**UNIVERSIDAD DEL ROSARIO**

**Formato de Sílabo**

**ESCUELA O FACULTAD: Escuela de Ingeniería, Ciencia y Tecnología**

**PROGRAMA O ÁREA: Matemáticas Aplicadas y Ciencias de la Computación**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Curso** | Introducción al procesamiento del lenguaje natural | | | | | |
| **Código** |  | | | | | |
| **Tipo de saber:** | Básico | **Complementario: X** | | | Formación integral | |
| **Tipo de curso:** | Obligatorio | | | **Electivo: X** | | |
| **Tipo de crédito:** | Horas de trabajo con acompañamiento directo del profesor:  48 | | Horas de trabajo independiente del estudiante:  96 | | | Total horas por periodo académico:  144 |
| **Número de Créditos:** | 1A+1C | | | | | |
| **Prerrequisitos** | Signal Processing and Feature Extraction, Inteligencia Artificial: representación y solución de problemas | | | | | |
| **Correquisitos** |  | | | | | |
| **Profesor** | Fabián Sánchez Salazar | | | | | |

**RESUMEN**

En este curso los estudiantes aprenderán los elementos básicos que conforman un sistema de procesamiento de lenguaje natural, así mismo como algunos métodos utilizados en este campo.

**PROPÓSITOS DE FORMACIÓN DEL CURSO**

Al final de este curso los estudiantes entenderán los algoritmos básicos utilizados para el procesamiento de lenguaje natural, así mismo como algunas de las aplicaciones de estas técnicas en sistemas reales.

**RESULTADOS DE APRENDIZAJE ESPERADOS (RAE)**

Al finalizar el curso el estudiante estará en la capacidad de:

* Definir las principales características de un sistema de procesamiento de lenguaje natural
* Conocer algunas técnicas básicas utilizadas en este campo.
* Describir lo que son sistemas de dialogo, y entender su implementación.
* Conocer las bases para el desarrollo de un sistema de procesamiento de lenguaje natural.

**CONTENIDOS**

1. Introducción al PLN

* Conceptos fundamentales
* Tipos de procesamiento de texto
* Conceptos básicos de lingüística
* Procesamiento básico de texto en Python.

2. Proceso de PLN

* Preparación de textos: Python NLTK
  + Tokenización
  + Lematización
  + Palabras de Parada
  + Etiquetado
* Modelos computables: Bag of Words, TF-IDF
* Procesamiento descriptivo de textos

3. Aprendizaje automático con textos (skinit-learn)

* Clasificación
* Agrupamiento

4. Redes Neuronales y PLN

* Introducción a redes neuronales
* Word2Vec
* Glove
* Fastext
* Modelos de lenguaje: RNR, LSTM.

**BIBLIOGRAFÍA**

**[1]** Jurafsky, D. & Martin J. (2000) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Upper Saddle River, NJ: Prentice-Hall.

**[2]** Manning, C. & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

[3] Bird, S. & Klein, E. & Loper, E. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O’Reilly.