

# Aprendizaje profundo

PREENTRENAMIENTO AUTOSUPERVISADO

---

Gibran Fuentes-Pineda

Octubre 2023

# Modelos de lenguaje y aprendizaje por transferencia: definición

- Un modelo de lenguaje es una distribución de probabilidad de una secuencia de  $T$  palabras

$$P(x^{[1]}, \dots, x^{[T]}) = P(x^{[1]}) \cdot \prod_{i=2}^T P(x^{[i]} | x^{[1]}, \dots, x^{[i-1]})$$

- Redes autoregresivas tipo Transformer se han usado como modelos del lenguaje
  - Usualmente entrenadas con tareas de aprendizaje no supervisado
  - Muy efectivas para el aprendizaje por transferencia

# Generative Pre-training Transformer (GPT)

- Decodificador con bloques Transformer (descarta codificador)

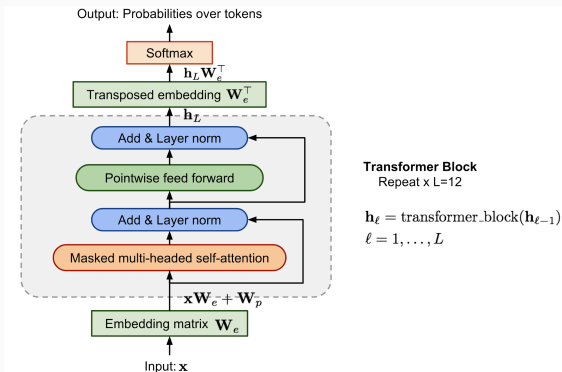
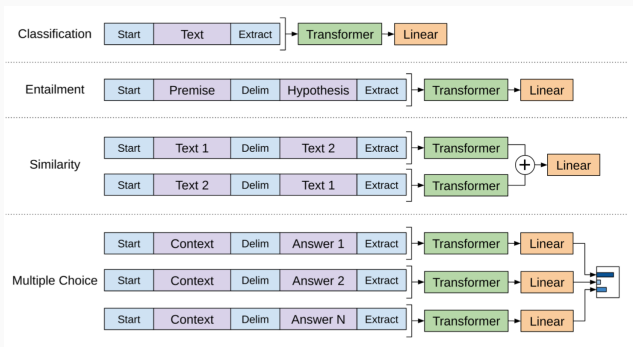


Imagen tomada de Lilian Weng. Generalized language models, 2019.

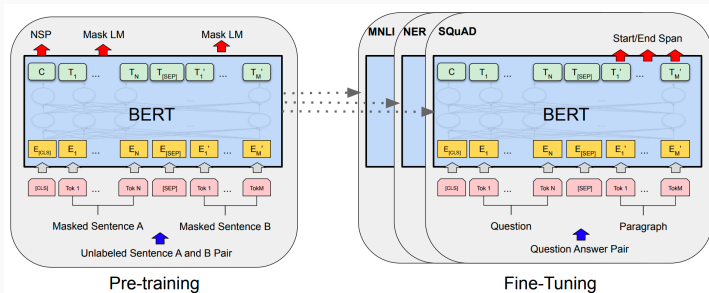
# Preentrenamiento de GPT

- Preentrenamiento autosupervisado generativo: predecir la siguiente palabra / *token*
- Aprendizaje por transferencia para distintas tareas de procesamiento del lenguaje natural



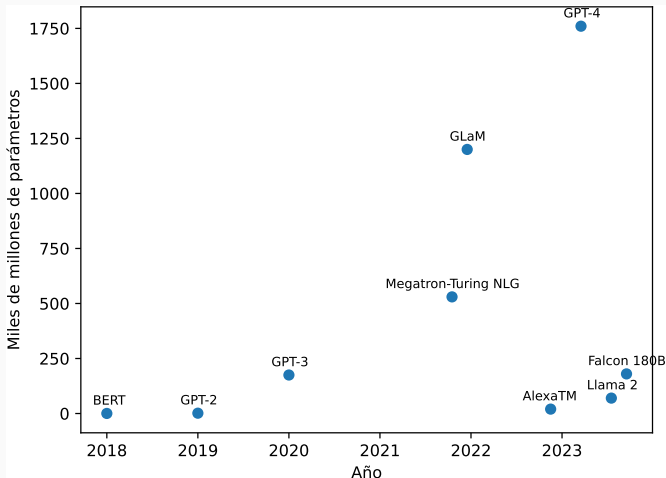
# Bidirectional Encoder Representations from Transformers (BERT)

- Codificador con bloques Transformer (sin decodificador)
- Preentrenamiento autosupervisado discriminativo
  1. Modelo de lenguaje enmascarado
  2. Predicción de la siguiente oración (NSP)
- Aprendizaje por transferencia para distintas tareas de procesamiento del lenguaje natural



# Modelos de lenguaje modernos

- Basados en arquitecturas profundas con bloques tipo Transformer



# Leyes de escalamiento

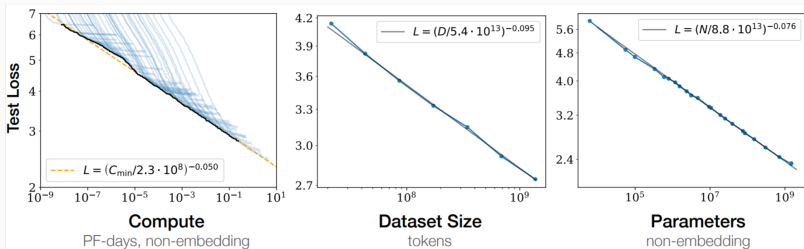


Figura tomada de Kaplan et al. Scaling Laws for Neural Language Models, 2020.