

FICHA DE LA ASIGNATURA:



MÁSTER
Lingüística y
Tecnologías

Semántica y Pragmática Computacional Optativa-4
créditos-

Equipo docente

Ana M^a Fernández-Pampillón Cesteros

Área de conocimiento: Lingüística General, Lingüística Computacional
apampi@ucm.es

Objetivos didácticos

El objetivo básico es conocer los fundamentos de la formalización de los niveles semántico y pragmático de la lengua, así como saber construir y evaluar herramientas de análisis lingüístico correspondientes a dichos niveles y utilizar y adaptar a nuevos objetivos las herramientas y los recursos ya existentes.

Este objetivo se desgrana en los siguientes objetivos específicos:

1. Conocer y saber utilizar los principales modelos y métodos de representación y procesamiento del lenguaje natural en sus niveles fonético-fonológico, léxico, morfológico, sintáctico, semántico, pragmático y discursivo.
2. Conocer y saber seleccionar críticamente el modelo de representación del conocimiento más adecuado a un problema o aplicación de la Lingüística computacional.
3. Saber seleccionar críticamente el modelo de representación del conocimiento más adecuado a un problema o aplicación de la Lingüística computacional.
4. Saber aplicar soluciones software para construir aplicaciones de la Lingüística computacional.
5. Saber utilizar soluciones (bibliotecas o repositorios) consolidadas para el procesamiento del lenguaje natural, la generación de representaciones gráficas, acceso y extracción de datos de la web y el almacenamiento y gestión de versiones de proyectos software.
6. Saber enumerar, definir y aplicar a tareas sencillas de procesamiento del lenguaje los principales modelos de redes neuronales para el Procesamiento del Lenguaje Natural (PLN).
7. Saber explicar cómo pueden contribuir las aplicaciones de la Lingüística computacional para mejorar la vida de las personas y para facilitar el progreso científico y tecnológico en beneficio de la humanidad.

Conocimientos y/o destrezas previas recomendadas

- Conocimiento de Python, gestión web y control de versiones
- capacidad de leer en inglés

- capacidad de autoaprendizaje y planificación
- capacidad de trabajar en equipo: “si quieres ir rápido, ve solo; si quieres ir lejos, ve acompañado”
- capacidad crítica: no dar nada por supuesto
- capacidad de aprender del error

Posibles asignaturas del Máster directamente relacionadas con ésta:

Enfoques Avanzados para la Lingüística Computacional
Aplicaciones de la Lingüística Computacional

Carga de trabajo/estudio prevista por semana para el alumno

6-8 horas de dedicación semanal, repartidas entre el estudio del material docente (35%), el análisis de material adicional (10%), la resolución de casos prácticos (35%), la comunicación con el tutor y los compañeros (10%), la asistencia a seminarios presenciales o en línea (10%).

Descripción general de la asignatura

Para responder a la cuestión de la comunicación hombre-máquina en lenguaje natural, la idea es, en primer lugar, conseguir representar computacionalmente (en una máquina) el conocimiento que tienen las personas en su cerebro y, en segundo lugar, conseguir que con ese conocimiento las máquinas puedan procesar, es decir entender y producir, el lenguaje natural. Esta asignatura aborda la formalización de los niveles semántico y pragmático del lenguaje natural.

De forma más concreta, se tratarán los cinco temas siguientes:

1. Representación del conocimiento
2. Representaciones basadas en la lógica
3. Representaciones basadas en estructuras de ranura-relleno
4. Representaciones basadas en semántica vectorial
5. Análisis semántico
6. Análisis del discurso

Estos temas se distribuyen en el siguiente cronograma temporal:

	Contenidos	Sesión síncrona
	Tema 1. Representación del conocimiento Tema 2. Representaciones basadas en la lógica	1
	Tema 2. Representaciones basados en de estructuras ranura-relleno	2
	Tema 3. Representaciones basadas en estructuras de ranura-relleno	3
	Tema 4. Representaciones basadas en semántica vectorial	4
	Tema 5. Análisis semántico	5
	Tema 6. Análisis del Discurso	6
	Sesión presencial	

Breve descripción de la Metodología(s) de aprendizaje(s) que se prevé utilizar:

El curso se impartirá en el Campus Virtual de la UCM. Se organizará en temas. Cada tema tendrá una duración aproximada de dos semanas y una carga de trabajo semanal de 6-8 horas. Cada semana se hará una propuesta temática de trabajo y se trabajará en tres partes:

1era. parte: presentación de la propuesta temática para que sirva de guía y punto de partida para el estudio.

2ª parte: estudio y actividades de consolidación, se realizarán a lo largo de la semana por parte de los alumnos y con la guía de los profesores. Consiste, básicamente, en adquirir los conceptos básicos del tema, realizar ejercicios en papel y plasmar este trabajo en un documento de “apuntes de Semántica y Pragmática computacional”.

3ª parte: resolución de casos prácticos, consiste en la resolución de un caso práctico propuesto. Implica el diseño en papel, la implementación en Python y la evaluación de su efectividad.

Las tareas que se realizarán cada tema son:

- Asistencia clase síncrona en las sesiones previstas en el cronograma
- Realización de los ejercicios, tests y casos prácticos planteados
- Síntesis del tema en el documento “Apuntes de Semántica y Pragmática computacional”

Se podrán realizar tutorías síncronas en la que el profesor atenderá las dudas que se planteen en foro. Para ello el alumno o los alumnos deben solicitarlas a través del foro, de forma que a la tutoría pueden asistir todos los alumnos que lo deseen. Después de resolver las dudas del alumno(s) que solicitó la tutoría, el resto de alumnos puede continuar planteando sus dudas.

Enumeración de las actividades de aprendizaje que se prevén utilizar para las sesiones virtuales:

Se utilizarán las actividades previstas en la memoria del título para aplicar los conocimientos adquiridos en la asignatura:

- Estudio individual del material básico
- Lectura y análisis de material complementario
- Resolución de casos prácticos
- Comunicación virtual con el profesor
- Foros y comunicación colaborativa
- Seminarios presenciales

Enumeración de las actividades de aprendizaje que se prevén utilizar para las sesiones presenciales:

- Resolución de casos teóricos o prácticos
- Resolución de dudas de los contenidos propuestos por el profesor

Procedimiento de evaluación

La asignatura se evalúa a partir de las actividades siguientes:

- Actividades evaluables del tema: 50% de la nota final
- Examen final presencial: 50% nota final. *Es imprescindible aprobar el examen final presencial para aprobar la asignatura*

Competencias

Las competencias específicas de la titulación, enfocadas a los niveles semántico y pragmático y al procesamiento de texto. Específicamente se desarrollarán las siguientes competencias:

CE8. Conocer los modelos y métodos más relevantes (simbólicos, estadísticos y biológicos) para el procesamiento automático de las lenguas y los fenómenos lingüísticos en sus diferentes niveles lingüísticos: fonético-fonológico, morfológico, sintáctico, semántico, pragmático y de discurso.

CE9. Saber describir las características del lenguaje humano que hacen difícil su tratamiento automático.

CE10. Saber utilizar con suficiente destreza los principales paquetes de programación para resolver tareas de procesamiento del lenguaje natural en formato texto.

CE11. Conocer el manejo de las herramientas software existentes para el procesamiento de las producciones lingüísticas en diferentes lenguas (ej. segmentadores, analizadores morfológicos, sintácticos, semánticos). CE12. Conocer los fundamentos teóricos y de implementación de las aplicaciones existentes de Lingüística Computacional (ej. agentes conversacionales, recuperación de información, extracción de entidades nombradas o generación de resúmenes).
Procedimiento para mostrar el progreso del alumno: Boletín de calificaciones de la asignatura virtual
Mecanismos de comunicación docente Foro, correo electrónico de la asignatura virtual y videoconferencia
Mecanismos de tutorización virtual: Foro, correo electrónico en la asignatura virtual y videoconferencia
Mecanismos de contacto -Foro de la asignatura virtual que podrá complementarse cuando sea necesario con sesiones de videoconferencia. Lo atenderán los profesores de la asignatura. -Correo electrónico institucional de los profesores en caso de que no tenga acceso al campus virtual -Además el alumno dispone de un servicio de ayuda para las incidencias informáticas de la Universidad en https://ssii.ucm.es/estudiante
Mecanismos de contacto para quejas y sugerencias de la asignatura: - El alumno debe presentar su queja, en primer lugar, al profesor, y hacerle cuantas sugerencias considere oportunas sobre la asignatura, tanto por vía de correo electrónico como por un buzón anónimo de “quejas y sugerencias” en la página de Presentación de la asignatura virtual. - El alumno se podrá dirigir también al Coordinador del Máster, así como al representante de alumnos en caso de que su queja o sugerencia no sea atendida. - Además, el máster dispone de un buzón de quejas y sugerencias en su página web atendido por el Coordinador del máster
Mecanismos para recoger la opinión de los alumnos sobre la asignatura: Participación en el programa Docencia de la UCM complementado con una encuesta anónima preparada por los profesores en la asignatura virtual
Requisitos técnicos especiales (no de campus virtual): No hay (sólo ordenador de gama media con conexión a internet)

Bibliografía básica:

JURAFSKY, D.; MARTIN J.H. (2009) Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. 3th Edition (borrador de la 3ª edición 2024 en: <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>)
 RICH, E.; KNIGHT, K. (1994). Inteligencia artificial (2a. ed.). McGraw-Hill. <https://ucm.on.worldcat.org/oclc/434069471>

Biblografía complementaria

EISENSTEIN, J. (2019) Introduction to Natural Language Processing. The MIT Press (versión borrador de 2018 en: <https://cseweb.ucsd.edu/~nnakashole/teaching/eisenstein-nov18.pdf>)
 GOLBERG, Y. (2015) A Primer on Neural Network Models for Natural Language Processing. <https://arxiv.org/abs/1510.00726>