Minería de Textos

Curso Académico: (2022 / 2023) **Fecha de revisión: 04/06/2021 14:15:03**

Departamento asignado a la asignatura:

Coordinador/a: SEGURA BEDMAR, ISABEL

Tipo: Optativa Créditos ECTS : 3.0

Curso : 4 Cuatrimestre :

OBJETIVOS

- Comprender los conceptos de Procesamiento de Lenguaje Natural (PLN) y Minería de Textos.
- Conocer las principales tareas de PLN y saber usar las principales librerías de PLN que dan soporte a estas tareas.
- Conocer e implementar los principales modelos de lenguaje para la representación de textos.
- Diseñar e implementar arquitecturas basadas en enfoques de conocimiento y aprendizaje automático para resolver aplicaciones del procesamiento de lenguaje natural.

Competencias básicas establecidas en el artículo 3 sobre competencias RD 1393/2007 modificado por el RD 861/2010 en el que se indica que se garantizarán, como mínimo las siguientes competencias básicas, en el caso del Grado, y aquellas otras que figuren en el Marco Español de Cualificaciones para la Educación Superior, MECES

Resultados del aprendizaje:

R1. Conocimiento y comprensión: Tener conocimientos básicos y la compresión de los fundamentos científicos y tecnológicos de la Ingeniería Informática, así como un conocimiento específicos de las ciencias de la computación, la ingeniería de computadores y sistemas de información.
R5 Aplicaciones de la Ingeniería: Los egresados serán capaces de aplicar su conocimiento y comprensión para resolver problemas, dirigir investigaciones y diseñar dispositivos o procesos del ámbito de la Ingeniería Informática de acuerdo con criterios de coste, calidad, seguridad, eficiencia, respeto por el medioambiente e implicaciones éticas. Estas habilidades incluyen el conocimiento, uso y limitaciones de sistemas informáticos, ingeniería de procesos, arquitecturas de computadores, modelos computacionales, equipos, trabajo práctico, bibliografía técnica y fuentes de información.

Competencias básicas y generales:

CGB3 - Capacidad para comprender y dominar los conceptos básicos de matemática discreta, lógica, algorítmica y complejidad computacional, y su aplicación para la resolución de problemas propios de la ingeniería.
CGO9 - Capacidad para resolver problemas con iniciativa, toma de decisiones, autonomía y creatividad. Capacidad para saber comunicar y transmitir los conocimientos, habilidades y destrezas de la profesión de Ingeniero Técnico en Informática.
CB3 - Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética
CB5 - Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

Competencias específicas:

CECRI5 - Conocimiento, administración y mantenimiento sistemas, servicios y aplicaciones informáticas.
CECRI6 - Conocimiento y aplicación de los procedimientos algorítmicos básicos de las tecnologías informáticas para diseñar

soluciones a problemas, analizando la idoneidad y complejidad de los algoritmos propuestos.
CECRI15 - Conocimiento y aplicación de los principios fundamentales y técnicas básicas de los sistemas inteligentes y su
aplicación práctica.

- CECC4. Capacidad para conocer los fundamentos, paradigmas y técnicas propias de los sistemas inteligentes y analizar, diseñar y construir sistemas, servicios y aplicaciones informáticas que utilicen dichas técnicas en cualquier ámbito de aplicación.
- CECC5. Capacidad para adquirir, obtener, formalizar y representar el conocimiento humano en una forma computable para la resolución de problemas mediante un sistema informático en cualquier ámbito de aplicación, particularmente los relacionados con aspectos de computación, percepción y actuación en ambientes o entornos inteligentes.
- CECC7. Capacidad para conocer y desarrollar técnicas de aprendizaje computacional y diseñar e implementar aplicaciones y sistemas que las utilicen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de grandes volúmenes de datos

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1. Introducción
2. Tareas básicas de PLN.
3. Modelos de Lenguaje para la representación de textos.
4. Clasificación de textos.
5. Reconocimiento de entidades.
6. Extracción de relaciones.
7. Similitud entre textos.

SISTEMA DE EVALUACIÓN

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

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

En particular, se tendrán que implementar cuatro casos prácticos:
1) Implementar y evaluar un prototipo para la clasificación de textos (33%).
2) Implementar y evaluar un prototipo para el reconocimiento de entidades (33%).
3) Implementar y evaluar un prototipo para la extracción de relaciones (33%).