uc3m Universidad Carlos III de Madrid

Procesamiento del lenguaje natural con aprendizaje profundo

Curso Académico: (2023 / 2024) Fecha de revisión: 15-05-2023

Departamento asignado a la asignatura: Departamento de Informática

Coordinador/a: SEGURA BEDMAR, ISABEL Tipo: Obligatoria Créditos ECTS: 3.0

Curso: 1 Cuatrimestre: 2

OBJETIVOS

- Diseñar y entrenar arquitecturas de redes neuronales profundas (redes neuronales convolucionales, redes neuronales recurrentes, LSTM, transformadores) y aplicarlas para abordar aplicaciones de Procesamiento de Lenguaje Natural (PLN) como la clasificación de texto, extracción de información, similitud de textos, generación automática de resúmenes y traducción automática.
- Estudiar, implementar y utilizar vectores de palabras entrenados con redes neuronales.

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

- 1. Introducción
- 2. Tareas básicas de PLN.
 - Modelos de word embeddings. Similitud de textos
- 3. Clasificación de Textos con Deep Learning (CNN; RNN, Transformers)
- 4. Extracción de Información.
- 5. Técnicas para el aumento de datos
- 6. Tareas de generación de texto.

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Se aplicará un modelo de clase invertida:

Cada semana, profesora publicará en aula global un Jupyter notebook que describa y contenga la implementación de una arquitectura profunda para la resolución de una determinada aplicación de PLN (por ejemplo, la detección de noticias falsas o la anonimización de notas clínicas). Los+as estudiantes deben estudiar y ejecutar estos notebooks antes de la clase.

Durante la clase, la profesora explicará el notebook y resolverá las posibles dudas. Una vez resueltas, la profesora propondrá propone posibles mejoras y planteará nuevos retos o problemas a resolver. Los estudiantes deben tratar de extender o adaptar las implementaciones de los notebooks para dar solución a los nuevos retos planteados por la profesora.

La metodología es práctica. Las clases tendrán una configuración de trabajo por parejas en aula informática.

El régimen de tutoría es de 2 horas semanales. Estas se podrán realizar de forma online o presencial.

SISTEMA DE EVALUACIÓN

EVALUACIÓN CONTINUA. Se valorarán las actividades planteadas durante las clases.

- Participación en clase.
- Desarrollo sistemas para resolver tareas de PLN (por ejemplo, clasificación de textos) con arquitecturas de deep learning.
- Participar en alguna competición de PLN.
- Escribir un artículo científico (extensión 4-8 páginas). El artículo científico deberá describir las arquitecturas, presentar y discutir los resultados obtenidos.

Peso porcentual del Examen Final:	0
Peso porcentual del resto de la evaluación:	100

BIBLIOGRAFÍA COMPLEMENTARIA

- Rudolph Russell DEEP LEARNING: Fundamentos del Aprendizaje Profundo para Principiantes, CreateSpace Independent Publishing Platform, 2018