Course guide
270225 - POE - Spoken and Written Language Processing

Last modified: 31/01/2025

Unit in charge: Barcelona School of Informatics
Teaching unit: 739 - TSC - Department of Signal Theory and Communications.

Degree: BACHELOR'S DEGREE IN DATA SCIENCE AND ENGINEERING (Syllabus 2017). (Compulsory subject).

Academic year: 2024 **ECTS Credits:** 6.0 **Languages:** Spanish

LECTURER

Coordinating lecturer: JOSE ADRIAN RODRIGUEZ FONOLLOSA

Others: Segon quadrimestre:
CARLOS ESCOLANO PEINADO - 11, 12, 13
JOSE ADRIAN RODRIGUEZ FONOLLOSA - 11, 12, 13

PRIOR SKILLS

College Calculus, Linear Algebra
Basic Probability and Statistics
Large programming experience in Pyhton
Machine Learning.
Introduction to Deep Learning

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CE5. Design and apply techniques of signal processing, choosing between different technological tools, including those of Artificial vision, speech recognition and multimedia data processing.
CE6. Build or use systems of processing and comprehension of written language, integrating it into other systems driven by the data. Design systems for searching textual or hypertextual information and analysis of social networks.

Generical:
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CG5. To be able to draw on fundamental knowledge and sound work methodologies acquired during the studies to adapt to the new technological scenarios of the future.

Transversal:
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
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.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.

Basic:
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

Lectures presenting new theoretical material and practical examples.
Theoretical and practical assignments grouped in subjects.
Research project, presented in written and oral form by the students.

LEARNING OBJECTIVES OF THE SUBJECT

1.Know the most important deep learning technologies of interest in the processing of oral and written language.
2.The student must know the most important applications of speech and language technology.
3.The student must be able to select the most appropriate speech and language technology for a particular task or application.
4.Develop innovative applications that use speech technology appropriately.
5.El alumno debe ser capaz de identificar los parámetros fundamentales de la voz en el dominio temporal y frecuencial
6.The student must know the most important mathematical and machine learning tools for the analysis of the voice as vector quantification (VQ), Gaussian mixture models (GMM) and hidden Markov models (HMM).
7.The student must know the techniques for statistical language modeling.

STUDY LOAD

Type	Hours	Percentage
Self study	90,0	60.00
Hours large group	45,0	30.00
Hours small group	15,0	10.00

Total learning time: 150 h

CONTENTS

Introduction to language and speech technologies and applications

Description:
Applications of oral and written language processing. Social impact.
Main blocks of a natural language processing system: speech recognition, natural language processing, text to speech conversion.
Language as a sequence of words. Vector representation of words. One-hot encoding versus continuous-space representations.
Word2vec: Continuous bag-of-words (CBOW) versus Continuous skip-gram. GloVe vectors. Structures and analogies in word vector representations.

Language Modeling

Description:
Statistical modeling based on N-grams.
Modeling with neural networks. Recurring networks Convolutional networks. Attention mechanisms: the Transformer.

Contextual language representations

Description:
General purpose language representations.
Unsupervised training. Unidirectional and bidirectional systems.
Main architectures: ULMfit, OpenAI GPT, ELMo, BERT, XLM. Applications.

Neural Machine Translation

Description:
Introduction to Machine Translation. Automatic quality evaluation: BLEU
Neural Machine Translation.

Introduction to automatic speech recognition

Description:
Pattern matching. Dynamic time warping.
Hidden Markov models. Isolated word recognition.
Large vocabulary continuous ASR: Acoustic modeling, Language modeling, Search.

Speech synthesis

Description:
Linguistic processing.
Prosody modeling.
Waveform generation.
Concatenation methods.

ACTIVITIES

Topic development: Introduction to speech and language technology and applications

Description:
Introduction to speech and language technology and applications.
Word vectors

Specific objectives:
2, 3

Related competencies :
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
CE5. Design and apply techniques of signal processing, choosing between different technological tools, including those of Artificial vision, speech recognition and multimedia data processing.
CE6. Build or use systems of processing and comprehension of written language, integrating it into other systems driven by the data. Design systems for searching textual or hypertextual information and analysis of social networks.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG5. To be able to draw on fundamental knowledge and sound work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
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.

Full-or-part-time: 18h
Theory classes: 6h
Laboratory classes: 2h
Self study: 10h

Topic development: Language Modeling

Specific objectives:
6

Related competencies :
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
CE5. Design and apply techniques of signal processing, choosing between different technological tools, including those of Artificial vision, speech recognition and multimedia data processing.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG5. To be able to draw on fundamental knowledge and sound work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
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.

Full-or-part-time: 18h
Theory classes: 6h
Laboratory classes: 2h
Self study: 10h

Topic development: Automatic Speech Recognition

Description:
Automatic Speech Recognition

Specific objectives:
5, 6, 7

Related competencies :
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
CE5. Design and apply techniques of signal processing, choosing between different technological tools, including those of Artificial vision, speech recognition and multimedia data processing.
CE6. Build or use systems of processing and comprehension of written language, integrating it into other systems driven by the data. Design systems for searching textual or hypertextual information and analysis of social networks.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG5. To be able to draw on fundamental knowledge and sound work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
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.

Full-or-part-time: 19h
Theory classes: 9h
Self study: 10h

Topic development: Speech Synthesis

Description:
Speech Synthesis

Specific objectives:
2

Related competencies :
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
CE5. Design and apply techniques of signal processing, choosing between different technological tools, including those of Artificial vision, speech recognition and multimedia data processing.
CE6. Build or use systems of processing and comprehension of written language, integrating it into other systems driven by the data. Design systems for searching textual or hypertextual information and analysis of social networks.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG5. To be able to draw on fundamental knowledge and sound work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
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.

Full-or-part-time: 16h
Theory classes: 6h
Self study: 10h

Topic development: Contextual language representations

Specific objectives:
1

Related competencies :
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
CE5. Design and apply techniques of signal processing, choosing between different technological tools, including those of Artificial vision, speech recognition and multimedia data processing.
CE6. Build or use systems of processing and comprehension of written language, integrating it into other systems driven by the data. Design systems for searching textual or hypertextual information and analysis of social networks.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG5. To be able to draw on fundamental knowledge and sound work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
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.

Full-or-part-time: 27h
Theory classes: 9h
Laboratory classes: 2h
Self study: 16h

Topic development: Neuronal Machine Translation

Description:
Neuronal Machine Translation

Full-or-part-time: 18h
Theory classes: 6h
Laboratory classes: 2h
Self study: 10h

Final project

Description:
Final project

Specific objectives:
3, 4

Related competencies :
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
CE5. Design and apply techniques of signal processing, choosing between different technological tools, including those of Artificial vision, speech recognition and multimedia data processing.
CE6. Build or use systems of processing and comprehension of written language, integrating it into other systems driven by the data. Design systems for searching textual or hypertextual information and analysis of social networks.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG5. To be able to draw on fundamental knowledge and sound work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
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.

Full-or-part-time: 34h
Theory classes: 3h
Laboratory classes: 7h
Self study: 24h

GRADING SYSTEM

Course evaluation is based on the following aspects:

- Two exams, a midterm exam, and the final exam, to assess the knowledge acquired but the student on the topics worked on in theory and practice sessions (60%)

- Evaluation of laboratory assignments: (40%)

BIBLIOGRAPHY

Basic:
- Huang, X.; Acero, A.; Hon, H.-W. Spoken language processing: a guide to theory, algorithm and system development. Upper Saddle River: Prentice Hall, 2001. ISBN 0130226165.
- Goodfellow, I.; Bengio, Y.; Courville, A. Deep learning. Cambridge, Massachusetts: The MIT Press, 2016. ISBN 9780262035613.
- Kamath, U.; Liu, J.; Whitaker, J. Deep learning for NLP and speech recognition. Cham: Springer, 2019. ISBN 9783030145958.

Complementary:
- Taylor, P. Text-to-speech synthesis. Cambridge, UK ; New York: Cambridge University Press, 2009. ISBN 9780521899277.