

---

**Profesor:** M. en I.  
Miguel Angel Camargo Rojas

**Horario:** miércoles a viernes de 14:30 a 16

**Email:** [macamargo@up.edu.mx](mailto:macamargo@up.edu.mx)

**Lugar:** Online (Zoom)

**Modalidad:** Teórico-Práctico

**Asesorías:** Vía Zoom (solicitar)

---

## OBJETIVOS

- Proporcionar una comprensión sólida de los fundamentos teóricos del procesamiento de lenguaje natural, incluyendo la sintaxis, la semántica y la pragmática del lenguaje.
- Capacitar a los estudiantes en el uso eficiente de módulos de Python de terceros, tales como NLTK y spaCy para el desarrollo y la implementación de soluciones prácticas en el campo del NLP, asegurando que adquieran habilidades aplicables en la industria y la investigación.

## METODOLOGÍA

Procesamiento de Lenguaje Natural es un curso teórico-práctico el cual se desarrollará mediante la explicación de los temas para la posterior implementación de cada solución en Python.

## PONDERACIÓN

Parcial #1 .....	33.3%
Parcial #2 .....	33.3%
Parcial #3 .....	33.3%

## FECHAS IMPORTANTES

Días de Asueto .....	Ninguno que nos afecte
Último día de clases .....	noviembre 25
Entrega Reporte Parcial #1.....	septiembre 4
Examen Parcial #1 .....	septiembre 5
Entrega Reporte Parcial #2.....	octubre 16
Examen Parcial #2 .....	octubre 17
Entrega Reporte Final .....	noviembre 28
Examen Departamental .....	diciembre 3

## POLÍTICAS DEL CURSO

- Mínimo aprobatorio: 6/10
- Tolerancia: 5 min\*
- No se permite el acceso a estudiantes no inscritos en la materia
- No desayunar mientras se toma la clase
- Mantener un ambiente cordial
- **No se aceptan entregas extemporáneas**
- Los exámenes se realizarán en la fecha y hora indicada.
- Las inasistencias no afectan la calificación. Sin embargo, no se aceptan trabajos extemporáneos.

## CÓDIGO DE ÉTICA

- Cero tolerancia ante el plagio
- Cero tolerancia a la copia académica. Los exámenes contarán con la siguiente leyenda. “Certifico que el trabajo realizado en este examen es estrictamente personal y reconozco que no cumplir las reglas o COMETER ACTOS DESHONESTOS en los exámenes puede resultar en la baja definitiva de la Universidad”

## SOFTWARE

- Python 3.x, Anaconda (Jupyter y Spyder) y/o VS Code y/o Google Colab

## TEMARIO

### 1. INTRODUCCIÓN

- 1.1. Historia del procesamiento de lenguaje natural
- 1.2. Conceptos lingüísticos
- 1.3. Trabajo basado en corpus
- 1.4. Herramientas básicas para el procesamiento de lenguaje natural

### 2. PROCESAMIENTO BÁSICO DE TEXTO

- 2.1. Expresiones regulares
- 2.2. Tokenización de Texto
- 2.3. Normalización y *Stemming*
- 2.4. Segmentación de Oraciones y Árboles de decisión

### 3. MODELADO DE LENGUAJE

- 3.1. N-grams
- 3.2. Probabilidades de N-grams
- 3.3. Perplejidad

### 4. CLASIFICACIÓN DE TEXTO

- 4.1. Etiquetado Gramatical
- 4.2. Análisis probabilístico del contexto gramatical
- 4.3. Análisis gramático probabilístico

### 5. APLICACIONES

- 5.1. Análisis de Sentimiento
  - 5.1.1. Naive Bayes para análisis de sentimiento

- 5.1.2. Lexicon de sentimientos
- 5.2. Alineación estadística y traducción automatizada
- 5.3. Agrupamiento y clasificación de textos
- 5.4. Recuperación de información
- 5.5. Categorización de texto

## BIBLIOGRAFÍA

- [1] D. Jurafsky y J. H. Martin, "Speech and Language Processing", 2da ed., Prentice Hall, 2009.
- [2] S. Bird, E. Klein, y E. Loper, "Natural Language Processing with Python", 1ra ed., O'Reilly Media, 2009. ★
- [3] R. Collobert et al., "Natural Language Processing (almost) from Scratch", 1ra ed., IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011.
- [4] J. Eisenstein, *Introduction to Natural Language Processing*. The MIT Press, 2019.
- [5] J. S. Coleman, "Natural Language Processing: A Very Short Introduction", 1ra ed., Oxford University Press, 2018.