BA in Management Program
Fall 2021
OPIM407 – Advanced Business Analytics

Instructor: M. Erdal Balaban, PhD.
Office: YBF xxxx
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	13:40 - 16:30	M	Online (Senchronize class)

Course Objective:

The main objective of this course is to prepare students to define and analyze data and to use advanced statistical learning methods (supervised & unsupervised methods) to make a decision. 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. Predictive data mining algorithms will be covered and business cases and problems will be solved with R. It is aimed to give students R's environment and the habits of problem solving and project development with a systematic approach named CRoss-Industry Standard Process for Data Mining (CRISP-DM).

Learning Outcomes:

Students who successfully complete this course are expected to be able to:

1. Understand business analytics, business intelligence, data analytics , data mining, big data and data science concepts,
2. Use R language commands as business analytics and use visual techniques to explain data and make better business decisions,
3. Use data mining (supervised and unsupervised learning algorithms) in solving real-life business problems,
4. Explain the assumptions of various algorithms such as k-Mean Clustering, Regressions, k-Nearest Neighbor, Naive Bayes Classifier, Decision Trees, Random Forest, Neurol Networks, Ensembling Learning, Association Rules.
5. Gain ability to write a comprehensive and well-organized report and prepare projects for real-world case study.

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/)

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 Alternative Cases

Case 1	Prediction of customer's situation (churn or not). https://www.kaggle.com/blastchar/telco-customer-churn
Case 2	Finding Bank Churners https://www.kaggle.com/sakshigoyal7/credit-card-customers
Case 3	Credit card Fraud https://www.kaggle.com/mlg-ulb/creditcardfraud
Case 4	İstanbul Stock Exchange https://archive.ics.uci.edu/ml/datasets/ISTANBUL+STOCK+EXCHANGE
Case 5	*Tennis Major Tournament Match Statistics* https://archive.ics.uci.edu/ml/datasets/Tennis+Major+Tournament+Match+Statistics OR any data related to BA as a case study can be used on web-sites: https://www.kaggle.com/datasets https://archive.ics.uci.edu/ml/datasets.php

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

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, 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:

Attandence & Participation	: 10%
Quizes & Homework Assignments	: 15%
Case Assignment	: 20%
Midterm Exam	: 25%
Term Project	: 30%

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:**	**September 27, 2021**
	Topic:	***Course Introduction****:* Summurizing course outline and rules. What Is Business Analytics and Business Intelligence?, What Is Data Mining?, Data Mining and Related Terms, Big Data, Data Science. CRoss-Industry Standard Process for Data Mining (CRISP-DM).
	Requirements:	Reading (before class) –Chapter 1, Optional textbook (Galit Shmueli, …) and Literature Review
Week 2	**Date:**	**October 3, 2021**
	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 and questions to discuss (before class),– Chapter 1-2-3, R for Everyone, OR Chapter 3, Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları, (Balaban & Kartal, 2018) **Assigned Homework #1**
Week 3	**Date:**	**October 10, 2021**
	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 and questions to discuss (before class) – Chapter 4-5-6, R for Everone **Homework #1 is due** **Assigned Homework #2**
Week 4	**Date:**	**October 17, 2021**
	Topic:	**Overview of the Data Mining Process & Perpormance Evaluation** 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 and questions to discuss (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:**	**October 24, 2021**
	Topic:	***Visualization:*** Base Graphics with ggplot2, Line Graphs, Bar Charts, Scatterplots, Boxplots, Histograms.
	Requirements:	Reading and questions to discuss (before class) –Chapter 7, R for everyone, OR Chapter 3- Optional textbook (Galit Shmueli,…) **Assigned Homework #3**
Week 6	**Date:**	**October 31, 2021**
	Topic:	***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 and questions to discuss (before class) –Chapter 18, R for Everone
Week 7	**Date:**	**November 7, 2021**
	Topic:	***Regression-based Forecasting (Numeric)*** Simple and Multiple Linear Regression and Evaluating Performance (Mean Absolute Error, Average Error, Mean Absolute Percentage Error, Root-Mean-Squared-Error)
	Requirements:	**Homework #3 is due**

Week 8	**Date:**	**November 14, 2021**
	Topic:	**Midterm Exam on Monday (multiple choice questions, R Scripts and BA concepts)**
	Requirements:	Open textbooks, notes, laptops for midterm exam. **Assign term project (Group Study).**
Week 9	**Date:**	**November 21, 2021**
	Topic:	***Predictive Analytics (Supervised Learning):*** k-Nearest Neighbor Algorithm *and Evaluating Predictive Performance* (Confusion Matrix, Accuracy, Error Rate, Sensitivity, Specificity, Precision, ROC Curve).
	Requirements:	Reading and questions to discuss (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, …) **Assigned Case Study** (Individual)
Week 10	**Date:**	**November 28, 2021**
	Topic:	***Predictive Analytics (Supervised Learning):*** Naive Bayes Classifier
	Requirements:	**Term Project Group Members (max. 4 students)**
Week 11	**Date:**	**December 5, 2021**
	Topic:	***Predictive Analytics (Supervised Learning):*** Decision Trees and Random Forest
	Requirements:	Reading and questions to discuss (before class) **Assigned Homework #4**
Week 12	**Date:**	**December 12, 2021**
	Topic:	***Predictive Analytics (Supervised Learning):*** Neural Networks
	Requirements:	Reading and questions to discuss (before class) **Homework #4 is due**
Week 13	**Date:**	**December 19, 2021**
	Topic:	***Predictive Analytics (Unsupervised Learning):*** Association Rules
	Requirements:	Reading and questions to discuss (before class) **Case Study Report is due**

Week 14	**Date:**	**December 26, 2021**
	Topic:	**Ensembling Learning** Majority Voting, Bagging, Stacking,Boosting
	Requirements:	Reading and questions to discuss (before class)
		January 10, 2022 Final Exam (Project Presentations & Final Report Submission) Term Project Presentations, final report