Course guide
270419 - PLH - Human Language Processing

Unit in charge: Barcelona School of Informatics
Teaching unit: 723 - CS - Department of Computer Science.
Degree: BACHELOR’S DEGREE IN ARTIFICIAL INTELLIGENCE (Syllabus 2021). (Compulsory subject).
Academic year: 2022
ECTS Credits: 6.0
Languages: Catalan

Coordinating lecturer: JORGE TURMO BORRÁS

Others:
Segon quadrimestre:
SALVADOR MEDINA HERRERA - 11, 12
JORGE TURMO BORRÁS - 11, 12

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CE27. To design and apply speech processing techniques, speech recognition and human language comprehension, with application in social artificial intelligence.
CE02. To master the basic concepts of discrete mathematics, logic, algorithmic and computational complexity, and its application to the automatic processing of information through computer systems. To be able to apply all these for solving problems.
CE14. To master the foundations, paradigms and techniques of intelligent systems and to analyze, designing and build computer systems, services and applications that use these techniques in any field of application, including robotics.
CE15. To acquire, formalize and represent human knowledge in a computable form for solving problems through a computer system in any field of application, particularly those related to aspects of computing, perception and performance in intelligent environments or environments.
CE16. To design and evaluate human-machine interfaces that guarantee the accessibility and usability of computer systems, services and applications.
CE17. To develop and evaluate interactive systems and presentation of complex information and its application to solving human-computer and human-robot interaction design problems.
CE18. To acquire and develop computational learning techniques and to design and implement applications and systems that use them, including those dedicated to the automatic extraction of information and knowledge from large volumes of data.

Generical:
CG5. Work in multidisciplinary teams and projects related to artificial intelligence and robotics, interacting fluently with engineers and professionals from other disciplines.
CG6. To identify opportunities for innovative applications of artificial intelligence and robotics in constantly evolving technological environments.
CG3. To define, evaluate and select hardware and software platforms for the development and execution of computer systems, services and applications in the field of artificial intelligence.
CG4. Reasoning, analyzing reality and designing algorithms and formulations that model it. To identify problems and construct valid algorithmic or mathematical solutions, eventually new, integrating the necessary multidisciplinary knowledge, evaluating different alternatives with a critical spirit, justifying the decisions taken, interpreting and synthesizing the results in the context of the application domain and establishing methodological generalizations based on specific applications.
CG7. To interpret and apply current legislation, as well as specifications, regulations and standards in the field of artificial intelligence.
CG8. Perform an ethical exercise of the profession in all its facets, applying ethical criteria in the design of systems, algorithms, experiments, use of data, in accordance with the ethical systems recommended by national and international organizations, with special emphasis on security, robustness, privacy, transparency, traceability, prevention of bias (race, gender, religion, territory, etc.) and respect for human rights.
CG9. To face new challenges with a broad vision of the possibilities of a professional career in the field of Artificial Intelligence. Develop the activity applying quality criteria and continuous improvement, and act rigorously in professional development. Adapt to organizational or technological changes. Work in situations of lack of information and / or with time and / or resource restrictions.
Transversal:
CT6. Autonomous Learning. Detect deficiencies in one’s own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT8. (ENG) Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.
CT2. Sustainability and Social Commitment. To know and understand the complexity of economic and social phenomena typical of the welfare society; Be able to relate well-being to globalization and sustainability; Achieve skills to use in a balanced and compatible way the technique, the technology, the economy and the sustainability.
CT1. Entrepreneurship and innovation. Know and understand the organization of a company and the sciences that govern its activity; Have the ability to understand labor standards and the relationships between planning, industrial and commercial strategies, quality and profit.

Basic:
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB5. That the students have developed those learning skills necessary to undertake later studies with a high degree of autonomy

TEACHING METHODOLOGY

LEARNING OBJECTIVES OF THE SUBJECT
1. To understand the fundamental theories and techniques associated with Natural Language Processing
2. To know the most relevant resources and applications of Natural Language Processing
3. To develop programs to solve particular tasks in the Natural Language Processing area

STUDY LOAD

<table>
<thead>
<tr>
<th>Type</th>
<th>Hours</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self study</td>
<td>90,0</td>
<td>60.00</td>
</tr>
<tr>
<td>Hours small group</td>
<td>30,0</td>
<td>20.00</td>
</tr>
<tr>
<td>Hours large group</td>
<td>30,0</td>
<td>20.00</td>
</tr>
</tbody>
</table>

Total learning time: 150 h

CONTENTS

- Natural Language Processing and its applications
- Techniques, resources and applications associated with word analysis
- Techniques, resources and applications associated with the analysis of word sequences
- Techniques, resources and applications associated with sentence analysis
Techniques and applications associated with the analysis of a text seen as a sequence of sentences

ACTIVITIES

Sesión introductoria

Specific objectives:
1, 2

Related competencies:
CG4. Reasoning, analyzing reality and designing algorithms and formulations that model it. To identify problems and construct valid algorithmic or mathematical solutions, eventually new, integrating the necessary multidisciplinary knowledge, evaluating different alternatives with a critical spirit, justifying the decisions taken, interpreting and synthesizing the results in the context of the application domain and establishing methodological generalizations based on specific applications.
CG3. To define, evaluate and select hardware and software platforms for the development and execution of computer systems, services and applications in the field of artificial intelligence.
CG6. To identify opportunities for innovative applications of artificial intelligence and robotics in constantly evolving technological environments.
CG5. Work in multidisciplinary teams and projects related to artificial intelligence and robotics, interacting fluently with engineers and professionals from other disciplines.
CE14. To master the foundations, paradigms and techniques of intelligent systems and to analyze, designing and build computer systems, services and applications that use these techniques in any field of application, including robotics.
CE15. To acquire, formalize and represent human knowledge in a computable form for solving problems through a computer system in any field of application, particularly those related to aspects of computing, perception and performance in intelligent environments or environments.
CE02. To master the basic concepts of discrete mathematics, logic, algorithmic and computational complexity, and its application to the automatic processing of information through computer systems. To be able to apply all these for solving problems.
CE18. To acquire and develop computational learning techniques and to design and implement applications and systems that use them, including those dedicated to the automatic extraction of information and knowledge from large volumes of data.
CE27. To design and apply speech processing techniques, speech recognition and human language comprehension, with application in social artificial intelligence.
CT8. (ENG) Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CB5. That the students have developed those learning skills necessary to undertake later studies with a high degree of autonomy
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB3. That students have the ability to gather and interpret relevant data (usually within their area of study) to make judgments that include a reflection on relevant social, scientific or ethical issues.

Full-or-part-time: 4h
Theory classes: 2h
Laboratory classes: 2h
Identificación de unidades lingüísticas en un documento

**Specific objectives:**
1, 2

**Related competencies:**
CG3. To define, evaluate and select hardware and software platforms for the development and execution of computer systems, services and applications in the field of artificial intelligence.
CG4. Reasoning, analyzing reality and designing algorithms and formulations that model it. To identify problems and construct valid algorithmic or mathematical solutions, eventually new, integrating the necessary multidisciplinary knowledge, evaluating different alternatives with a critical spirit, justifying the decisions taken, interpreting and synthesizing the results in the context of the application domain and establishing methodological generalizations based on specific applications.
CG6. To identify opportunities for innovative applications of artificial intelligence and robotics in constantly evolving technological environments.
CG5. Work in multidisciplinary teams and projects related to artificial intelligence and robotics, interacting fluently with engineers and professionals from other disciplines.
CE18. To acquire and develop computational learning techniques and to design and implement applications and systems that use them, including those dedicated to the automatic extraction of information and knowledge from large volumes of data.
CE15. To acquire, formalize and represent human knowledge in a computable form for solving problems through a computer system in any field of application, particularly those related to aspects of computing, perception and performance in intelligent environments or environments.
CE02. To master the basic concepts of discrete mathematics, logic, algorithmic and computational complexity, and its application to the automatic processing of information through computer systems. To be able to apply these for solving problems.
CE14. To master the foundations, paradigms and techniques of intelligent systems and to analyze, designing and build computer systems, services and applications that use these techniques in any field of application, including robotics.
CE27. To design and apply speech processing techniques, speech recognition and human language comprehension, with application in the social field of artificial intelligence.
CT8. (ENG) Perspectiva de género. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.
CT6. Autonomous Learning. Detect deficiencies in one’s own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CB5. That the students have developed those learning skills necessary to undertake later studies with a high degree of autonomy
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.

**Full-or-part-time:** 2h
Theory classes: 2h
Bloc de tratamiento de una palabra

Specific objectives:
1, 2

Related competencies:
CG4. Reasoning, analyzing reality and designing algorithms and formulations that model it. To identify problems and construct valid algorithmic or mathematical solutions, eventually new, integrating the necessary multidisciplinary knowledge, evaluating different alternatives with a critical spirit, justifying the decisions taken, interpreting and synthesizing the results in the context of the application domain and establishing methodological generalizations based on specific applications.

CG3. To define, evaluate and select hardware and software platforms for the development and execution of computer systems, services and applications in the field of artificial intelligence.

CG6. To identify opportunities for innovative applications of artificial intelligence and robotics in constantly evolving technological environments.

CG5. Work in multidisciplinary teams and projects related to artificial intelligence and robotics, interacting fluently with engineers and professionals from other disciplines.

CE14. To master the foundations, paradigms and techniques of intelligent systems and to analyze, designing and build computer systems, services and applications that use these techniques in any field of application, including robotics.

CE15. To acquire, formalize and represent human knowledge in a computable form for solving problems through a computer system in any field of application, particularly those related to aspects of computing, perception and performance in intelligent environments or environments.

CE02. To master the basic concepts of discrete mathematics, logic, algorithmic and computational complexity, and its application to the automatic processing of information through computer systems. To be able to apply all these for solving problems.

CE18. To acquire and develop computational learning techniques and to design and implement applications and systems that use them, including those dedicated to the automatic extraction of information and knowledge from large volumes of data.

CE27. To design and apply speech processing techniques, speech recognition and human language comprehension, with application in social artificial intelligence.

CT8. (ENG) Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

CT6. Autonomous Learning. Detect deficiencies in one’s own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.

CB5. That the students have developed those learning skills necessary to undertake later studies with a high degree of autonomy
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.

Full-or-part-time: 11h
Theory classes: 8h
Practical classes: 3h
Bloc de tratamiento de secuencias de palabras con significado

Specific objectives:
1, 2

Related competencies:
CG4. Reasoning, analyzing reality and designing algorithms and formulations that model it. To identify problems and construct valid algorithmic or mathematical solutions, eventually new, integrating the necessary multidisciplinary knowledge, evaluating different alternatives with a critical spirit, justifying the decisions taken, interpreting and synthesizing the results in the context of the application domain and establishing methodological generalizations based on specific applications.
CG3. To define, evaluate and select hardware and software platforms for the development and execution of computer systems, services and applications in the field of artificial intelligence.
CG6. To identify opportunities for innovative applications of artificial intelligence and robotics in constantly evolving technological environments.
CG5. Work in multidisciplinary teams and projects related to artificial intelligence and robotics, interacting fluently with engineers and professionals from other disciplines.
CG18. To acquire and develop computational learning techniques and to design and implement applications and systems that use them, including those dedicated to the automatic extraction of information and knowledge from large volumes of data.
CE15. To acquire, formalize and represent human knowledge in a computable form for solving problems through a computer system in any field of application, particularly those related to aspects of computing, perception and performance in intelligent environments or environments.
CE14. To master the foundations, paradigms and techniques of intelligent systems and to analyze, designing and build computer systems, services and applications that use these techniques in any field of application, including robotics.
CE02. To master the basic concepts of discrete mathematics, logic, algorithmic and computational complexity, and its application to the automatic processing of information through computer systems. To be able to apply all these for solving problems.
CT8. (ENG) Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.
CT6. Autonomous Learning. Detect deficiencies in one’s own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CB5. That the students have developed those learning skills necessary to undertake later studies with a high degree of autonomy
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.

Full-or-part-time: 7h
Theory classes: 5h
Practical classes: 2h

Bloc de tratamiento de una frase aislada

Full-or-part-time: 6h
Theory classes: 4h
Practical classes: 2h

Bloc de tratamiento de un texto como secuencia de frases

Full-or-part-time: 2h
Theory classes: 1h 30m
Practical classes: 0h 30m
<table>
<thead>
<tr>
<th>Práctica 1</th>
<th>Full-or-part-time: 4h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Laboratory classes: 4h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Práctica 2</th>
<th>Full-or-part-time: 8h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Laboratory classes: 8h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Práctica 3</th>
<th>Full-or-part-time: 8h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Laboratory classes: 8h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Práctica 4</th>
<th>Full-or-part-time: 8h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Laboratory classes: 8h</td>
</tr>
</tbody>
</table>
Prácticas

Specific objectives:
3

Related competencies:
CG8. Perform an ethical exercise of the profession in all its facets, applying ethical criteria in the design of systems, algorithms, experiments, use of data, in accordance with the ethical systems recommended by national and international organizations, with special emphasis on security, robustness, privacy, transparency, traceability, prevention of bias (race, gender, religion, territory, etc.) and respect for human rights.
CG7. To interpret and apply current legislation, as well as specifications, regulations and standards in the field of artificial intelligence.
CG9. To face new challenges with a broad vision of the possibilities of a professional career in the field of Artificial Intelligence. Develop the activity applying quality criteria and continuous improvement, and act rigorously in professional development. Adapt to organizational or technological changes. Work in situations of lack of information and / or with time and / or resource restrictions.
CG5. Work in multidisciplinary teams and projects related to artificial intelligence and robotics, interacting fluently with engineers and professionals from other disciplines.

CE02. To master the basic concepts of discrete mathematics, logic, algorithmic and computational complexity, and its application to the automatic processing of information through computer systems. To be able to apply all these for solving problems.
CE16. To design and evaluate human-machine interfaces that guarantee the accessibility and usability of computer systems, services and applications.
CE18. To acquire and develop computational learning techniques and to design and implement applications and systems that use them, including those dedicated to the automatic extraction of information and knowledge from large volumes of data.
CE14. To master the foundations, paradigms and techniques of intelligent systems and to analyze, designing and build computer systems, services and applications that use these techniques in any field of application, including robotics.
CE17. To develop and evaluate interactive systems and presentation of complex information and its application to solving human-computer and human-robot interaction design problems.

Full-or-part-time: 45h
Self study: 45h
Exam

Specific objectives:
1, 2

Related competencies:
CG4. Reasoning, analyzing reality and designing algorithms and formulations that model it. To identify problems and construct valid algorithmic or mathematical solutions, eventually new, integrating the necessary multidisciplinary knowledge, evaluating different alternatives with a critical spirit, justifying the decisions taken, interpreting and synthesizing the results in the context of the application domain and establishing methodological generalizations based on specific applications.
CG3. To define, evaluate and select hardware and software platforms for the development and execution of computer systems, services and applications in the field of artificial intelligence.
CG6. To identify opportunities for innovative applications of artificial intelligence and robotics in constantly evolving technological environments.
CG5. Work in multidisciplinary teams and projects related to artificial intelligence and robotics, interacting fluently with engineers and professionals from other disciplines.
CE18. To acquire and develop computational learning techniques and to design and implement applications and systems that use them, including those dedicated to the automatic extraction of information and knowledge from large volumes of data.
CE15. To acquire, formalize and represent human knowledge in a computable form for solving problems through a computer system in any field of application, particularly those related to aspects of computing, perception and performance in intelligent environments or environments.
CE14. To master the foundations, paradigms and techniques of intelligent systems and to analyze, designing and build computer systems, services and applications that use these techniques in any field of application, including robotics.
CE02. To master the basic concepts of discrete mathematics, logic, algorithmic and computational complexity, and its application to the automatic processing of information through computer systems. To be able to apply all these for solving problems.
CE27. To design and apply speech processing techniques, speech recognition and human language comprehension, with application in social artificial intelligence.
CT8. (ENG) Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.
CT6. Autonomous Learning. Detect deficiencies in one’s own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CB5. That the students have developed those learning skills necessary to undertake later studies with a high degree of autonomy
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.

Full-or-part-time: 45h
Self study: 45h

GRADING SYSTEM

NEX: final exam grade
NLAB: average grade of laboratory practices
NF: final grade of the course

\[ NF = 0.5 \times NEX + 0.5 \times NLAB \]

Reassessment
Only those students who had previously taken the final exam and failed it can take the reassessment exam.
BIBLIOGRAPHY

Basic: