



1. Datos Generales de la asignatura

Nombre de la asignatura:	Análisis computacional del lenguaje
Clave de la asignatura:	IAD-2402
SATCA¹:	2-3-5
Carrera:	Ingeniería en Inteligencia Artificial

2. Presentación

Caracterización de la asignatura

La asignatura, Análisis Computacional del Lenguaje, aporta al perfil de egreso del ingeniero o ingeniera en Inteligencia Artificial los conocimientos, habilidades y destrezas requeridas para desarrollar sistemas inteligentes capaces de analizar, interpretar y generar texto en lenguaje humano de manera efectiva. La importancia de la asignatura radica en el hecho de mejorar la comunicación entre humanos y computadoras y potenciar el desarrollo de tecnologías que faciliten la interacción con las máquinas de manera natural y eficiente.

En un mundo cada vez más digitalizado, el desarrollo de este tipo de sistemas a partir de grandes volúmenes de datos no estructurados es esencial para empresas, organizaciones y profesionales que buscan mantenerse relevantes y competitivos en el mercado a través de la creación de productos y servicios innovadores, la mejora de la experiencia de usuario y la toma de decisiones informadas basadas en el análisis de datos textuales, lo que proporciona una ventaja competitiva significativa. La asignatura de Análisis computacional del lenguaje se relaciona estrechamente con las asignaturas de Modelos de Aprendizaje Automático, Aprendizaje Automático Profundo y Big Data y sus aplicaciones. Las primeras dos otorgan los fundamentos necesarios para el uso de algoritmos de aprendizaje automático y profundo para la generación de modelos capaces de reconocer patrones en texto con el objetivo de facilitar tareas como traducción automática, recuperación de información, clasificación de texto, análisis de sentimientos, entre otras. Con respecto a Big Data y sus aplicaciones, esta proporciona las habilidades y técnicas necesarias para trabajar con grandes volúmenes de datos textuales para extraer información relevante y significativa de estos.

Para abordar la asignatura de Análisis Computacional del Lenguaje se organiza el temario agrupando los contenidos conceptuales en cuatro temas principales. El Tema 1 proporciona una introducción al análisis computacional del lenguaje, también conocido como Procesamiento del Lenguaje Natural (PLN), sus antecedentes, los niveles lingüísticos, su relación con la Inteligencia Artificial (IA) y los recursos disponibles para llevar a cabo tareas de análisis computacional del lenguaje. En el Tema 2 se analiza el flujo de trabajo del análisis computacional del lenguaje iniciando por la extracción de datos, el preprocesamiento de texto, y el aumento y transformación de datos.

¹ Sistema de Asignación y Transferencia de Créditos Académicos



El Tema 3 aborda la generación y uso de modelos de lenguaje a través de algoritmos clásicos de aprendizaje automático, aprendizaje profundo y transformadores. Finalmente, el Tema 4 aborda algunas de las aplicaciones del análisis computacional del lenguaje como la generación automática de texto y de resúmenes, al análisis de sentimiento, agentes conversaciones y asistentes virtuales.

Intención didáctica

La extensión de los contenidos antes descritos debe ser adecuada de tal forma que sea posible cubrir los conceptos fundamentales de cada tema, así como proporcionar ejemplos prácticos y casos de estudio que permitan una comprensión más profunda. El enfoque de la asignatura debe ser práctico y orientado al desarrollo de proyectos que permitan a los estudiantes aplicar los conceptos aprendidos bajo un paradigma colaborativo y dar solución a problemas de la vida real.

El enfoque de la asignatura debe permitir a los estudiantes desarrollar competencias tales como resolución de problemas complejos, trabajo colaborativo, pensamiento crítico y analítico, comunicación efectiva, aprendizaje autónomo y la adaptabilidad y flexibilidad en el uso de diferentes tecnologías y herramientas de desarrollo en el ámbito del análisis computacional del lenguaje. Por su parte, el docente debe actuar como facilitador de aprendizaje, proporcionando orientación y apoyo durante el proceso de aprendizaje, ejemplos claros y casos de estudio que permitan a los estudiantes comprender los conceptos teóricos y su aplicación práctica y retroalimentación constructiva que permita un desarrollo continuo de habilidades y competencias. Debe fomentar la participación, discusión e intercambio de ideas de los estudiantes en el desarrollo de proyectos para la solución de problemas de la vida real. Además, debe estar actualizado con las últimas tendencias y tecnologías en el campo de análisis computacional del lenguaje para enriquecer la experiencia educativa y preparar a los estudiantes para los desafíos del mundo real.

3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
Tecnológico Nacional de México del 4 al 6 de marzo de 2024.	Representantes de los Institutos Tecnológicos de: Celaya, Chihuahua, Iztapalapa III, La Paz, Matehuala, Mérida, Minatitlán, Querétaro, Saltillo, Tijuana. Institutos Tecnológico Superior de Teziutlán. Tecnológico de Estudios Superiores de Ixtapaluca.	Propuesta sintética de la carrera de Ingeniería en Inteligencia Artificial.



Tecnológico Nacional de México del 22 al 26 de abril de 2024	Representantes de los Institutos Tecnológicos de: Celaya, Chihuahua, Iztapalapa III, La Paz, Matehuala, Mérida, Minatitlán, Querétaro, Saltillo, Tijuana. Institutos Tecnológico Superior de Teziutlán, Tecnológico de Estudios Superiores de Ixtapaluca.	Diseño y/o desarrollo curricular de la carrera de Ingeniería en Inteligencia Artificial.
Tecnológico Nacional de México del 27 al 31 de mayo de 2024.	Representantes de los Institutos Tecnológicos de: Celaya, La Paz, Matehuala, Mérida, Minatitlán.	Consolidación curricular de la carrera de Ingeniería en Inteligencia Artificial.

4. Competencia(s) a desarrollar

Competencia(s) específica(s) de la asignatura
Desarrolla sistemas inteligentes capaces de analizar, interpretar y generar texto en lenguaje humano de manera efectiva mediante el uso de técnicas y tecnologías actuales de análisis computacional del lenguaje.

5. Competencias previas

<ul style="list-style-type: none"> ● Identifica los principios fundamentales de estructuras de datos para aplicarlos en la organización y manipulación eficiente de grandes cantidades de texto y datos lingüísticos. ● Comprende los conceptos básicos de aprendizaje automático para la generación de modelos. ● Maneja herramientas de desarrollo de software y entornos de programación, demostrando habilidades para escribir, probar y depurar códigos en diferentes lenguajes de programación. ● Comprende los conceptos fundamentales de Big Data, reconociendo su importancia en el almacenamiento y recuperación de información.
--

6. Temario

No.	Temas	Subtemas
1	Introducción al análisis computacional del lenguaje	1.1. Antecedentes. 1.2. Niveles lingüísticos. 1.2.1. Fonológico. 1.2.2. Morfológico. 1.2.3. Léxico. 1.2.4. Sintáctico. 1.2.5. Semántico. 1.2.6. Discurso. 1.2.7. Pragmático.



		<ul style="list-style-type: none"> 1.3. Análisis computacional del lenguaje e Inteligencia Artificial. 1.4. Recursos para el análisis computacional del lenguaje.
2	Flujo de trabajo del análisis computacional del lenguaje	<ul style="list-style-type: none"> 2.1. Extracción de datos. 2.2. Preprocesamiento de texto. <ul style="list-style-type: none"> 2.2.1. Tokenización. 2.2.2. Derivación y lematización. 2.2.3. Eliminación de palabras vacías. 2.2.4. Análisis de dependencias. 2.2.5. Reconocimiento de entidades nombradas. 2.2.6. Manejo de emojis. 2.3. Aumento de datos. 2.4. Transformación de datos. <ul style="list-style-type: none"> 2.4.1. Vectorización basada en frecuencia. 2.4.2. Bolsa de palabras. 2.4.3. Incrustaciones de palabras.
3	Modelos de lenguaje	<ul style="list-style-type: none"> 3.1. Modelos clásicos de aprendizaje automático para análisis de texto. <ul style="list-style-type: none"> 3.1.1. Agrupamiento. 3.1.2. Clasificación. 3.2. Aprendizaje profundo para análisis de texto. <ul style="list-style-type: none"> 3.2.1. Redes Neuronales Convolucionales (CNN). 3.2.2. Redes Neuronales Recurrentes (RNN) 3.2.3. Memoria Larga a Corto Plazo (LSTM) 3.2.4. LSTM Bidireccional (BiLSTM). 3.3. Transformadores. <ul style="list-style-type: none"> 3.3.1. Arquitectura codificador-decodificador 3.3.2. Mecanismos de atención.
4	Aplicaciones de análisis computacional del lenguaje	<ul style="list-style-type: none"> 4.1. Generación automática de texto. 4.2. Generación automática de resúmenes. 4.3. Análisis de sentimientos. 4.4. Agentes conversacionales. 4.5. Asistentes virtuales.



7. Actividades de aprendizaje de los temas

1. Introducción al análisis computacional del lenguaje	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Conoce la historia, niveles lingüísticos y recursos del Procesamiento de Lenguaje Natural, así como su relación con la Inteligencia Artificial.</p> <p><i>Genéricas:</i></p> <ul style="list-style-type: none"> ● Capacidad de análisis y síntesis. ● Capacidad de comunicación oral y escrita. ● Habilidad para buscar, procesar y analizar información procedente de fuentes diversas. ● Capacidad de investigación. 	<ul style="list-style-type: none"> ● Desarrollar una línea del tiempo del análisis computacional del lenguaje. ● Diseñar una infografía sobre los niveles lingüísticos. ● Investigar y presentar en plenaria la relación del análisis computacional del lenguaje con la Inteligencia Artificial. ● Investigar y analizar diferentes recursos para el desarrollo de sistemas basados en análisis computacional del lenguaje.
2. Flujo de trabajo del análisis computacional del lenguaje	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Conoce y aplica el flujo de trabajo del análisis computacional del lenguaje.</p> <p><i>Genéricas:</i></p> <ul style="list-style-type: none"> ● Capacidad de análisis y síntesis. ● Capacidad de comunicación oral y escrita. ● Habilidad para buscar, procesar y analizar información procedente de fuentes diversas. ● Capacidad de aplicar los conocimientos en la práctica. 	<ul style="list-style-type: none"> ● Implementa tareas de extracción de datos de fuentes no estructuradas. ● Realiza tareas de preprocesamiento de texto mediante tecnologías actuales. ● Implementa técnicas de aumento de datos para diversificar y mejorar conjuntos de datos existentes. ● Aplica técnicas de transformación de datos mediante tecnologías de vanguardia.
3. Modelos de lenguaje	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Genera modelos de lenguaje basados en aprendizaje automático, aprendizaje profundo y transformadores.</p> <p><i>Genéricas:</i></p> <ul style="list-style-type: none"> ● Capacidad de análisis y síntesis. ● Capacidad de comunicación oral y escrita. 	<ul style="list-style-type: none"> ● Genera modelos de lenguaje mediante algoritmos de aprendizaje automático. ● Genera modelos de lenguaje mediante técnicas de aprendizaje profundo. ● Analiza la arquitectura codificador-decodificador y los mecanismos de atención en transformadores. ● Implementa modelos pre entrenados de lenguaje basados en transformadores.



<ul style="list-style-type: none"> ● Habilidad para buscar, procesar y analizar información procedente de fuentes diversas. ● Capacidad de aplicar los conocimientos en la práctica. ● Capacidad de trabajo en equipo. 	
4. Aplicaciones de análisis computacional del lenguaje	
Competencias	Actividades de aprendizaje
<p><i>Específica(s):</i> Conoce las aplicaciones de análisis computacional del lenguaje y su uso en la Inteligencia Artificial.</p> <p><i>Genéricas:</i></p> <ul style="list-style-type: none"> ● Capacidad de análisis y síntesis. ● Capacidad de comunicación oral y escrita. ● Habilidad para buscar, procesar y analizar información procedente de fuentes diversas. ● Capacidad de aplicar los conocimientos en la práctica. ● Capacidad de trabajo en equipo. 	<ul style="list-style-type: none"> ● Implementa mecanismos de generación automática de texto y resúmenes. ● Desarrolla un sistema para el análisis de sentimientos en un contexto específico. ● Crea un agente conversacional para un dominio específico mediante tecnologías actuales. ● Desarrolla y presenta en equipo y en plenaria una aplicación basada en análisis computacional del lenguaje que dé solución a una problemática de la vida real.

8. Práctica(s)

<p>Con el objetivo de vincular y fortalecer el aprendizaje del saber con el saber hacer se recomienda la elaboración y desarrollo de las siguientes prácticas:</p> <p>Tema 2</p> <ul style="list-style-type: none"> ● Desarrollo de un módulo de software basado en técnicas de Web Scrapping. ● Diseño y desarrollo de un sistema para el preprocesamiento de texto. ● Implementación de técnicas de transformación de datos a conjuntos de datos disponibles en Internet. <p>Tema 3</p> <ul style="list-style-type: none"> ● Desarrollo de modelos de lenguaje basados en algoritmos de aprendizaje automático. ● Creación de modelos de lenguaje basados en redes neuronales. ● Implementación de modelos de lenguaje pre entrenados basados en transformadores para tareas de traducción automática y/o generación automática de resúmenes. <p>Tema 4</p> <ul style="list-style-type: none"> ● Diseño y desarrollo de un agente conversacional mediante tecnologías Low-code y no-code. ● Desarrollo de un módulo de análisis de sentimientos mediante tecnologías análisis computacional del lenguaje de vanguardia.
--



9. Proyecto de asignatura

El objetivo del proyecto que planteé el docente que imparta esta asignatura, es demostrar el desarrollo y alcance de la(s) competencia(s) de la asignatura, considerando las siguientes fases:

Fundamentación: marco referencial (teórico, conceptual, contextual, legal) en el cual se fundamenta el proyecto de acuerdo con un diagnóstico realizado, mismo que permite a los estudiantes lograr la comprensión de la realidad o situación objeto de estudio para definir un proceso de intervención o hacer el diseño de un modelo.

Planeación: con base en el diagnóstico en esta fase se realiza el diseño del proyecto por parte de los estudiantes con asesoría del docente; implica planificar un proceso: de intervención empresarial, social o comunitario, el diseño de un modelo, entre otros, según el tipo de proyecto, las actividades a realizar los recursos requeridos y el cronograma de trabajo.

Ejecución: consiste en el desarrollo de la planeación del proyecto realizada por parte de los estudiantes con asesoría del docente, es decir en la intervención (social, empresarial), o construcción del modelo propuesto según el tipo de proyecto, es la fase de mayor duración que implica el desempeño de las competencias genéricas y específicas a desarrollar.

Evaluación: es la fase final que aplica un juicio de valor en el contexto laboral-profesión, social e investigativo, ésta se debe realizar a través del reconocimiento de logros y aspectos a mejorar se estará promoviendo el concepto de “evaluación para la mejora continua”, la metacognición, el desarrollo del pensamiento crítico y reflexivo en los estudiantes.

10. Evaluación por competencias

Para evaluar las actividades de aprendizaje se recomienda solicitar:

- Mapas conceptuales o mentales
- Reporte de investigación
- Reportes de prácticas
- Prototipos
- Tablas comparativas
- Estudio de casos
- Exposiciones en clase
- Portafolio de evidencias, entre otros.

Para verificar el nivel del logro de las competencias del estudiante se recomienda utilizar:

- Listas de cotejo.
- Guías de observación
- Rúbricas, entre otros.



11. Fuentes de información

1. Azunre, P. (2021). Transfer Learning for Natural Language Processing. Estados Unidos: Manning.
2. Lee, R. S. T. (2023). Natural Language Processing: A Textbook with Python Implementation. Alemania: Springer Nature Singapore.
3. Patel, A. A., Arasanipalai, A. U. (2021). Applied Natural Language Processing in the Enterprise. O'Reilly Media.
4. Raaijmakers, S. (2022). Deep Learning for Natural Language Processing. Reino Unido: Manning.
5. Singh, J. (2023). Natural Language Processing in the Real World: Text Processing, Analytics, and Classification. Estados Unidos: CRC Press.
6. Tunstall, L., Werra, L. V., Wolf, T. (2022). Natural Language Processing with Transformers. Estados Unidos: O'Reilly Media.