

PROGRAMA DE DESARROLLO DE LA ASIGNATURA

Denominación De La Asignatura: Procesamiento de Lenguaje Natural	Código: 64491107	Créditos: 3
Nombre del Profesor: Miguel Angel Rippe Espinosa	Grupo: 1 y 2	Período Académico: 2026-1S
Dispositivo Pedagógico Indicado en el Sílabo:		Horario
<ul style="list-style-type: none"> • Seminario Teórico-Práctico. 		
Recursos educativos y herramientas de apoyo: <ul style="list-style-type: none"> • Presentaciones con notebooks. • Espacio de atención a estudiantes. • Lecturas recomendadas. • Talleres aplicados 		
Formulación del problema general del cual se ocupará el curso: <p>El procesamiento de lenguaje natural es una rama de la inteligencia artificial, las ciencias de la computación y la lingüística que se ocupa del procesamiento (entendimiento y generación) del texto (lenguaje natural) por las maquinas, gracias a software de inteligencia artificial. En este curso se tratarán algunos fundamentos teóricos y prácticos de esta ciencia explorando modelos lingüísticos lógicos y probabilísticos con el fin de extraer conocimiento de bases de datos textuales.</p> <p>Adicionalmente, se hará mucho énfasis en todos lo relacionado con los LLM Large Language Models como ChatGPT entre otros</p> <p>Se recomienda que el estudiante tenga algunos conocimientos de aprendizaje automático y redes neuronales. Así mismo, requiere algunas habilidades elementales de programación en Python.</p>		

Método de evaluación y recomendaciones:

Dado que la asignatura tiene 3 créditos académicos y que cada crédito corresponde a 48 horas de trabajo, el tiempo total que se debe dedicar a esta asignatura durante el semestre es 144 horas. Por otro lado, como cada semana se realizan una sesión de 3 horas y el semestre consta de 16 semanas, de las 144 horas, 48 horas son de clase. Así, las 96 horas restantes corresponden a trabajo autónomo de cada estudiante, es decir, cada estudiante debe dedicar durante el semestre 96 horas de estudio por fuera de las clases. Con base en esto, y considerando que la apropiación de los conceptos estudiados en esta asignatura depende fuertemente del trabajo autónomo, se recomienda distribuir esas 96 horas de la forma más equitativa posible a lo largo del semestre.

Durante el desarrollo del curso se llevarán a cabo 3 tipos de actividades evaluativas, cada una con un peso específico en la calificación final.

- *Actividad 1. Proyecto aplicado de clase (60%)*

Descripción: El proyecto debe ser un caso de uso implementado de aplicación de NLP para resolver algún problema de empresa o persona. La calificación del proyecto está distribuida en tres partes: 1) Definición del caso de uso (20%), 2) Sustentación de Avance (20%) y 3) Presentación Final (20%) al final de la asignatura.

- *Actividad 2. Casos de clase (30%).*

Descripción. Durante el semestre se desarrollarán varios casos de aplicación, los cuales consistirán en ejercicios que el estudiante debe desarrollar en una sesión de clase, donde aplicara los conceptos estudiados en clase. El profesor entregará un enunciado al inicio de clase y el estudiante deberá desarrollarlo durante la sesión.

- *Actividad 3. Participación en clase y asistencia (10%).*

Recuerde que:

- El desarrollo del proyecto puede ser en grupo de máximo 3 estudiantes.
- Los casos se deben desarrollar de forma individual y serán siempre desarrollados en Python.
- La participación en clase se evaluará por la cantidad de aportes con preguntas o conceptos nuevos a la clase.
- Se permite y se motiva el uso de Inteligencias Artificiales Generativas en clase

Bibliografía:

Texto de referencia:

- Documentación NLTK <https://www.nltk.org/>
- *Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems.* por Sowmya Vajjala , Bodhisattwa Majumder, et ál. | 7 julio 2020. O'Reilly
- *Nueva gramática básica de la lengua española.* RAE
- Hardeniya, Perkins, Chopra, Joshi and Mathur. *Natural Language Processing: Python and NLTK.* Packt Publishing Ltd. 2016
- Deng and Liu. *Deep Learning in Natural Language Processing.* Springer Nature Singapore Pte Ltd. 2018
- Jacob Eisenstein. *Natural Language Processing.* MIT Press, 2019.
- Sahalan, Hassanien and Tolba. *Intelligent Natural Language Processing: Trends and Applications.* Springer International Publishing AG. 2018.
- *Best Resources to Learn Natural Language Processing in 2021* <https://www.mltut.com/best-resources-to-learn-natural-language-processing/>

Lecturas recomendadas:

- Mikolov, Tomas, Chen, Kai, Corrado, Greg, Dean, Jeffrey: Efficient estimation of word representations in vector space , *arXiv preprint arXiv:1301.3781* , 2013
- Schnabel, Tobias, Labutov, Igor, Mimno, David, Joachims, Thorsten: Evaluation methods for unsupervised word embeddings , *Proceedings of the 2015 conference on empirical methods in natural language processing* , 298–307, 2015
- Kübler, Sandra, McDonald, Ryan, Nivre, Joakim: Dependency Parsing , *Synthesis Lectures on Human Language Technologies* 2(1), 1-127, 2009
- Las que se encuentran en el material

Cursos Online recomendados:

CS224n: Natural Language Processing with Deep Learning. Stanford / Winter 2020. <http://web.stanford.edu/class/cs224n/>
NLP - Natural Language Processing with Python. Udemy. <https://www.udemy.com/course/nlp-natural-language-processing-with-python/>

PROGRAMACIÓN SEMANAL DEL CURSO

NO	Sesión	Tema	Lectura/ Video
1	06/02/26	Introducción general y manipulación básica de texto Comprender qué es el NLP, retos actuales, aplicaciones. Hacer repaso intensivo sobre manipulación básica de texto (strings en Python, expresiones regulares).	La Siguiente Gran Revolución: NLP (Procesamiento del Lenguaje Natural) Material de clase
2	13/02/26	Preprocesamiento I Tokenización, lowercasing, stopwords	Material de clase
3	20/02/26	Preprocesamiento II TF-IDF, visualización de corpus	Material de clase
4	27/02/26	Exploración semántica y sintáctica POS tagging, estructuras gramaticales, spaCy Introducción a corpus etiquetados	Material de clase
5	06/03/26	Modelos supervisados Clasificación binaria y multiclas	Material de clase
6	13/03/26	Caso de Estudio 1 Análisis de sentimientos Práctica en clase con dataset real (reseñas, tweets)	Material de clase
7	20/03/26	Aprendizaje no supervisado Clustering de texto, LDA Visualización con t-SNE o UMAP	Material de clase
8	27/03/26	Word Embeddings Word2Vec, analogías, similitud	Material de clase
9	03/04/26	Semana de Santa	
10	10/04/26	Redes neuronales simples Introducción a Keras, feedforward básico	Material de clase
11	17/04/26	RNNs RNN básicas, problemas de memoria, SimpleRNN	Material de clase

12	24/04/26	LSTM y generación de texto	Material de clase
13	01/05/26	Introducción a Transformers Arquitectura, atención, paper "Attention is All You Need" Preentrenamiento y fine-tuning	Material de clase
14	08/05/26	Aplicaciones con Transformers Carga y uso de modelos de HuggingFace	Material de clase
15	15/05/26	BERT, GPT y clasificación Comparación entre tareas con BERT y generación con GPT	Material de clase
16	22/05/26	LangChain y agentes RAG, pipelines con LLMs Introducción a LangChain	Material de clase
17	29/05/26	Proyecto Final	Material de clase

**UNIVERSIDAD CENTRAL
VICERRECTORÍA ACADÉMICA**

PROGRAMA PARA EL DESARROLLO DE LA ASIGNATURA

Criterios para evaluar la adquisición de competencias de los estudiantes:		Medios (acciones y producciones de los estudiantes a los que se aplicarán los criterios):
Finalidad 1:	<p>Competencia para enfrentar problemas prácticos usando procesamiento de lenguaje natural para resolverlos</p> <p>Esta competencia se evidencia a través de realizaciones en las cuales:</p> <ul style="list-style-type: none"> • La construcción de un proyecto como finalidad del espacio formativo donde se aplican los fundamentos conceptuales del NLP. • La definición clara del caso de uso de aplicación. • El desarrollo básico del caso de aplicación o caso de uso. 	<p>Materiales escritos:</p> <p>El estudiante debe realizar la propuesta de proyecto teniendo en cuenta los lineamientos dados para su elaboración. Y desarrollará la implementación de todo o parte del caso de uso durante el semestre.</p>
Finalidad 2:	<p>Competencia para entender los conceptos de los modelos matemáticos y computables numéricos que existen para el NLP</p> <p>Esta competencia se evidencia a través de realizaciones en las cuales:</p> <ul style="list-style-type: none"> • El estudiante propone preguntas o temas de discusión en clase. • Responde a preguntas y cuestionarios propuestos por el profesor en las diferentes sesiones. 	<p>Materiales escritos:</p> <p>La adquisición de esta competencia se evidenciará a través de la participación en clase y las respuestas teóricas presentadas por el profesor en las diferentes sesiones. Estas tareas permitirán a los estudiantes demostrar su comprensión y aplicación de los conceptos, herramientas y metodologías vistas lo largo de su proceso formativo. Esta participación debe reflejar una postura bien fundamentada, sustentada en argumentos académicos y científicos. Los estudiantes deberán:</p> <ul style="list-style-type: none"> • <i>Contribuir con Opiniones Informadas: presentar opiniones y perspectivas basadas en un análisis riguroso de la literatura y los datos relevantes.</i> • <i>Interactuar en Debates y Discusiones: involucrarse en debates constructivos, planteando preguntas y respuestas que fomenten un diálogo académico enriquecedor.</i> • <i>Aplicar el Pensamiento Crítico: evaluar y criticar de manera reflexiva las ideas y argumentos presentados por sus compañeros y por los instructores.</i> <p>Esta participación no solo demostrará la competencia adquirida, sino que también contribuirá al desarrollo de habilidades analíticas y argumentativas</p>

Finalidad 3:	<p>Competencia ética</p> <p>Esta competencia se evidencia a través de realizaciones en las cuales:</p> <ul style="list-style-type: none"> ● Se identifica el tipo de información que se manipulará en el desarrollo de la clase y su proyecto. ● Se reconoce la normatividad del uso de los datos personales teniendo en cuenta la ley de protección de datos en Colombia. 	<p>Realizar una búsqueda rigurosa de antecedentes que den guía y sustento a la propuesta proyecto de clase. En ese sentido, el estudiante debe citar y referenciar cualquier idea y/o resultado obtenido por otros autores a fin de dar el respectivo reconocimiento.</p> <p>Por otra parte, el estudiante debe considerar el realizar un acceso, tratamiento y análisis de datos desde un marco ético y legal.</p>
--------------	---	---

Firma del Profesor

Firma del Director del Programa

Fecha: Febrero 01 de 2026
