

**TRANSFORMERS**  
**CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS, A.C.**  
**DR. ADRIÁN PASTOR LÓPEZ MONROY**  
**DR. FERNANDO SÁNCHEZ VEGA**

CICLO  
**SEMESTRE 3**

CLAVE DE LA ASIGNATURA  
**####**

### OBJETIVO(S) GENERAL(ES) DE LA ASIGNATURA

El Transformer es una arquitectura de *Deep Learning* que permite construir Redes Neuronales altamente efectivas y eficaces. Desde su llegada en 2017 en aplicaciones de Procesamiento de Lenguaje Natural ha dominado el estado del arte en aplicaciones que van desde Traducción Máquina, Análisis de Sentimiento, Inferencia en Texto, etc. El Transformer también ha mostrado resultados sobresalientes en áreas tales como Visión por Computadora y Procesamiento de Señales mostrando potencial al igualar o superar a otras arquitecturas profundas de referencia tales como las Redes Neuronales Convolucionales y Recurrentes. El objetivo general de este curso es proporcionar a los estudiantes conocimientos fundamentales y avanzados del mundo de los *Transformers*, y que actualmente 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, salud, lingüística, robótica, visión por computadora, procesamiento de señales, etc.

Para tomar esta materia es prerequisite ser avanzado en el uso de programación en Python/PyTorch y tener alguno de los siguientes perfiles:

1. **Perfil 1:** Haber tomado el curso Procesamiento de Lenguaje Natural en el CIMAT-Gto.
2. **Perfil 2:** Haber tomado Aprendizaje Máquina 1 o Reconocimiento Estadístico de Patrones u Optimización 1 en el CIMAT-Gto; idealmente las tres materias anteriores.
3. **Perfil 3:** Tener experiencia previa y fuerte trabajando con una amplia variedad de arquitecturas de Redes Neuronales con aplicaciones en Visión, Robótica o Procesamiento de Lenguaje Natural utilizando PyTorch o Tensorflow.

### TEMAS Y SUBTEMAS

#### I. **Introducción**

- a. Incrustaciones contextualizadas profundas (Embeddings)
- b. Modelos de Secuencia a Secuencia (Seq2Seq)
- c. Mecanismos y tipos de Atención
- d. Transformer Básico
  - a. Incrustaciones Posicionales (Positional Encoding)
  - b. Auto Atención (Self-Attention)

c. Cabezas y Atención Múltiple (Multi Head Attention)

## II. Transformers en Texto

- a. Bidirectional Encoder Representations from Transformers (BERT)
- b. Robustly Optimized BERT Pretraining Approach (RoBERTa)
- c. Aplicaciones: Traducción Máquina, Question Answering, ChatBots
- d. El REFORMER
- e. Pattern Exploiting Training and ALBERT
- f. Arquitecturas GPT2, GPT3 y XLNET
- g. Modelo T5-large y BART, ELECTRA

## III. Transformers en Visión, Audio y Multimodales

- a. Transformers en otras modalidades
- b. Multimodal BERT (MMBT)
- c. Multimodal Transformer for Text, Audio, Video and more
- d. Visual and Language BERT (ViLBERT)
- e. Visual Transformer: ViT

## IV. Adaptaciones en Transformers y Técnicas Avanzadas

- f. Estado del Arte en Transformers
- g. Modelos Destilados
- h. Estrategias de Fusión
- i. Técnicas Avanzadas de entrenamiento

## ACTIVIDADES DE APRENDIZAJE

Cursos híbridos (videos pregrabados, clase online y presencial), desarrollo de proyectos y tareas, lectura de publicaciones especializadas, y preparación de presentaciones y proyectos.

CRITERIOS Y PROCEDIMIENTOS DE EVALUACIÓN Y ACREDITACIÓN: Tareas: 40%; Quizzes: 20%; Proyecto Final: 40%

## 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
5. *Transformers for Natural Language Processing: Build innovative deep neural network architectures for NLP with Python, PyTorch, TensorFlow, BERT, RoBERTa, and more*. Denis Rothman. Packt Publishing; Edición 1st (29 enero 2021)