BUSINESS INTELLIGENCE AND ANALYTICS

IE University

Semester: 2º
Category: OPTIONAL
Number of credits: 6.0
Language: English

PREREQUISITES

A basic understanding of mathematics and statistics, e.g., how to calculate a mean value, what is a p-value, etc, and a basic understanding of software programming concepts, e.g., data types, loop structures, if conditions, will be useful for this course.

SUBJECT DESCRIPTION

The interaction between companies and their customers has changed dramatically in recent years. Customers and prospective customers want to interact with companies on their own terms, and a customer's continued business is no longer a guarantee. Thus, companies have realized that they need to understand their customers better and to be able to respond to various customer needs in a timely fashion. Business analytics is the use of information technologies for analyzing data to help managers make better decisions about their business and how they serve customers. The innovative use of business analytics technologies forms a powerful basis for competitive advantage in today's networked economy. The purpose of this course is to explain how data-driven business analytics technologies, such as data mining, can help in many important business applications. This course begins by covering these topics at a basic, fundamental level for those who have little or no experience with these technologies. This course builds on this foundation to provide a comprehensive exploration of a variety of data analytics techniques.

The course will be divided into 3 main blocks:

- Data mining Process and Data Requirements and Visualization
- Fitting Models to Data and Models Evaluations
- Other Data mining problems and techniques

OBJECTIVES AND SKILLS

Understand and learn about various exploratory, predictive, and causal data analytic techniques, such as classification and regression trees, clustering, nearest neighbor approaches, association rule mining, etc. Learn how the above business analytics techniques are applied in a variety of business applications and organizational settings. Understand the process of introducing data analytics technologies into the business environment.

All of this includes:

- Building the case for business analytics technologies.
- Collecting relevant data and visualizing different types of data.

- Learning different techniques and tools for business data analytics
- Using R for data mining and analysis
- Applying data analytics techniques to various business problems.

METHODOLOGY

The course will progress alternating lectures and labs. In lecture sessions, I will cover the theory behind different techniques and algorithms. In the lab sessions, we will walk through demonstrations of how to apply and evaluate the output from the techniques and algorithms, using R and it also gives the opportunity for you to apply what you have learnt in class into practice.

Teaching methodology	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	20.0 %	30 hours
Discussions	3.33 %	5 hours
Exercises	20.0 %	30 hours
Group work	26.67 %	40 hours
Other individual studying	30.0 %	45 hours
TOTAL	100.0 %	150 hours

PROGRAM

SESSION 1

Course Introduction and Introduction to Business Analytics

- Introduction to the course
- Prediction & Causality
- Data Mining Process
- Types of Machine Learning problems & some applications

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R

(Chapter 1 & 2) (See Bibliography)

SESSION 2

(Continue with session 1)

SESSION 3

Data & Data Visualization I

- Data types, data collection, and data quality
- Other data-related issues: data privacy and data bias

SESSION 4

Lab: Introduction to R

- Introduction to R, importing data, data cleaning, and data summary
- You should endeavor to install R & Rstudio before our first class meeting. However, if you

encounter problems, this will be a good time to mention the issues you are having to me. I will use this time to also walk you through the R Studio interface.

SESSION 5

Data & Data Visualization II
- Common Data Visualization Tools

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 3) (See Bibliography)

SESSION 6

Lab: Data Visualization II
- Basic R Data Visualization

SESSION 7

Complete the interactive case.

Technical note & tutorials: Machine Learning (IST020075-U-ENG-HTM)

SESSION 8

Fitting models to data: Classification I
- Classification Tree
- Ensemble Method: Random Forest, Bagging

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 9 & 13) (See Bibliography)

SESSION 9

Fitting models to data: Classification II
- We will continue to explore other cassification methods such as logistic regression and K-Nearest Neighbours

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 7 and 10) (See Bibliography)

SESSION 10

Lab: Classification I
- Classification Tree in R

SESSION 11

Lab: Classification II
- How to apply different classification methods in R such as logistic regression and K-nearest Neighbours

SESSION 12

Fitting models to data: Numeric Prediction

- Overview of Numeric prediction
- Basic numeric prediction methods: regression, regression tree, k-NN Regression

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 6, 7 and 9) (See Bibliography)

SESSION 13

Lab: Numeric Prediction.
Quiz 1

SESSION 14

Avoid Overfitting and Model Performance Evaluation

- An overview of model fit and evaluation
- Generalization and overfit
- Holdout and Cross-Validation
- Evaluation Measurements: Accuracy, AUC, etc.

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 5) (See Bibliography)

SESSION 15

Continue with session 14

SESSION 16

Lab: Model Fit & Evaluation 1

- Apply Evaluation Methods in R

SESSION 17

Lab: Model Fit & Evaluation 2

- Practice different Evaluation Methods in R
- Submit your Team for the group assignment project

SESSION 18

Cluster Analysis

- Clustering Methods: k-means clustering, hierarchical clustering, and their applications

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 15) (See Bibliography)

SESSION 19

Lab: Cluster Analysis

- Practice Clustering problems in R

SESSION 20

Neural Network

   - An overview of Neural Network analysis and its application

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 11) (See Bibliography)

SESSION 21

Lab: Neural Network

   - Practice Neural Network Problems in R.

SESSION 22

Text Mining

Basic of Working with Textual Data and Analysis

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 20) (See Bibliography)

SESSION 23

Lab: Text Mining

   - Practice Text Mining Problems in R

Quiz 2

SESSION 24

Association Rules

   - Association Rules algorithm and its application

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 14) (See Bibliography)

SESSION 25

Group Project Feedback Session

SESSION 26

Social Network Analysis

   - Network data and representation, Social Network Analysis, and its application

Book Chapters: Data mining for business analytics : concepts, techniques, and applications in R (Chapter 19) (See Bibliography)

SESSION 27

Other advanced topics

   - Other advanced topics: Big data

SESSION 28

Complete the interactive case

Technical note & tutorials: Big Data (IST020061-U-ENG-HTM)

SESSION 29

Exam Review
We will spend the next-to-last session on exam preparation. We will go through exam formats and some practice problems.

SESSION 30

FINAL EXAM

The final exam will be worth 30% of your final grade. This is a closed book, closed notes, paper, and pencil exam. You are allowed/encouraged to bring a calculator. However, you are not allowed to use/turn on your smartphone or laptop at any time during the exam, e.g., as a calculator. The exam will take place in the regular class location, at the regularly scheduled time for Session 30.

BIBLIOGRAPHY

Compulsory

- Galit Shmueli, Peter C. Bruce, Inbal Yahav, Nitin R. Patel, Kenneth C.
Lichtendahl Jr.. (2017). *Data mining for business analytics : concepts, techniques, and applications in R.* Wiley. ISBN 9781118879337 (Digital)

EVALUATION CRITERIA

The evaluation of this class includes:

Criteria	Percentage	Comments
Individual Work	10 %	lab exercises completion
Workgroups	30 %	
Final Exam	30 %	
Class Participation	10 %	Class attendance
Intermediate Tests	20 %	2 quizzes distributed during the course

RE-SIT / RE-TAKE POLICY

Each student has four (4) chances to pass any given course distributed over two (2) consecutive academic years. Each academic year consists of two calls: one (1) ordinary call (during the semester when the course is taking place); and one (1) extraordinary call (or "re-sit") in June/July.

Students who do not comply with the 70% attendance requirement in each subject during the semester will automatically fail both calls (ordinary and extraordinary) for that Academic Year and have to re-take the course (i.e., re-enroll) during the next Academic Year.

The Extraordinary Call Evaluation criteria will be subject to the following rules:

- Students failing the course in the ordinary call (during the semester) will have to re-sit evaluation for the course in June / July (except those students who do not comply with the attendance rule, and therefore will not have that opportunity, since they will fail both calls and must directly re- enroll in the course during the next Academic Year).

- It is not permitted to change the format nor the date of the extraordinary call exams or deadlines under any circumstance. All extraordinary call evaluation dates will be announced in advance and must be taken into consideration before planning the summer (e.g. internships, trips, holidays, etc.)
- The June/July re-sit will consist of a comprehensive evaluation of the course. Your final grade for the course will depend on the performance in this exam or evaluation only. I.e., continuous evaluation over the semester (e.g. participation, quizzes, projects and/or other grade components over the semester) will not be taken into consideration on the extraordinary call. Students will have to achieve the minimum passing grade of 5 and the maximum grade will be capped at 8.0 (out of 10.0) – i.e., "notable" in the extraordinary call.
- Re-takers: Students who failed the subject on a previous Academic Year and are now re-enrolled as re-takers in a course will need to check the syllabus of the assigned professor, as well as contact the professor individually, regarding the specific evaluation criteria for them as re-takers in the course during that semester (ordinary call of that Academic Year). The maximum grade that may be obtained as a retaker during the ordinary call (i.e., the 3 rd call) is 10.0 (out of 10.0).

After exams and other assessments are graded by the professor (on either the ordinary or extraordinary call), students will have a possibility to attend a review session (whether it be a final exam, a final project, or the final overall grade in a given course). Please be available to attend the session in order to clarify any concerns you might have regarding your grade. Your professor will inform you about the time and place of the review session.

! Students failing more than 18 ECTS credits after the June/July re-sits will be asked to leave the Program. Please, make sure to prepare yourself well for the exams in order to pass your failed subjects.

! In case you decide to skip the opportunity to re-sit for an exam or evaluation during the June/July extraordinary call, you will need to enroll in that course again for the next Academic Year as a re-taker, and pay the corresponding tuition fees. As you know, students have a total of four (4) allowed calls to pass a given subject or course, in order to remain in the program.