MATHEMATICS (LM39)

(Lecce - Università degli Studi)

Teaching DATA MINING

GenCod A004898

Owner professor MARCO PULIMENO

Teaching in italian DATA MINING

Teaching DATA MINING

SSD code ING-INF/05

Reference course MATHEMATICS

Course type Laurea Magistrale

Credits 6.0

Teaching hours Front activity hours: 42.0

For enrolled in 2022/2023

Taught in 2022/2023

Course year 1

Language ENGLISH

Curriculum TEORICO-MODELLISTICO

Location Lecce

Semester Second Semester

Exam type Oral

Assessment Final grade

Course timetable https://easyroom.unisalento.it/Orario

BRIEF COURSE DESCRIPTION

The course provides a modern introduction to data mining, which spans techniques, algorithms and methodologies for discovering structure, patterns and relationships in data sets (typically, large ones) and making predictions. Applications of data mining are already happening all around us, and, when they are done well, sometimes they even go unnoticed. For instance, how does the Google web search work? How does Shazam recognize a song? How does Netflix recommend movies to its users? The principles of data mining provide answers to these and others questions. Data mining overlaps the fields of computer science, statistical machine learning and data bases. The course aims at providing the students with the knowldedge required to explore, analyze and leverage available data in order to turn the data into valuable and actionable information for a company, for instance, in order to facilitate a decision-making process.

REQUIREMENTS

Calculus. Probability theory. Linear Algebra. Programming skills.

COURSE AIMS

The course describes methods and models for the analysis of large amounts of data. Students will gain a solid background with a broad spectrum of basic knowledge related to data mining:

- the students will acquire the basic cognitive tools to think analytically, creatively and critically, and have the abstraction and problem-solving skills needed to cope with complex systems;
- they will gain solid knowledge of data mining models and methodologies;
- they will be able to work on large data collections, including heterogeneous and produced at high speed data, in order to carry out in-depth thematic analyses, drawing on this knowledge to improve the decision-making process.

TEACHING METHODOLOGY

Frontal lessons using slides made available to students and classroom exercises

ASSESSMENT TYPE	Oral exam. During the exam the student is asked to illustrate theoretical topics in order to verify his/her knowledge and understanding of the selected topics. The student must demonstrate adequate knowledge and understanding of the issues presented or indicated, applying in a relevant manner the theories and conceptual models covered by the study programme.
REFERENCE TEXT BOOKS	Mining of Massive Datasets J. Leskovec, A. Rajaraman and J. Ullman Freely availableonline: http://www.mmds.org Data Mining and Analysis M. J. Zaki and W. Meira Freely available online: https://dataminingbook.info