COURSE DESCRIPTION – ACADEMIC YEAR 2021/2022

Course title	**Data Mining and Decision Making**
Course code	76417
Scientific sector	INF/01
Degree	Bachelor in Informatics and Management of Digital Business (L-31)
Semester	1+2
Year	3
Credits	12
Modular	Yes

Total lecturing hours	80
Total lab hours	40
Attendance	Attendance is not compulsory, but non-attending students have to contact the lecturers at the start of the course to agree on the modalities of the independent study.
Prerequisites	Basic programming concepts
Course page	https://ole.unibz.it/ and https://teams.microsoft.com/

Specific educational objectives	The course belongs to the type "caratterizzante - informatica". The course is designed to acquire professional skills and knowledge useful when exploring datasets. It introduces the whole data mining workflow, from data ingestion to analysis, making insights which are essential for data-driven decision making. **Module 1: Introduction to Data Mining** In this module, the students will learn how to organize and analyze data by writing programs. More specifically, the students will practically learn how to import, manipulate, analyze, visualize, and model a dataset. The students will also get familiar with libraries that can be effectively used for data preparation, mining, analytics, and visualization. The student will be able to get insights from the data and make data-driven decisions, learning how to avoid common pitfalls that can mislead the analysis. These concepts are explored through projects and case studies, using the Python programming language, following the best practices of reproducible research. **Module 2: Data-driven decision making** In this module, students will learn more advanced data mining methods, to make inference on the data and create regression and classification models. Decision theory and human decision-making methods are combined with artificial intelligence, machine learning and deep learning to address data-intensive, data-driven decision making. These concepts are explored through projects and case studies, using the KNIME analytics platform, to manage complex data-intensive scenarios. The students will also learn how to integrate Python and Keras in KNIME.

Module 1	**Introduction to Data Mining**
Module code	76417A
Module scientific sector	INF/01
Lecturer	Antonio Liotta
Contact	Piazza Domenicani 3, antonio.liotta@unibz.it, +390471016049
Scientific sector of lecturer	ING-INF/05
Teaching language	English
Office hours	Wednesdays 10:30-12:30, to be arranged beforehand by email.
Lecturing assistant (if any)	--
Contact LA	--
Office hours LA	--
Credits	6
Lecturing hours	40
Lab hours	20
List of topics	• Introduction to Knowledge Discovery in Data • Programming for Data Science • Data quality and data preparation • Data Mining tasks and algorithms • Methods and techniques for data analysis, visualization and decision support • Projects/Case studies on data-driven decision making
Teaching format	Frontal lectures, lab assignments, project work.

Module 2	**Data-driven Decision Making**
Module code	76417B
Module scientific sector	INF/01
Lecturer	Giuseppe Di Fatta
Contact	TBD
Scientific sector of lecturer	ING-INF/05
Teaching language	English
Office hours	To be arranged beforehand by email.
Lecturing assistant (if any)	--
Contact LA	--
Office hours LA	--
Credits	6
Lecturing hours	40
Lab hours	20
List of topics	• Decision Theory and Human Decision Making • Introduction to Artificial Intelligence • Machine Learning and Deep Learning algorithms • AI frameworks and tools • Ethical and social implications of AI • Projects/Case studies on AI-driven decision making
Teaching format	Frontal lectures, lab assignments, project work.

Learning outcomes	Knowledge and understanding: • D.12 - Know methodologies for data analysis, machine learning and their application to decision making in the business context. Applying knowledge and understanding:

	• D2.3 - Ability to analyse business problems and to develop proposals for solutions with the help of IT tools. • D2.6 - Ability to design, describe and present IT solutions to policy makers. • D2.9 - Ability to support the management of IT departments and software companies by providing information as needed. • D2.11 - Ability to analyse large amounts of data on economic facts and processes. Making judgments • D3.1 - Ability to collect and interpret data useful for forming independent judgments on IT and economic aspects of information systems. Communication skills • D4.5 - Ability to collaborate in interdisciplinary teams to achieve IT objectives. Learning skills • D5.3 - Ability to follow rapid technological developments and to learn about innovative aspects of the latest generation of information technology and systems.

Assessment	The exam modalities are the same for both the attending and the non-attending students. Project work (70% of the final grade) and oral exam (30% of the final grade). All project works must have been submitted, at the very latest, 15 days ahead of the oral exam. In case of a positive mark, the projects will count for all 3 regular exam sessions.
Assessment language	English
Assessment Typology	Collegial
Evaluation criteria and criteria for awarding marks	Project work (70% of the final grade) and oral exam (30% of the final grade). • Relevant for project work: clarity of presentation, ability to gain useful and novel insights from data, creativity, critical thinking, ability to adhere to reproducible research best practices. • Ability to use Python to employ (understand, recall and use) data analytics methods in practical settings, starting from data collection, preparation, exploration tasks and going to coherent and insightful data analysis and visualization. • Ability to employ and choose a range of machine learning methods to make inference on the data and create regression and classification models. • Ability to evaluate machine learning models in the context of data-intensive, data-driven decision making. • Ability to implement and critically evaluate complete data science workflows in KNIME, integrating Python and Keras.

Required readings	Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it • *Introduction to Data Mining*, by Pan-Ning Tang, M. Steinbach, A. Karpatne, V. Kumar. Pearson Education Ltd (2nd Edition, 2020). • *Python Data Science Handbook*, by Jake VanderPlas. O'Reilly Media (1st Edition, 2016).
Supplementary readings	• *Fundamentals of Data Visualization*. Wilke. Available online
Software used	Jupyter Notebook (for Python programing) https://jupyter.org/ KNIME https://www.knime.com/