

26-27

GUÍA DE ESTUDIO PÚBLICA



ACCESO INTELIGENTE A LA INFORMACIÓN

CÓDIGO 31101019

UNED

26-27

ACCESO INTELIGENTE A LA INFORMACIÓN
CÓDIGO 31101019

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	ACCESO INTELIGENTE A LA INFORMACIÓN
Código	31101019
Curso académico	2026/2027
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

Tipo	Optativa
Cuatrimestre	Anual
Créditos/horas totales	6/150
Horas de estudio teórico	75
Horas de prácticas	75
Horas complementarias	0

Esta asignatura tiene por objetivo estudiar la Recuperación de Información y la Búsqueda de Respuestas como dos grandes áreas que tratan sobre el acceso a la información desde dos premisas diferentes. Para cada una de ellas se introduce su problemática y se aborda desde una perspectiva histórica hasta la irrupción de las técnicas basadas en los últimos modelos de IA Generativa.

Reseña del Profesorado:

LÓPEZ OSTENERO, FERNANDO

Fernando López es Doctor en Ingeniería Industrial por la UNED y Profesor Colaborador en el Departamento de Lenguajes y Sistemas Informáticos de la UNED, donde forma parte del grupo de investigación NLP&IR de la UNED.

Su investigación en el ámbito del Procesamiento del Lenguaje Natural se centra en la recuperación de información y ha participado en la organización de varias tareas de evaluación internacional.

BENITO SANTOS, ALEJANDRO

Alejandro Benito-Santos es Doctor en Ingeniería Informática por la Universidad de Salamanca y Profesor Ayudante Doctor en el Departamento de Lenguajes y Sistemas Informáticos de la UNED, donde forma parte del grupo NLP &IR.

Su investigación se centra en el Procesamiento del Lenguaje Natural y la Recuperación de Información, con especial interés en la evaluación de modelos de lenguaje, la similitud semántica y la interpretabilidad. Ha publicado en foros internacionales de referencia como ACL, EMNLP y Knowledge-Based Systems.

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

Para realizar la asignatura es necesario tener un nivel de lectura fluido del inglés. 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	FERNANDO LOPEZ OSTENERO (Coordinador/a de asignatura)
Correo Electrónico	flopez@lsi.uned.es
Teléfono	91398-7793
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	LENGUAJES Y SISTEMAS INFORMÁTICOS

Nombre y Apellidos	ALVARO RODRIGO YUSTE
Correo Electrónico	alvarory@lsi.uned.es
Teléfono	91398-9693
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	LENGUAJES Y SISTEMAS INFORMÁTICOS

Nombre y Apellidos	ALEJANDRO BENITO SANTOS
Correo Electrónico	al.benito@lsi.uned.es
Teléfono	91398-6484
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	LENGUAJES Y SISTEMAS INFORMÁTICOS

HORARIO DE ATENCIÓN AL ESTUDIANTE

La tutorización de los estudiantes se llevará a cabo a través de la plataforma de enseñanza virtual de posgrado de la UNED.

D. Fernando López Ostenero: Asistencia al estudiante lunes de 10:00 a 14:00 horas.

D. Alejandro Benito Santos: Asistencia al estudiante martes de 16:00 a 20:00 horas.

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

Tema 1: Recuperación de Información.

1. Definición de la tarea
2. Recuperación basada en el léxico
 - Sistemas de pesado
 - Indexación
3. Evaluación
 - Métricas de evaluación
 - Colecciones
4. Recursos dependientes del idioma
 - Lematizadores
 - Stemmers
 - Segmentación de compuestos y palabras
5. Recuperación de Información multilingüe
 - Recursos para Recuperación de Información multilingüe
 - Enfoques para la traducción
 - Fusión documental
6. Evolución de la investigación en el área.

Tema 2: Recuperación de Información Semántica.

1. Limitaciones del matching léxico
2. Representaciones semánticas
 - Representaciones distribuidas
 - Embeddings de palabras
 - Embeddings de frases y documentos
3. Recuperación basada en embeddings
 - Recuperación densa

- Búsqueda vectorial
- Indexación semántica
- 4. Combinación de enfoques léxicos y semánticos
- Recuperación híbrida
- Reranking
- Fusión de resultados
- 5. Evolución reciente de la recuperación semántica

Tema 3: Búsqueda de Respuestas.

1. Definición de la tarea
2. Evaluación de sistemas de Question Answering
 - Métricas de evaluación
 - Análisis de errores
 - Colecciones y foros de evaluación
3. Arquitectura de sistemas de Question Answering
 - Análisis de preguntas
 - Recuperación de documentos
 - Extracción o generación de respuestas
 - Validación de respuestas
4. Clasificación de Sistemas de búsqueda automática de respuestas
 - Nivel de utilización de técnicas de PLN
 - Taxonomía de Moldovan
5. Tipos de preguntas y respuestas
 - Clasificación de los diferentes tipos y subtipos de preguntas
6. Enfoques anteriores al Deep Learning
7. Enfoques basados en Deep Learning
 - Modelos tipo encoder para QA: Bert, RoBERTa, etc.

Tema 4: Sistemas de Búsqueda Automática de Respuestas.

1. Modelos generativos para acceso a la información
 - Modelos seq2seq
 - Modelos decoder-only
 - Large Language Models como sistemas de acceso a la información
 - Capacidades y limitaciones de los modelos generativos
2. Fundamentos de Retrieval-Augmented Generation (RAG)
 - Motivación y arquitectura general

- Integración de recuperación y generación
 - RAG frente a generación sin recuperación
 - Casos de uso de los sistemas RAG
3. Construcción del contexto en sistemas RAG
- Segmentación documental
 - Selección y ordenación de pasajes
 - Ventana de contexto
 - Query parsing y reformulación de consultas
 - Generación condicionada por evidencia
 - Diseño de sistemas de acceso generativo
 - Prompt augmentation
 - Grounding y atribución de evidencia
 - Control de la generación
 - Limitaciones derivadas del contexto recuperado
4. Evaluación de sistemas generativos de acceso a la información
- Limitaciones de las métricas clásicas de Question Answering
 - Evaluación automática y evaluación humana
 - Faithfulness, factualidad y grounding
 - Análisis de errores en sistemas RAG
5. Sistemas RAG en entornos reales
- Latencia, coste y escalabilidad
 - Actualización de la base de conocimiento
 - Monitorización y logging
 - Retos de robustez y reproducibilidad
6. Evolución reciente del acceso generativo a la información

METODOLOGÍA

Los estudiantes realizan dos tipos de actividades en esta asignatura: una tarea para cada uno de los temas que componen la asignatura. Todas ellas estarán disponibles dentro del material de estudio correspondiente a cada tema.

Para finalizar, se realizará una práctica final sobre las líneas de investigación tratadas en la asignatura.

Todo el material de estudio está disponible en el entorno virtual del posgrado, y toda la interacción entre profesores y estudiantes se puede llevar a cabo en este entorno.

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

En esta asignatura no se realiza un examen presencial.

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

En cada tema de la asignatura habrá que realizar una tarea sobre los contenidos del tema.

El enunciado de cada tarea será específico para cada curso y estará disponible en el curso virtual.

Criterios de evaluación

Los criterios de evaluación dependerán de cada tarea. Estarán disponibles en el curso virtual de la asignatura.

Ponderación de la PEC en la nota final La tarea de cada tema supondrá un 15% de la calificación final, por lo que la calificación de todas ellas supone un 60% de la calificación de la asignatura.

Fecha aproximada de entrega Las fechas aproximadas de entrega de cada tarea se encuentran indicadas en el plan de trabajo de la asignatura.

Comentarios y observaciones

OTRAS ACTIVIDADES EVALUABLES

¿Hay otra/s actividad/es evaluable/s? Si,no presencial

Descripción

Existe una práctica final de la asignatura. El enunciado de esta práctica estará disponible en el curso virtual de la asignatura.

Criterios de evaluación

Ponderación en la nota final La práctica final supondrá un 40% de la calificación final.

Fecha aproximada de entrega

Las fechas aproximadas de entrega se detallarán en el curso virtual

Comentarios y observaciones

¿CÓMO SE OBTIENE LA NOTA FINAL?

Para aprobar la asignatura se deberán entregar todas las tareas de los temas, más la práctica final. Los pesos de las tareas de los temas y la práctica final serán los siguientes:

Tema 1: 15%

Tema 2: 15%

Tema 3: 15%

Tema 4: 15%

Práctica Final: 40%

BIBLIOGRAFÍA BÁSICA

Dentro del curso virtual se aportarán referencias a lecturas que cubren los objetivos marcados en cada tema.

BIBLIOGRAFÍA COMPLEMENTARIA

Dentro del curso virtual se aportarán referencias a lecturas que cubren los objetivos marcados en cada tema.

RECURSOS DE APOYO Y WEBGRAFÍA

El campus virtual de posgrados de la UNED, proporcionará interfaz de interacción entre el estudiante y sus profesores. La plataforma de Cursos Virtuales de la UNED permite impartir y recibir formación, gestionar y compartir documentos, crear y participar en comunidades temáticas, así como realizar proyectos online.

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.