Chapter 18, R for Everone **Assign Homework #4**
Week 8	**Date:** Topic: Requirements:	**November 25-26, 2020** **Midterm Exam on Wednesday** Textbooks, notes, laptops
Week 9	**Date:** Topic:	**December 2-3, 2020** *Regression-based Forecasting (Numeric)* Simple and Multiple Linear Regression
	Requirements:	**Homework #4 is due** **Assigned term project**

Week 10	Date:	December 9-10, 2020
	Topic:	*Predictive Analytics (Supervised Learning):*
		k-Nearest Neighbor Algorithm
	Requirements:	**Assign case study #1,** Chapter 4.1, Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları, (Balaban & Kartal
Week 11	Date:	December 16-17, 2020
	Topic:	*Predictive Analytics (Supervised Learning):*
		Naive Bayes Classifier
	Requirements:	**Case study #1 is due**
		Assign case study #2, Chapter 4.2, Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları, (Balaban & Kartal, 2018)
Week 12	Date:	December 23-24, 2020
	Topic:	*Predictive Analytics (Supervised Learning):*
		Logistic Regression
	Requirements:	**Case study #2 is due**
		Assign case study #3, Chapter 4.3, Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları, (Balaban & Kartal, 2018)
Week 13	Date:	December 30-31, 2020
	Topic:	*Predictive Analytics (Supervised Learning):*
		Decision Trees
	Requirements:	**Case study #3 is due**
		Assign case study #4, Chapter 4.3, Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları, (Balaban & Kartal, 2018)
Week 14	Date:	January 6-7, 2021
	Topic:	*Predictive Analytics (Supervised Learning):*
		Artificial Neural Networks.
	Requirements:	**Case study #4 is due**
Week 15	Date:	January 13, 2021
	Topic:	Term Project (Individual)
	Requirements:	**Term Project Final Report**