



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN LINGÜÍSTICA



Propuesta de programa para impartir asignaturas en

(INDICAR EL PROGRAMA QUE CORRESPONDA)

NB No se considerarán propuestas que no incluyan toda la información solicitada

MAESTRÍA EN LINGÜÍSTICA APLICADA

MAESTRÍA EN LINGÜÍSTICA HISPÁNICA

TÍTULO DE LA ASIGNATURA

Aplicaciones de Python a problemas del lenguaje natural

PROFESOR/A

Gemma Bel Enguix

CLAVE	SEMESTRE 2025-2	CAMPO DE CONOCIMIENTO Lingüística computacional		NÚMERO DE CRÉDITOS 8
<input type="radio"/> Optativa		HORAS	HORAS POR SEMANA	HORAS AL SEMESTRE
MODALIDAD		TEORÍA	PRÁCTICA	64
		38	26	
<input type="radio"/> Curso		DURACIÓN DEL CURSO		<i>semestral</i>

ACTIVIDAD ACADÉMICA CON SERIACIÓN SUBSECUENTE

ACTIVIDAD ACADÉMICA CON SERIACIÓN ANTECEDENTE

OBJETIVO(S) DEL CURSO O SEMINARIO

Objetivo general

Introducción de métodos de tratamiento computacional del lenguaje basados en Python para desarrollar técnicas de apoyo en lingüística.

Objetivos específicos

<ul style="list-style-type: none"> • Conocimiento básico del lenguaje de programación Python aplicado a las necesidades de los lingüistas. • Familiarización con técnicas, librerías y funciones orientadas al tratamiento automático de textos. • Iniciación a métodos de análisis, comprensión y generación computacional de textos. Aproximación a la extracción de información y ciencia de datos basada en Python. • Experimentación en la extracción, almacenamiento y análisis de ejemplos de comunicación textual y multimodal extraídos de la web. 			
ÍNDICE TEMÁTICO			
			Horas
UNIDAD	TEMA	Teóricas	Prácticas
1	Conceptos generales de Python y visualización	8	6
2	Herramientas de análisis de texto y vectorización	8	5
3	Aproximación al estudio del lenguaje de la red y las plataformas de redes sociales	8	5
4	Métodos de análisis de sentimientos	8	5
5	Aprendizaje automático	6	5
TOTAL DE HORAS		38	26
SUMA TOTAL DE HORAS:		64	

CONTENIDO TEMÁTICO DESGLOSADO

UNIDAD	TEMA/SUBTEMAS
1	Conceptos generales de Python
1	Redes léxicas y lexicón
1	SpaCy, Stanza, NLTK, FreeLing
1	Matplotlib, gráficas y análisis de resultados
2	Métodos de preprocesamiento de texto
2	Formas de vectorización léxica
2	Introducción a los vectores de palabras: word2vec, Glove, BERT
3	La web como corpus y el lenguaje teclado
3	Twitter (X), Facebook, reddit, TikTok. El texto en la red y la comunicación multimodal.
4	Análisis de sentimientos: Polaridad vs. Emociones
4	Métodos basados en diccionarios
4	Métodos basados en aprendizaje automático
5	Principales formulaciones, métodos y algoritmos de aprendizaje automático
5	Formulas de evaluación
5	Algoritmos clásicos: SVM, LR, KNN

<p>BIBLIOGRAFÍA BÁSICA</p> <p>Allen, J. (1995) Natural Language Understanding, The Benjamin/Cummings Publishing Company. Redwood City, Calif. (Second Edition).</p> <p>Bird, S. and Klein, E. and Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media ISBN: 9780596555719</p> <p>Gatto, M. (2014). Web As Corpus: Theory and Practice. Blomsbury.</p> <p>Jurafsky, Daniel & James H. Martin. (2008). Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition (2nd edition), Prentice Hall.</p> <p>Manning, C. & Schütze, H. (1999). Foundations of Statistical Natural Language Processing, MIT Press. Cambridge, MA.</p> <p>YUS, F. (2011). Cyberpragmatics: internet-mediated Communication in Context. John Benjamins.</p>	
<p>BIBLIOGRAFÍA COMPLEMENTARIA</p> <p>Charniak, E., Statistical Language Learning, The MIT Press, Cambridge (Mass.), 1997.</p> <p>Cowie, J, Wilks, Y. (1996), Information Extraction.</p> <p>Hutchins J, Somers H. (1985), Introducción a la traducción automática, Antonio Machado</p> <p>Manning, C., Prabhakar, R. & Schütze, H. (2998), Introduction to Information Retrieval. Cambridge University Press</p> <p>Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. http://arxiv.org/abs/1301.3781</p> <p>Mitkov, R, ed. (2019), The Oxford Handbook of Computational Linguistics, Oxford UP, Oxford, 2003. Second Edition.</p>	
<p>ESTRATEGIAS DIDÁCTICAS</p> <p>Exposición oral por parte del profesor ()</p> <p>Exposición audiovisual a cargo del profesor (x)</p> <p>Exposición oral por parte de los alumnos (x)</p> <p>Ejercicios dentro de clase (x)</p> <p>Ejercicios fuera del aula (x)</p> <p>Lecturas obligatorias ()</p>	<p>MECANISMOS DE EVALUACIÓN DEL APRENDIZAJE DE LOS ALUMNOS</p> <p>Exámenes Parciales ()</p> <p>Examen final escrito ()</p> <p>Trabajos y tareas fuera del aula (x)</p> <p>Exposición en seminarios por los alumnos (x)</p> <p>Participación en clase (x)</p>

Trabajo de Investigación ()	Asistencia (x)
Otro	Seminario ()
	Trabajo de investigación ()
NB: En los cursos obligatorios o que constituyen el primer acercamiento a un tema, la presentación de los temas debe descansar preponderantemente en el profesor.	Otro
<p>CRITERIOS DE EVALUACIÓN</p> <p>Los principales criterios de evaluación serán:</p> <ul style="list-style-type: none"> • Asistencia y participación en clase (40%) • Realización de trabajos y ejercicios prácticos (40%) • Participación en clase (20%) <p>La asignatura no tiene exámenes puesto que se realiza una evaluación continua de los estudiantes.</p>	
<p>NOTA SOBRE LA ASIGNATURA</p> <p>Para cursar la asignatura sería mejor tener un conocimiento básico de Python. En caso de no tenerlo, se requerirá un trabajo adicional.</p>	

PROPUESTAS DE HORARIO PARA IMPARTIR EL CURSO:

2- Martes de 10 a 14

FECHA: 10 NOVIEMBRE 2024