



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



Unidad de Competencia						
Procesamiento de Lenguaje Natural (NLP)						
Cuatrimestre	Fecha de elaboración			Modalidad	Tipo	Área de formación
7	DD	MM	AÑO	• Escolarizada Opción: ○ Presencial	• Curso	Ciencia de datos, aprendizaje automático y modelos avanzados
	27	11	2025			
Perfil docente						
• Formación: ○ Ingeniería, Lingüística Computacional, Ciencia de Datos, Computación o áreas afines. ○ Posgrado deseable en NLP, IA, Deep Learning, Lingüística o Ciencias Cognitivas. • Experiencia profesional: ○ Mínimo 3 años desarrollando modelos NLP. ○ Dominio de Python, NLTK, spaCy, HuggingFace Transformers. ○ Experiencia con embeddings, modelos preentrenados, clasificación de texto, análisis de sentimiento o chatbots. • Dominio de inglés: ○ B2 o superior para lectura de papers y documentación técnica • Competencias digitales:						
Horas con Mediación Docente Cuatrimestrales		Horas con Mediación Docente semana/mes		Horas de Actividad Independiente	Horas de Práctica de Campo	Créditos
HT	HP	HT	HP	64	0	8
32	32	2	2			
Horas de docencia total de la unidad de competencia: 64						
Propósito general:				Presentación:		
El estudiantado implementa, entrena y evalúa modelos de				El procesamiento de lenguaje natural es una disciplina fundamental		



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



procesamiento de lenguaje natural mediante técnicas clásicas y modernas (embeddings, transformers, modelos preentrenados), para resolver tareas como clasificación, análisis de sentimiento, extracción de información y generación básica de lenguaje.		dentro de la IA moderna. La asignatura cubre desde la tokenización y preprocesamiento hasta técnicas avanzadas basadas en transformers, embeddings contextualizados y modelos preentrenados. Se conecta directamente con Deep Learning, Redes Neuronales, Machine Learning II, IA Generativa y LLMs, Auditoría Algorítmica y proyectos integradores.	
Competencias			
Genéricas		Disciplinares	
• Pensamiento crítico y computacional. • Capacidad de análisis. • Comunicación técnica. • Trabajo colaborativo.		• Procesa texto mediante técnicas clásicas y profundas. • Implementa modelos NLP modernos. • Evalúa modelos en tareas de lenguaje.	
Profesionales			
• Diseña sistemas conversacionales básicos. • Aplica modelos preentrenados y embeddings. • Integra soluciones NLP en aplicaciones reales.			
Producto final			
Sistema de Procesamiento de Lenguaje Natural, que incluya: Pipeline de preprocesamiento, Modelo implementado (clásico o basado en transformers), Métricas de evaluación, Análisis del comportamiento del modelo, Documentación técnica e Integración básica mediante API o notebook interactiva			
Enseñanza con perspectiva de género, interculturalidad y Derechos Humanos			
Con la intención de propiciar una enseñanza con perspectiva de género, interculturalidad y derechos humanos, la o el docente que imparta esta unidad de competencia deberá:			
• Eliminar y rechazar actitudes y comportamientos sexistas, clasistas, racistas y todo tipo de estereotipos que supongan discriminación y violencia en los ámbitos de enseñanza-aprendizaje (virtual y presencial), a partir del ejemplo en su práctica docente. • Promover que el desarrollo del programa educativo y material a utilizar se base en el respeto, protección, promoción y garantía de los derechos humanos, de la dignidad e integridad de todas las personas. • Promover la eliminación de estereotipos o cualquier otra práctica de discriminación o denigración de la persona, que pueda presentarse			



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



mediante el uso del lenguaje, el empleo de ejemplos, la diferenciación entre grupos, origen étnico o cualquier otra circunstancia.

- Respetar la identidad de género de las personas en todo momento y nombrarlas tal y como ellas lo deseen.
- Respetar y promover la igualdad de trato y de oportunidades entre las y los estudiantes, favoreciendo acciones afirmativas en beneficio de los grupos en situación de vulnerabilidad, incluso en las afectaciones personales provocadas por la pandemia COVID-19
- Promover y usar lenguaje incluyente y no sexista tanto en el salón de clases como en los trabajos, tesis y otros documentos que elabore la comunidad estudiantil.

Estos elementos son motivados por la información asequible en los programas de estudio de la Licenciatura en Derecho, Modalidad Educativa Presencial (sistema escolarizado), Tomo II, de la Universidad Nacional Autónoma de México (UNAM 2020).



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



Nivel Unidad/Subcompetencias	
Nombre de la Subcompetencia 1	Elementos de la subcompetencia 1
Fundamentos de NLP y Preprocesamiento de Texto.	Conocimientos: • Tokenización, stemming, lematización • Limpieza y normalización de texto • Representaciones: Bag of Words, TF-IDF • Introducción a corpus y datasets de texto
Número de semanas	
4 semanas	
Propósito de la subcompetencia	Habilidades: • Limpieza de texto • Tokenización y normalización • Construcción de representaciones clásicas Valores y actitudes: • Ética en manejo de datos lingüísticos • Responsabilidad
El estudiantado comprende y aplica técnicas clásicas para preparar texto en tareas de NLP.	
Estrategias Didácticas	Recursos didácticos
<i>Las estrategias y recursos didácticos consideran la inclusión y perspectiva de género, así como la diversidad cultural.</i>	
• Laboratorios con NLTK/spaCy • Procesamiento de corpus • Mini-proyecto	• NLTK • spaCy • Python

Nombre de la Subcompetencia 2	Elementos de la subcompetencia 2
Embeddings y Modelos Clásicos de NLP	Conocimientos: • Word embeddings: Word2Vec, GloVe • Contextualización básica
Número de semanas	
4 semanas	



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



Propósito de la subcompetencia	• Clasificadores clásicos: Naive Bayes, SVM, Logistic Regression • Extracción de características lingüísticas
El estudiantado utiliza embeddings y modelos clásicos para tareas de clasificación y análisis de texto.	Habilidades: • Rigor técnico • Honestidad en análisis
	Valores y actitudes: • Ética en manejo de datos lingüísticos • Responsabilidad
Estrategias Didácticas	Recursos didácticos
<i>Las estrategias y recursos didácticos consideran la inclusión y perspectiva de género, así como la diversidad cultural.</i>	
• Laboratorio Word2Vec • Clasificación de texto • Evaluación comparativa	• Gensim • scikit-learn • Datasets de texto

Nombre de la Subcompetencia 3	Elementos de la subcompetencia 3
Transformers y Modelos Preentrenados	Conocimientos: • Arquitectura Transformer • Preentrenamiento y fine-tuning • Modelos como BERT, RoBERTa, DistilBERT • HuggingFace Transformers
Número de semanas	
4 semanas	Habilidades: • Fine-tuning con HuggingFace • Manejo de tokenizadores avanzados • Interpretación de resultados
Propósito de la subcompetencia	
El estudiantado implementa modelos preentrenados mediante fine-tuning para tareas específicas de NLP.	



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



	Valores y actitudes: • Responsabilidad ética (sesgos lingüísticos) • Profesionalista
Estrategias Didácticas	Recursos didácticos
<i>Las estrategias y recursos didácticos consideran la inclusión y perspectiva de género, así como la diversidad cultural.</i>	
• Laboratorio transformer • Entrenamiento en GPU • Comparación de arquitecturas	• Transformers (HuggingFace) • PyTorch / TensorFlow • GPU local o cloud

Nombre de la Subcompetencia 4	Elementos de la subcompetencia 4
Aplicaciones de NLP y Sistemas Conversacionales Iniciales	Conocimientos: • Análisis de sentimiento • Extracción de información (NER) • Similaridad semántica • Introducción a chatbots basados en reglas y modelos Habilidades: • Construcción de NER básico • Implementación de sentiment analysis • Integración mediante API Valores y actitudes: • Ética y responsabilidad social • Respeto a privacidad y datos sensibles
Número de semanas	
4 semanas	
Propósito de la subcompetencia	
El estudiantado diseña aplicaciones NLP reales, integrando modelos entrenados y flujos de procesamiento.	
Estrategias Didácticas	Recursos didácticos
Las estrategias y recursos didácticos consideran la inclusión y perspectiva de género, así como la diversidad cultural.	



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



- Proyectos aplicados
- Integración con API Flask/FastAPI
- Presentación de resultados

- Jupyter
- FastAPI / Flask
- HuggingFace

Evaluación

La evaluación de los aprendizajes se desarrollará de forma continua durante el proceso de enseñanza-aprendizaje a través de los siguientes momentos:

Evaluación diagnóstica: Revisión de conceptos de Machine Learning II y Deep Learning.

Evaluación formativa: Laboratorios por unidad, Miniproyectos, Reportes parciales, Retroalimentación continua

Evaluación sumativa: Sistema NLP completo, Examen práctico y Portafolio digital

Estrategia de evaluación de los aprendizajes:

Estrategias que el estudiante desarrolla para la elaboración de las evidencias de aprendizaje.

- Mapas Conceptuales
- Mapas Mentales
- Ensayos
- Exposiciones
- Cuadros Sinópticos
- Esquemas
- Cuadros comparativos
- Mnemotecnias
- Maquetas
- Prototipos
- Cuestionarios
- Audios
- Videos
- Wikis
- Webquest

Instrumentos de Evaluación:

- Rúbricas
- Listas de cotejo
- Revisión de código
- Evaluaciones prácticas



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



• Examen • Diseño de actividades didácticas • Diseño de materiales didácticos • Discusión dirigida • Diagramas de flujo	
Referencias	
<i>Las referencias consideran la inclusión y perspectiva de género, así como la diversidad cultural.</i>	
Básicas:	Complementarias:
• Jurafsky, D. & Martin, J. (2023). Speech and Language Processing. • Goldberg, Y. (2020). Neural Network Methods for NLP. • HuggingFace Transformers Documentation.	• Mikolov et al., Word2Vec Papers • Devlin et al., BERT PaperspaCy Documentation
Videos:	
Bases de Datos:	
• CONRICYT • Kaggle • HuggingFace Datasets	