



UNIVERSIDAD  
DE  
CÓRDOBA

# Deep Learning in Sentiment Analysis: Recent Architectures

Aprendizaje Profundo

**Carlos Checa Moreno**

26507320G

i02chmoc@uco.es

Curso Académico 24/25



# ÍNDICE GENERAL

---

<b>CAPÍTULO 1</b>	<b>INTRODUCCIÓN .....</b>	<b>3</b>
<b>CAPÍTULO 2</b>	<b>FLUJO DE TRABAJO EN ANÁLISIS DE SENTIMIENTOS .....</b>	<b>4</b>
<b>CAPÍTULO 3</b>	<b>ARQUITECTURAS DE ANÁLISIS PROFUNDO EN ANÁLISIS DE SENTIMIENTOS</b>	<b>6</b>
3.1	MECANISMO DE ATENCIÓN .....	6
3.1.1	<i>Transformers</i> .....	7
3.2	REDES NEURONALES RECURRENTES .....	8
3.2.1	<i>Long Short-Term Memory (LSTM)</i> .....	9
3.2.2	<i>Gated Recurrent Unit (GRU)</i> .....	10
3.2.3	<i>Convolutional Neural Networks (CNN)</i> .....	11
<b>CAPÍTULO 4</b>	<b>DESAFÍOS ABIERTOS DE INVESTIGACIÓN .....</b>	<b>12</b>
<b>CAPÍTULO 5</b>	<b>CONCLUSIÓN .....</b>	<b>14</b>
<b>BIBLIOGRAFÍA.....</b>		<b>15</b>

# ÍNDICE DE FIGURAS

---

<b>FIGURA 1.</b> SUPERVISED SENTIMENT ANALYSIS WORKFLOW WITH DEEP LEARNING .....	4
<b>FIGURA 2.</b> ATTENTION MECHANISM WITH BI-DIRECTIONAL LSTM ENCODER FOR TEXT CLASSIFICATION.....	7
<b>FIGURA 3.</b> TRANSFORMER ARCHITECTURE WITH MULTI-HEADED SELF-ATTENTION.....	8
<b>FIGURA 4.</b> AUTOCONEXIÓN EN RED NEURONAL [3] .....	9
<b>FIGURA 5.</b> LONG-SHORT-TERM MEMORY CELL ARCHITECTURE FOR TEXT CLASSIFICATION .....	10
<b>FIGURA 6.</b> GATED RECURRENT UNIT ARCHITECTURE FOR TEXT CLASSIFICATION .....	10
<b>FIGURA 7.</b> TEXT CLASSIFICATION FLOW WITH CONVOLUTIONAL NEURAL NETWORK. ....	11

# Capítulo 1

## INTRODUCCIÓN

---

En la actualidad se está generando una enorme cantidad de texto no estructurado cargado de opiniones a través de diversas plataformas como redes sociales, foros o páginas de comercio online. Extraer valor real de esta información se ha convertido en un reto tanto para la industria como para el ámbito académico. En este contexto, el análisis de sentimiento ha cobrado gran relevancia dentro del procesamiento del lenguaje natural (NLP), ya que permite clasificar e identificar información subjetiva, siendo especialmente útil para entender la opinión pública, analizar productos o servicios, e incluso detectar fenómenos como el discurso de odio, el ciberacoso o la propagación de desinformación.

Además, este tipo de análisis se ha mostrado eficaz a la hora de predecir tendencias e incluso resultados electorales, superando en ocasiones a las encuestas tradicionales.

Por otro lado, los avances recientes en redes neuronales artificiales han supuesto un fuerte impulso para todo este campo. Gracias al aumento en la capacidad de procesamiento, el acceso a grandes volúmenes de datos y la mejora en las arquitecturas de aprendizaje profundo, se ha logrado un rendimiento sin precedentes en áreas clave como el reconocimiento de voz, la visión por computador y el propio análisis de lenguaje.

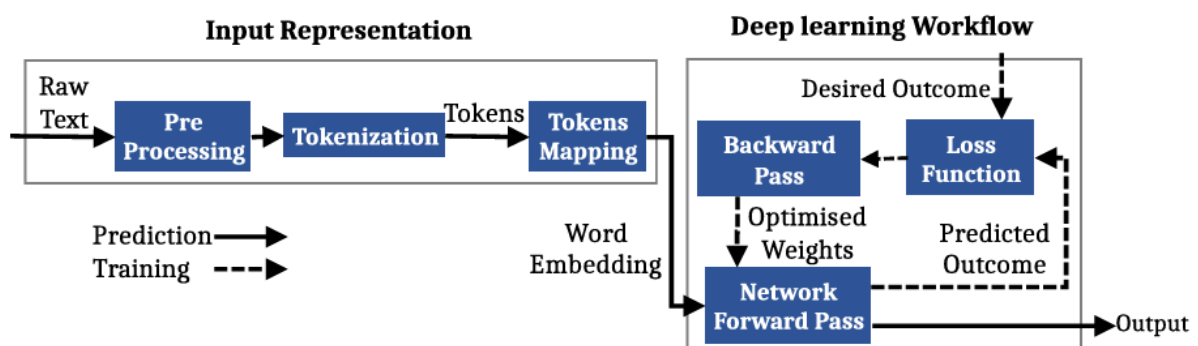
En este trabajo se explorarán las arquitecturas más utilizadas para el análisis de sentimientos revisando el artículo *Deep Learning in Sentiment Analysis: Recent Architectures* (Tariq Abdullah y Ahmed Ahmet, 2022).

## Capítulo 2

# FLUJO DE TRABAJO EN ANÁLISIS DE SENTIMIENTOS

Antes de que una red neuronal artificial pueda procesar datos de entrada en forma de texto, es necesario transformar dicho texto en una representación numérica conocida como word embeddings. Este proceso se realiza en tres etapas principales: un preprocesamiento opcional del texto, la tokenización (es decir, dividir el texto en palabras o frases clave) y finalmente la conversión de esos tokens en vectores numéricos.

Una vez obtenidos los word embeddings, estos se introducen en la red mediante una pasada hacia adelante (forward pass), que permite generar una predicción. En el caso del aprendizaje supervisado, esta predicción se compara con el resultado deseado para calcular el error, el cual se utiliza posteriormente en el proceso de backpropagation, ajustando los parámetros de la red.



**Figura 1.** Supervised sentiment analysis workflow with deep learning

El uso de word embeddings ha supuesto un avance notable en el campo del procesamiento del lenguaje natural, superando en rendimiento a métodos que no los emplean. Estos pueden ser tanto monolingües como multilingües. Entre los enfoques monolingües más conocidos destacan word2vec, GloVe y FastText, modelos que capturan relaciones semánticas y contextuales mediante diferentes técnicas de entrenamiento. En el ámbito multilingüe, se busca representar palabras de distintos idiomas en un espacio común. Como ejemplos tenemos modelos como VecMap y MUSE.

## Capítulo 3

# ARQUITECTURAS DE ANÁLISIS PROFUNDO EN ANÁLISIS DE SENTIMIENTOS

---

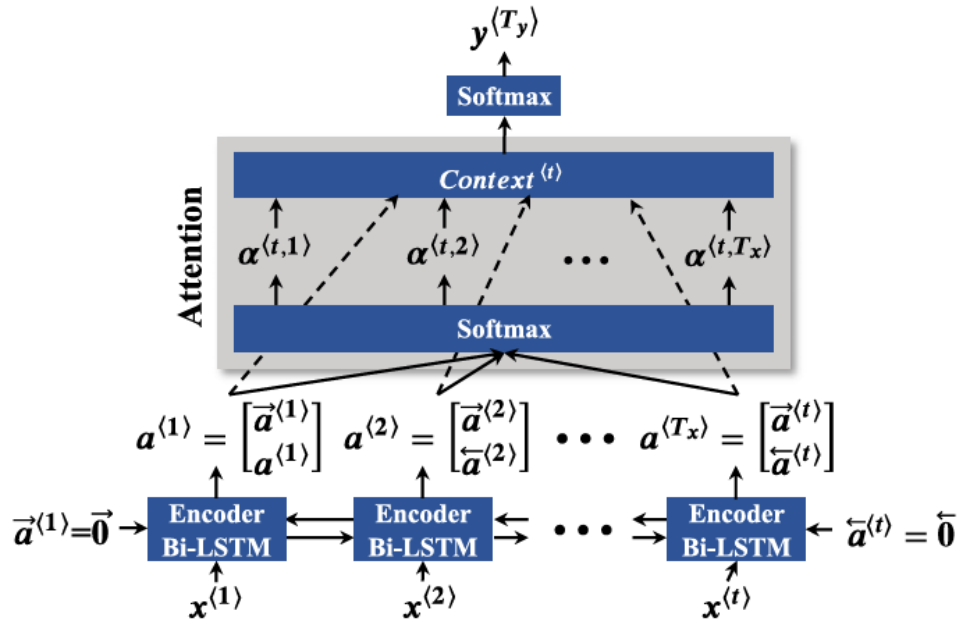
En este capítulo se presentan las principales arquitecturas de deep learning utilizadas para el análisis de sentimientos.

### 3.1 Mecanismo de Atención

El mecanismo de attention ha cobrado gran relevancia en tareas de traducción automática y se ha extendido ampliamente en el procesamiento de lenguaje natural. Su principal ventaja frente a arquitecturas tradicionales es que permite al modelo centrarse en la información más relevante, calculando una combinación ponderada de las características de entrada. Existen diversos tipos de attention como la Content-based, additive, location-based, general, dot-product y scaled dot-product. siendo esta última la base del modelo Transformer.

En el análisis de sentimiento, se utiliza comúnmente la attention basada en ubicación, aplicada sobre codificadores como LSTM bidireccionales. Este mecanismo calcula pesos de atención mediante funciones softmax y genera un vector de contexto que resume la información más importante de la secuencia, optimizando sus parámetros durante el proceso de entrenamiento.

---



**Figura 2.** Attention mechanism with bi-directional LSTM encoder for text classification

### 3.1.1 Transformers

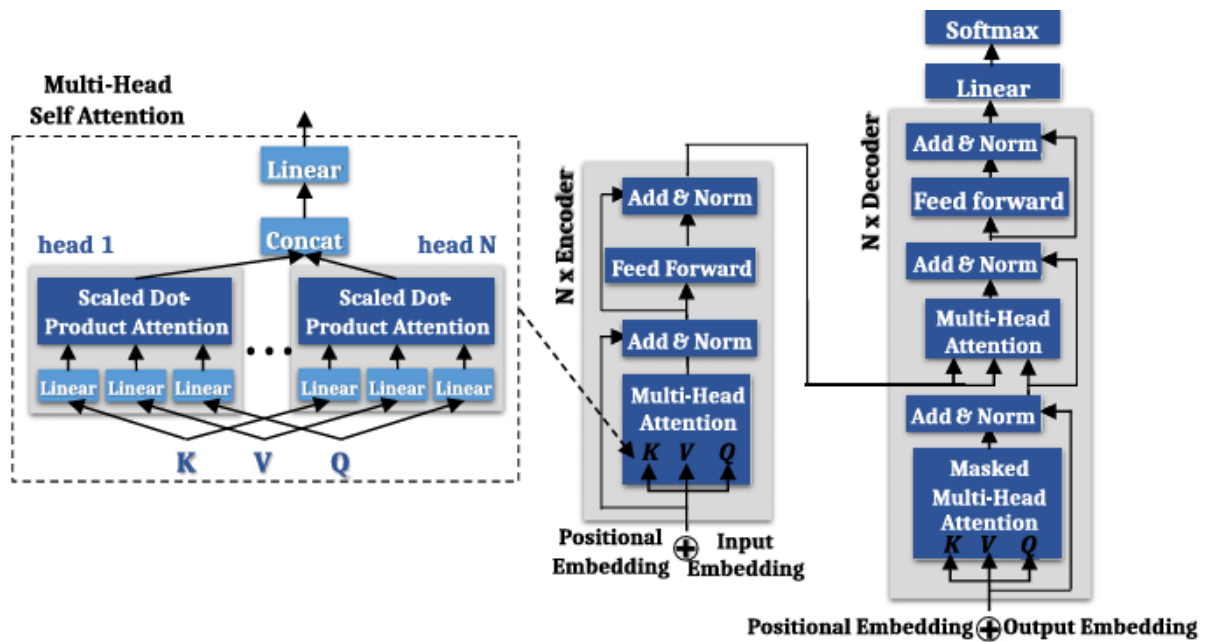
Basados en el mecanismo de multi-headed self-attention, estos modelos fueron diseñados originalmente para tareas de modelado del lenguaje y se han convertido en el estándar en numerosos contextos, incluyendo comprensión del lenguaje natural, traducción automática o reconocimiento de voz.

En lugar de leer el texto de manera secuencial, como hacen las redes neuronales recurrentes (RNN), los Transformers pueden mirar todo el texto de una vez, identificando las relaciones entre las palabras en diferentes partes del texto sin importar el orden en que aparecen.

Su funcionamiento se basa en una etapa de pre-entrenamiento no supervisado sobre grandes corpus generales, seguida de una fase de ajuste fino para tareas específicas. En el pre-entrenamiento se utilizan técnicas como el enmascaramiento de palabras, predicción de la siguiente oración o modelado de permutaciones. La arquitectura de un Transformer consiste en capas apiladas de codificadores y decodificadores, donde cada capa emplea mecanismos de atención escalada, conexiones residuales y normalización.



Una de las principales ventajas de los Transformers frente a modelos recurrentes es su capacidad para capturar dependencias a largo plazo dentro del texto, gracias a caminos más cortos y eficientes entre los tokens de entrada. Además, su complejidad computacional por capa es significativamente menor, lo que los hace especialmente adecuados para procesar grandes volúmenes de datos secuenciales con una mayor granularidad y precisión.

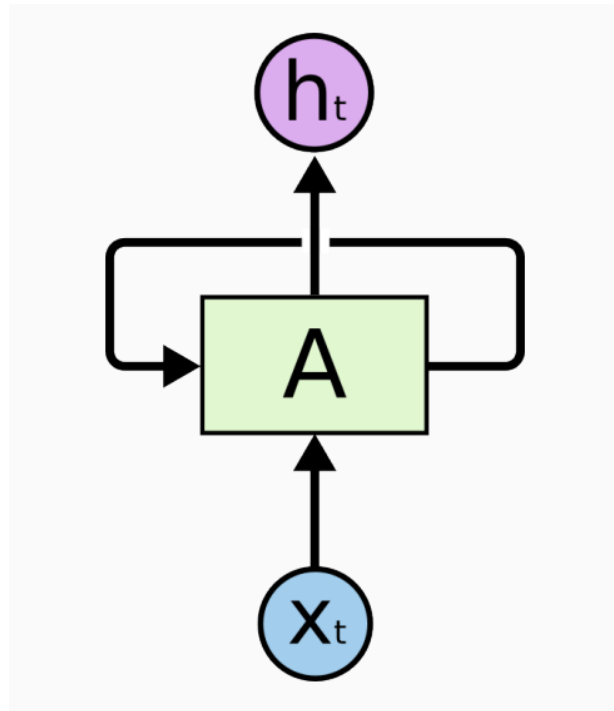


**Figura 3.** Transformer architecture with Multi-Headed Self-Attention

## 3.2 Redes Neuronales Recurrentes

Este tipo de redes se convirtieron en la opción más utilizada para la comprensión del lenguaje natural, traducción automática, reconocimiento de voz, reconocimiento de video y análisis de secuencias de ADN antes de la introducción de los Transformers [2].

Estas son una familia de redes neuronales enfocadas en aprender datos secuenciales, que dependen del tiempo y son persistentes. Se caracterizan por que sus capas ocultas tienen como entradas sus propias salidas.



**Figura 4.** Autoconexión en Red Neuronal [3]

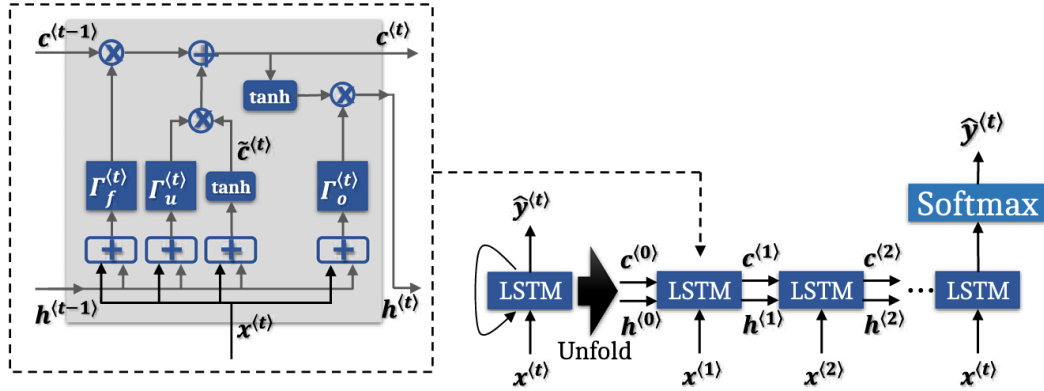
Estas autoconexiones son especialmente interesantes cuando se trabaja con lenguaje, ya que el significado de una oración no puede entenderse únicamente palabra por palabra, sino que requiere del análisis del contexto general.

Sin embargo, a pesar de sus ventajas, las RNN presentan algunas limitaciones técnicas importantes, como la dificultad para aprender dependencias a largo plazo debido al problema del desvanecimiento o explosión del gradiente durante el entrenamiento. Para resolver estos inconvenientes, surgieron variantes mejoradas como las LSTM (Long Short-Term Memory) y GRU (Gated Recurrent Units), que incorporan mecanismos de memoria más sofisticados para retener información relevante durante más tiempo.

### 3.2.1 Long Short-Term Memory (LSTM)

Las LSTM son una variante de las RNN que resuelven el problema de los gradientes que desaparecen o explotan, que ocurre cuando las RNN tienen dificultades para aprender de secuencias largas. Las LSTM incluyen una celda de memoria adicional y mecanismos de puertas para controlar qué información recordar y olvidar. Estas puertas incluyen:

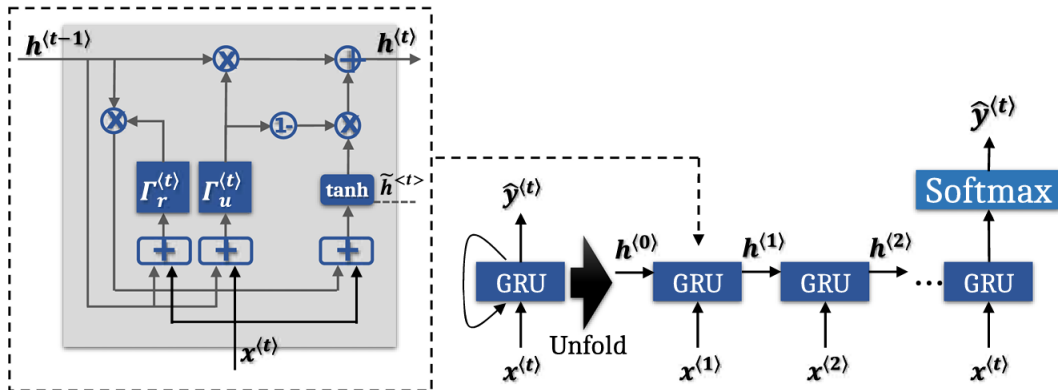
- Puerta de olvido: Decide qué información se descarta.
- Puerta de actualización: Decide cuánto de la nueva información se guarda.
- Puerta de salida: Determina qué parte de la memoria se usa para la salida final.



**Figura 5.** Long-Short-Term Memory cell architecture for text classification

### 3.2.2 Gated Recurrent Unit (GRU)

El GRU es una versión simplificada de las LSTM, usando solo dos puertas (de reinicio y actualización) y no mantiene una celda de memoria separada. La arquitectura de GRU es más simple, lo que puede ser útil en redes neuronales más profundas y reduce la necesidad de grandes cantidades de datos para generalizar.

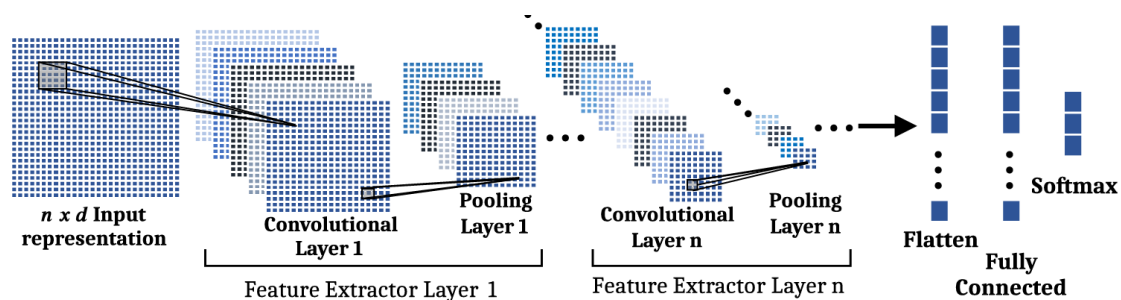


**Figura 6.** Gated Recurrent Unit architecture for text classification

### 3.2.3 Convolutional Neural Networks (CNN)

Estas redes aplican filtros convolucionales a las representaciones de palabras (embeddings) para extraer características importantes y luego clasifican el texto utilizando capas totalmente conectadas y una función de activación no lineal.

Concretamente cada entrada de texto se representa como una secuencia de palabras convertida en una matriz de embeddings de tamaño  $n \times d$ , donde  $n$  es el número de palabras y  $d$  es la dimensión del embedding. Esta matriz se pasa por una capa convolucional, donde un filtro de tamaño  $h \times d$  se desliza sobre la matriz. En cada posición, se realiza una multiplicación elemento a elemento, se suma un término de sesgo, y luego se aplica una función de activación no lineal. Las representaciones extraídas se aplanan y se conectan completamente antes de pasar a un clasificador Softmax para predecir el sentimiento.



**Figura 7.** Text classification flow with Convolutional Neural Network.

## Capítulo 4

# DESAFÍOS ABIERTOS DE INVESTIGACIÓN

---

La revisión de análisis de sentimiento fino, grueso, entre dominios y multilingüe muestra avances graduales, especialmente tras la adopción de modelos tipo Transformer. Sin embargo, persisten varios desafíos importantes:

1. Ambigüedad, dependencias a largo plazo, anáfora, sarcasmo y negación. La ambigüedad y el sarcasmo son especialmente problemáticos, generando muchos errores en tareas con clase neutral como verdad base.
  2. Falta de invariancia entre dominios en las arquitecturas evaluadas. Aunque los modelos preentrenados han mejorado el análisis entre dominios, su rendimiento sigue por debajo de los modelos entrenados y evaluados en el mismo dominio.
  3. Rendimiento multilingüe deficiente, especialmente en idiomas con pocos recursos. La mayoría de los enfoques se centran en unos pocos idiomas, y se necesitan modelos verdaderamente multilingües para una aplicación más global.
  4. El aprendizaje zero-shot y few-shot aún representa un gran desafío, tanto en contextos multilingües como multidominio, aunque tiene un alto potencial para producir modelos generalistas.
-

5. Desajuste entre lenguaje y dominio en análisis de sentimientos multilingüe. Este problema híbrido requiere cerrar la brecha entre idioma y contexto, combinando estrategias entre dominios y lenguas.
  6. Modelos preentrenados extremadamente grandes plantean barreras por su alto costo computacional, lo que limita su acceso a instituciones con menos recursos. Es crucial investigar arquitecturas más eficientes para democratizar el acceso a la inteligencia artificial avanzada.
-

## Capítulo 5

# CONCLUSIÓN

---

Este artículo presenta una revisión detallada y actualizada sobre las tareas de análisis de sentimientos. La revisión abarca tendencias recientes en arquitecturas de aprendizaje profundo, enfocándose en trabajos publicados entre el 1 de julio de 2017 y el 1 de julio de 2022.

Antes del auge de los modelos Transformer, las arquitecturas más comunes eran las CNN, RNN y los mecanismos de atención, utilizadas para extraer características semánticas y sintácticas a partir de word embeddings.

Sin embargo, las tendencias actuales muestran un alejamiento de estas arquitecturas tradicionales y una adopción generalizada de los modelos Transformer bidireccionales y contextuales. Estos modelos requieren menos ajuste posterior (fine-tuning) y son más eficaces para tareas de lenguaje natural gracias a su preentrenamiento con grandes corpus de datos generales.

---

# BIBLIOGRAFÍA

---

- [1] Tariq Abdullah and Ahmed Ahmet. 2022. Deep Learning in Sentiment Analysis: Recent Architectures. *ACM Comput. Surv.* 55, 8, Article 159 (December 2022), 37 pages. <https://dl.acm.org/doi/10.1145/3548772>
- [2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, ÅÅ. Kaiser, and I. Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*
- [3] Gutiérrez, P. A., & Sánchez Monedero, J. (2025, abril 8). Aprendizaje profundo: Introducción a las redes neuronales recurrentes. Departamento de Ciencia de la Computación e Inteligencia Artificial, Universidad de Córdoba. [https://moodle.uco.es/m2425/pluginfile.php/336828/mod\\_resource/content/1/AP\\_Tema3\\_Red esRecurrentes.pdf](https://moodle.uco.es/m2425/pluginfile.php/336828/mod_resource/content/1/AP_Tema3_Red esRecurrentes.pdf)