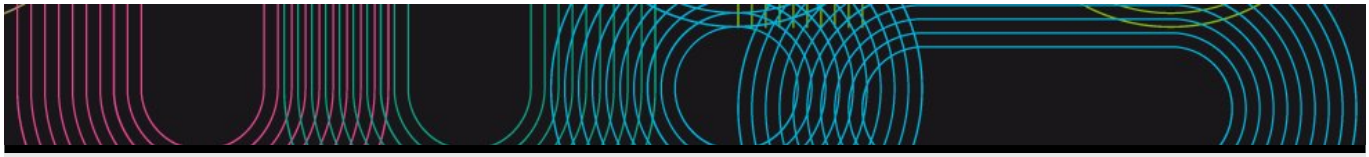

IDENTIFYING DATA				
Natural language understanding				
Subject	Natural language understanding			
Code	O06M193V01104			
Study programme	Máster universitario en Inteligencia artificial			
Descriptors	ECTS Credits	Choose	Year	Quadmester
	6	Mandatory	1st	1st
Teaching language	English			
Department				
Coordinator	Darriba Bilbao, Víctor Manuel			
Lecturers	Darriba Bilbao, Víctor Manuel			
E-mail	darriba@uvigo.es			
Web	http://guiadocente.udc.es/guia_docent/index.php?centre=614&ensenyament=614544&consulta=assignatures&any_academic=2023_24			
General description	The course introduces the basic concepts and techniques associated with natural language processing, the starting point for the design of information exploitation and dialogue environments based on human language, both at the lexical and syntactic, semantic and pragmatic levels. The objective is to introduce the student to the complexity inherent to the analysis of human natural language, mainly associated to the ambiguity and contextual dependencies it presents, and to the design of data structures and algorithms that allow its practical treatment.			

Training and Learning Results	
Code	
A1	CB6 - Possess and understand knowledge that provides a basis or opportunity to be original in the development and/or application of ideas, often in a research context
A2	CB7 - Students should be able to apply their acquired knowledge and problem-solving skills in new or unfamiliar environments within broader (or multidisciplinary) contexts related to their area of study.
A5	CB10 - That students possess the learning skills that will enable them to continue studying in a manner that will be largely self-directed or autonomous.
B1	Maintain and extend sound theoretical approaches to enable the introduction and exploitation of new and advanced technologies in the field of Artificial Intelligence.
B3	Search and select useful information needed to solve complex problems, handling with fluency the bibliographic sources of the field.
B4	Elaborate adequately and with certain originality written compositions or motivated arguments, write plans, work projects, scientific articles and formulate reasonable hypotheses in the field.
C1	Understanding and mastering techniques for text processing in natural language
C2	Understanding and mastery of the fundamentals and techniques of semantic processing of linked, structured, and unstructured documents, and of the representation of their content.
C3	Understanding and knowledge of the techniques of representation and processing of knowledge through ontologies, graphs, and RDF, as well as the tools associated with them.
D2	Master the oral and written expression and comprehension of a foreign language.
D3	Utilizar las herramientas básicas de las tecnologías de la información y las comunicaciones (TIC) necesarias para el ejercicio de su profesión y para el aprendizaje a lo largo de su vida.
D7	Develop the ability to work in interdisciplinary or transdisciplinary teams to offer proposals that contribute to sustainable environmental, economic, political and social development.
D8	Value the importance of research, innovation and technological development in the socioeconomic and cultural progress of society.

Expected results from this subject	
Expected results from this subject	Training and Learning Results

To know, understand and analyze the formal representation of diverse lexical, syntactic and semantic phenomena of natural language.	A1 A5 B1 B3 B4 C1 D2 D8
To know, understand and know how to use the technologies, frameworks and libraries for the construction of natural language processing systems.	A1 A2 A5 B3 B4 C1 C2 D2 D3 D7
To design, implement and know how to use algorithms and data structures to treat and support the various phenomena characteristic of natural language.	A1 A2 A5 B1 B3 B4 C1 C2 C3 D2 D3 D7 D8
To know, understand and analyze natural language processing techniques for processing and disambiguation at the lexical, syntactic and semantic levels.	A1 A2 A5 B1 B3 B4 C1 C2 C3 D2 D3 D7 D8
To know and understand the problems posed by ambiguity and imprecision in natural language data sources and techniques to solve them.	A1 A2 A5 B1 B3 B4 C1 C3 D2 D3 D7 D8

Contents

Topic	
Introduction.	Levels of analysis. Ambiguity and contextual dependencies.
Lexical analysis.	Segmentation. Dictionaries and thesauri. Part-of-speech tagging.

Syntactic parsing.	Algebraic grammars. Mildly context-sensitive grammars. Dependency grammars. Probabilistic grammars.
Semantic parsing.	Lexical semantics. Semantic dependencies. Semantic graphs.

Planning

	Class hours	Hours outside the classroom	Total hours
Lecturing	21	21	42
Laboratory practical	14	48	62
Problem solving	9	25	34
Objective questions exam	3	9	12

*The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.

Methodologies

	Description
Lecturing	Theoretical classes, in which the content of each topic is exposed. The student will have copies of the slides beforehand and the professor will promote an active attitude, asking questions to clarify specific aspects and leaving open questions for the student's reflection.
Laboratory practical	Practical classes with the use of computers, which allow the student to familiarize himself/herself from a practical point of view with the issues presented in the theoretical classes.
Problem solving	Problem-based learning, seminars, case studies and projects.

Personalized assistance

Methodologies	Description
Lecturing	The teachers will attend the students in individualized mentoring sessions, dedicated to the orientation in the study and to the resolution of doubts on the contents, duties and activities of the course.
Laboratory practical	The teachers will attend the students in individualized mentoring sessions, dedicated to the orientation in the study and to the resolution of doubts on the contents, duties and activities of the course.
Problem solving	The teachers will attend the students in individualized mentoring sessions, dedicated to the orientation in the study and to the resolution of doubts on the contents, duties and activities of the course.

Assessment

	Description	Qualification	Training and Learning Results			
Laboratory practical	The delivery of the practicals must be done within the deadline established in the virtual campus and must follow the specifications indicated in the statement for both presentation and defense.	50	A1 A2 A5	B3 B4	C1 C2 C3	D2 D3 D7 D8
Objective questions exam	Compulsory realization. The mastery of the theoretical and operative knowledge of the subject will be evaluated.	50	A1 A2	B1	C1 C2 C3	D2

Other comments on the Evaluation

EVALUATION CRITERIA FOR ALL STUDENTS IN ALL OPPORTUNITIES

Students must achieve at least 40% of the maximum grade for each part (theory, practice) and in any case the sum of both parts must reach a 5 to pass the course. If any of the above requirements is not met, the grade for the course will be established according to the lowest grade obtained.

In case of not reaching the minimum grade in one of the parts, the student will have a second opportunity in which only the delivery of that part will be required.

The delivery of the practicals must be done within the deadline established in the virtual campus and must follow the specifications indicated in the statement for both its presentation and defense.

The student who submits all the compulsory practicals or attends the objective test in the official evaluation period will be considered "Presented".

In the case of fraudulent performance of exercises or tests, the Regulations for the evaluation of students' academic performance and review of qualifications will be applied. In application of the corresponding regulations on plagiarism, the total or partial copy of any practical or theory exercise will result in failure in both opportunities of the course, with a grade of 0.0 in both cases.

EXAM DATES

The official exam dates for the different opportunities, will be published on the ESEI website: https://esei.uvigo.es/docencia/exames/

CONSULTATION/REQUEST OF TUTORING SESSIONS

Tutoring sessions schedules can be consulted through the faculty's personal page, available at https://esei.uvigo.es/docencia/profesorado/

Sources of information

Basic Bibliography

Manning, C., & Schutze, H., **Foundations of statistical natural language processing**, 978-0262133609, 1, MIT Press, 1999

Goldberg, Y., **Neural network methods for natural language processing. Synthesis lectures on human language technologies**, 978-1627052986, 1, Morgan Claypool, 2017

Eisenstein, J., **Introduction to Natural Language Processing**, 978-0262042840, 1, MIT Press, 2019

Jurafsky, D. & Martin, J. H., **Speech and Language Processing**, 978-0131873216, 3 (draft), https://web.stanford.edu/~jurafsky/slp3/, 2022

Jurafsky, D. & Martin, J. H., **Speech and Language Processing**, 978-0131873216, 2, Prentice Hall, 2008

Indurkhya, N. & Damerau, F.J. (Eds.), **Handbook of Natural Language Processing**, 978-1420085921, 2, Routledge, 2010

Complementary Bibliography

Chollet, F., **Keras: The python deep learning library**, Astrophysics Source Code Library, 2018

Russell, S., Norvig, P., **Artificial Intelligence: A Modern Approach**, 978-0134610993, 4, Pearson, 2022

Manning, C.D., Raghavan, P., Schütze, H., **Introduction to Information Retrieval**, 978-0521865715, 1, Cambridge University Press, 2008

Kübler, S., McDonald, R., & Nivre, J., **Dependency Parsing. Synthesis lectures on human language technologies**, 978-1598295962, 1, Morgan Claypool, 2009

Recommendations

Subjects that continue the syllabus

Web intelligence and semantic technologies/O06M193V01205
Language modelling/O06M193V01204
Text mining/O06M193V01302

Subjects that are recommended to be taken simultaneously

Machine learning I/O06M193V01105

Other comments

All students are reminded of the prohibition of the use of mobile devices in exercises, practices and exams, in compliance with article 13.2.d) of the Statute of the University Student, regarding the duties of the university student body, which establishes the duty to "Refrain from using or cooperating in fraudulent procedures in assessment tests, in the work carried out or in official university documents."