Mondragon Unibertsitatea

Bachelor's degree
Business Data Analytics

Subject

Fundamentals of Mathematics and Statistics II

Year: 2 **Credits**: 12 ECTS **Language**: Spanish

Competencies

Core competences:

CB1. Students have demonstrated knowledge and understanding in an area of study that builds on the foundation of general secondary education and is usually at a level that, while relying on advanced textbooks, also includes certain elements involving cutting-edge knowledge in their field of study

CB2. Students are able to apply their knowledge to their work or vocation in a professional manner and possess the skills that are typically demonstrated through the development and defence of arguments and problem solving within their field of study.

General competences:

CG1. Analytical thinking and ability to respond to complex challenges.

Specific competences:

CE3. Have statistical and mathematical skills applied to data science.

Learning outcomes

RA1. Be able to analyse, logically and from different perspectives and disciplines, the complex challenges they face. To this end, identify which are the key and important aspects of the problem, analyse them and, after a critical and reasoned consideration, propose, in a reasoned manner, different creative and transformative alternatives/solutions.

RA10. Have statistical and mathematical skills applied to data science.

Syllabus

Time series

- Introduction
- White noise (WN) and random walk (RW)
- Autocorrelation function (ACF)
- Autoregression Models (AR) and Moving Averages (MA)
- ARMA models
- ARIMA models
- VAR model

Linear regression estimation methods

- Ordinary Least Squares
- Maximum Likelihood
- Gradient Descent
- Goodness-of-fit

General linear model estimation method

- Ordinary Least Squares

Logistic regression estimation method

- Iterative Reweighted Least Squares

Linear discriminant analysis estimation method

- A priori probability method (Bayes theorem)

Support vector machine estimation method

- Maximisation of the distance around the hyperplane separating the classes (Lagrange Multipliers)

Linear Algebra

- Singular Value Decomposition (SVD)
- Principal Component Analysis (PCA)

Probability theory

- Discrete probability
- Conditional discrete probability

Information theory

- Entropy Concept
- Information gain
- Information ratio
- Gini Impurity
- Consistency

Discretisation of continuous variables

- Chi-Merge
- MDLP

- CAIM

Algorithms based on decision trees (regression and classification)

- CHAID
- ID3
- CART
- C4.5
- C5.0
- Joint learning methods (ensemble learning: bagging, boosting, stacking)

Variable selection

- Filter
- Wrapper
- Embedded

Optimisation

- Introduction to Combinatorial Optimisation
- Local search algorithms
- Population-based algorithms
- Evaluation of optimisation algorithms
- Introduction to multi-objective optimisation

Introduction to neural networks

- Perceptron

Training activities

The training activities planned for this module are the following:

- Challenge-based learning (3 ECTS)
- Teamwork (1 ECTS)
- Workshops (4 ECTS)
- Online resources (0.5 ECTS)
- Reflection (0.5 ECTS)
- Individual work (3 ECTS)

Assessment system

Assessment will be by means of the continuous assessment system, providing constant feedback to both teachers and students on the learning process throughout the academic period:

- Learning activities involving the presentation of knowledge and individual study may be assessed by means of oral and/or written tests, which will account for a maximum of 60% of the final mark.
- The training activities aimed at acquiring the practical skills of the subjects will be assessed through the completion of various activities (assignments, case studies,

challenges, etc.) accounting for at least 40% of the final mark.

Details of the assessment and marking will be made explicit in the annual academic planning of the subjects, in accordance with the teachers responsible and the determining factors of each course.

Bibliography

- Peña, D. (2010). Análisis de Series Temporales, Alianza.
- Luis M. Merino, Evangelina Santos (2006), Álgebra Lineal con Métodos Elementales, Paraninfo.
- David J. C. MacKay (2003), *Information Theory, Inference and Learning Algorithms*, *Cambridge University Press*.
- A. Diaz (1996), Optimización Heurística y Redes Neuronales en Dirección de Operaciones e Ingeniería, Paraninfo.
- C. R. Reeves (1993), Modern Heuristic Techniques for Combinatorial Problems, Blackwell Scientific Publications.
- Enrique J. Carmona, Severino Fernández (2019), Fundamentos de la Computación Evolutiva, Marcombo.
- Aurelien Geron (2020), Aprende *Machine Learning* con *Scikit-Learn*, *Keras* y *Tensorflow*, Anaya Multimedia.
- François Chollet (2017), *Deep Learning with Python*, *Manning Publications*.