UNIVERSITÀ CATTOLICA del Sacro Cuore

Applied statistics and big data (business analytics)

PROF. GABRIELE CANTALUPPI; PROF. CORRADO ROSSI

Course aims and intended learning outcomes

Make participants able to implement the main statistical learning methods, used to analyse the enormous amounts of data kept in various corporate repositories, in order to support decision-making processes.

Particular attention will be paid to methodological and practical aspects, through illustration of real cases with the aid of tools widely used in the corporate world.

On successful completion of the course participants will be expected to possess:

1. knowledge of concepts, terms and methods of the most used statistical learning and business analytics techniques and grasp of their strengths and weaknesses (DD1- Knowledge and understanding);
2. ability to correctly apply statistical learning methods and business analytics techniques to real economics and management problems (DD2- Applying knowledge and understanding);
3. quantitative thinking addressed to make independent judgements, driven by application of statistical learning methods and business analytics techniques (DD3- Making judgements);
4. ability to present statistical learning and business analytics arguments and the conclusions from them, by means of the extraction of qualitative information from data, with clarity and accuracy and in forms that are suitable for the audiences being addressed, both orally and in writing (DD4-Communication);
5. mastery of statistical learning methods and business analytics techniques, rigorous reasoning and data-driven decision-making, useful for quantitative analyses in other courses of the curriculum, as well as for analyses required in careers in all fields involving management of data (DD5- Lifelong learning skills).

Course content

MODULE I: *Statistical modelling* (Prof. Gabriele Cantaluppi)

1. Introduction to Statistical Learning. Assessing model accuracy.
2. Resampling methods. Cross-validation and bootstrap.
3. Variable selection in regression problems. Best subset selection and stepwise selection.
4. Shrinkage methods. The ridge regression and the lasso. Application: predicting payment defaults of a credit card company.
5. Tree based methods. Regression trees and classification trees. Applications: explaining car seats sales; explaining housing values.
6. The PLS algorithm. Application: customer satisfaction analysis in the mobile phone industry.

MODULE II: *Leveraging Analytics and Big Data to build competitive advantage in the Marketplace* (Prof. Corrado Rossi)

1. Set the stage: strategy and approach to beat the competition using data insights.
2. Transforming business decisions, from emotions to data driven organization.
3. Real Cases: examples of usage of analytics and big data in the world of business in different industries (retail, banking, insurance, …).
4. Business value from the dark data: cognitive computing.
5. Developing a Roadmap to apply insights to improve business performance.

UNIVERSITÀ CATTOLICA del Sacro Cuore

Reading list

For the I module

G. JAMES-D. WITTEN-T. HASTIE-R. TIBSHIRANI, An Introduction to Statistical Learning, Springer, New York, 2015, http://www-bcf.usc.edu/~gareth/ISL.

L. FAHRMEIR-TH. KNEIB-S. LANG-B. MARX, Regression. Models, Methods and Applications, Springer, New York, 2013.

G.A. MARCOULIDES (EDITED BY), Modern Methods for Business Research, Lawrence Erlbaum Associates, London, 1998 (chapter 10).

For the II module

T. DAVENPORT-J. HARRIS, Competing on Analytics, Harvard Business Review Press, Boston Massachusetts, 2007.

Advised texts

G. CANTALUPPI, An Introduction to R, EDUCatt, 2016.

G. CANTALUPPI, Computational Laboratory for Economics with R, EDUCatt, 2016.

Teaching method

Lectures integrated with illustrations of business, economic and financial real cases.

Assessment method and criteria

Module I: A written exam consisting of three open-ended questions on the statistical modelling subjects (final weight 40/60). Available time: 60 minutes.

Module II: Written exam in which the students are requested to develop an approach and a solution to manage the improvement of business performances for a company, leveraging applied statistics and big data (final weight 20/60). Available: 60 minutes.

Aim of the exam is to assess reasoning analytic abilities on the course subjects. Language properties and presentation communication abilities are also assessed.

Notes and prerequisites

Students enrolling in this course should ideally have a basic understanding of Statistics with regard to data analysis, probability and inference, at the level of the combined courses Statistics/‘Statistica I’ and Applied Statistics/‘Statistica applicata’ taught in this University, and a basic knowledge of the software R.