



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA ACADÉMICA**  
**DIRECCIÓN DE EDUCACIÓN SUPERIOR**



**PROGRAMA SINTÉTICO**

<b>UNIDAD ACADÉMICA:</b> ESCUELA SUPERIOR DE CÓMPUTO (ESCOM), UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA, CAMPUS ZACATECAS (UPIIZ)	
<b>PROGRAMA ACADÉMICO:</b> Ingeniería en Sistemas Computacionales	
<b>UNIDAD DE APRENDIZAJE:</b> Natural language processing	<b>SEMESTRE:</b> VII <b>PLAN DE ESTUDIOS:</b> 2020

<b>PROPÓSITO DE LA UNIDAD DE APRENDIZAJE</b>				
Aplica el procesamiento de lenguaje natural a partir de los procesos de normalización de textos, extracción de información, clasificación de textos, análisis de opinión y sistemas de diálogo.				
<b>CONTENIDOS:</b>	I. Contextualización y normalización del texto II. Representación del texto y extracción de información III. Clasificación y agrupamiento de textos IV. Análisis de opinión y detección de emociones V. Sistemas de diálogo			
<b>ORIENTACIÓN DIDÁCTICA:</b>	<b>Métodos de enseñanza</b>		<b>Estrategias de aprendizaje</b>	
	a) Inductivo		a) Estudio de Casos	
	b) Deductivo	X	b) Aprendizaje Basado en Problemas	
	c) Analógico		c) Aprendizaje Orientado a Proyectos	X
	d) Heurístico	X	d)	
<b>EVALUACIÓN Y ACREDITACIÓN:</b>	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	X		
<b>BIBLIOGRAFÍA BÁSICA:</b>	<b>Autor(es)</b>	<b>Año</b>	<b>Título del documento</b>	<b>Editorial / ISBN</b>
	Aggarwal, C.	2018	Machine learning for text	Springer/ 9783319735306
	Jurafsky, D. y Martin, J.	2021	Speech and Language Processing	Pearson Prentice Hall/ 9780131873216
	*Manning, C. y Schütze, H.	1999	Foundations of statistical natural language processing	MIT press/ 9780262133609
	Sarkar, D.	2019	Text analytics with Python: a practitioner's guide to natural language processing	Apress/ 9781484243534
	Zhai, C. y Massung, S.	2016	Text data management and analysis: a practical introduction to information retrieval and text mining	Morgan & Claypool/ 978-1970001167

\*Bibliografía clásica



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA ACADÉMICA**  
**DIRECCIÓN DE EDUCACIÓN SUPERIOR**



**PROGRAMA DE ESTUDIOS**

**UNIDAD DE APRENDIZAJE:** Natural language processing

**HOJA 2 DE 8**

<b>UNIDAD ACADÉMICA:</b> ESCUELA SUPERIOR DE CÓMPUTO (ESCOM), UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA, CAMPUS ZACATECAS (UPIIZ)		
<b>PROGRAMA ACADÉMICO:</b> Ingeniería en Sistemas Computacionales		
<b>SEMESTRE:</b> VII <b>PLAN DE ESTUDIOS:</b> 2020	<b>ÁREA DE FORMACIÓN:</b> Profesional	<b>MODALIDAD:</b> Escolarizada
<b>TIPO DE UNIDAD DE APRENDIZAJE:</b> Teórica- práctica/ Optativa		
<b>VIGENTE A PARTIR DE:</b> Enero 2023	<b>CRÉDITOS:</b> <b>TEPIC:</b> 7.5 <b>SATCA:</b> 6.3	
<b>INTENCIÓN EDUCATIVA</b> La unidad de aprendizaje contribuye al perfil de egreso de la Ingeniería en Sistemas Computacionales proporcionándole los fundamentos teórico-prácticos que le permitan desarrollar software que maneje texto mediante técnicas de procesamiento de lenguaje natural. Asimismo, desarrolla habilidades transversales como el aprendizaje autónomo, el trabajo en equipo, la capacidad analítica y la resolución de problemas.  Esta unidad de aprendizaje se relaciona de manera antecedente con Machine Learning e Inteligencia artificial; de forma lateral con Trabajo Terminal I; y de forma consecuente no tiene relación.		
<b>PROPÓSITO DE LA UNIDAD DE APRENDIZAJE</b> Aplica el procesamiento de lenguaje natural a partir de los procesos de normalización de textos, extracción de información, clasificación de textos, análisis de opinión y sistemas de diálogo.		

<b>TIEMPOS ASIGNADOS</b>  <b>HORAS TEORÍA/SEMANA:</b> 3.0  <b>HORAS PRÁCTICA/SEMANA:</b> 1.5  <b>HORAS TEORÍA/SEMESTRE:</b> 54.0  <b>HORAS PRÁCTICA/SEMESTRE:</b> 27.0  <b>HORAS APRENDIZAJE AUTÓNOMO:</b> 24.0  <b>HORAS TOTALES/SEMESTRE:</b> 81.0	<b>UNIDAD DE APRENDIZAJE REDISEÑADA POR:</b> Academia de Ciencias de la Computación  <b>REVISADA POR:</b>  M. en C. Iván Giovanni Mosso García <b>Subdirector Académico ESCOM</b>  M. en C. Roberto Oswaldo Cruz Lejía <b>Subdirector Académico UPIIZ</b>  <b>APROBADA POR:</b> Consejo Técnico Consultivo Escolar  M. en C. Andrés Ortigoza Campos <b>Presidente ESCOM</b> 22/11/2022  Dr. Fernando Flores Mejía <b>Presidente del CTCE de UPIIZ</b> 27/06/2022	<b>APROBADO POR:</b> Comisión de Programas Académicos del Consejo General Consultivo del IPN.  24/11/2022  <b>AUTORIZADO Y VALIDADO POR:</b>  Mtro. Mauricio Igor Jasso Zaranda <b>Secretario Académico</b>
--	---	--



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA ACADÉMICA**  
**DIRECCIÓN DE EDUCACIÓN SUPERIOR**



UNIDAD DE APRENDIZAJE: Natural language processing

HOJA 3 DE 8

UNIDAD TEMÁTICA I Contextualización y normalización del texto	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
<b>UNIDAD DE COMPETENCIA</b>  Aplica procesos de normalización a textos con base en la segmentación, eliminación y lematización de palabras.	1.1. Contextualización 1.1.1. Sistemas y aplicaciones de lenguaje natural 1.1.2. Fundamentos lingüísticos para el procesamiento de lenguaje natural 1.1.3. Fundamentos matemáticos para el procesamiento de lenguaje natural 1.1.4. Lenguajes de programación y herramientas para el procesamiento de lenguaje natural	3.0		1.0
	1.2. Normalización de textos 1.2.1. Segmentación de palabras y oraciones 1.2.2. Eliminación de elementos no relevantes 1.2.3. Lematización y enraizamiento 1.2.4. Etiquetado con categorías gramaticales	4.5	3.0	1.5
	Subtotal	7.5	3.0	2.5

UNIDAD TEMÁTICA II Representación del texto y extracción de información	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
<b>UNIDAD DE COMPETENCIA</b>  Genera diferentes representaciones del texto a partir del modelo de espacio vectorial y las técnicas de extracción de información.	2.1. Representación de textos para su procesamiento semántico 2.1.1. Modelo de bolsa de palabras 2.1.2. Modelo de espacio vectorial	1.5		
	2.2. Características utilizadas en la representación de textos 2.2.1. Presencia (binarización) y frecuencia de palabras 2.2.2. Frecuencia de términos e inversa de frecuencia de documentos (TF-IDF) 2.2.3. Word embeddings	3.0	3.0	1.0
	2.3. Técnicas para la extracción de información 2.3.1. Expresiones regulares 2.3.2. Análisis sintáctico superficial y profundo 2.3.3. Frecuencia y relevancia de las palabras	3.0		2.0
	2.4. Extracción de información en textos 2.4.1. Palabras similares y asociaciones entre palabras 2.4.2. Palabras clave 2.4.3. Tópicos 2.4.4. Resumen automático 2.4.5. Entidades nombradas	4.5	3.0	2.0
	2.5. Generación de respuestas a preguntas	1.5		1.0
	Subtotal	13.5	6.0	6.0



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA ACADÉMICA**  
**DIRECCIÓN DE EDUCACIÓN SUPERIOR**



**UNIDAD DE APRENDIZAJE:** Natural language processing

**HOJA:** 4 **DE** 8

UNIDAD TEMÁTICA III Clasificación y agrupamiento de textos	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
<b>UNIDAD DE COMPETENCIA</b>  Implementa la agrupación y clasificación de textos a partir de los algoritmos de aprendizaje automático.	3.1. Clasificación de textos 3.1.1. Descripción de la tarea de clasificación 3.1.2. Representación del texto para tareas de clasificación 3.1.3. El problema de la dispersión 3.1.4. Reducción de dimensionalidad 3.1.5. Métricas para la evaluación de la clasificación de texto	3.0		1.5
	3.2. Aplicación de algoritmos de aprendizaje automático para la clasificación de texto 3.2.1. Regresión logística 3.2.2. Naive Bayes 3.2.3. Máquinas de soporte vectorial	3.0	3.0	1.0
	3.3. Agrupamiento de texto 3.3.1. Descripción de la tarea de agrupamiento 3.3.2. Agrupamiento duro 3.3.3. Agrupamiento suave 3.3.4. Agrupamiento jerárquico 3.3.5. Métricas para la evaluación del agrupamiento de texto	3.0	1.5	1.0
	3.4. Aplicación de algoritmos de agrupamiento 3.4.1. K-medias 3.4.2. DBSCAN	3.0	3.0	1.0
	Subtotal	12.0	7.5	4.5



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA ACADÉMICA**  
**DIRECCIÓN DE EDUCACIÓN SUPERIOR**



UNIDAD DE APRENDIZAJE: Natural language processing

HOJA: 5 DE 8

UNIDAD TEMÁTICA IV Análisis de opinión y detección de emociones	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
<b>UNIDAD DE COMPETENCIA</b>  Implementa algoritmos de clasificación de opiniones y detección de emociones a partir de métodos de aprendizaje automático y basado en léxico.	4.1. Aspectos básicos del análisis de opinión y detección de emociones	1.5		1.0
	4.1.1. Planteamiento del problema			
	4.1.2. Enfoques del análisis de opinión (basado en diccionarios o léxico, supervisado)			
	4.1.3. Enfoques de la detección de emociones (basado en diccionarios o léxico, supervisado)			
	4.2. Análisis de opinión	4.5	3.0	2.0
4.2.1. Clasificación de opinión basada en diccionarios				
4.2.2. Clasificación de opinión basada en algoritmos de aprendizaje automático				
4.3. Detección de emociones	4.5	3.0	2.0	
4.3.1. Modelo basados en léxico de dominio específico				
4.3.2. Modelo basados en léxico de dominio general				
4.3.3. Modelo basados en aprendizaje automático				
	Subtotal	10.5	6.0	5.0

UNIDAD TEMÁTICA V Sistemas de diálogo	CONTENIDO	HORAS CON DOCENTE		HRS AA
		T	P	
<b>UNIDAD DE COMPETENCIA</b>  Diseña sistemas de diálogo entre la máquina y el usuario con base en los diferentes modelos y aplicaciones.	5.1. Características de los sistemas de diálogo	3.0		1.0
	5.1.1. Tipo de comunicación sistema-usuario			
	5.1.2. Obtención de información			
	5.1.3. Particularidades de los usuarios			
	5.1.4. Base de conocimiento			
	5.2. Tipos de políticas	1.5		1.0
	5.2.1. Basadas en reglas			
	5.2.2. Basadas en la recuperación			
	5.2.3. Generativas			
	5.2.4. Basadas en gráficos de conocimiento			
	5.3. Diseño de los sistemas de diálogo	3.0	4.5	3.0
	5.3.1. Modelo del entorno			
	5.3.2. Modelo de la tarea			
	5.3.3. Modelo del usuario			
	5.3.4. Bases de conocimiento			
5.3.5. Técnicas de diseño (intuición y corpus)				
5.4. Aplicación de los sistemas de diálogo	3.0		1.0	
5.4.1. Sistemas de información				
5.4.2. Sistemas pregunta-respuesta				
5.4.3. Sistemas multimodales				
5.4.4. Sistemas de transacciones				
5.4.5. Traducción de la lengua oral				
	Subtotal	10.5	4.5	6.0



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA ACADÉMICA**  
**DIRECCIÓN DE EDUCACIÓN SUPERIOR**



**UNIDAD DE APRENDIZAJE:** Natural language processing

**HOJA:** 6 **DE** 8

ESTRATEGIAS DE APRENDIZAJE	EVALUACIÓN DE LOS APRENDIZAJES
<p><b>Estrategia de Aprendizaje Orientado a Proyectos</b></p> <p>El alumno desarrollará las siguientes actividades:</p> <ol style="list-style-type: none"> <li>Desarrollo de conceptos teóricos e indagación documental</li> <li>Exposición de resultados de la indagación documental</li> <li>Resolución de casos de estudio en equipo e individual de los temas que sean vistos en clase</li> <li>Desarrollo de proyecto donde se implementen los procesos de lenguaje natural:               <ol style="list-style-type: none"> <li>Normalización del texto</li> <li>Representación del texto</li> <li>Clasificación del texto</li> <li>Análisis y detección de emociones en el texto</li> </ol> </li> <li>Exposición del proyecto</li> <li>Uso de software para procesamiento de textos en lenguaje natural</li> <li>Realización de prácticas</li> </ol>	<p>Evaluación diagnóstica</p> <p>Portafolio de evidencias:</p> <ol style="list-style-type: none"> <li>Reporte de indagación</li> <li>Presentación digital sobre resultados de indagación</li> <li>Reporte de casos resueltos</li> <li>Reporte de proyecto</li> <li>Presentación digital de proyecto</li> <li>Reporte de uso de software</li> <li>Reporte de prácticas</li> <li>Evaluación escrita</li> </ol>

RELACIÓN DE PRÁCTICAS			
PRÁCTICA No.	NOMBRE DE LA PRÁCTICA	UNIDADES TEMÁTICAS	LUGAR DE REALIZACIÓN
1	Normalización de textos	I	Sala de cómputo
2	Representación de texto	II	
3	Extracción de palabras similares y asociaciones entre palabras	II	
4	Extracción de palabras clave y entidades nombradas	II	
5	Clasificación de texto	III	
6	Agrupamiento de texto	III	
7	Análisis de polaridad	IV	
8	Detección de emociones	IV	
9	Diseño de un chatbot	V	
		<b>TOTAL DE HORAS:</b>	27.0



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA ACADÉMICA**  
**DIRECCIÓN DE EDUCACIÓN SUPERIOR**



UNIDAD DE APRENDIZAJE: Natural language processing

HOJA: 7 DE 8

Bibliografía												
Tipo	Autor(es)	Año	Título del documento	Editorial/ISBN	Documento							
					Libro	Antología	Otros					
B	Aggarwal, C. C.	2018	Machine learning for text	Springer/ 9783319735306	X							
B	Jurafsky, D. & Martin, J.	2021	Speech and Language Processing	Pearson Prentice Hall/ 9780131873216	X							
B	*Manning, C. & Schütze, H.	1999	Foundations of statistical natural language processing	MIT press/ 978-0262133609	X							
C	Martí, M. A.	2003	Tecnologías del lenguaje	Editorial UOC/ 9788484298809	X							
C	Minker, W. & Bennacef, S.	2004	Speech and Human-Machine Dialog	Springer/ 978-0387572376	X							
B	Sarkar, D.	2019	Text analytics with Python: a practitioner's guide to natural language processing	Apress/ 9781484243534	X							
B	Zhai, C. & Massung, S.	2016	Text data management and analysis: a practical introduction to information retrieval and text mining	Morgan & Claypool/ 978-1970001167	X							
Recursos digitales												
Autor, año, título y Dirección Electrónica					Texto	Simulador	Imagen	Tutorial	Video	Presentación	Diccionario	Otro
Stanford University. (2022). Stanford CS224U: Natural Language Understanding Spring 2021. <a href="https://www.youtube.com/playlist?list=PLoROMvodv4rPt5D0zs3YhbWSZA8Q_DyiJ">https://www.youtube.com/playlist?list=PLoROMvodv4rPt5D0zs3YhbWSZA8Q_DyiJ</a>									X			
University of Michigan. (2019). Natural Language Processing (FULL COURSE). <a href="https://www.youtube.com/playlist?list=PLLssT5z_Dsk8BdawOVCaTCO99Ya58ryR">https://www.youtube.com/playlist?list=PLLssT5z_Dsk8BdawOVCaTCO99Ya58ryR</a>									X			



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA ACADÉMICA**  
**DIRECCIÓN DE EDUCACIÓN SUPERIOR**



**UNIDAD DE APRENDIZAJE:** Natural language processing

**HOJA:** 8 **DE** 8

**PERFIL DOCENTE:** Ingeniería en Sistemas Computacionales o afín, y/o Maestría en Ciencias de la Computación o afín.

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

**ELABORÓ**

**REVISÓ**

**AUTORIZÓ**

---

Dr. Joel Omar Juárez Gambino  
**Coordinador**

---

M. en C. Alfonso Sánchez Aguilar  
**Participante**

---

M. en C. Roberto Oswaldo Cruz  
Lejía  
**Subdirector Académico UPIIZ**

---

M. en C. Andrés Ortigoza Campos  
**Director ESCOM**

---

M. en C. Miguel Ángel Rodríguez  
Castillo  
**Participante**

---

M. en C. Iván Giovanni Mosso  
García  
**Subdirector Académico  
ESCOM**

---

Dr. Fernando Flores Mejía  
**Director UPIIZ**