Guía docente
270225 - POE - Procesado del Lenguaje Oral y Escrito

Última modificación: 30/01/2024

Unidad responsable: Facultad de Informática de Barcelona
Unidad que imparte: 739 - TSC - Departamento de Teoría de la Señal y Comunicaciones.

Titulación: GRADO EN CIENCIA E INGENIERÍA DE DATOS (Plan 2017). (Asignatura obligatoria).

Curso: 2023 **Créditos ECTS:** 6.0 **Idiomas:** Castellano, Inglés

PROFESORADO

Profesorado responsable: JOSE ADRIAN RODRIGUEZ FONOLLOSA

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

CAPACIDADES PREVIAS

Cálculo y álgebra lineal
Probabilidad básica y estadística
Nivel alto de programación en Python
Aprendizaje automático.
Introducción al Deep Learning

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:
CE5. Diseñar y aplicar técnicas de procesado de señal, eligiendo entre distintas herramientas tecnológicas, incluidas las de visión Artificial, de reconocimiento del lenguaje hablado y las de tratamiento de datos multimedia.
CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.

Genéricas:
CG1. Concebir sistemas computacionales que integren datos de procedencias y formas muy diversas, creen con ellos modelos matemáticos, razonen sobre dichos modelos y actúen en consecuencia, aprendiendo de la experiencia.
CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.
CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.
CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.

Transversales:
CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.
CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.
CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.

Básicas:
CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

METODOLOGÍAS DOCENTES

Sesiones de teoría intercalando la exposición de nuevo material teórico con ejemplos prácticos.
Tareas teóricas y prácticas agrupadas en temas.
Proyecto final presentado en forma escrita y oral por los estudiantes.

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

1.Conocer las tecnologías más importantes de aprendizaje profundo de interés en el procesado del lenguaje oral y escrito.
2.El alumno debe conocer las aplicaciones más importantes de la tecnología del habla y del lenguaje.
3.El alumno ha de ser capaz de seleccionar la tecnología del habla y del lenguaje más adecuada para una determinada tarea o aplicación.
4.Desarrollar aplicaciones que utilicen la tecnología del habla de froma adecuada e innovadora.
5.The student should be able to identify the fundamental parameters of the voice in the temporal and frequency domain
6.El alumno debe conocer las herramientas matemáticas de modelado y aprendizaje automático más útiles en el análisis de la voz como la cuantificación vectorial, los modelos de mezclas de Gaussianas y los modelos ocultos de Markov.
7.El alumno debe conocer las técnicas de modelado estadístico del lenguaje.

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

Tipo	Horas	Porcentaje
Horas aprendizaje autónomo	90,0	60.00
Horas grupo grande	45,0	30.00
Horas grupo pequeño	15,0	10.00

Dedicación total: 150 h

CONTENIDOS

Introducción a la tecnologías del habla y del lenguaje y sus aplicaciones

Descripción:
Aplicaciones del procesado del lenguaje oral y escrito. Impacto social.
Bloques principales de un sistema de procesado del habla: reconocimiento del habla, procesado del lenguaje natural, conversión de texto a voz.
El lenguaje como secuencia de palabras. Representación vectorial de palabras. Codificación one-hot. Representaciones vectoriales en un espacio continuo.
Word2vec: Continuous bag-of-words versus Continuous skip-gram. Vectores GloVe. Estructuras y analogías en las representaciones vectoriales de palabras.

Modelado del lenguaje

Descripción:
Modelado estadístico basado en N-gramas.
Modelado con redes neuronales. Redes recurrentes. Redes convolucionales. Sistemas de atención: el Transformer.

Representaciones contextuales del lenguaje

Descripción:
Pattern matching. Dynamic time warping.
Modelos ocultos de Markov: reconocimiento de palabras aisladas
Reconocimiento de habla continua: modelado acústico, modelado de lenguaje, búsqueda.

Traducción Automática Neuronal

Descripción:
Introducción a la traducción automática. Evaluación automatiza de la calidad: BLEU
Traducción Automática con redes neuronales.

Fundamentos del reconocimiento automático de la voz

Descripción:
Pattern matching. Dynamic time warping.
Modelos ocultos de Markov: reconocimiento de palabras aisladas
Reconocimiento de habla continua: modelado acústico, modelado de lenguaje, búsqueda.

Síntesis de voz

Descripción:
Procesador lingüístico.
Modelado de la prosodia.
Generación de formas de onda.
Métodos de concatenación.

ACTIVIDADES

Desarrollo del tema: Introducción a la tecnologia del habla y del lenguaje y aplicaciones

Descripción:
Introducción a la tecnologia del habla y del lenguaje y aplicaciones.
Word vectors

Objetivos específicos:
2, 3

Competencias relacionadas:
CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.
CG1. Concebir sistemas computacionales que integren datos de procedencias y formas muy diversas, creen con ellos modelos matemáticos, razonen sobre dichos modelos y actúen en consecuencia, aprendiendo de la experiencia.
CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.
CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.
CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.
CE5. Diseñar y aplicar técnicas de procesado de señal, eligiendo entre distintas herramientas tecnológicas, incluidas las de visión Artificial, de reconocimiento del lenguaje hablado y las de tratamiento de datos multimedia.
CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.
CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.
CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.
CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

Dedicación: 18h
Grupo grande/Teoría: 6h
Grupo pequeño/Laboratorio: 2h
Aprendizaje autónomo: 10h

Desarrollo del tema: Modelado del lenguaje

Objetivos específicos:
6

Competencias relacionadas:
CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.
CG1. Concebir sistemas computacionales que integren datos de procedencias y formas muy diversas, creen con ellos modelos matemáticos, razonen sobre dichos modelos y actúen en consecuencia, aprendiendo de la experiencia.
CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.
CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.
CE5. Diseñar y aplicar técnicas de procesado de señal, eligiendo entre distintas herramientas tecnológicas, incluidas las de visión Artificial, de reconocimiento del lenguaje hablado y las de tratamiento de datos multimedia.
CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.
CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.
CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.
CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

Dedicación: 18h
Grupo grande/Teoría: 6h
Grupo pequeño/Laboratorio: 2h
Aprendizaje autónomo: 10h

Desarrollo del tema: Reconocimiento automático del habla

Descripción:
Reconocimiento automático del habla

Objetivos específicos:
5, 6, 7

Competencias relacionadas:
CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.
CG1. Concebir sistemas computacionales que integren datos de procedencias y formas muy diversas, creen con ellos modelos matemáticos, razonen sobre dichos modelos y actúen en consecuencia, aprendiendo de la experiencia.
CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.
CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.
CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.
CE5. Diseñar y aplicar técnicas de procesado de señal, eligiendo entre distintas herramientas tecnológicas, incluidas las de visión Artificial, de reconocimiento del lenguaje hablado y las de tratamiento de datos multimedia.
CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.
CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.
CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.
CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

Dedicación: 19h
Grupo grande/Teoría: 9h
Aprendizaje autónomo: 10h

Desarrollo del tema: Síntesis de voz

Descripción:
Síntesis de voz

Objetivos específicos:
2

Competencias relacionadas:
CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.
CG1. Concebir sistemas computacionales que integren datos de procedencias y formas muy diversas, creen con ellos modelos matemáticos, razonen sobre dichos modelos y actúen en consecuencia, aprendiendo de la experiencia.
CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.
CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.
CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.
CE5. Diseñar y aplicar técnicas de procesado de señal, eligiendo entre distintas herramientas tecnológicas, incluidas las de visión Artificial, de reconocimiento del lenguaje hablado y las de tratamiento de datos multimedia.
CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.
CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.
CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.
CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

Dedicación: 16h
Grupo grande/Teoría: 6h
Aprendizaje autónomo: 10h

Desarrollo del tema: Representaciones contextuales del lenguaje

Objetivos específicos:
1

Competencias relacionadas:
CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.
CG1. Concebir sistemas computacionales que integren datos de procedencias y formas muy diversas, creen con ellos modelos matemáticos, razonen sobre dichos modelos y actúen en consecuencia, aprendiendo de la experiencia.
CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.
CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.
CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.
CE5. Diseñar y aplicar técnicas de procesado de señal, eligiendo entre distintas herramientas tecnológicas, incluidas las de visión Artificial, de reconocimiento del lenguaje hablado y las de tratamiento de datos multimedia.
CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.
CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.
CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.
CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

Dedicación: 27h
Grupo grande/Teoría: 9h
Grupo pequeño/Laboratorio: 2h
Aprendizaje autónomo: 16h

Desarrollo del tema: Traducción Automática Neuronal

Descripción:
Traducción Automática Neuronal

Dedicación: 18h
Grupo grande/Teoría: 6h
Grupo pequeño/Laboratorio: 2h
Aprendizaje autónomo: 10h

Proyecto final

Descripción:
Proyecto final

Objetivos específicos:
3, 4

Competencias relacionadas:
CG5. Poder recurrir a conocimientos fundamentales y metodologías de trabajo sólidas adquiridos durante los estudios para adaptarse a los nuevos escenarios tecnológicos del futuro.
CG1. Concebir sistemas computacionales que integren datos de procedencias y formas muy diversas, creen con ellos modelos matemáticos, razonen sobre dichos modelos y actúen en consecuencia, aprendiendo de la experiencia.
CG2. Elegir y aplicar los métodos y técnicas más adecuados a un problema definido por datos que representen un reto por su volumen, velocidad, variedad o heterogeneidad, incluidos métodos informáticos, matemáticos, estadísticos y de procesado de la señal.
CG4. Identificar oportunidades para aplicaciones innovadoras orientadas a datos en entornos tecnológicos en continua evolución.
CE6. Construir o utilizar sistemas de procesado y comprensión del lenguaje escrito, integrándolo en otros sistemas dirigidos por los datos. Diseñar sistemas de búsqueda de información textual o hipertextual y de análisis de redes sociales.
CE5. Diseñar y aplicar técnicas de procesado de señal, eligiendo entre distintas herramientas tecnológicas, incluidas las de visión Artificial, de reconocimiento del lenguaje hablado y las de tratamiento de datos multimedia.
CT5. Uso solvente de los recursos de información. Gestionar la adquisición, la estructuración, el análisis y la visualización de datos e información en el ámbito de especialidad y valorar de forma crítica los resultados de dicha gestión.
CT6. Aprendizaje autónomo. Detectar deficiencias en el propio conocimiento y superarlas mediante la reflexión crítica y la elección de la mejor actuación para ampliar dicho conocimiento.
CT7. Tercera lengua. Conocer una tercera lengua, preferentemente el inglés, con un nivel adecuado oral y escrito y en consonancia con las necesidades que tendrán los titulados y tituladas.
CB4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

Dedicación: 34h
Grupo grande/Teoría: 3h
Grupo pequeño/Laboratorio: 7h
Aprendizaje autónomo: 24h

SISTEMA DE CALIFICACIÓN

La evaluación de la asignatura se basa en los siguientes aspectos:

- Dos controles, un control intermedio y el examen final, para evaluar de forma individual los conocimientos adquiridos sobre los temas trabajados en las sesiones de teoría y las prácticas (60%)

- Evaluación de los trabajos de laboratorio: (40%)

BIBLIOGRAFÍA

Básica:
- 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 [en línea]. Cambridge, Massachusetts: The MIT Press, 2016 [Consulta: 07/03/2024]. Disponible a: https://www.deeplearningbook.org/. ISBN 9780262035613.
- Kamath, U.; Liu, J.; Whitaker, J. Deep learning for NLP and speech recognition. Cham: Springer, 2019. ISBN 9783030145958.

Complementaria:

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