25-26

GUÍA DE ESTUDIO PÚBLICA

TÉCNICAS BASADAS EN GRAFOS APLICADOS AL PROCESAMIENTO DEL LENGUAJE

CÓDIGO 31101305

UNED

25-26

TÉCNICAS BASADAS EN GRAFOS APLICADOS AL PROCESAMIENTO DEL LENGUAJE

CÓDIGO 31101305

ÍNDICE

PRESENTACIÓN Y CONTEXTUALIZACIÓN
REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA
EQUIPO DOCENTE
HORARIO DE ATENCIÓN AL ESTUDIANTE
COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE
RESULTADOS DE APRENDIZAJE
CONTENIDOS
METODOLOGÍA
SISTEMA DE EVALUACIÓN
BIBLIOGRAFÍA BÁSICA
BIBLIOGRAFÍA COMPLEMENTARIA
RECURSOS DE APOYO Y WEBGRAFÍA
IGUALDAD DE GÉNERO

Nombre de la asignatura	TÉCNICAS BASADAS EN GRAFOS APLICADOS AL PROCESAMIENTO DEL LENGUAJE
Código	31101305
Curso académico	2025/2026
Título en que se imparte	MÁSTER UNIVERSITARIO EN TECNOLOGÍAS DEL LENGUAJE
Tipo	CONTENIDOS
Nº ETCS	6
Horas	150
Periodo	ANUAL
Idiomas en que se imparte	CASTELLANO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

En muchas aplicaciones del Procesamiento del Lenguaje Natural las entidades pueden representarse de forma natural como nodos de un grafo y las relaciones entre ellas como enlace entre dichos nodos.
Se ha desarrollado una amplia investigación que ha demostrado que las representaciones basadas en grafos de las unidades lingüísticas, como palabras, oraciones y documentos da lugar a nuevas y eficientes soluciones en numerosas tareas de procesamiento del lenguaje. Estas aplicaciones van desde el etiquetado léxico, el análisis sintáctico y la desambiguación del sentido de las palabras, a la extracción de información, o la generación de resúmenes.
Así mismo, la teoría de redes, tiene una gran relevancia en los sistemas de recuperación información, que ha permitido grandes avances en el desarrollo de aplicaciones en la web, que van desde el desarrollo de buscadores hasta el análisis de redes sociales.
En esta asignatura se dan a conocer los conceptos básicos de las redes y su representación matemática en forma de grafos, así como los algoritmos más utilizados en el procesamiento de grafos. Se presentan y utilizan herramientas existentes para la construcción, análisis y visualización de grafos. A continuación se describen las aplicaciones más importantes de la teoría de grafos a la recuperación de información en la web y al procesamiento del lenguaje natural.
Ficha técnica:

•Tipo: Optativa

•Duración: Anual

•Créditos Totales y Horas: 6 / 150

•Horas de estudio teórico: 70

•Horas de trabajo práctico: 70

•Horas de actividades complementarias: 10

La asignatura de caracter optativo "Técnicas basadas en grafos aplicados al procesamiento del lenguaje" pertenece al Máster Universitario en "Tecnologías del Lenguaje" y se imparte por profesores del Departamento de "Lenguajes y Sistemas Informáticos" de la ETSI Informática de la UNED. Esta asignatura es de carácter anual con una carga de 6 ECTS.
Esta asignatura, al ser optativa, puede ser cursada aisladamente, aunque el estudiante se beneficiaría si hubiera cursado previamente o curse en paralelo la asignatura de

Fundamentos del procesamiento lingüístico.

La asignatura aporta al perfil profesional del máster en varios aspectos fundamentales:

Comprensión avanzada del procesamiento del lenguaje natural (PLN): Los grafos son una representación poderosa para modelar relaciones y estructuras de datos complejas. Al aplicar técnicas basadas en grafos al PLN, los estudiantes desarrollan una comprensión más profunda de las relaciones entre distintos elementos lingüísticos (palabras, oraciones, documentos, etc.).

Desarrollo de habilidades de modelado y visualización: Trabajar con grafos implica la creación, manipulación y visualización de estructuras de datos complejas. Los estudiantes adquieren habilidades en el modelado de problemas del mundo real como grafos, lo que les permite representar de manera efectiva sistemas lingüísticos complejos. Además, aprenden a visualizar y analizar estos modelos, lo que facilita la identificación de patrones y la toma de decisiones informadas en el procesamiento del lenguaje.

Aplicación en sistemas de búsqueda y recomendación: Muchos sistemas de búsqueda y recomendación dependen de modelos basados en grafos para comprender las relaciones entre los elementos del lenguaje. Los estudiantes adquieren habilidades para diseñar y desarrollar algoritmos de búsqueda y recomendación más efectivos, lo que mejora su capacidad para construir sistemas inteligentes que puedan entender y responder a las consultas de los usuarios de manera más precisa.

Preparación para investigaciones avanzadas: el uso de técnicas basadas en grafos representa un área de investigación activa. Al estudiar esta asignatura, los estudiantes están mejor preparados para participar en investigaciones avanzadas en PLN, lo que puede incluir el desarrollo de nuevos algoritmos, la mejora de técnicas existentes y la aplicación de enfoques innovadores para resolver problemas complejos en el procesamiento del lenguaje.

Reseña del Profesorado:

ARAUJO SERNA, LOURDES:

Forma parte del grupo NLP&IR de la UNED. Ha desarrollado en universidades públicas diversa actividad docente relacionada con los lenguajes de programación y la algoritmia. En la actualidad investiga en procesamiento del lenguaje natural, recuperación de información y en su aplicación a diversas áreas como el dominio médico y la educación. Ha dirigido diversas tesis doctorales y proyectos de investigación en estos temas.

e.mail: lurdes@lsi.uned.es

MARTÍNEZ ROMO, JUAN:

Forma parte del grupo NLP&IR de la UNED. Ha impartido docencia en temáticas como los lenguajes de programación, tecnologías web, motores de búsqueda, programación de dispositivos móviles y procesamiento del lenguaje natural (PLN). Sus intereses principales de investigación son el PLN, recuperación de información con adversario, informática biomédica y tecnologías educativas. Ha dirigido varias tesis doctorales y proyectos de

investigación en estos temas.

e.mail: juaner@lsi.uned.es

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

Se requieren conocimiento de algoritmos, programación y matemáticas propios de un titulado en Informática o en un área afin.
Esta asignatura puede ser cursada aisladamente, aunque el estudiante se beneficiaría si hubiera cursado previamente o curse en paralelo la asignatura de *Fundamentos del procesamiento lingüístico*.

EQUIPO DOCENTE

Nombre y Apellidos	M LOURDES ARAUJO SERNA
Correo Electrónico	lurdes@lsi.uned.es
Teléfono	91398-7318
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	LENGUAJES Y SISTEMAS INFORMÁTICOS

Nombre y Apellidos	JUAN MARTINEZ ROMO
Correo Electrónico	juaner@lsi.uned.es
Teléfono	91398-9378
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	LENGUAJES Y SISTEMAS INFORMÁTICOS

HORARIO DE ATENCIÓN AL ESTUDIANTE

La tutorización de los alumnos se llevará a cabo a través de la plataforma de e-Learning Alf, por teléfono y por correo electrónico:

•Lourdes Araujo Serna (coordinadora)
email: lurdes@lsi.uned.es
Tfno: 913987318
Horario: Jueves de 10:00 a 14:00 horas.
ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento: LENGUAJES Y SISTEMAS INFORMÁTICOS
calle Juan del Rosal 16, Madrid 28040

•Juan Martínez Romo
email: juaner@lsi.uned.es
Tfno: 913989378
Horario: Miércoles y Jueves de 11:00 a 13:00 horas.
ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA

Departamento: LENGUAJES Y SISTEMAS INFORMÁTICOS
calle Juan del Rosal 16, Madrid 28040

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

COMPETENCIAS

C1 Comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.
C2 Abstracción, análisis, síntesis y relación de ideas.
C3 Capacidad crítica y de decisión.
C4 Capacidad de estudio y autoaprendizaje
C5 Capacidad creativa y de investigación.
C6 Habilidades sociales para el trabajo en equipo
C7 Capacidad de estudio de los sistemas y aproximaciones existentes y para distinguir las aproximaciones más efectivas.
C8 Capacidad para detectar carencias en el estado actual de la ciencia y la tecnología.
C9 Capacidad para proponer nuevas aproximaciones que de solución a las carencias detectadas.

RESULTADOS DE APRENDIZAJE

CONOCIMIENTOS O CONTENIDOS

CO1 Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
CO2 Capacidad de comprender y manejar de forma básica los aspectos más importantes relacionados con los lenguajes y sistemas informáticos en general, y, de manera especial, en los siguientes ámbitos: Tecnologías del lenguaje y de acceso a la información en web.

HABILIDADES O DESTREZAS

H1 Capacidad de aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios relacionados con su área de estudio.
H2 Capacidad de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.
H3 Poseer las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.
H4 Capacidad de especificar, diseñar, implementar y evaluar tanto cualitativa como cuantitativamente los modelos y sistemas propuestos.
H5 Capacidad para proponer y llevar a cabo experimentos con la metodología adecuada como para poder extraer conclusiones y determinar nuevas líneas de actuación e

investigación.

COMPETENCIAS

C1 Comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.
C2 Abstracción, análisis, síntesis y relación de ideas.
C3 Capacidad crítica y de decisión.
C4 Capacidad de estudio y autoaprendizaje.
C5 Capacidad creativa y de investigación.
C6 Habilidades sociales para el trabajo en equipo.
C7 Capacidad de estudio de los sistemas y aproximaciones existentes y para distinguir las aproximaciones más efectivas.
C8 Capacidad para detectar carencias en el estado actual de la ciencia y la tecnología.
C9 Capacidad para proponer nuevas aproximaciones que de solución a las carencias detectadas.

CONTENIDOS

Introducción

Es un tema es introductorio y motiva al estudio de la asignatura con diversos ejemplos.

Definiciones y conceptos básicos

En este capítulo se proporciona una introducción a las redes que pueden representar las relaciones entre elementos del mundo real. Para ello se estudian los distintos tipos de redes y sus propiedades principales. Se describen los modelos más importantes de redes y se presentan ejemplos.

Algoritmos para grafos

En este capítulo se presentan diversos algoritmos para grafos. En primer lugar se describen formas alternativas de representación de grafos y su relación con el tamaño de los grafos.
Se presentan algoritmos para distintos tipos de recorridos de los grafos, para la identificación del árbol de recubrimiento y para la búsqueda del camino mínimo.
Se presentan también algoritmos de partición de grafos, de identificación de comunidades y de paseos aleatorios (o random walks).

Algoritmos basados en Grafos para Recuperación de Información

En este capítulo se describe la aplicación de grafos a problemas relacionados con la recuperación de información en la web. Se presentan herramientas de visualización de grafos. Se describen dos de los algoritmos más utilizados para establecer rankings entre páginas web: Hits y Pagerank. Se describen también algunos algoritmos de clustering y clasificación de textos basados en grafos.

Algoritmos basados en grafos para procesamiento de lenguaje natural.

En este capítulo se describe con se aplican los grafos en distintos problemas fundamentales del procesamiento del lenguaje: etiquetado léxico, análisis sintáctico, desambiguación del sentido de las palabras, inducción del sentido de las palabras, identificación de temas y pesado de términos.

METODOLOGÍA

La metodología es la general del programa de máster; junto a las actividades y enlaces con fuentes de información externas, existe material didáctico propio preparado por el equipo docente. Se trata de una metodología adaptada a las directrices del EEES, de acuerdo con el documento del IUED.

La asignatura no tiene clases presenciales. Los contenidos teóricos se impartirán a distancia, de acuerdo con las normas y estructuras de soporte telemático de la enseñanza en la UNED. El material docente incluye un resumen de los contenidos de cada tema y distintos tipos de actividades relacionadas con la consulta bibliográfica y la utilización de herramientas asociadas a las tecnologías y técnicas presentadas en el curso. Tratándose de un master orientado a la investigación, las actividades de aprendizaje se estructuran en torno al estado del arte en cada una de las materias del curso y a los problemas en los que se van a focalizar las prácticas que el alumno deberá realizar.

SISTEMA DE EVALUACIÓN

TIPO DE PRIMERA PRUEBA PRESENCIAL

Tipo de examen — No hay prueba presencial

TIPO DE SEGUNDA PRUEBA PRESENCIAL

Tipo de examen2 — No hay prueba presencial

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad: No

Descripción

No hay prueba final. La evaluación se basa en la realización de diversas tareas.

Criterios de evaluación

Ponderación de la prueba presencial y/o los trabajos en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

PRUEBAS DE EVALUACIÓN CONTINUA (PEC)

¿Hay PEC?: Si,PEC no presencial

Descripción

Las actividades prácticas programadas dentro de la asignatura son:
Análisis y visualización de grafos.
Identificación de comunidades
Comparativa de los algoritmos Hits y Pagerank
Extracción de palabras clave en documentos mediantes técnicas basadas en grafos
La estructura y características de cada entrega se detallan en los enunciados.
En la evaluación se tiene en cuenta que se cumplan todos los requisitos indicados en el enunciado y también las aportaciones adicionales que pueda hacer el estudiante sobre el tema tratado.
Las fechas de entrega se reparten a partir de enero, teniendo que entregar un trabajo aproximadamente entre cada mes y mes y medio.
Las prácticas opcionales se entregan a principio de junio.

Criterios de evaluación

La realización correcta y en plazo de las cuatro las tareas obligatorias asegurará una nota de APROBADO en la asignatura. La nota máxima que se puede obtener con las tareas obligatorias, dependiendo de la calidad de las entregas, es de NOTABLE.
Notas superiores se obtienen con las tareas opcionales asociados a los 3 últimos temas.

Ponderación de la PEC en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

Las prácticas obligatorias tienen un plazo de entrega fi jo que se establece en cada enunciado. La fecha de entrega de la primera tarea es en enero. La fecha de la última tarea obligatoria es a mediados de mayo. Los estudiantes disponen del tiempo restante para corregir las tareas que no hayan aprobado y para realizar las tareas opcionales si lo desean.
Los estudiantes que no entreguen las tareas en el plazo establecido para la convocatoria de junio tendrán otro plazo de entrega a principios de septiembre para la convocatoria extraordinaria.

OTRAS ACTIVIDADES EVALUABLES

¿Hay otra/s actividad/es evaluable/s? No

Descripción

Criterios de evaluación

Ponderación en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

¿CÓMO SE OBTIENE LA NOTA FINAL?

En esta asignatura se realiza evaluación continua. La evaluación se llevará a cabo a partir de las siguientes pruebas:
Realización de una serie de tareas obligatorias a lo largo del curso.
Realización de alguna tarea opcional para aumentar la calificación.
Aquellos alumnos que deseen obtener una mayor calificación, podrán elegir uno de entre los trabajos optativos que se irán proponiendo. En estos casos la calificación final dependerá de la calidad del trabajo realizado.
Las tareas obligatorias se deberán entregar en los plazos que se vayan indicando. La no entrega de las tareas en el plazo previsto supondrá suspender. El trabajo optativo para subir nota también tendrá una fecha límite de entrega.
Si se han aprobado todas las tareas obligatorias, la nota final se calcula como la media de la nota de dichas tareas ponderada por 0.8. Las tareas opcionales pueden subir la nota hasta 10 en función de la calidad del trabajo.

BIBLIOGRAFÍA BÁSICA

ISBN(13): 9780521896139
Título: GRAPH-BASED NATURAL LANGUAGE PROCESSING AND INFORMATION RETRIEVAL primera edición
Autor/es: Dragomir Radev; Rada Mihalcea
Editorial: CAMBRIDGE UNIVERSITY PRESS

Este libro aborda los temas centrales del curso. Sin embargo no es imprescindible adquirirlo. En el curso virtual pueden encontrar la información necesaria para seguir el curso.

BIBLIOGRAFÍA COMPLEMENTARIA

ISBN(13): 9780521195331
Título: NETWORKS, CROWDS, AND MARKETS 2010 edición
Autor/es: Jon Kleinberg; David Easley
Editorial: : CAMBRIDGE UNIVERSITY

Libro disponible en la página de uno de los autores

https://www.cs.cornell.edu/home/kleinber/networks-book/

RECURSOS DE APOYO Y WEBGRAFÍA

La plataforma de e-Learning Alf proporcionará el adecuado interfaz de interacción entre el alumno y sus profesores. Alf es una plataforma de e-Learning y colaboración que permite impartir y recibir formación, gestionar y compartir documentos, crear y participar en comunidades temáticas, así como realizar proyectos online.
Se ofrecerán las herramientas necesarias para que, tanto el equipo docente como el alumnado, encuentren la manera de compaginar tanto el trabajo individual como el aprendizaje cooperativo.

IGUALDAD DE GÉNERO

En coherencia con el valor asumido de la igualdad de género, todas las denominaciones que en esta Guía hacen referencia a órganos de gobierno unipersonales, de representación, o miembros de la comunidad universitaria y se efectúan en género masculino, cuando no se hayan sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el sexo del titular que los desempeñe.