



FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Programación orientada a PLN
Código	E000014383
Título	Graduado o Graduada en Traducción, Interpretación y Tecnologías del Lenguaje por la Universidad Pontificia Comillas
Impartido en	Grado en Traducción, Interpretación y Tecnologías del Lenguaje [Segundo Curso]
Nivel	Reglada Grado Europeo
Cuatrimestre	Semestral
Créditos	6,0 ECTS
Carácter	Obligatoria (Grado)
Departamento / Área	Departamento de Telemática y Computación
Descriptor	Esta asignatura introduce al alumno en el conocimiento de la Inteligencia Artificial como herramienta orientada al análisis y el tratamiento automático del lenguaje humano. Para ello, el alumno deberá adquirir una comprensión general de conceptos como representación del conocimiento, aprendizaje automático y modelos de lenguaje, con su posterior aplicación a tareas prácticas en el ámbito del procesamiento del lenguaje natural.

Datos del profesorado	
Profesor	
Nombre	Blanca Hernández Pardo
Departamento / Área	Departamento de Traducción e Interpretación y Comunicación Multilingüe
Despacho	Cantoblanco [D-303]
Correo electrónico	bhpardo@comillas.edu
Teléfono	4426
Profesor	
Nombre	Luis Alarcón Massó
Departamento / Área	Departamento de Organización Industrial
Correo electrónico	lalarcon@comillas.edu

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura	
Aportación al perfil profesional de la titulación	
<p>En el perfil profesional del graduado en Grado en Traducción, Interpretación y Tecnologías del Lenguaje, esta asignatura introduce al alumno en el conocimiento de la Inteligencia Artificial como herramienta orientada al análisis y tratamiento automático del lenguaje humano. Para ello, el alumno deberá adquirir una comprensión general de conceptos como representación del conocimiento, aprendizaje automático y modelos de lenguaje, con su posterior aplicación a tareas prácticas en el ámbito del procesamiento del lenguaje natural. En este curso se ha seleccionado el lenguaje de programación Python, el cual, por su sencillez y versatilidad, permite al estudiante</p>	



experimentar con técnicas básicas de IA aplicadas a textos, como la clasificación, extracción de información y evaluación de modelos lingüísticos.

Prerrequisitos

No hay.

Competencias - Objetivos

Competencias

- CN 7 - Conoce los fundamentos de la programación en su aplicación al procesamiento del lenguaje natural para la mediación interlingüística.
- HB 1 - Puede aplicar teorías y conceptos lingüísticos a la comprensión y análisis de la lengua y evaluar la calidad de la comunicación en diferentes contextos.
- HB 5 - Puede trabajar en equipo con otros profesionales del lenguaje, la comunicación y el PLN.
- HB 6 - Puede adaptarse a los cambios tecnológicos y a las nuevas tendencias en el campo del multilingüismo, y desarrollar habilidades de aprendizaje autónomo y continuo en dicho ámbito.
- HBO 1 - Puede aplicar herramientas y tecnologías lingüísticas, incluyendo recursos digitales, al análisis y procesamiento del lenguaje natural.
- HBO 6 - Puede buscar y escoger herramientas y técnicas específicas para el uso de aplicaciones de inteligencia artificial en los sectores de servicios lingüísticos e interlingüísticos (traducción automática, tradumática, chat bots, web scraping, etc.)
- HBO 7 - Puede utilizar con eficacia lenguajes de programación y herramientas de desarrollo en los ámbitos lingüístico y multilingüístico.
- CTO 2 - Aplica técnicas de programación orientada a problemas lingüísticos y desarrolla soluciones informáticas para los diversos sectores de los servicios lingüísticos e interlingüísticos.
- CTO 3 - Contribuye al desarrollo de proyectos en el campo de la lingüística computacional y la inteligencia artificial, aplicando metodologías científicas y herramientas tecnológicas como minería de datos o elaboración y tratamiento de datasets.

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

1. Repaso de estructuras básicas en Python orientadas a la manipulación de texto
2. Expresiones regulares y preprocesamiento de texto
3. Tokenización, lematización, stemming
4. Etiquetado POS y análisis morfosintáctico de textos
5. Análisis de corpus: lectura de archivos, segmentación y recuento de frecuencias
6. Traducción automática y evaluación con métricas automáticas
7. Extracción terminológica y alineación

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

La metodología se basa en el aprendizaje por simulación profesional. Cada unidad parte de un caso práctico donde una empresa ficticia o real requiere una solución basada en PLN. El alumno desarrolla scripts gradualmente, reflexiona sobre los resultados e implementa



mejoras. Se utiliza Google Colab como entorno de trabajo colaborativo y accesible.

Metodología Presencial: Actividades

Clases magistrales expositivas y participativas:

- El profesor combinará exposición de los contenidos teóricos y con ejemplos prácticos, tanto matemáticos como de programación.
- El alumno dispondrá de algunos ejemplos prácticos de código, generado dentro y fuera del aula por el profesor.

CN 10, HB 6, CN 7, CN 8

Ejercicios prácticos y resolución de problemas:

El alumno resolverá problemas planteados por el profesor de forma presencial, fomentando las dinámicas de trabajo cooperativo.

CN 10, HB 1, HB 5, HB 6, CN 7, CN 8

Sesiones prácticas con uso de software:

- Se dedicarán las sesiones prácticas a resolver dudas de la práctica semanal y a que los alumnos finalicen su práctica.
- Las prácticas se graduarán por dificultad que los alumnos irán implementando conforme vayan finalizando cada hito.

CN 10, HB 1, HB 5, HB 6, CN 7, CN 8

Metodología No presencial: Actividades

Ejercicios prácticos y resolución de problemas:

El alumno dispondrá de problemas concretos enfocados a asimilarlos conceptos explicados teóricos en la sesión anterior de teoría para desarrollar de forma no presencial.

CN 10, HB 1, HB 5, HB 6, CN 7, CN 8

Sesiones prácticas con uso de software:

- Una vez liberada la práctica semanal después de la sesión de teoría correspondiente, el alumno trabajará sobre ella de forma no presencial.
- En el aula se seguirá trabajando sobre la teoría y los primeros casos prácticos y se cubrirán los hitos planteados de forma progresiva.

CN 10, HB 1, HB 5, HB 6, CN 7, CN 8

Estudio personal:

El objetivo principal del trabajo no presencial es llegar a entender y comprender los conceptos teóricos de la asignatura, así como ser capaz de poner en práctica estos conocimientos para resolver los diferentes tipos de problemas.

CN 10, HB 1, HB 5, HB 6, CN 7, CN 8

RESUMEN HORAS DE TRABAJO DEL ALUMNO

Horas Presenciales			
Lecciones de carácter expositivo	Ejercicios prácticos/resolución de problemas	Trabajos individuales/grupales	Actividades de evaluación continua del rendimiento



<u>10</u>	<u>30</u>	<u>10</u>	<u>10</u>
Horas No Presenciales			
Ejercicios prácticos/resolución de problemas	Sesiones prácticas	Trabajos	Estudio personal y documentación
<u>30</u>	<u>30</u>	<u>10</u>	<u>20</u>
CRÉDITOS ECTS: 6 (150 horas)			

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Actividades de evaluación	Criterios de evaluación	Peso
Exámenes: Examen Final	Examen Final: se evaluará el pensamiento computacional y abstracto para la resolución de problemas.	50 %
Sesiones prácticas: Retos colaborativos. Trabajos no presenciales. Prácticas	Realización de varias prácticas a lo largo del curso y retos planteados en sesiones colaborativas.	40 %
Asistencia y participación	Asistencia y aportación razonada y documentada a la hora de resolver ejercicios prácticos en el aula.	10 %

Calificaciones

Es obligatoria para todos los alumnos la asistencia a clase y a las actividades docentes presenciales. Para poder presentarse al examen final es requisito no haber faltado injustificadamente a más de un tercio de las clases y haber entregado todas las tareas obligatorias. De no cumplir este requisito, el alumno perderá las convocatorias ordinaria y extraordinaria (art. 93.1 del Reglamento General).

El incurrir en una falta académica grave, como es el plagio de materiales previamente publicados, el copiar en su examen u otra actividad evaluada, o no respetar las normas de convivencia básicas puede llevar a la apertura de un expediente sancionador y la pérdida de dos convocatorias.

Convocatoria ordinaria:

Los porcentajes que compondrán la nota final de la asignatura en convocatoria ordinaria se muestran en el anterior apartado de evaluación.

Convocatoria extraordinaria:

La calificación de la asignatura coincidirá íntegramente con la nota del examen oficial en dicha convocatoria.

Alumnos que han cumplido escolaridad y con dispensa de escolaridad:

El alumno debe informar a la profesora mediante correo electrónico al comienzo del curso (hasta un máximo de 20 días naturales desde el inicio de la asignatura) sobre su escolaridad cumplida o dispensa de escolaridad.

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

El uso de IA para crear trabajos completos o partes relevantes, sin citar la fuente o la expresamente en la descripción del trabajo, será considerado plagio y regulado conf Universidad.

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python*. O'Reilly Media.

Hammond, M. (2020). *Python for linguists*. Cambridge University Press.

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2022). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 82, 3713-3744. <https://doi.org/10.1007/s11042-022-13428-4>

Rojo, G. (2021). *Introducción a la lingüística de corpus en español*. Routledge. ISBN 978-0-367-63584-8. DOI: 10.4324/9781003119760.

Tsujii, J. (2021). Natural Language Processing and Computational Linguistics. *Computational Linguistics*, 47(4), 707-734. https://doi.org/10.1162/coli_a_00413



Sangacha-Tapia, L. M., Celi, R. J., Acosta-Guzmán, I. L., & Varela-Tapia, E. A. (2023). *Inteligencia Artificial Aplicada a Procesamiento de Lenguaje Natural (NLP) con Python y Machine Learning*. Editorial Grupo AEA. ISBN: 9789942651433.

Vajjala, S., Majumder, B., Gupta, A., & Surana, H. (2020). *Practical natural language processing: A comprehensive guide to building real-world NLP systems* (1st ed.). O'Reilly Media.

Bibliografía Complementaria

Herramientas de PLN: spaCy, NLTK y Stanza

Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). spaCy: Industrial-strength natural language processing in Python. Zenodo. <https://doi.org/10.5281/zenodo.1212303>

Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C. D. (2020). Stanza: A Python natural language processing toolkit for many human languages. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 101-108. <https://aclanthology.org/2020.acl-demos.14/>

Spring, R., & Johnson, M. (2022). A comparison of the LCA, NLTK and SpaCy tools for analyzing writing complexity in second language learners. *System*, 108, 102846. <https://doi.org/10.1016/j.system.2022.102846>

Lingüística Computacional aplicada a Traducción

Joty, S., Guzmán, F., Màrquez, L., & Nakov, P. (2017). Discourse structure in machine translation evaluation. *Computational Linguistics*, 43(4), 683-722. https://doi.org/10.1162/coli_a_00297

Marie, B., Fujita, A., & Rubino, R. (2021). Scientific credibility of machine translation research: A meta-evaluation of 769 papers. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, 7297-7306. <https://aclanthology.org/2021.acl-long.566/>

Moghe, N., Arora, R., Bansal, S., & Søgaard, A. (2024). Machine translation meta evaluation through translation accuracy challenge sets. *Computational Linguistics*, 50(1), 1-45. https://doi.org/10.1162/coli_a_00489

Evaluación de Traducción Automática y Métricas

Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002). BLEU: A method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 311-318. <https://aclanthology.org/P02-1040/>

López Caro, Á. (2023). *Machine Translation evaluation metrics benchmarking*. Universidad de Barcelona. Disponible en: https://diposit.ub.edu/dspace/bitstream/2445/214303/1/tfm_lopez_caro_alvaro.pdf

Qian, S., Sindhujan, A., Kabra, M., Kanojia, D., Orăsan, C., Ranasinghe, T., & Blain, F. (2024). What do Large Language Models Need for Machine Translation Evaluation? *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. <https://arxiv.org/html/2410.03278v1>



Corpus Paralelos y Análisis Contrastivo

Hallebeek, J. (s.f.). El Corpus paralelo. *Revista de Lingüística Teórica y Aplicada*, 24, 147-164. Disponible en: <https://core.ac.uk/download/pdf/16355934.pdf>

Wang, A. (2022). Corpus linguistics for translation and contrastive studies: A guide for research. *Journal of Computer-Assisted Language Learning Research*, 13(3), 313-318. <https://doi.org/10.1075/jicb.20014.wan>

Hartmann, R. R. K. (1996). Contrastive textology and corpus linguistics: On the value of parallel texts. *Language Sciences*, 18(3-4), 947-957. [https://doi.org/10.1016/S0388-0001\(96\)00056-3](https://doi.org/10.1016/S0388-0001(96)00056-3)

Pešková, J. (2023). El corpus como herramienta para investigar rasgos específicos de los textos traducidos: El empobrecimiento léxico. En *Traducción y sostenibilidad cultural* (pp. 351-358). Ediciones Universidad de Salamanca. Disponible en: <https://eusal.es/eusal/catalog/download/978-84-1311-056-1/4503/939-1>

Extracción Terminológica Automática

Marciniak, M., Rychlik, P., & Mykowiecka, A. (2023). TermoUD — A language-independent terminology extraction tool. *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, 178-186. <https://aclanthology.org/2023.eacl-demo.21/>

Nazar, R. (2023). A lightweight statistical method for terminology extraction. *Journal of Computer-Assisted Linguistic Research*, 6(1), 45-72. <https://polipapers.upv.es/index.php/jclr/article/view/20427>

Hätty, A., Tannert, S., & Heid, U. (2020). Predicting degrees of technicality in automatic terminology extraction. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2883-2889. <https://aclanthology.org/2020.acl-main.258/>

Herramientas de Análisis Morfológico y Sintáctico

FreeLing - Analizador morfosintáctico multilingüe. TALP Research Center, Universitat Politècnica de Catalunya. <http://nlp.lsi.upc.edu/freeling/>

PORCUS - Parsing de Oraciones Realizado por Computadoras de la Universidad de Santiago. <https://www.tecling.com/porcus/>

Exea - Analizador morfológico de la Biblioteca Virtual Miguel de Cervantes. <https://exea.csic.es/es/analizador>

Recursos y Herramientas en español

Datos.gob.es (2025). 10 Librerías populares de procesamiento del lenguaje natural. <https://datos.gob.es/es/blog/10-librerias-populares-de-procesamiento-del-lenguaje-natural>

Phrase (2025). Traducción asistida por ordenador ¿Qué es una herramienta CAT? <https://phrase.com/es/blog/posts/cat-tools/>



Revistas especializadas en Lingüística Computacional y Traducción

Computational Linguistics - MIT Press. <https://direct.mit.edu/coli>

Machine Translation - Springer. <https://link.springer.com/journal/10590>

Natural Language Processing - Cambridge University Press. <https://www.cambridge.org/core/journals/natural-language-processing>

Language Resources and Evaluation - Springer. <https://link.springer.com/journal/10579>

ACM Transactions on Asian Language Information Processing - ACM. <https://dl.acm.org/journal/talip>

Tradumàtica - Universitat Autònoma de Barcelona. <https://revistes.uab.cat/tradumatica/>

Sendebar - Universidad de Granada. <https://revistaseug.ugr.es/index.php/sendebar>

En cumplimiento de la normativa vigente en materia de **protección de datos de carácter personal**, le informamos y recordamos que puede consultar los aspectos relativos a privacidad y protección de datos que ha aceptado en su matrícula entrando en esta web y pulsando "descargar"

[https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792](https://servicios.upcomillas.es/sedeelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792)