Procesamiento del lenguaje natural

Curso Académico: (2019 / 2020) **Fecha de revisión: 15-05-2020**

Departamento asignado a la asignatura: Departamento de Teoría de la Señal y Comunicaciones

Coordinador/a: MARTÍNEZ OLMOS, PABLO

Tipo: Optativa Créditos ECTS : 3.0

Curso : Cuatrimestre :

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

El curso requiere los siguientes conocimientos básicos:
- Cálculo
- Programación
- Estadística

OBJETIVOS

- Conocer las técnicas básicas de pre-procesado de texto.
- Utilizar herramientas software para pre-procesado de texto.
- Conocer las técnicas de modelado de tópicos.
- Utilizar herramientas software de modelado de tópicos en cuerpos de documentos.
- Utilizar modelos de tópicos para extracción de información.
- Aprender a manejar y entrenar modelos de representación semántica en un espacio vectorial.
- Aprender a entrenar modelos de lenguaje utilizando redes neuronales recursivas.
- Conocer estructuras básicas de traducción basadas en redes neuronales recursivas.
- Utilizar herramientas de optimización para construir modelos de lenguaje con redes neuronales recursivas.

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

- Técnicas de preprocesado de documentos
- Modelos de Tópicos
- Redes Neuronales y Redes Neuronales Recurrentes (RNNs)
- Modelos de lenguaje con RNNs
- Modelos Secuencia a Secuencia y Traducción Automática
- Modelos de Atención

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Se proponen tres tipos de actividades formativas:
- clases de teoría,
- prácticas de laboratorio.

Los créditos ECTS incluyen en todos los casos la parte correspondiente de trabajo personal o en equipo por parte del alumno.

CLASES DE TEORÍA (1.5 ECTS)
Las clases de teoría serán lecciones magistrales en pizarra apoyadas en el uso de transparencias y otros medios audiovisuales (e.g., simulaciones por ordenador). Es importante destacar que estas clases requerirán iniciativa y trabajo personal y en grupo por parte del alumno (habrá conceptos que deberán estudiar personalmente a partir de algunas indicaciones, casos particulares de tendrán que desarrollar, etc.)

PRÁCTICAS (1.5 ECTS)
Dada la naturaleza de los modernos algoritmos de tratamiento de señales, resultan fundamentales para la formación el alumno las prácticas de laboratorio en las que se simulan/implementan (empleando software específico) los algoritmos cuyo diseño y análisis se discute en las clases teórica.
Estas prácticas, por una parte, permiten al estudiante aplicar los conocimientos teóricos adquiridos para ensayar soluciones prácticas, de modo que pueda consolidar y analizar de forma crítica tales conocimientos.

SISTEMA DE EVALUACIÓN

La evaluación continua (80%) se basará en proyectos de programación para implementar las herramientas vistas en clase y una presentación de artículos recientes en el campo.

Habrá un examen final que contará un 20%.

Peso porcentual del Examen Final:	20
Peso porcentual del resto de la evaluación:	80

BIBLIOGRAFÍA BÁSICA

- Cristopher Bishop Pattern Recognition and Machine Learning, Springer, 2006

- Ian Goodfellow and Yoshua Bengio and Aaron Courville Deep Learning, MIT Press, 2017

- Steven Bird, Ewan Klein, Edward Loper Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit, O'Reilly, 2009