

Macroentrenamiento en
Inteligencia Artificial MeIA **2025**

Módulo 2 - Especialización en Procesamiento
de Lenguaje Natural
Presentado por: Dr. Miguel Ángel Álvarez
Carmona

Junio 2025

Contenido

1. Introducción
2. Objetivos del módulo
3. Temario detallado del módulo
 - a. Fundamentos avanzados de PLN
 - b. Procesamiento de texto con aprendizaje automático
 - c. Modelos de lenguaje y embeddings semánticos
 - d. Modelos de aprendizaje profundo para PLN
 - e. Modelos Transformer
 - f. Aplicaciones prácticas del PLN
4. Requisitos para los participantes
5. Materiales, equipo y software necesarios
6. Metodología de enseñanza y evaluación
7. Calendario tentativo para la impartición del módulo
8. Modalidades de participación
9. Listado de integrantes

Introducción

Este módulo está diseñado para profundizar en el **Procesamiento de Lenguaje Natural (PLN)** utilizando técnicas avanzadas de inteligencia artificial. Se explorarán modelos de lenguaje modernos, embeddings semánticos, redes neuronales aplicadas a PLN y casos de uso prácticos en diversas industrias. El curso se impartirá en Google Colab, permitiendo a los participantes experimentar con modelos de IA en un entorno práctico sin necesidad de configuraciones adicionales.

Objetivos del Módulo

- Explorar los fundamentos avanzados de PLN.
- Implementar modelos de lenguaje y embeddings semánticos.
- Aplicar técnicas de aprendizaje automático para procesamiento de texto.
- Entrenar y evaluar modelos de aprendizaje profundo para PLN.
- Estudiar en profundidad los modelos Transformer.
- Desarrollar aplicaciones reales utilizando IA para PLN.

Temario detallado del módulo

1 Fundamentos avanzados de PLN

- Preprocesamiento de texto avanzado
- Técnicas de tokenización y lematización

2 Procesamiento de texto con aprendizaje automático

- Métodos clásicos de clasificación de texto (Naive Bayes, SVM, Random Forest)
- Extracción de características con TF-IDF y Bag of Words
- Reducción de dimensionalidad para PLN

3 Modelos de lenguaje y embeddings semánticos

- Word2Vec, GloVe y FastText
- Modelos contextuales: BERT, GPT y sus variantes

4 Modelos de aprendizaje profundo para PLN

- Redes neuronales recurrentes (RNN, LSTM, GRU)
- Modelos avanzados para PLN

5 Modelos Transformer

- Arquitectura de los Transformers
- Funcionamiento del mecanismo de atención
- Evaluación y ajuste de modelos preentrenados
- Aplicaciones de Transformers en PLN

6 Aplicaciones prácticas del PLN

- Agrupación de textos
- Análisis de sentimientos
- Generación de texto automática
- Traducción automática

Requisitos para los participantes

- Conocimientos básicos en programación (Python)
- Familiaridad con conceptos de IA y aprendizaje automático
- Acceso a una computadora con conexión a Internet
- Cuenta de Google para el uso de Google Colab
- Librerías de Python: NumPy, Pandas, Matplotlib, scikit-learn, TensorFlow, PyTorch, spaCy, NLTK, Hugging Face Transformers

Metodología de enseñanza y evaluación

- Metodología:
 - Clases teórico-prácticas en Google Colab
 - Implementación de modelos en ejercicios guiados
 - Mini-proyectos en PLN
- Evaluación:
 - Cuestionarios cortos
 - Evaluaciones prácticas
 - Proyecto final aplicando técnicas de PLN.

Calendario tentativo para la impartición del módulo

Día	Tema	Duración
1	Fundamentos avanzados de PLN	4 horas
2	Procesamiento de texto con aprendizaje automático	4 horas
3	Modelos de lenguaje y embeddings semánticos	4 horas
4	Modelos de aprendizaje profundo para PLN	4 horas
5	Modelos Transformer y aplicaciones prácticas del PLN	4 horas

Modalidades de participación

Por equipo

Listado de integrantes

- Miguel Ángel Álvarez Carmona (Representante)
- Centro de Investigación en Matemáticas
- Doctor en Ciencias Computacionales.
- IA y Procesamiento de Lenguaje Natural.

- Ángel Ramón Aranda Campos
- Centro de Investigación en Matemáticas
- Doctor en Ciencias con orientación en Ciencias de la Computación
- Inteligencia Artificial

- Ángel Díaz Pacheco.
- Universidad de Guanajuato
- Doctor en Ciencias Computacionales.
- Inteligencia Artificial.