SUBJECT TEACHING GUIDE

269 - SEMANTICS, CONNECTED DATA AND TEXT DATA MINING

University Master's Degree in Data Science

Academic year 2023-2024

1. IDENTIFYING DATA					
Degree	University Master's Degree in Data Science			Type and Year	Optional. Year 1
Faculty	Faculty of Sciences				
Discipline	INTELLIGENCE IN DATA SCIENCE				
Course unit title and code	269 - SEMANTICS, CONNECTED DATA AND TEXT DATA MINING				
Number of ECTS credits allocated	4	Term		Semester based (2)	
Web	https://moodle.unican.es/course/view.php?id=12850				
Language of instruction	Spanish	English Friendly	Yes	Mode of delivery	Face-to-face

Department	DPTO. MATEMATICAS, ESTADISTICA Y COMPUTACION
Name of lecturer	DOMINGO GOMEZ PEREZ
E-mail	domingo.gomez@unican.es
Office	Facultad de Ciencias. Planta: + 3. DESPACHO DOMINGO GOMEZ PEREZ (3005)
Other lecturers	JOSE JAVIER RAMASCO SUQUIA ALEJANDRO VILLAR FERNANDEZ FERNANDO AGUILAR GOMEZ

3.1 LEARNING OUTCOMES

- - Understand complementary methods for the analysis of large masses of unstructured data, entering in the field of text mining (and Web mining)
- Understand the fundamentals of the representation and analysis of data with complex networks

4. OBJECTIVES

At the end of the class, the students should be able to apply different methodologies and techniques of automatic learning in a critical way in real problems, including the mining of texts and Web mining. A second objective, of a practical nature, is to provide students with the necessary capabilities and standard tools that would allow them to independently carry out data analytics projects.

6. COURSE ORGANIZATION

CONTENTS	
1	Semantic networks
2	Ontologies and ontology learning
3	Linked data
4	Analysis of complex networks
5	Web and text mining

7. ASSESSMENT METHODS AND CRITERIA

Description	Type	Final Eval.	Reassessn	%
Assessment of reports and written work	Work	No	Yes	50,00
Final exam	Written exam	Yes	No	50,00
TOTAL				100,00

Observations

Observations for part-time students

Part-time students have two options: follow the continuous evaluation (they can answer the questions through moodle) or carry out a project (like their classmates) that they will have to present and defend orally if the number of participants allows it. If they decide not to follow the continuous evaluation, the percentages would change: 60% written work and 40% final exam.

8. BIBLIOGRAPHY AND TEACHING MATERIALS

BASIC

Sholom M. Weiss, Nitin Indurkhya, Tong Zhang, Fred Damerau. Text mining: predictive methods for analyzing unstructured information. Springer Science+Business Media (2005)

Juan Antonio Pastor Sánchez. Tecnologías de la web semántica. UOC (2012)

Toby Segaran, Colin Evans, Jamie Taylor. Programming the semantic web. O'Reilly (2009)