



Guía docente

230362 - DLSL - Aprendizaje Profundo para el Habla y el Language

Última modificación: 06/05/2019

Unidad responsable: Escuela Técnica Superior de Ingeniería de Telecomunicación de Barcelona

Unidad que imparte: 739 - TSC - Departamento de Teoría de la Señal y Comunicaciones.

Titulación: MÁSTER UNIVERSITARIO EN INGENIERÍA DE TELECOMUNICACIÓN (Plan 2013). (Asignatura optativa).

Curso: 2019

Créditos ECTS: 2.5

Idiomas: Inglés

PROFESORADO

Profesorado responsable: Giró Nieto, Xavier

Otros:
Bonafonte Cavez, Antonio Jesus
Rodriguez Fonollosa, Jose Adrian
Ruiz Costa-Jussa, Marta
Hernando Pericas, Francisco Javier
Pascual, Santiago
Sayrol Clois, Elisa
Giró Nieto, Xavier

CAPACIDADES PREVIAS

Se aconsejan unas bases mínimas de aprendizaje automático y procesado de la señal. Los estudiantes desarrollarán sus proyectos en Python, así que se recomienda un contacto previo con este lenguaje de programación.

COMPETENCIAS DE LA TITULACIÓN A LAS QUE CONTRIBUYE LA ASIGNATURA

Específicas:

CE1. Capacidad para aplicar métodos de la teoría de la información, la modulación adaptativa y codificación de canal, así como técnicas avanzadas de procesado digital de señal a los sistemas de comunicaciones y audiovisuales.

Transversales:

CT3. TRABAJO EN EQUIPO: Ser capaz de trabajar como miembro de un equipo interdisciplinar, ya sea como un miembro más o realizando tareas de dirección, con la finalidad de contribuir a desarrollar proyectos con pragmatismo y sentido de la responsabilidad, asumiendo compromisos teniendo en cuenta los recursos disponibles.

CT4. 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.

CT5. 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.

METODOLOGÍAS DOCENTES

- Clases magistrales
- Clases de aplicación
- Trabajo en equip
- Trabajo en equip (a distancia)

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

El objetivo del curso es entrenar a los estudiantes en métodos de aprendizaje profundo para la voz y el lenguaje. Las Redes Neuronales Recurrentes serán presentadas y analizadas en detalle para entender el potencial de estas herramientas en el estado del arte del procesado de series temporales. Se describirán buenas prácticas en la ingeniería y consejos de escalabilidad para poder resolver tareas como la traducción automática, el reconocimiento de voz, la síntesis de voz o la respuesta a preguntas. Las sesiones prácticas proporcionarán habilidades para que los asistentes sean competentes en herramientas contemporáneas para el análisis de datos.

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

Tipo	Horas	Porcentaje
Horas grupo pequeño	10,0	16.00
Horas aprendizaje autónomo	42,5	68.00
Horas grupo grande	10,0	16.00

Dedicación total: 62.5 h

CONTENIDOS

Redes Neuronales Avanzadas

Descripción:

- Arquitecturas: LSTM, GRU, recursivas, Conv-LSTM, redes de memoria dinámica, TDNN, redes de autopista...
- Entrenamiento: Bases de datos, propagación hacia atrás, optimización, adversarios...
- Aprendizaje: supervisado/no-supervisado, continuo...
- Visualización
- Modelos de atención
- Inmersiones: seq2seq, skip-thought vectors...
- Conjuntos de modelos

Objetivos específicos:

Al final del curso, los estudiantes estarán familiarizados con técnicas del estado del arte basadas en arquitecturas de aprendizaje profundo.

Dedicación: 16 h

Grupo grande/Teoría: 6h

Actividades dirigidas: 8h

Aprendizaje autónomo: 2h 30m

Aplicaciones a la Voz y el Lenguaje

Descripción:

- Procesado del lenguaje natural
- Traducción automática
- Reconocimiento del habla
- Reconocimiento de locutores
- Síntesis del habla
- Multimodal: lenguaje i visión.
- Entornos de desarrollo y herramientas: TensorFlow, Keras, Kaldi

Dedicación: 50 h

Grupo grande/Teoría: 6h

Actividades dirigidas: 10h

Aprendizaje autónomo: 34h



SISTEMA DE CALIFICACIÓN

Teoría: 30%
Práctica: 60%
Asistencia: 10%

BIBLIOGRAFÍA

Básica:

- Socher, Richard. Deep Learning for Natural Language Processing [en línea]. Stanford University, 2016 [Consulta: 05/09/2016]. Disponible a: <http://cs224d.stanford.edu/>.
- Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning [en línea]. MIT Press, 2017 [Consulta: 05/09/2016]. Disponible a: <http://www.deeplearningbook.org/>. ISBN 0262035618.

Complementaria:

- Joan Bruna. Topics Course on Deep Learning [en línea]. UC Berkeley, 2016 [Consulta: 05/09/2016]. Disponible a: <https://github.com/joanbruna/stat212b>.
- Dhruv Batra. Deep Learning for Perception [en línea]. Blacksburg, VA, USA: Virginia Tech, 2016 [Consulta: 02/05/2020]. Disponible a: <https://computing.ece.vt.edu/~f15ece6504/>.
- Aaron Courville and Yoshua Bengio. Deep Learning Summer School [en línea]. Montreal, Quebec: CIFAR, ICRA, CRM, 2016 [Consulta: 05/09/2016]. Disponible a: <https://sites.google.com/site/deeplearningsummerschool2016/home>.
- Luong, T.; Cho, T.; Manning, C. Neural Machine Translation [en línea]. ACL Tutorials, 2016 [Consulta: 05/09/2016]. Disponible a: <https://sites.google.com/site/acl16nmt/>.
- Giró-i-Nieto, X.; Sayrol, E.; Salvador, A.; Torres, J.; Mohedano, E.; McGuinness, K. Deep Learning for Computer Vision [en línea]. Barcelona: UPC, 2016 [Consulta: 05/09/2016]. Disponible a: <http://imatge-upc.github.io/telecombcn-2016-dlcv/>.

RECURSOS

Otros recursos:

Los detalles de este curso estan disponibles y actualizados en linea en: <https://telecombcn-dl.github.io/2017-dlsl/>