**PEC5: ESTADO DEL ARTE EN DEEP LEARNING**

**EDUARDO MORA GONZÁLEZ**

El artículo titulado "*Attention Is All You Need*" fue escrito *por Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser* e *Illia Polosukhin* en 2017. Este artículo representa un hito significativo en el campo del procesamiento del lenguaje natural.

El enfoque principal del artículo se centra en la presentación de la arquitectura de redes neuronales conocida como *Transformer*. Esta arquitectura introduce el mecanismo de atención como un componente fundamental en lugar de depender de las técnicas tradicionales basadas en Redes Neuronales Recurrentes (RNN). La atención, en este contexto, se refiere a la capacidad del modelo para asignar diferentes niveles de importancia a diferentes partes de la secuencia de entrada, permitiendo una mayor capacidad de modelado y captura de relaciones a largo plazo.

El enfoque propuesto en el artículo representa una novedad importante en el campo. Los autores demuestran que el modelo *Transformer* supera en rendimiento a los enfoques tradicionales basados en RNN en tareas de procesamiento del lenguaje natural. Esta arquitectura, basada únicamente en el mecanismo de atención, se considera un avance significativo y ha generado un gran interés en la comunidad científica.

El artículo presenta una descripción detallada de la arquitectura del modelo *Transformer*. En lugar de depender de las estructuras recurrentes, este modelo se basa en capas de atención que operan tanto en el nivel de palabras como en el nivel de subpalabras. El mecanismo de atención permite que el modelo se enfoque en diferentes partes de la secuencia de entrada en función de su relevancia contextual, lo que lleva a una mejor capacidad de representación y comprensión del lenguaje.

Además de describir la arquitectura del modelo, el artículo también proporciona resultados experimentales comparando el rendimiento del *Transformer* con los enfoques basados en RNN en diferentes conjuntos de datos y tareas de procesamiento del lenguaje natural. Estos experimentos demuestran consistentemente la superioridad del modelo *Transformer* en términos de calidad de traducción y capacidad de modelado del lenguaje.

En conclusión, el artículo "*Attention Is All You Need*" destaca la arquitectura *Transformer* como una alternativa altamente efectiva a los enfoques tradicionales basados en RNN en el campo del procesamiento del lenguaje natural. La propuesta de utilizar el mecanismo de atención como elemento central en el diseño de modelos ha generado un impacto significativo en el campo y ha abierto nuevas posibilidades para el desarrollo de modelos de aprendizaje automático en el procesamiento del lenguaje natural.