

DATA SCIENCE FUNDAMENTALS

Extracting the value that can be hidden in data to support decision-making processes. Challenges include analysis, capture, data curation, search, sharing, storage, transfer, visualization, querying, updating and information privacy. The course will review the different phases of the project to extract knowledge from databases with special emphasis on the extraction of predictive models.

➤ **Course duration: 10 contact hours**

➤ **Dates: 19, 20 & 21 July 2023**

■ **Recommended background knowledge for students:**

- Data Bases
- Basics of Programming

Meet Our INSTRUCTORS



Dr. Ernestina Menasalvas is full professor of data mining at the Department of Computer Systems Languages and Software Engineering, School of Computer Science, Universidad Politécnica de Madrid (UPM). She holds a PhD in Computer Science. She is currently a member of the MIDAS Data Mining and Group at the Center for Biotechnology and databases and data mining professor at UPM. Her research interests include data mining project development, recently focusing on data mining in the medical field.



Dr. Alejandro Rodríguez-González is associate professor at the Department of Computer Languages and Systems and Software Engineering at the Universidad Politécnica de Madrid and he holds a PhD in Computer Science. His main research interests are Semantic Web, Artificial Intelligence and the Biomedical informatics field, especially the development medical diagnosis systems, medical knowledge representation, the extraction of knowledge from different sources (text, social media, etc.) and the analysis of social media impact on e-Health, and he has participated in competitive projects in related fields.

LEARNING GOALS

■ **Learn**

Project development methodologies



Transform data into knowledge

■ **Understand**

Basic AI methods



Supervised and unsupervised modelling, data preparation and comprehension, validation



■ **Gain**

Insight

Practical applications of the technology in the real world

SYLLABUS

MODULES

01

Session I

1. Introduction to Big Data Analytics and the Cognitive Era
2. The Big Data Value Chain
3. Methodologies for Data Analytics project development
4. Data Science projects: phases and tasks

Session II

1. Type of Data Mining problems:
 - a. Clustering,
 - b. association,
 - c. classification
2. Tools and techniques

02

Students will be given a dataset and they will have to extract patterns out of it by using some data analytics tool (R, WEKA, Knime, ...)

03

1. Industry talk
2. Question and answer session

Meet Our INDUSTRY SPEAKER



Dr. Juan Miguel Gómez Berbís is associate professor at the Department of Computer Science in the Universidad Carlos III de Madrid. He received his PhD from the Digital Enterprise Research Institute (DERI) at National University of Ireland, Galway. He received his MBA from IE Business School, an MS in Software Engineering from the Swiss Federal Institute of Technology (EPFL) in Lausanne (Switzerland) and an MS in Telecommunications Engineering from the Universidad Politécnica de Madrid (UPM). He is now leading H2020 initiatives and ITEA3 major flagship projects.