



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



PROGRAMA SINTÉTICO

UNIDAD ACADÉMICA: ESCUELA SUPERIOR DE CÓMPUTO, UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA CAMPUS COAHUILA, UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA CAMPUS TLAXCALA

PROGRAMA ACADÉMICO: Ingeniería en Inteligencia Artificial

UNIDAD DE APRENDIZAJE: Tecnologías de lenguaje natural

SEMESTRE: V

PROPÓSITO DE LA UNIDAD DE APRENDIZAJE:

Aplica las tecnologías de lenguaje natural a partir de los procesos de normalización de textos, procesamiento semántico de textos, extracción de información, análisis de personalidad y opinión, de diálogo y traducción automática.

CONTENIDOS:	I. Contextualización y normalización de textos II. Procesamiento semántico de textos y extracción de información III. Análisis de personalidad y opinión IV. Sistemas de diálogo V. Traducción automática			
ORIENTACIÓN DIDÁCTICA:	Métodos de enseñanza		Estrategias de aprendizaje	
	a) Inductivo		a) Estudio de casos	X
	b) Deductivo	X	b) Aprendizaje basado en problemas	
	c) Analógico		c) Aprendizaje orientado proyectos	X
	d) Heurístico	X		
EVALUACIÓN Y ACREDITACIÓN:	Diagnóstica	X	Saberes Previamente Adquiridos	X
	Solución de casos	X	Organizadores gráficos	
	Problemas resueltos		Problemarios	
	Reporte de proyectos	X	Exposiciones	X
	Reportes de indagación	X	Otras evidencias a evaluar: Reporte de uso de software	
	Reportes de prácticas	X		
	Evaluación escrita			
BIBLIOGRAFÍA BÁSICA:	Autor(es)	Año	Título del documento	Editorial / ISBN
	Bird, S., Klein, E., Loper, E.	2009	<i>Natural Language Processing with Python</i>	O'Reilly/ 9780596516499
	Jurafsky, D., Martin, J.	2008	<i>Speech and Language Processing</i>	Pearson Prentice Hall / 9780131873216
	Poibeau T.	2017	<i>Machine Translation</i>	MIT Press Essential Knowledge series/ 9780262534215
	Shevat, A.	2017	<i>Designing Bots</i>	O'Reilly/9781491974827
	Zhai, C., Massung, S.	2016	<i>Text Data Management and Analysis</i>	ACM and Morgan & Claypool Publishers/ 9781970001167



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



PROGRAMA DE ESTUDIOS

UNIDAD DE APRENDIZAJE: Tecnologías de lenguaje natural

HOJA 2 DE 9

UNIDAD ACADÉMICA: ESCUELA SUPERIOR DE CÓMPUTO, UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA CAMPUS COAHUILA, UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA CAMPUS TLAXCALA

PROGRAMA ACADÉMICO: Ingeniería en Inteligencia Artificial

SEMESTRE:
V

ÁREA DE FORMACIÓN:
Profesional

MODALIDAD:
Escolarizada

TIPO DE UNIDAD DE APRENDIZAJE:
Teórica-Práctica/Obligatoria

VIGENTE A PARTIR DE:
Enero 2022

CRÉDITOS:

TEPIC: 7.5

SATCA: 6.7

INTENCIÓN EDUCATIVA

La unidad de aprendizaje de Tecnologías de lenguaje natural contribuye al perfil de Ingeniero en Inteligencia Artificial proporcionándole los fundamentos teórico-prácticos para la implementación de procesos de normalización de textos, procesamiento semántico de textos, extracción de información, análisis de personalidad y opinión, de diálogo y traducción automática. Asimismo, fomenta las habilidades transversales de trabajo en equipo, comunicación efectiva, ética, creatividad e ingenio.

Esta unidad de aprendizaje se relaciona de manera antecedente con Algoritmos y estructuras de datos, Probabilidad y estadística, Matemáticas avanzadas para la ingeniería y Fundamentos de inteligencia Artificial; de manera lateral con Aprendizaje de máquina; y de manera consecuente con Metodología de la investigación y divulgación científica, Reconocimiento de voz.

PROPÓSITO DE LA UNIDAD DE APRENDIZAJE

Aplica las tecnologías de lenguaje natural a partir de los procesos de normalización de textos, procesamiento semántico de textos, extracción de información, análisis de personalidad y opinión, de diálogo y traducción automática.

TIEMPOS ASIGNADOS

HORAS TEORÍA/SEMANA: 3.0

HORAS PRÁCTICA/SEMANA: 1.5

HORAS TEORÍA/SEMESTRE: 54.0

HORAS PRÁCTICA/SEMESTRE: 27.0

HORAS APRENDIZAJE AUTÓNOMO: 32.0

HORAS TOTALES/SEMESTRE: 81.0

**UNIDAD DE APRENDIZAJE
DISEÑADA POR:**

Comisión de Diseño del Programa Académico.

APROBADO POR:

Comisión de Programas Académicos del H. Consejo General Consultivo del IPN.

22/10/2020

**AUTORIZADO Y VALIDADO
POR:**

Ing. Juan Manuel Velázquez Peto
Director de Educación Superior



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Tecnologías de lenguaje natural

HOJA 3 DE 9

UNIDAD TEMÁTICA I Contextualización y normalización de textos	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Aplica procesos de normalización a textos en lenguaje natural.	1.1 Contextualización	3.0		3.0
	1.1.1 Sistemas y aplicaciones de lenguaje natural			
	1.1.2 Fundamentos lingüísticos para tecnologías de lenguaje			
	1.1.3 Fundamentos matemáticos para tecnologías de lenguaje			
	1.1.4 Lenguajes de programación, librerías y software para tecnologías de lenguaje			
	1.2 Normalización de textos	6.0	3.5	1.5
	1.2.1 Segmentación en palabras y oraciones			
	1.2.2 Eliminación de elementos no relevantes			
	1.2.3 Etiquetado con categorías gramaticales			
	Subtotal	9.0	3.5	4.5

UNIDAD TEMÁTICA II Procesamiento semántico de textos y extracción de información	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Extrae aspectos semánticos e información a partir de la representación vectorial de textos, análisis sintáctico y expresiones regulares.	2.1 Representación vectorial de textos para su procesamiento semántico	3.0		2.0
	2.1.1 Modelo de bolsa de palabras			
	2.1.2 Modelo de espacio vectorial para textos			
	2.2 Selección de características en vectores	4.5	3.5	2.0
	2.2.1 Frecuencia y probabilidad de palabras			
	2.2.2 Funciones de frecuencia de término (TF)			
	2.2.3 Frecuencia inversa del documento (IDF)			
	2.2.4 Técnicas de mapeo de palabras a vectores numéricos (word embeddings)			
	2.3 Extracción de aspectos semánticos a partir de textos	6.0	3.5	1.5
	2.3.1 Palabras similares			
	2.3.2 Asociaciones entre palabras			
	2.3.3 Terminología			
	2.3.4 Palabras clave			
	2.4 Extracción de información a partir de textos	4.5	4.5	2.0
	2.4.1 Análisis sintáctico superficial y profundo para extracción de entidades e información.			
	2.4.2 Expresiones regulares para extracción de información			
	2.4.3 Resumen automático			
	2.5 Generación de respuestas a preguntas	1.5		1.5
	Subtotal	19.5	11.5	9.0



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Tecnologías de lenguaje natural

HOJA 4 DE 9

UNIDAD TEMÁTICA III Análisis de personalidad y opinión	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Analiza aspectos de personalidad en textos a partir de rasgos y detección de polaridad.	3.1 Detección y análisis de aspectos de personalidad en textos	6.0	3.5	3.5
	3.1.1 Edad			
	3.1.2 Sexo			
	3.1.3 Rasgos personales: extraversión, neurosis, simpatía, diligencia, franqueza			
	3.2 Análisis de opinión	3.0	3.5	2.5
	3.2.1 Detección de polaridad de opinión con base en diccionarios de polaridad			
	3.2.2 Detección de polaridad de opinión usando algoritmos de aprendizaje de máquina			
	Subtotal	9.0	7.0	6.0

UNIDAD TEMÁTICA IV Sistemas de diálogo	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Implementa sistemas de dialogo a partir de fundamentos de texto escrito y reconocimiento de voz.	4.1 Tipos de sistemas dialogo	0.5		1.5
	4.2. Componentes de sistemas diálogo	0.5		
	4.3 Herramientas de implementación de sistemas de dialogo	0.5		
	4.4. Agente conversacional basado en texto escrito	3.0	1.5	3.0
	4.5. Agente conversacional basado en reconocimiento de voz y generación de habla	4.5	3.5	3.5
	4.5.1 Fundamentos de reconocimiento de voz			
	4.5.2 Generación de habla			
	Subtotal	9.0	5.0	8.0



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Tecnologías de lenguaje natural

HOJA 5 DE 9

UNIDAD TEMÁTICA V Traducción automática	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
UNIDAD DE COMPETENCIA Diferencia los sistemas de traducción automática a partir de sus características.	5.1 Tipos de sistemas de traducción automática 5.1.1 Basados en reglas 5.1.2 Basados en estadística 5.1.3 Basados en ejemplos 5.1.4 Basados en redes neuronales 5.1.5 Híbridos	4.0	0.0	1.5
	5.2. Caracterización de sistemas existentes de traducción automática 5.2.1 Google Translate 5.2.2 Microsoft Translator 5.2.3 IBM Watson Language Translator 5.2.4 Bing Translator 5.2.5 Otros	3.5		3.0
	Subtotal	7.5	0.0	4.5



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Tecnologías de lenguaje natural

HOJA 6 DE 9

ESTRATEGIAS DE APRENDIZAJE	EVALUACIÓN DE LOS APRENDIZAJES
Estrategia de Aprendizaje Orientado a Proyectos El alumno desarrollará las siguientes actividades: <ol style="list-style-type: none">1. Desarrollo de conceptos teóricos e indagación documental.2. Presentación oral de los resultados de indagación documental.3. Realización de prácticas en equipo e individual de los temas que sean vistos en clase que le permita el análisis de los temas.4. Resolución de casos de estudio en equipo e individual de los temas que sean vistos en clase que le permita el análisis de los temas.5. Desarrollo de un proyecto donde se implemente los procesos de lenguaje natural vistos en clase.6. Presentación oral del proyecto7. Uso de software para procesamiento de textos en lenguaje natural.	Evaluación diagnóstica. Portafolio de evidencias: <ol style="list-style-type: none">1. Reporte de indagación2. Exposición de resultados de indagación3. Reporte de prácticas4. Reporte de resolución de casos de estudio5. Reporte de proyecto6. Exposición de proyecto7. Reporte de uso de software

RELACIÓN DE PRÁCTICAS			
PRÁCTICA No.	NOMBRE DE LA PRÁCTICA	UNIDADES TEMÁTICAS	LUGAR DE REALIZACIÓN
1	Normalización de textos	I	Sala de cómputo
2	Generación de representación vectorial de textos	II	
3	Extracción de palabras similares	II	
4	Extracción de asociaciones entre palabras	II	
5	Extracción de terminología y palabras clave	II	
6	Extracción de información	II	
7	Generación de resumen	II	
8	Análisis de aspectos de personalidad	III	
9	Análisis de polaridad en opinión	III	
10	Implementación de un agente conversacional	IV	
11	Caracterización de un sistema existente de traducción automática	V	
		TOTAL DE HORAS 27.0	



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Tecnologías de lenguaje natural

HOJA 7 DE 9

Bibliografía							
Tipo	Autor(es)	Año	Título del documento	Editorial	Documento		
					Libro	Antología	Otros
C	Aggarwal, C. C.	2018	Machine Learning for Text	Cham: Springer International Publishing/ 9783319735306	X		
B	Bird, S., Klein, E., Loper, E.	2009	Natural Language Processing with Python	O'Reilly Media/ 9780596516499	X		
C	Ganegedara, T.	2018	Natural Language Processing with TensorFlow: Teach Language to Machines Using Python's Deep Learning Library	Packt Publishing Ltd./ 9781788478311	X		
B	Jurafsky, D., Martin, J.	2008	Speech and Language Processing	Pearson Prentice Hall / 9780131873216	X		
C	Kamath, U., Liu, J., Whitaker, J.	2019	Deep learning for NLP and Speech Recognition	Springer/ 9783030145989	X		
B	Poibeau T.	2017	Machine Translation	MIT Press Essential Knowledge series/ 9780262534215	X		
C	Sarkar, D.	2019	Text Analytics with Python: A Practitioner's Guide to Natural Language Processing	Apress/ 9781484243534	X		
B	Shevat, A.	2017	Designing Bots	O'Reilly Media/9781491974827	X		
C	Vajjala, S., Et Al	2020	Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP System	O'Reilly Media/9781492054054	X		
B	Zhai, C., Massung, S.	2016	Text Data Management and Analysis	ACM and Morgan & Claypool Publishers/ 9781970001167	X		



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Tecnologías de lenguaje natural

HOJA 8 DE 9

Recursos digitales								
Autor, año, título y Dirección Electrónica	T e x t o	S i m u l a d o r	I m a g e n	T u t o r i a l	V i d e o	P r e s e n t a c i ó n	D i c c i o n a r i o	O t r o
¿Qué es la traducción automática? Recuperado el 29 de septiembre de 2020 de: https://www.systransoft.com/es/systran/tecnologia/que-es-la-traduccion-automatica/ .	X							
¿Quién habla? Agente conversacional vs chatbot vs asistente virtual Recuperado el 29 de septiembre de 2020 de: https://searchdatacenter.techtarget.com/es/cronica/Quien-habla-Agente-conversacional-vs-chatbot-vs-asistente-virtual .	X							
Aplicaciones de procesamiento de lenguaje natural. Recuperado el 29 de septiembre de 2020 de: https://core.ac.uk/download/pdf/18586869.pdf .	X							
El Procesamiento de Lenguaje Natural (PLN) en la producción de noticias. Recuperado el 29 de septiembre de 2020 de: https://www.youtube.com/watch?v=pPmpHyT_fCA .	X							
Minería de textos y sus aplicaciones. Recuperado el 29 de septiembre de 2020 de: http://www.semanticwebbuilder.org.mx/es_mx/swb/Mineria_de_textos_y_sus_aplicaciones .								
Procesamiento de lenguaje natural. Recuperado el 29 de septiembre de 2020 de: https://www.sas.com/es_ar/insights/analytics/what-is-natural-language-processing-nlp.html .								
Resumen automático. Recuperado el 29 de septiembre de 2020 de: https://www.meaningcloud.com/es/productos/resumen-automatico .	X							
Sistemas de diálogo hablado y multimodal. Recuperado el 29 de septiembre de 2020 de: http://www.ugr.es/~rlopezc/sistemas_dialogo.htm .	X							



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA ACADÉMICA
DIRECCIÓN DE EDUCACIÓN SUPERIOR



UNIDAD DE APRENDIZAJE: Tecnologías de lenguaje natural

HOJA 9 DE 9

PERFIL DOCENTE: Ingeniería, Licenciatura y/o Maestría en Ciencias de Computación o afines

EXPERIENCIA PROFESIONAL	CONOCIMIENTOS	HABILIDADES DIDÁCTICAS	ACTITUDES
Preferentemente dos años en la profesión en el área de Inteligencia Artificial y/o Procesamiento de lenguaje natural y dos años de docencia a nivel superior.	En Inteligencia Artificial y Procesamiento de lenguaje natural. Del Modelo Educativo Institucional (MEI).	Discursivas Cognoscitivas Metodológicas De conducción del grupo Para evaluar	Responsabilidad Tolerancia Honestidad Respeto Paciencia Disciplina Constancia Compromiso social e Institucional

ELABORÓ

REVISÓ

AUTORIZÓ

Dra. Olga Kolesnikova
Profesora Coordinadora

Dr. Joel Omar Juárez Gambino
Profesor Colaborador

M. en C. Andrés Ortigoza Campos
Director ESCOM

Dr. Eric Manel Rosales Peña
Alfaro
Profesor Colaborador

M. en C. Iván Giovanni Mosso
García
Subdirección Académica

Ing. Carlos Alberto Paredes Treviño
Director UPIIC