PRACTICE ANALYTICS (EMaCS-03-08)				
DEGREE PROGRAM:		Master in Computer Science for the Human-Centric and Sustainable Industry		
SEMESTER: Third	**TYPE:** Basic	**CREDITS:** 7.5 ECTS	**WORKLOAD:** 187.5 hours	**MENTORING:** 4 hours/week
LANGUAGE: English				

OBJECTIVES	
General	• To understand the roles of different jobs in data science and related areas • To understand the methodologies and steps of a complete cycle of a Data Analysis project • To apply the method through a project, from raw data to deployment
Specific	• Understand the key stages of the life cycle of a data analysis process • Understand the importance of following an appropriate methodology in a data analysis process • Follow CRISP-DM method in a data analysis process with real data • Apply pre-processing methods and data visualization methods to raw data • Select, adapt and apply machine learning techniques to create detection, classification or prediction models • Evaluate the quality of models created, qualitatively and quantitatively • Prepare deployment of the models

SUSTAINABILITY

The course "Practice Analytics" contributes to sustainability by teaching students to apply efficient and sustainable methodologies in data analysis projects. By following the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology, students learn to structure their projects effectively with a lower environmental impact. Additionally, addressing the scalability of Analytics to Big Data and data stream management encourages efficiency and sustainable handling of large datasets. Data analysis projects can also be directed towards sustainability-related issues, such as energy efficiency, resource management, or data-driven decision-making for more sustainable practices.

RESILIENCE AND HUMAN-CENTRIC DEVELOPMENT

The "Practice Analytics" course contributes to resilience and human-centric development by providing students with practical skills and specific knowledge required for roles in the field of data science. By understanding and applying different methodologies in data analysis projects, students develop critical skills to address challenges and adapt to changing demands in the field. Training in team and project management also contributes to human-centric development by fostering communication and teamwork skills, fundamental aspects for success in professional environments and for creating solutions that address the needs of individuals and communities.

SUBJECT MATTER

1. Understand the main stages of the data analysis process;
2. Follow the CRISP-DM methodology in a real data analysis process;
3. Select, adapt and apply learning techniques to create classification, detection or prediction models;
4. Critically and quantitatively assess the quality of the deployed models;
5. Scale Analytics to Big Data;
6. Data streams;
7. Data sharing.

COMPETENCES

C2. BROWSING, SEARCHING AND FILTERING DATA, INFORMATION AND DIGITAL CONTENT
C3 MANAGING AND EVALUATING DATA, INFORMATION AND DIGITAL CONTENT
C4. INTEGRATING AND RE-ELABORATING INFORMATION and DIGITAL CONTENT
C6 USING MACHINE LEARNING AND A.I. TECHNIQUES
C12 IDENTIFYING NEEDS AND TECHNOLOGICAL RESPONSES
C15 MANAGING SYSTEMS and/or PROJECTS

<table>
<tr><th colspan="2">LEARNING OUTCOMES</th></tr>
<tr><td>Knowledge</td><td>• Know the different types of jobs in data science.
• Know different methodologies to apply in data science projects.
• Know adequate methods to deploy machine learning models.</td></tr>
<tr><td>Skills</td><td>• Develop the capacity to organise and manage teams for data science projects.
• Develop the capacity to correctly apply CRISP-DM to successfully complete data science projects.</td></tr>
<tr><td>Attitudes/values</td><td>• Develop and practice communication and team work skills.</td></tr>
<tr><th colspan="2">TEACHING METHODS</th></tr>
<tr><td colspan="2">• Presentation and discussion of the topics.
• Presentation and discussion of assignments and examples.
• Invited talks.
• Data science project with raw data, from business understanding to deployment.</td></tr>
<tr><th colspan="2">EVALUATION</th></tr>
<tr><td colspan="2">Evaluation is mostly practical, divided as follows:
• 5% - Class attendance
• 5% - Quiz in class
• 90% - Project, with application of CRISP-DM methodology divided into five deliverables:
o 15% - Business understanding
o 20% - Data understanding
o 20% - Data preparation
o 20% - Modelling
o 15% - Evaluation and Deployment
Students can submit each deliverable only once.</td></tr>
<tr><th colspan="2">PRECONDITIONS</th></tr>
<tr><td colspan="2">None</td></tr>
<tr><td>DEPARTMENT</td><td>Department of Informatics Engineering and Systems</td></tr>
<tr><td>LECTURERS</td><td>Mateus Mendes</td></tr>
<tr><td>LITERATURE</td><td>• Data Science (The MIT Press Essential Knowledge series), John D. Kelleher and Brendan Tierney, MIT Press 2018, ISBN-13: 978-0262535434.
• Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2, Sebastian Raschka and Vahid Mirjalili, Packt Publishing 2019; 3rd edition.
• Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning: The MIT Press, 2016, 800 pp, ISBN: 0262035618
• John Kelleher, Brian Mac Namee, and Aoife D’Arcy; Fundamentals of Machine Learning for Predictive Data Analytics; The MIT Press, 2015, ISBN: 9780262044691
• Aurélien Géron; Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow; O'Reilly Media 2019; ISBN-13 : 978-1492032649
• Ian Pointer; Programming PyTorch for Deep Learning: Creating and Deploying Deep Learning Applications 1st Ed.; O'Reilly Media 2019.
• Other resources available online, namely python manuals, sk-learn, tensorflow, pytorch, keras and onnx.</td></tr>
</table>