

PROCESAMIENTO DE LENGUAJE NATURAL
CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS, A.C.
DR. ADRIÁN PASTOR LÓPEZ MONROY

CICLO

SEMESTRE 2

CLAVE DE LA ASIGNATURA

####

OBJETIVO(S) GENERAL(ES) DE LA ASIGNATURA

El lenguaje humano es un medio de comunicación muy más eficaz, y a su vez extremadamente complejo. Uno de los retos a resolver en esta era del conocimiento y la información es el tratamiento del lenguaje por medios automáticos. El objetivo general de este curso es proporcionar a los estudiantes conocimientos fundamentales y avanzados del Procesamiento del Lenguaje Natural y a las herramientas disponibles actualmente, que representan un nicho de oportunidad para el desarrollo de investigación y trabajos de alto impacto en ciencias de la computación con aplicación a una amplia variedad de áreas incluyendo, ciencias sociales, humanidades, lingüística, etc.

TEMAS Y SUBTEMAS

I. Fundamentos

- a. Procesamiento de Texto, Normalización y Segmentación
- b. Recursos Léxicos y Ontologías
- c. Etiquetado y Extracción de Información en Texto
- d. Información Sintáctica y Semántica

II. Minería de Textos

- a. Clasificación Supervisada y Patrones Lingüísticos
 - a. Bolsas de Términos
 - b. Esquemas de Pesado Automáticos
 - c. Análisis de Sentimientos y Estilo
- b. Representaciones Basadas en Conceptos
 - a. Análisis Semántico Latente (LSA) [PCA para Texto]
 - b. Asignación Latente de Dirichlet (LDA)
 - c. Representaciones Distribucionales de Términos

III. Modelos de Lenguaje

- a. Modelos de Lenguaje Estadísticos
- b. Modelos de Lenguaje Neuronales
- c. Incrustaciones de Términos (Embeddings)

- a. Word2Vec
- b. Glove
- c. FastText

IV. Aprendizaje Profundo para Procesamiento de Lenguaje Natural

- d. Redes Neuronales Convolucionales
- e. Redes Neuronales Recurrentes
 - a. Gated Recurrent Units (GRUs)
 - b. Long Short Term Memories (LSTMs)
- f. Modelos Secuencia a Secuencia (Sequence2Sequence)
- g. Mecanismos de Atención en Redes Neuronales
- h. Transformadores (Transformers)
- i. Codificadores Bidireccionales de Transformadores (BERT)
- j. Tópicos Tendencia en Procesamiento de Lenguaje Natural
 - a. Generación Adversaria de Texto
 - b. Ataques Adversarios
 - c. Transferencia de Conocimiento

ACTIVIDADES DE APRENDIZAJE

Cursos presenciales

Desarrollo de proyectos

Lectura de publicaciones especializadas

Preparación de presentaciones

CRITERIOS Y PROCEDIMIENTOS DE EVALUACIÓN Y ACREDITACIÓN

Tareas: 40%;

Exámenes (2): 30%;

Presentación Final: 30%

BIBLIOGRAFÍA

1. *Speech and Language Processing*, Daniel Jurafsky, James H Martin. Pearson
2. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*, Steven Bird, Ewan Klein, Edward Loper. O'Reilly
3. *Deep Learning*. Ian Goodfellow, Yoshua Bengio, Aaron Courville. MIT Press
4. *Deep Learning with Python*. Francois Chollet. Manning