

Jornadas Técnicas Industrial Track 4.0

• DIHBU
Digital Innovation Hub Industry 4.0

BURGOS, 24 DE OCTUBRE 2023



Humanos en el LOOP del paradigma IA. Quién es Quién



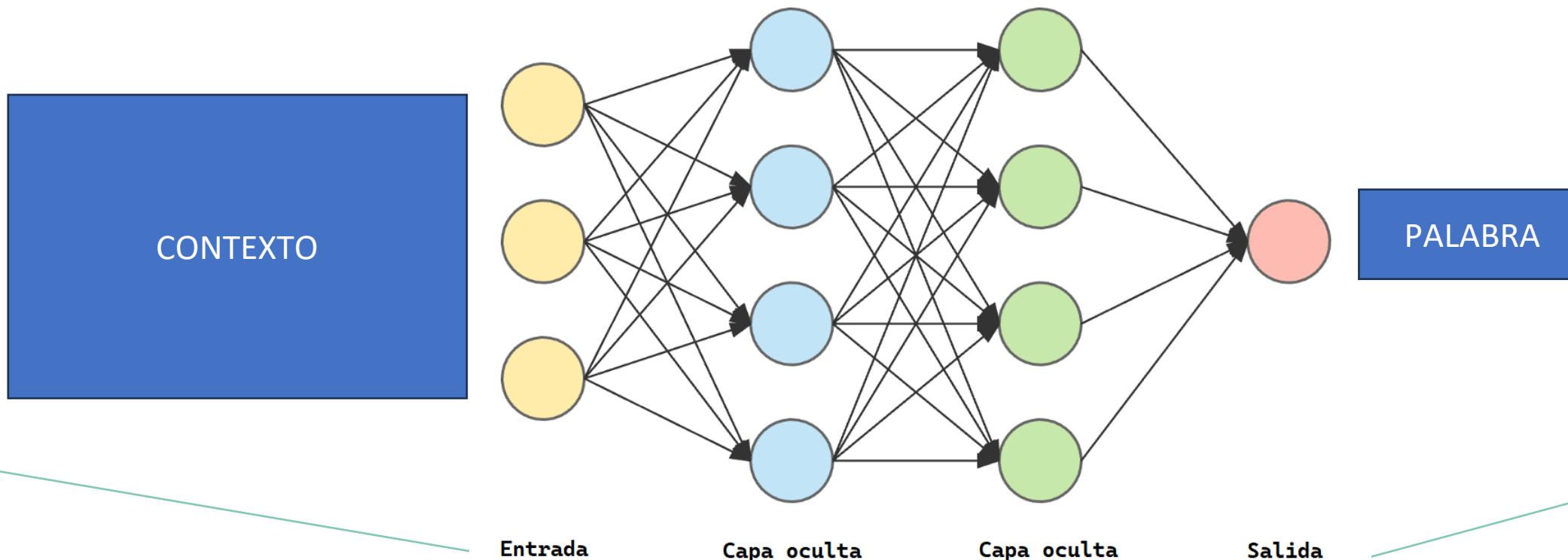
© Javier_Sedano 2023

Dr. Javier Sedano
Subdirector y Director I + D de ITCL



¿Qué son?

Redes neuronales profundas (Deep Learning) aplicadas a predicción de la siguiente palabra



¿Qué son? Ejemplo

Redes neuronales profundas (Deep Learning) aplicadas a predicción de la siguiente palabra

Un **modelo de lenguaje colosal**¹ o **LLM** (siglas en inglés para *Large Language Model*), también llamado **modelo de lenguaje de gran tamaño**, es un **modelo de lenguaje** que consta de una **red neuronal** con muchos parámetros (normalmente miles de millones o más), entrenados en grandes cantidades de texto sin

¿Qué son? Ejemplo

Redes neuronales profundas (Deep Learning) aplicadas a predicción de la siguiente palabra

Un **modelo de lenguaje colosal**¹ o LLM (siglas en inglés para *Large Language Model*), también llamado **modelo de lenguaje de gran tamaño**, es un **modelo de lenguaje** que consta de una **red neuronal** con muchos parámetros (normalmente miles de millones o más), entrenados en grandes cantidades de texto **sin etiquetar**

¿Qué son? Ejemplo

Redes neuronales profundas (Deep Learning) aplicadas a predicción de la siguiente palabra

Un **modelo de lenguaje colosal**¹ o LLM (siglas en inglés para *Large Language Model*), también llamado **modelo de lenguaje de gran tamaño**, es un **modelo de lenguaje** que consta de una **red neuronal** con muchos parámetros (normalmente miles de millones o más), entrenados en grandes cantidades de texto sin etiquetar mediante

¿Qué son? Ejemplo

Redes neuronales profundas (Deep Learning) aplicadas a predicción de la siguiente palabra

Un **modelo de lenguaje colosal**¹ o LLM (siglas en inglés para *Large Language Model*), también llamado **modelo de lenguaje de gran tamaño**, es un **modelo de lenguaje** que consta de una **red neuronal** con muchos parámetros (normalmente miles de millones o más), entrenados en grandes cantidades de texto sin etiquetar mediante **aprendizaje**

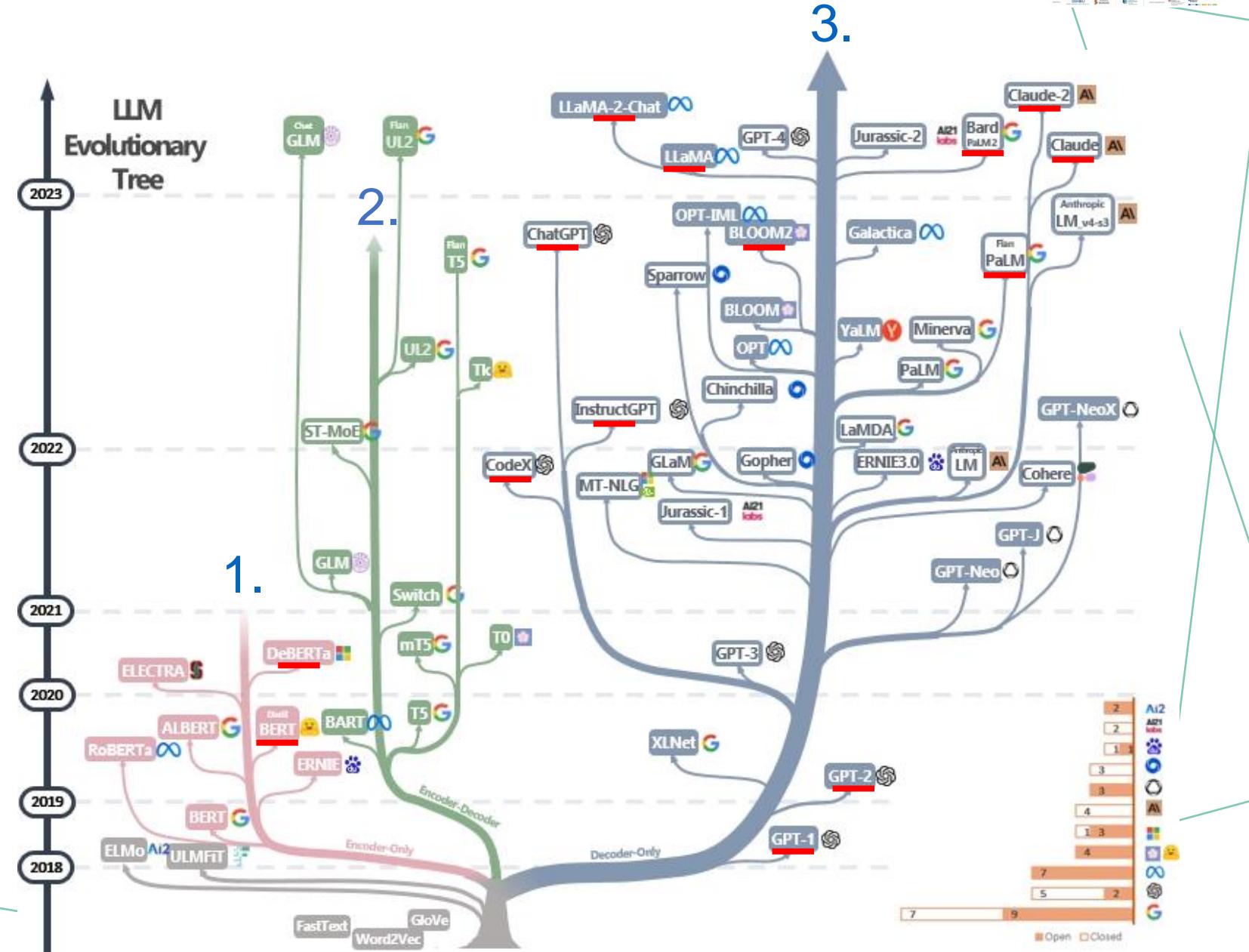
¿Qué son? Ejemplo

Redes neuronales profundas (Deep Learning) aplicadas a predicción de la siguiente palabra

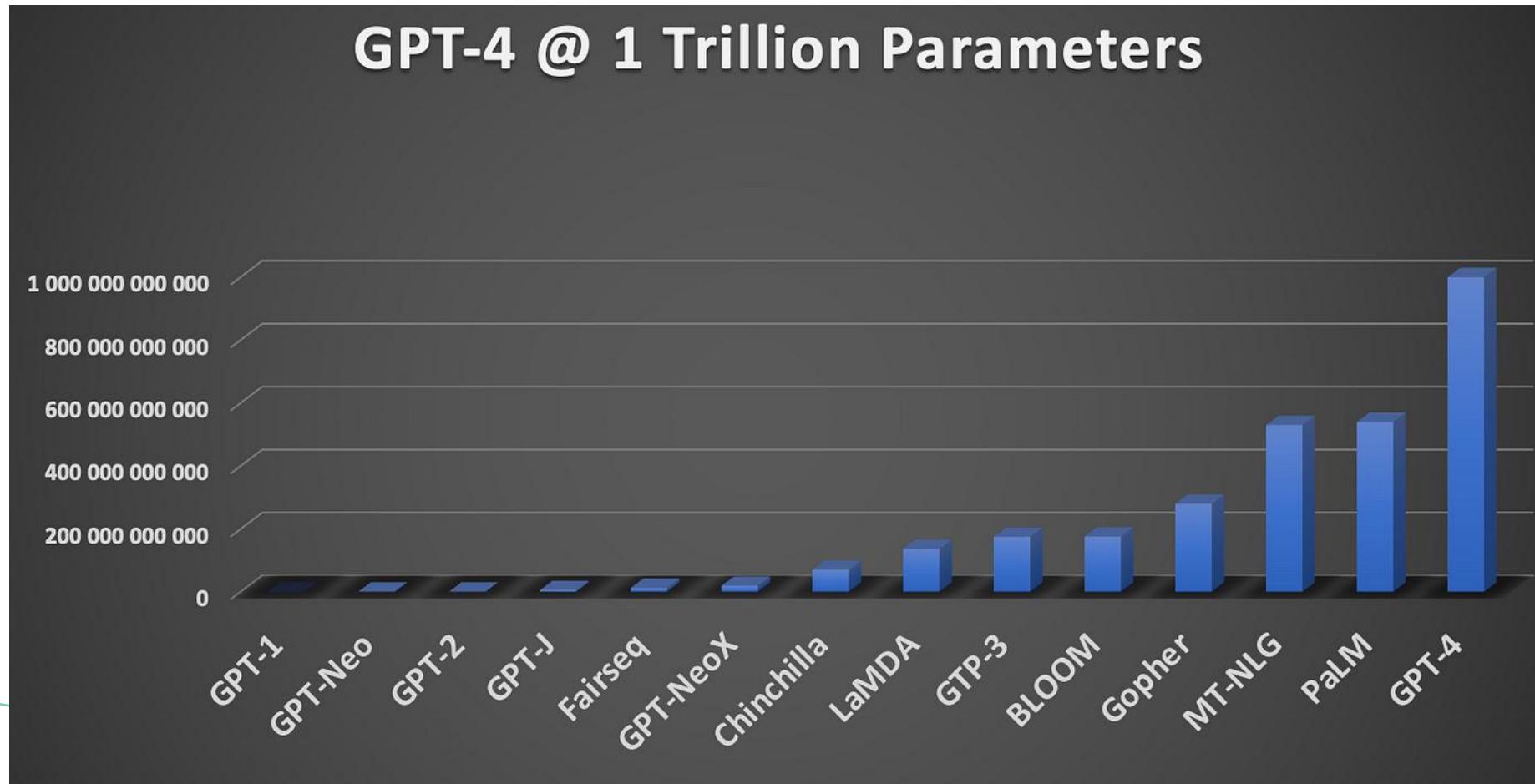
Un **modelo de lenguaje colosal**¹ o LLM (siglas en inglés para *Large Language Model*), también llamado **modelo de lenguaje de gran tamaño**, es un **modelo de lenguaje** que consta de una **red neuronal** con muchos parámetros (normalmente miles de millones o más), entrenados en grandes cantidades de texto sin etiquetar mediante **aprendizaje autosupervisado**

Historia

1. Encoder only
2. Encoder-decoder
3. Decoder only

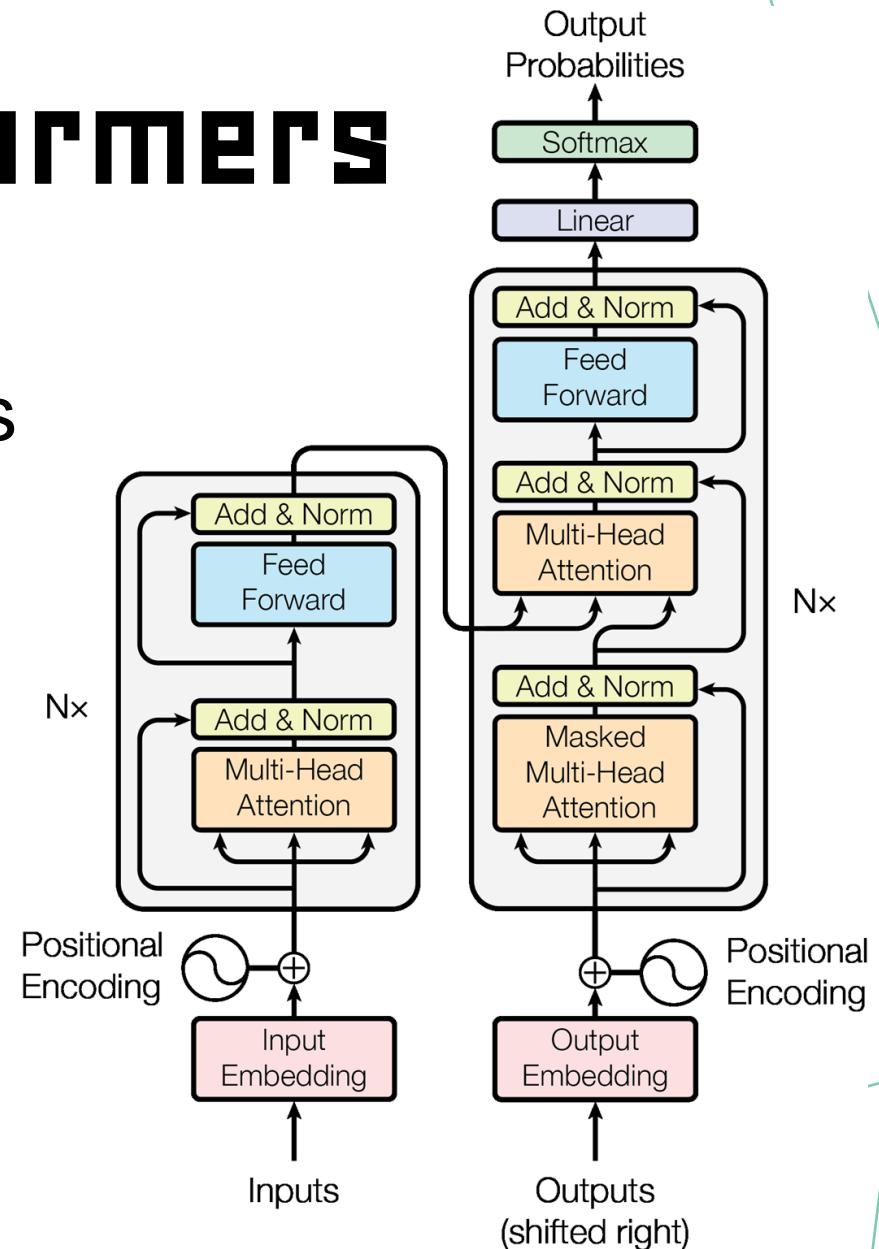


Historia



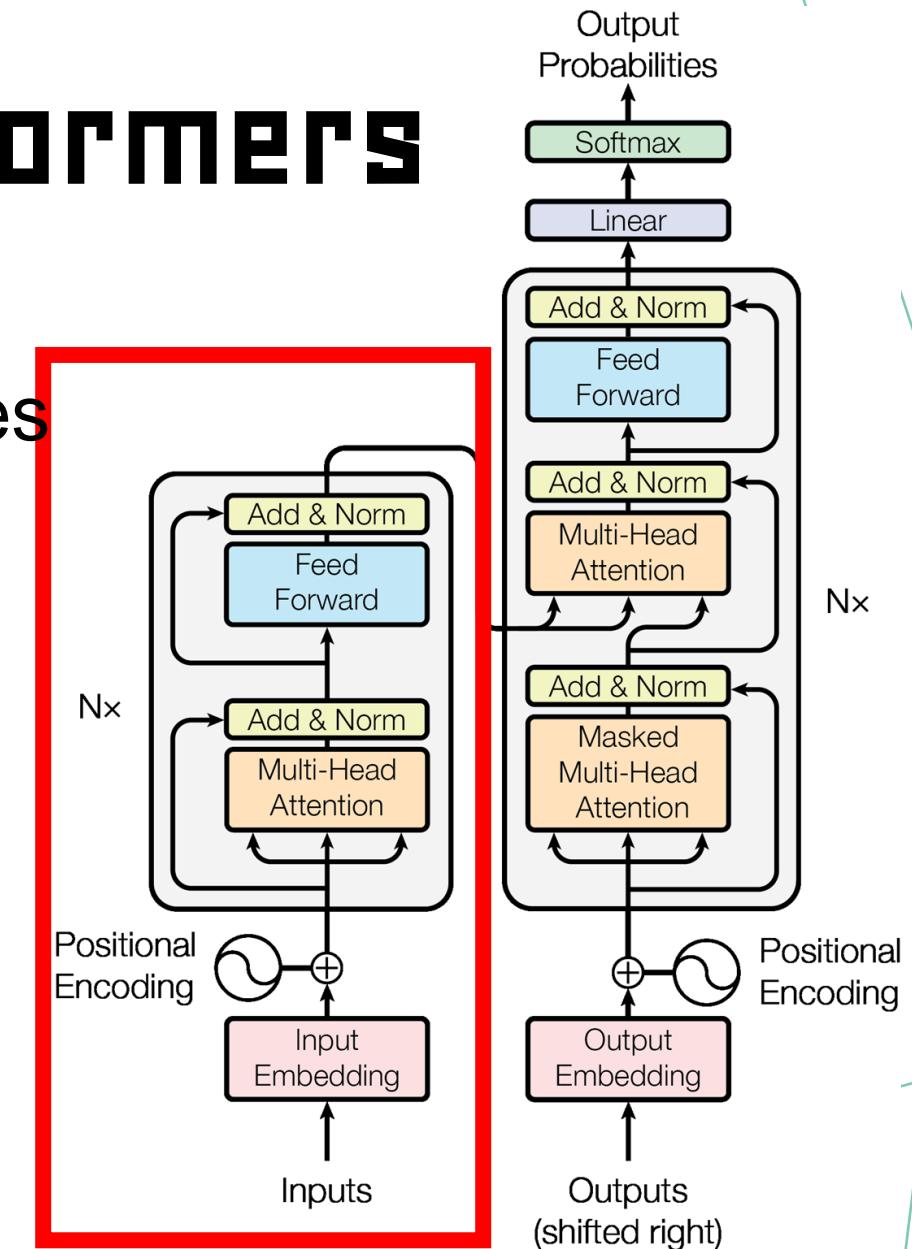
Tecnología: Transformers

- Arquitectura auto-regresiva
⇒ Evolución de las recurrentes como las LSTM



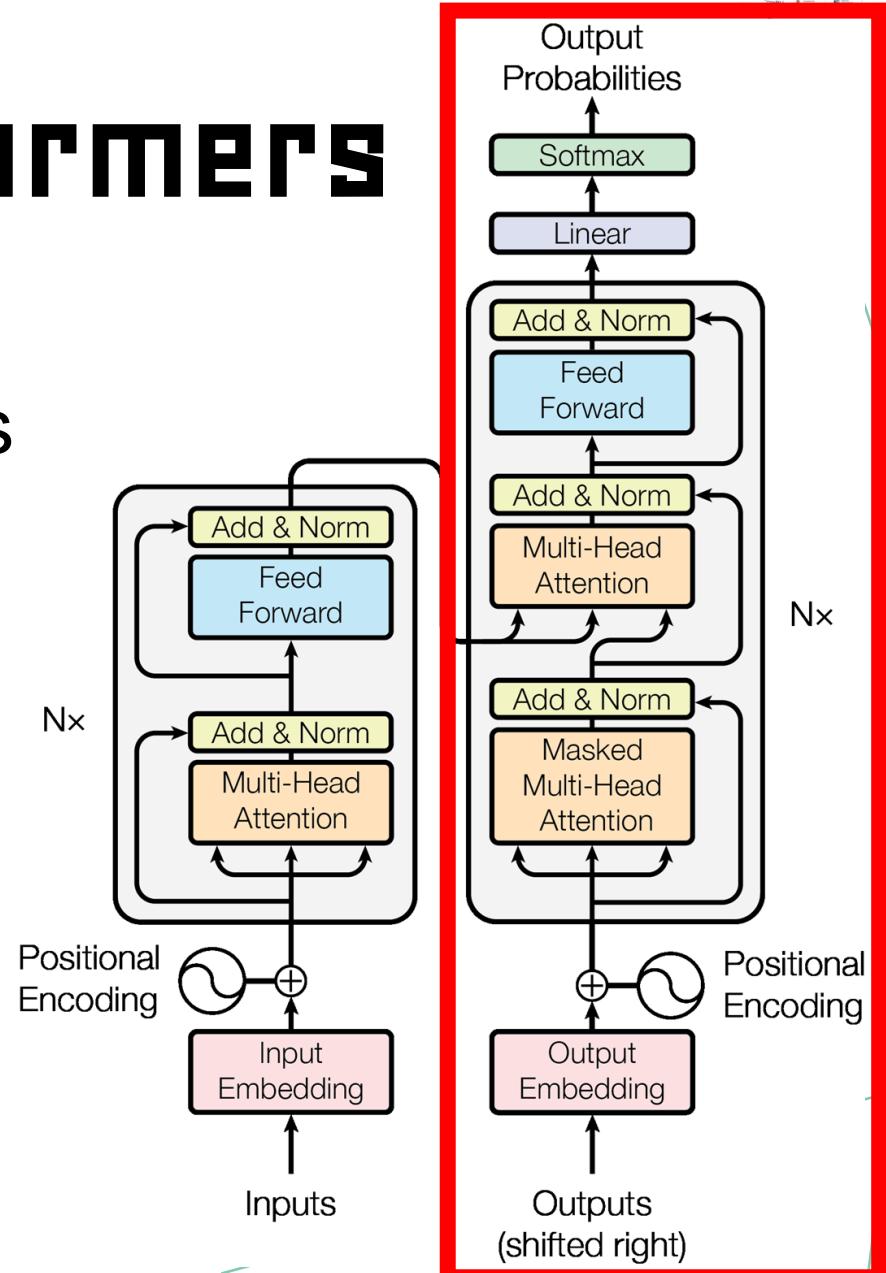
Tecnología: Transformers

- Arquitectura auto-regresiva
⇒ Evolución de las recurrentes como las LSTM
- Encoder



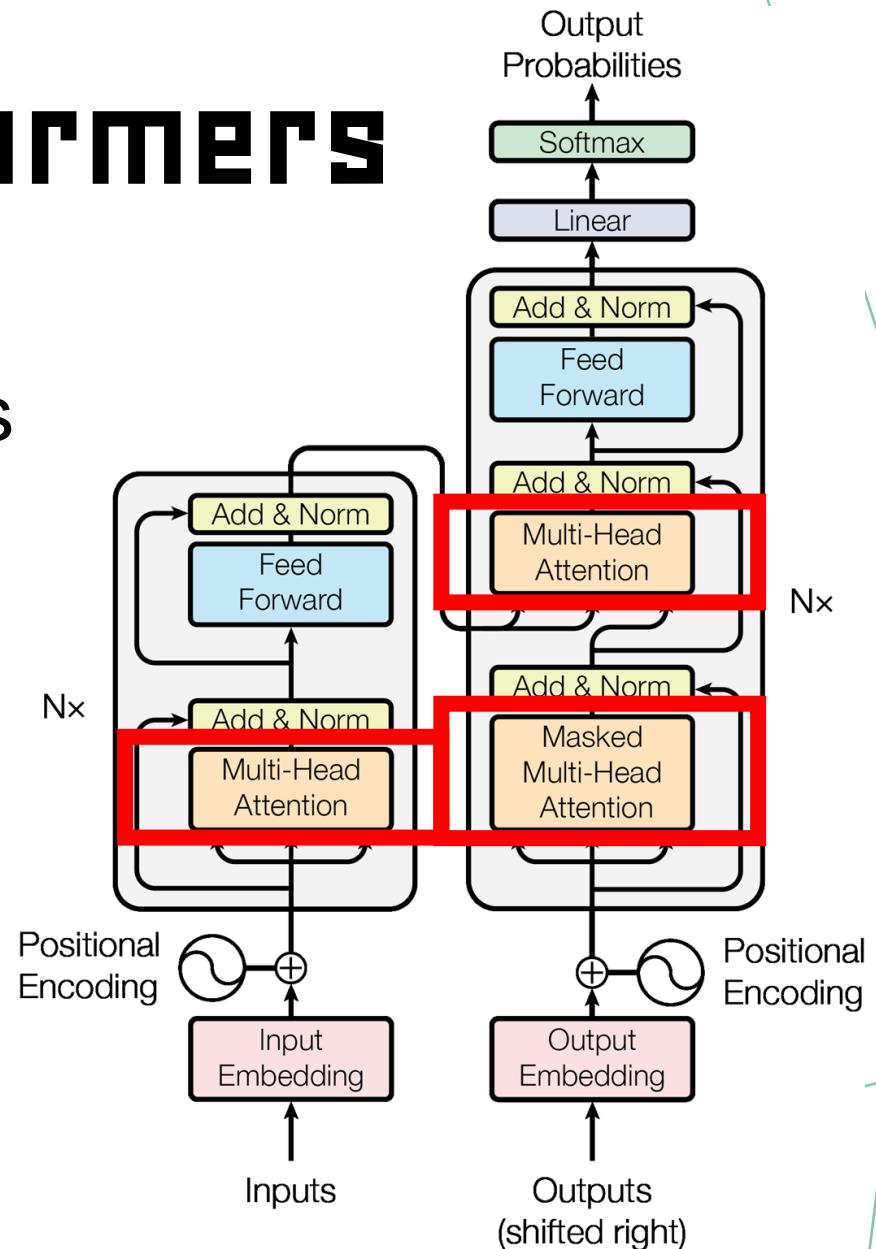
Tecnología: Transformers

- Arquitectura auto-regresiva
⇒ Evolución de las recurrentes como las LSTM
- Encoder
- Decoder



Tecnología: Transformers

- Arquitectura auto-regresiva
⇒ Evolución de las recurrentes como las LSTM
- Encoder
- Decoder
- Capa de atención



Modelos Estado del Arte



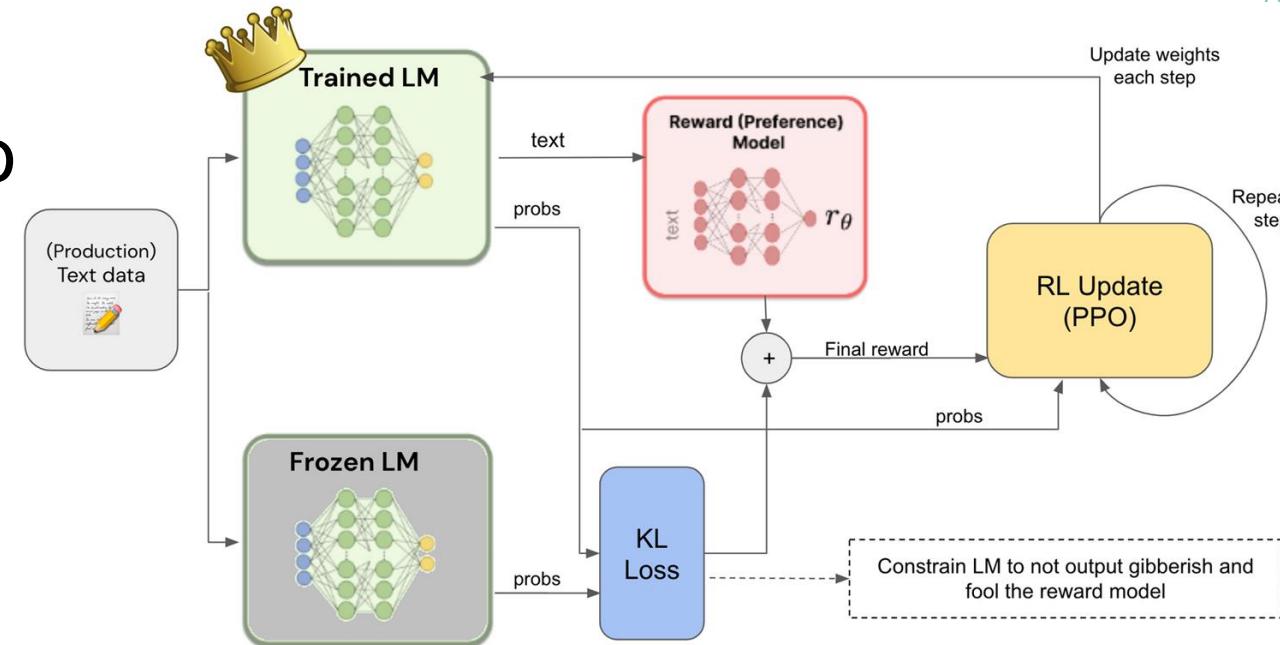
PaLM 2



¿y Chat GPT?

Reinforcement Learning from Human Feedback [RLHF]

- Problema \Rightarrow Del LLM puede salir cualquier cosa, ¿cómo lo controlamos?
- Se le enseña mediante Aprendizaje por refuerzo



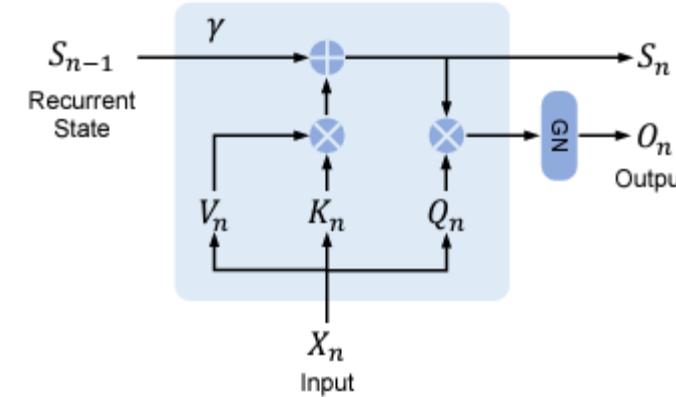
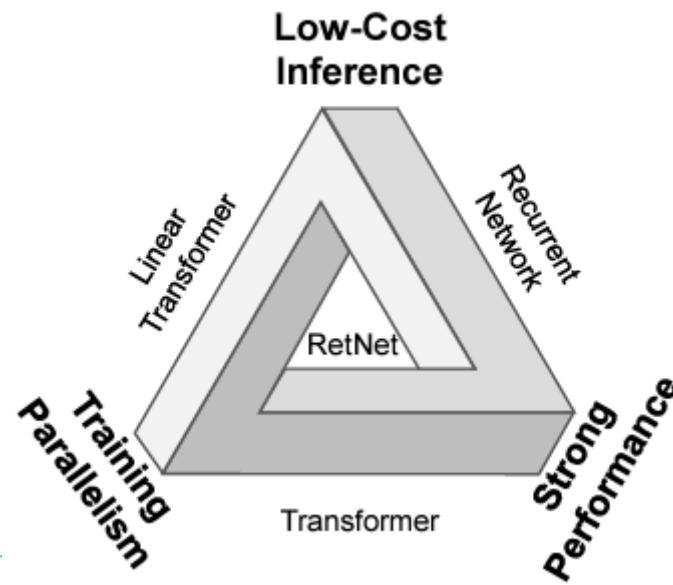
Reinforcement Learning from Human Feedback [RLHF]

- Problema \Rightarrow Del LLM puede salir cualquier cosa, ¿cómo lo controlamos?
- Se le enseña mediante Aprendizaje por refuerzo
- Se genera un modelo alineado



Competidores

- Transformers desde 2017
- Retnet a partir de 2023: <https://arxiv.org/abs/2307.08621>



Gracias

Dr. Javier Sedano
javier.sedano@itcl.es

