19-20

MÁSTER UNIVERSITARIO EN LAS TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN EN LA ENSEÑANZA Y EL TRATAMIENTO DE LENGUAS

GUÍA DE ESTUDIO PÚBLICA



PROCESAMIENTO DEL LENGUAJE NATURAL

CÓDIGO 24409117

UNED

19-20

PROCESAMIENTO DEL LENGUAJE NATURAL

CÓDIGO 24409117

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

ADENDA AL SISTEMA DE EVALUACIÓN CON MOTIVO DE LA PANDEMIA COVID 19

Nombre de la asignatura	PROCESAMIENTO DEL LENGUAJE NATURAL
Código	24409117
Curso académico	2019/2020
Título en que se imparte	MÁSTER UNIVERSITARIO EN LAS TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN EN LA ENSEÑANZA Y EL TRATAMIENTO DE LENGUAS
Tipo	CONTENIDOS
Nº ETCS	5
Horas	125.0
Periodo	SEMESTRE 1
Idiomas en que se imparte	CASTELLANO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

La convergencia entre la lingüística y la informática ha resultado en un nuevo perfil profesional e investigador que requiere una formación específica. En sentido estricto, cualquier actividad que implique un análisis o generación de la lengua utilizando el ordenador puede considerarse como procesamiento del lenguaje natural (PLN). No obstante, mientras los lingüistas prefieren hablar de la lingüística computacional como un área de conocimiento dentro de la lingüística aplicada, los informáticos consideran el procesamiento del lenguaje natural como una rama de la inteligencia artificial debido a la posibilidad de desarrollar sistemas de computación que simulen algún aspecto de la capacidad lingüística del ser humano. En este contexto, Procesamiento del Lenguaje Natural es una asignatura teórico-práctica dirigida principalmente a lingüistas y que se centra en dos grandes áreas temáticas:

1. Los fundamentos teóricos básicos del enfoque simbólico en la modelización del conocimiento lingüístico para la construcción de sistemas computacionales que analicen y/o generen textos en lenguaje natural;

2. Las principales aplicaciones de este campo de investigación que ayuden a derribar las barreras comunicativas en la sociedad multilingüe de la información y el conocimiento.

Procesamiento del Lenguaje Natural es una asignatura de 5 créditos ECTS que pertenece al módulo de contenidos formativos propios del máster *Las TIC en la enseñanza y el tratamiento de las lenguas*. La asignatura no pretende que los alumnos desarrollen aplicaciones informáticas, sino que trata de dotar a los lingüistas del conocimiento necesario para empezar a trabajar conjuntamente con informáticos en proyectos de ingeniería lingüística. En general, esta asignatura posibilita la formación tanto de profesionales que deseen organizar y coordinar proyectos lingüísticos dentro del marco de las tecnologías del lenguaje, como de investigadores que deseen realizar avances científicos en alguno de los aspectos teóricos del procesamiento del lenguaje natural. Además, permite a los estudiantes conocer a fondo FunGramKB (www.fungramkb.com), una base de conocimiento léxico-conceptual diseñada para tareas de PLN, en especial aquellas relacionadas con el procesamiento de texto, como la traducción automática, la extracción y recuperación de

información, etc.

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

A pesar de que *Procesamiento del Lenguaje Natural* no está orientada hacia el tratamiento informatizado de una lengua concreta, el alumno deberá ser capaz de leer y comprender con relativa facilidad textos en inglés, ya que muchas de las lecturas propuestas en esta asignatura están escritas en esta lengua.
Igualmente, es muy recomendable que el alumno posea unos sólidos conocimientos sobre aspectos básicos de lingüística teórica, en concreto sobre semántica y sintaxis.
Por otra parte, esta asignatura no exige al alumno conocimiento alguno en programación informática, aunque se le iniciará en la representación formal del conocimiento.

EQUIPO DOCENTE

Nombre y Apellidos	ANSELMO PEÑAS PADILLA (Coordinador de asignatura)
Correo Electrónico	anselmo@lsi.uned.es
Teléfono	91398-7750
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	LENGUAJES Y SISTEMAS INFORMÁTICOS

Nombre y Apellidos	ALBERTO PEREZ GARCIA-PLAZA
Correo Electrónico	alberto.perez@lsi.uned.es
Teléfono	91398-8412
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	LENGUAJES Y SISTEMAS INFORMÁTICOS

HORARIO DE ATENCIÓN AL ESTUDIANTE

Se podrá contactar con el equipo docente de la asignatura a través de los siguientes medios:
•Correo electrónico: pedrou@invi.uned.es (canal de comunicación preferente)
•Número de teléfono: (+34) 91 398 6819 // 6006
•Foros de la plataforma ALF
•Dirección postal: Departamento de Filología Extranjera y su Lingüistica Facultad de Filología; Senda del Rey, 7 28040 Madrid
•Guardias: Miércoles (10:00 - 14:00) y Jueves (15:00 - 19:00)
El horario de atención telefónica al alumno se indicará en la plataforma virtual a principio del curso.

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

Competencias Básicas:

CB6 - 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

CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CB8 - Que los estudiantes sean capaces 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

CB9 - Que los estudiantes sepan 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

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

Competencias Generales:

CG04 - Adquirir competencias cognitivas superiores en análisis y síntesis

CG05 - Ser capaz de aplicar los conocimientos a la práctica

CG06 - Resolver problemas en entornos nuevos o poco conocidos

CG12 - Adquirir competencias en el uso de las TIC

CG14 - Desarrollar competencias en la gestión y organización de la información y la documentación

CG15 - Desarrollar competencias en la recolección y análisis de datos, el manejo de bases de datos y su presentación

CG20 - Adquirir compromiso ético (por ejemplo, en la realización de trabajos sin plagios, etc.)

CG21 - Adquirir ética profesional (abarca también la ética como investigador)

Competencias Específicas:

CE07 - Evaluar el trabajo cooperativo entre lingüistas e informáticos dentro de la lingüística computacional.

CE11 - Diseñar e implementar un programa y un informe de investigación para resolver una tarea de extracción de información y generación de resúmenes.

CE12 - Diseñar e implementar un trabajo de investigación relativo a la red semántica que esté basado en la utilización de las TIC.

CE13 - Desarrollar las destrezas analíticas necesarias para la construcción de un corpus en formato digital.

CE14 - Adquirir las competencias necesarias para manejar de forma eficaz las herramientas computacionales para el análisis léxico.

CE15 - Reflexionar sobre los modelos segmentales y prosódicos adecuados para los

módulos lingüísticos en los proyectos de tecnologías del habla.
CE16 - Identificar sus propias carencias formativas y sus necesidades de especialización individual en metodologías fonéticas como la transcripción y notación de corpus.
CE17 - Familiarizarse con las herramientas disponibles para el traductor profesional de forma aislada e integrada en la mesa del traductor.
CE18 - Reconocer la importancia del uso de las TIC en el estudio de la diversidad lingüística y la multiculturalidad.
CE19 - Apreciar, manejar y combinar las diferentes técnicas de investigación según el tipo de corpus con el que se trabaje.
CE20 - Diseñar e implementar un trabajo de investigación dentro de la lingüística diacrónica que esté basado en la utilización de las TIC.

RESULTADOS DE APRENDIZAJE

Al finalizar el curso, el alumno será capaz de:

1. Analizar los principales logros del procesamiento del lenguaje natural.
2. Identificar y exponer los principales desafíos a los que se enfrenta actualmente el investigador en el procesamiento del lenguaje natural.
3. Evaluar el trabajo cooperativo entre lingüistas e informáticos en el campo del procesamiento del lenguaje natural.
4. Justificar el papel del lingüista en un proyecto de ingeniería lingüística.
5. Describir la formación teórica que precisa el investigador en procesamiento del lenguaje natural.
6. Reflexionar de forma crítica sobre el contenido de artículos científicos relacionados con el procesamiento del lenguaje natural.
7. Redactar un proyecto de investigación en del campo de la ingeniería lingüística

CONTENIDOS

1. Introducción al Procesamiento del Lenguaje Natural

2. Recursos lingüísticos para el PLN

3. Tratamiento computacional de la sintaxis

4. Tratamiento computacional de la semántica

5. Recuperación y extracción de información

6. Minería de opinión

METODOLOGÍA

La asignatura adoptará una metodología de enseñanza a distancia que garantice la interacción entre el profesor y la los miembros de la clase, y así favorezcer el trabajo autónomo. Con el fin de consolidar el proceso de aprendizaje, esta metodología se centrará tanto en los contenidos teóricos mediante la lectura de artículos científicos, como en el desarrollo de actividades prácticas basadas en dichas lecturas. Más concretamente, hemos asignado a cada tema de la asignatura una serie de lecturas obligatorias que presentan con claridad los contenidos teórico-prácticos sobre el procesamiento lingüístico y sus aplicaciones informáticas. A partir de estas lecturas, los alumnos demostrarán la asimilación de los contenidos mediante la elaboración de tres prácticas y la redacción de un proyecto personal. A continuación, describimos brevemente estos tipos de prueba:

- Las prácticas ofrecerán a la clase la oportunidad de participar en tareas específicas diseñadas a partir de un escenario real de ingeniería del conocimiento. Éstas permitirán a los miembros de la clase demostrar sus destrezas en el campo del tratamiento computacional de la sintaxis y la semántica.
- El proyecto personal consistirá en la redacción de un trabajo que no sólo estimule la reflexión crítica sobre el material estudiado, sino que además fomente la toma de decisiones motivadas, la previsión de problemas potenciales y la propuesta de posibles soluciones desde una perspectiva investigadora.

SISTEMA DE EVALUACIÓN

TIPO DE PRUEBA PRESENCIAL

Tipo de examen	No hay prueba presencial

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad	No
Descripción	

La **evaluación global** de la asignatura está basada en la superación de cuatro prácticas. En concreto:

3 tres prácticas de evaluación continua (NO PRESENCIAL)

1 práctica final de mayor extensión que las anteriores (NO PRESENCIAL)

La prueba final supondrá un 40% de la nota global de la asignatura y consistirá en la elaboración de un trabajo crítico sobre alguno de los aspectos de contenido de entre los que se hayan tratado en la asignatura. El tema o temas de investigación será propuestos por el profesor desde la apertura del curso. Esta práctica final no se dirige tanto a la solución de problemas concretos, sino más bien a la indagación bibliográfica de material relevante según el criterio del alumno, así como a la investigación y aportación de contenidos innovadores. El resto de pruebas de evaluación continua aparecen detalladas en los siguientes apartados de esta guía (ver abajo).

La fecha aproximada de entrega de la "práctica 4" será en las dos o tres últimas semanas de curso. Esta fecha, no obstante, podrá sufrir leves alteraciones según las necesidades del año concreto. La fecha concreta de entrega se especificará en la plataforma ALF desde el comienzo del curso.

No hay ninguna variación con respecto al planteamiento general en el caso de los estudiantes en centros penitenciarios o para las convocatorias extraordinarias.

Criterios de evaluación

Los principales criterios de evaluación:

Aportación de contenidos correctos según lo estudiado durante el curso

Aportación de información nueva encontrada por los miembros de la clase

Originalidad e innovación de la propuesta realizada por los miembros de la clase

Idoneidad formal: corrección ortográfica, gramatical y de estilo

Idoneidad de formato: adaptación a las instrucciones de formato que se indiquen

Ideoneidad de extensión del trabajo según los márgenes establecidos

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

Además de superar la **práctica final ("práctica 4")**, cuya información aparece detallada en apartados anteriores de esta guía (ver arriba) y en el apartado final (ver abajo), la evaluación de la asignatura también comprende la entrega de los trabajos de **evaluación continua** que se detallan a continuación:

Práctica 1: 20%

Práctica 2: 20%

Practica 3: 20%

Las fechas aproximadas de entrega de la "prácticas 1, 2 y 3" serán, respectivamente, la tercera semana de noviembre; la tercera semana de diciembre; y la primera semana de febrero. Estas fechas, no obstante, podrá sufrir leves alteraciones según las necesidades del año concreto. Las fechas concretas de entrega se especificarán en la plataforma ALF desde el comienzo del curso.

No hay ninguna variación con respecto al planteamiento general en el caso de los estudiantes en centros penitenciarios o para las convocatorias extraordinarias.

Criterios de evaluación

Los principales criterios de evaluación:

Aportación de contenidos correctos según lo estudiado en cada tema concreto

Idoneidad formal: corrección ortográfica, gramatical y de estilo

Idoneidad de formato: adaptación a las instrucciones de formato que se indiquen

Ponderación de la PEC en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

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?

La **evaluación global** de la asignatura *Procesamiento del Lenguaje Natural* se hará según el siguiente plan de prácticas, todas ellas de tipo NO PRESENCIAL:

Práctica 1: 20%

Práctica 2: 20%

Practica 3: 20%

Práctica 4: 40%

IMPORTANTE (nota mínima): para superar la asignatura, la práctica final ("práctica 4") deberá superarse con una nota igual o superior a 5 sobre 10. Esta nota mínima será requisito indispensable para computar el resto de prácticas del curso, tanto en la convocatoria ordinaria como extraordinaria.

IMPORTANTE (fechas de entrega): la fecha máxima de entrega se indicará en la plataforma ALF desde la apertura del curso. Es importante recordar que es necesario ajustarse a las fechas marcadas desde el inicio para superar la asignatura.

IMPORTANTE (convocatoria extraordinaria): tanto el baremo como el sistema de evaluación serán los mismos en las convocatoria ordinaria y extraordinaria, es decir, cuatro tareas relacionadas con las lecturas del curso y otras lecturas que deberán recopilarse individualmente. Es importante recordar que, mientras que lo esencial del sistema de trabajo y de evaluación se mantienen en ambas convocatorias, sí podría haber cambios en el contenido de algunos de los ejercicios propuestos para la convocatoria extraordinaria. En relación con este punto, se informará a la clase de dichos cambios mediante la plataforma ALF y/o mediante correo electrónico una vez terminado el curso y publicadas las notas finales. Aquellas prácticas que se hayan superado correctamente durante el curso se guardarán para la convocatoria extraordinaria y se evaluarán según el baremo propuesto. Por tanto, sólo deberán completarse aquéllas que no hubieran sido entregadas a tiempo durante la convocatoria ordinaria.

IMPORTANTE (aclaración): no hay ninguna variación con respecto al planteamiento general en el caso de los estudiantes en centros penitenciarios o para las convocatorias extraordinarias.

BIBLIOGRAFÍA BÁSICA

Cambria, E. y B. White (2014). "Jumping NLP curves: A review of Natural Language Processing research". *IEEE Computational Intelligence Magazine*, Mayo 2014: 48-57

Gómez Guinovart, X. (2000). “Perspectivas de la lingüística computacional”. *Novática: Revista de la Asociación de Técnicos de Informática*, número especial del 25 aniversario (Horizonte 2025), 145: 85-87.

Llisterri, J. (2003). "Lingüística y tecnologías del lenguaje". *Lynx. Panorámica de Estudios Lingüísticos* 2: 9-71.

Mairal Usón, R. y F. Cortés Rodríguez (2006). “An overview of Role and Reference Grammar”. En R. Mairal, M. A. Escobar Álvarez, M. S. Peña Cervel, E. Samaniego Fernández y F. Cortés Rodríguez (eds.) *Current Trends in Linguistic Theory*. Madrid: Universidad Nacional de Educación a Distancia, 97-176.
Peñalver-Martínez, I., R. Valencia-García y F. García-Sánchez (2011). "Minería de Opiniones basada en características guiada por Ontologías". *Procesamiento del Lenguaje Natural*, 46: 91-98. Sociedad Española para el Procesamiento del Lenguaje Natural, España.
Periñán Pascual, C. (2005). “Procesamiento del lenguaje natural: de lingüista a ingeniero del conocimiento”. En I. K. Brady, M. Navarro Coy y C. Periñán Pascual (eds.) *Nuevas Tendencias en Lingüística Aplicada*. Murcia: Quaderna, 293-317.
Periñán Pascual, C. y F. Arcas Túnez (2010). “The architecture of FunGramKB”. *Proceedings of the 7th International Conference on Language Resources and Evaluation*. European Language Resources Association (ELRA). Valletta (Malta), 2667-2674.
Periñán Pascual, C. y R. Mairal Usón (2009). “Bringing Role and Reference Grammar to natural language understanding”. *Procesamiento del Lenguaje Natural* 43: 265-273.
Periñán Pascual, C. y R. Mairal Usón (2010). “La gramática de COREL: un lenguaje de representación conceptual”. *Onomázein* 21: 11-45.
Rodríguez, H. (2000). “Técnicas básicas en el tratamiento informático de la lengua”. *Quark. Ciencia, Medicina, Comunicación y Cultura* 19: 26-34.

BIBLIOGRAFÍA COMPLEMENTARIA

Dale, R., H. Moisl y H. Somers (eds.) (2000): *Handbook of Natural Language Processing*, Nueva York, Marcel Dekker.
Mitkov, R. (ed.) (2003): *The Oxford Handbook of Computational Linguistics*, Oxford, Oxford University Press.

RECURSOS DE APOYO Y WEBGRAFÍA

La plataforma de aprendizaje en Internet permitirá no solo la interacción entre los alumnos y el profesor, sino también el acceso a las lecturas obligatorias y a las instrucciones para la realización de los trabajos propuestos. Igualmente, el foro de debate y el correo electrónico se convierten en instrumentos imprescindibles para la consulta y la resolución de dudas relacionadas con el trabajo del alumno.

ADENDA AL SISTEMA DE EVALUACIÓN CON MOTIVO DE LA PANDEMIA COVID 19

https://app.uned.es/evacaldos/asignatura/adendasig/24409117

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.