Course guide
270708 - HLE - Human Language Engineering

Last modified: 21/07/2022

Unit in charge:	Barcelona School of Informatics
Teaching unit:	723 - CS - Department of Computer Science.

Degree:	MASTER'S DEGREE IN ARTIFICIAL INTELLIGENCE (Syllabus 2017). (Optional subject).
	MASTER'S DEGREE IN DATA SCIENCE (Syllabus 2021). (Optional subject).

Academic year: 2022   **ECTS Credits:** 4.5   **Languages:** English

LECTURER

Coordinating lecturer:	SALVADOR MEDINA HERRERA
Others:	Primer quadrimestre:
	SALVADOR MEDINA HERRERA - 10

PRIOR SKILLS

- Introductory concepts and methods of Natural Language Processing.

- Programming.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CEA3. Capability to understand the basic operation principles of Machine Learning main techniques, and to know how to use on the environment of an intelligent system or service.
CEA5. Capability to understand the basic operation principles of Natural Language Processing main techniques, and to know how to use in the environment of an intelligent system or service.
CEA7. Capability to understand the problems, and the solutions to problems in the professional practice of Artificial Intelligence application in business and industry environment.
CEP3. Capacity for applying Artificial Intelligence techniques in technological and industrial environments to improve quality and productivity.
CEP4. Capability to design, write and report about computer science projects in the specific area of ??Artificial Intelligence.

Generical:
CG1. Capability to plan, design and implement products, processes, services and facilities in all areas of Artificial Intelligence.

Transversal:
CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

CT4. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.

CT6. REASONING: Capability to evaluate and analyze on a reasoned and critical way about situations, projects, proposals, reports and scientific-technical surveys. Capability to argue the reasons that explain or justify such situations, proposals, etc..
CT7. ANALISIS Y SINTESIS: Capability to analyze and solve complex technical problems.

Basic:

CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

CB8. Capability to communicate their conclusions, and the knowledge and rationale underpinning these, to both skilled and unskilled public in a clear and unambiguous way.

CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

TEACHING METHODOLOGY

This course will build on different teaching methodology (TM) aspects, including:

TM1: theoretical lecture sessions
TM2: practical sessions with invited speakers from the industry
TM3: oral presentations of the students

LEARNING OBJECTIVES OF THE SUBJECT

1.Learning the current trends of Human Language Engineering and further challenges.
2.Learning knowledge and tools required to develop Human Language Engineering applications in the selected areas (Information Extraction, Machine Translation and Dialogue Systems), and comparison criteria.
3.Development of criteria to identity problems to be solved using Human Language Engineering.
4.Application of the acquired knowledge to specific real problems.
6.Understanding the potential applications of Human Language Engineering in the business environment.

STUDY LOAD

Type	Hours	Percentage
Hours large group	40,5	36.00
Self study	72,0	64.00

Total learning time: 112.5 h

CONTENTS

Course Introduction

Description:
Presentation of the course: aims, plan and structure.
General overview of the range of applications associated with language engineering. Currents trends.

Information Extraction

Description:
Entity and Relation extraction. Event and Time extraction. Sentiment and Affect extraction. Summarisation.

Machine Translation

Description:
Classical MT. Statistical MT. Resources and models for MT. MT Evaluation.

Dialogue Systems

Description:
Question Answering. Conversational Agents. Chatbots. Virtual Assistants.

ACTIVITIES

Course Introduction

Description:
Presentation of the course: aims, plan and structure.
General overview of the range of applications associated with language engineering. Currents trends.

Specific objectives:
1

Related competencies :
CG1. Capability to plan, design and implement products, processes, services and facilities in all areas of Artificial Intelligence.
CEA3. Capability to understand the basic operation principles of Machine Learning main techniques, and to know how to use on the environment of an intelligent system or service.
CEA7. Capability to understand the problems, and the solutions to problems in the professional practice of Artificial Intelligence application in business and industry environment.
CEA5. Capability to understand the basic operation principles of Natural Language Processing main techniques, and to know how to use in the environment of an intelligent system or service.
CT4. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.

CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

Full-or-part-time: 8h 42m
Theory classes: 2h
Practical classes: 1h
Self study: 5h 42m

Information Extraction

Description:
Entity and Relation extraction. Event and Time extraction. Sentiment and Affect extraction. Summarisation. Considering open and restricted domains. Considering monolingual and crosslingual scenarios.

Specific objectives:
2, 3, 6

Related competencies :
CG1. Capability to plan, design and implement products, processes, services and facilities in all areas of Artificial Intelligence.
CEA3. Capability to understand the basic operation principles of Machine Learning main techniques, and to know how to use on the environment of an intelligent system or service.
CEA7. Capability to understand the problems, and the solutions to problems in the professional practice of Artificial Intelligence application in business and industry environment.
CEA5. Capability to understand the basic operation principles of Natural Language Processing main techniques, and to know how to use in the environment of an intelligent system or service.
CEP3. Capacity for applying Artificial Intelligence techniques in technological and industrial environments to improve quality and productivity.
CT7. ANALISIS Y SINTESIS: Capability to analyze and solve complex technical problems.
CT6. REASONING: Capability to evaluate and analyze on a reasoned and critical way about situations, projects, proposals, reports and scientific-technical surveys. Capability to argue the reasons that explain or justify such situations, proposals, etc..
CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

CT4. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.

CB8. Capability to communicate their conclusions, and the knowledge and rationale underpinning these, to both skilled and unskilled public in a clear and unambiguous way.

Full-or-part-time: 26h
Theory classes: 6h
Practical classes: 3h
Self study: 17h

Machine Translation

Description:
Classical MT. Statistical MT. Resources and models for MT. MT Evaluation.
Text-to-text translation. Speech-to-speech translation.

Specific objectives:
2, 3, 6

Related competencies :
CG1. Capability to plan, design and implement products, processes, services and facilities in all areas of Artificial Intelligence.
CEA3. Capability to understand the basic operation principles of Machine Learning main techniques, and to know how to use on the environment of an intelligent system or service.
CEA7. Capability to understand the problems, and the solutions to problems in the professional practice of Artificial Intelligence application in business and industry environment.
CEA5. Capability to understand the basic operation principles of Natural Language Processing main techniques, and to know how to use in the environment of an intelligent system or service.
CEP3. Capacity for applying Artificial Intelligence techniques in technological and industrial environments to improve quality and productivity.
CT7. ANALISIS Y SINTESIS: Capability to analyze and solve complex technical problems.
CT6. REASONING: Capability to evaluate and analyze on a reasoned and critical way about situations, projects, proposals, reports and scientific-technical surveys. Capability to argue the reasons that explain or justify such situations, proposals, etc..
CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

CT4. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.

CB8. Capability to communicate their conclusions, and the knowledge and rationale underpinning these, to both skilled and unskilled public in a clear and unambiguous way.

Full-or-part-time: 34h 36m
Theory classes: 8h
Practical classes: 4h
Self study: 22h 36m

Dialogue Systems

Description:
Question Answering. Conversational Agents. Chatbots. Virtual Assistants.

Specific objectives:
2, 3, 6

Related competencies :
CG1. Capability to plan, design and implement products, processes, services and facilities in all areas of Artificial Intelligence.
CEA3. Capability to understand the basic operation principles of Machine Learning main techniques, and to know how to use on the environment of an intelligent system or service.
CEA7. Capability to understand the problems, and the solutions to problems in the professional practice of Artificial Intelligence application in business and industry environment.
CEA5. Capability to understand the basic operation principles of Natural Language Processing main techniques, and to know how to use in the environment of an intelligent system or service.
CEP3. Capacity for applying Artificial Intelligence techniques in technological and industrial environments to improve quality and productivity.
CT7. ANALISIS Y SINTESIS: Capability to analyze and solve complex technical problems.
CT6. REASONING: Capability to evaluate and analyze on a reasoned and critical way about situations, projects, proposals, reports and scientific-technical surveys. Capability to argue the reasons that explain or justify such situations, proposals, etc..
CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

CT4. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.

CB8. Capability to communicate their conclusions, and the knowledge and rationale underpinning these, to both skilled and unskilled public in a clear and unambiguous way.

Full-or-part-time: 26h
Theory classes: 6h
Practical classes: 3h
Self study: 17h

Industrial presentations

Description:
Industrial presentations: learning current applications of human language engineering.

Specific objectives:
3, 4

Related competencies :
CG1. Capability to plan, design and implement products, processes, services and facilities in all areas of Artificial Intelligence.
CEA3. Capability to understand the basic operation principles of Machine Learning main techniques, and to know how to use on the environment of an intelligent system or service.
CEA7. Capability to understand the problems, and the solutions to problems in the professional practice of Artificial Intelligence application in business and industry environment.
CEP4. Capability to design, write and report about computer science projects in the specific area of ??Artificial Intelligence.
CEA5. Capability to understand the basic operation principles of Natural Language Processing main techniques, and to know how to use in the environment of an intelligent system or service.
CEP3. Capacity for applying Artificial Intelligence techniques in technological and industrial environments to improve quality and productivity.
CT7. ANALISIS Y SINTESIS: Capability to analyze and solve complex technical problems.
CT6. REASONING: Capability to evaluate and analyze on a reasoned and critical way about situations, projects, proposals, reports and scientific-technical surveys. Capability to argue the reasons that explain or justify such situations, proposals, etc..
CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

Full-or-part-time: 8h 42m
Theory classes: 2h
Practical classes: 1h
Self study: 5h 42m

Students presentations

Description:
Students presentations: oral presentation of the study or development carried on.

Specific objectives:
3, 4

Related competencies :
CG1. Capability to plan, design and implement products, processes, services and facilities in all areas of Artificial Intelligence.
CEA3. Capability to understand the basic operation principles of Machine Learning main techniques, and to know how to use on the environment of an intelligent system or service.
CEA7. Capability to understand the problems, and the solutions to problems in the professional practice of Artificial Intelligence application in business and industry environment.
CEP4. Capability to design, write and report about computer science projects in the specific area of ??Artificial Intelligence.
CEA5. Capability to understand the basic operation principles of Natural Language Processing main techniques, and to know how to use in the environment of an intelligent system or service.
CEP3. Capacity for applying Artificial Intelligence techniques in technological and industrial environments to improve quality and productivity.
CT7. ANALISIS Y SINTESIS: Capability to analyze and solve complex technical problems.
CT6. REASONING: Capability to evaluate and analyze on a reasoned and critical way about situations, projects, proposals, reports and scientific-technical surveys. Capability to argue the reasons that explain or justify such situations, proposals, etc..
CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

Full-or-part-time: 8h 42m
Theory classes: 2h
Practical classes: 1h
Self study: 5h 42m

GRADING SYSTEM

The students must deliver a report from three practical sessions where invited speakers from the industry explain actual applications of HLE. Only three industrial sessions will be subject of reporting, even in the case that some additional sessions would be scheduled.

Each report is the 10% of the final mark.

For the other 70% of the mark, each student will choose one option among:

1. Deep study of a specific HLE application or a comparative study of HLE applications
2. Development of a HLE application
3. Development of a proposal to solve a specific real challenge

In all the cases, a preliminar deliverable will be required (10%), as well as a final report (50%), and an oral presentation (10%).

BIBLIOGRAPHY

Basic:
- Jurafsky, D.; Martin, J.H. Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. 2nd ed. Prentice Hall, 2008. ISBN 9332518416.
- Jurafsky, D.; Martin, J.H. Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. 3rd.edition draft. 2019.