

T4 Introducción a los LLMs

Índice

1. Introducción
2. Antecedentes
3. GPT, BERT, T5 y ChatGPT
4. Pre-entrenamiento
5. Post-entrenamiento
6. Evaluación

1 Introducción

Large language models (LLMs): redes grandes para la generación de secuencias de tokens, principalmente basadas en la arquitectura Transformer (decoder-only)

Foundation model (FM): modelo pre-entrenado con datos masivos para poder aplicarlo en muchos casos de uso

- **FMs típicos:** LLMs que responden a prompts en lenguaje natural generando texto, imágenes, audio, vídeo, código, etc.
- **FMs avanzados:** solo grandes labs como Google, Meta, Alibaba y OpenAI disponen de recursos para entrenarlos
- **FMs frontera:** FMs avanzados con capacidades peligrosas para la seguridad pública

Andrej Karpathy: autoridad en IA que produce recursos educativos muy interesantes

- [Deep Dive into LLMs like ChatGPT \(3h 31min\):](#) fundamentos internos de los LLMs
- [How I use LLMs \(2h 11min\):](#) guía práctica con ejemplos de uso
- [The busy person's intro to LLMs \(1h\):](#) vídeo paralelo opcional, más antiguo
- [Zero to Hero:](#) playlist de su canal más técnico
- [Eureka Labs:](#) empresa de IA educativa cuyo primer curso promete ser el mejor curso de IA
- [Software in the era of AI](#) código (1.0) → pesos red (2.0) → prompts LLM (3.0)

Hugging Face (HF): imprescindible para estar al día

- [Curso HF sobre LLMs](#) LLMs usando librerías HF
- [Catálogo de cursos HF](#) MCP, post-training, agents, RL, CV, audio, etc.
- [Curso de Maxime Labonne:](#) edición actualizada (enero de 2025)

Objetivo: introducir antecedentes, LLMs pioneros, pre-entrenamiento, post-entrenamiento y evaluación

2 Antecedentes

Modelo de lenguaje (probabilístico): distribución de probabilidad sobre las cadenas (de palabras) de un lenguaje

Claude Shannon (1951): entropía del inglés tras preguntar por la letra siguiente a $n - 1$ previas, F_n

$$F_0 := - \sum_{i=1}^{26} \frac{1}{26} \log_2 \frac{1}{26} = \log_2 26 = 4.7 \text{ bits} \quad F_1 = 4.14 \quad F_2 = 3.56 \quad F_3 = 3.3$$

1980s y 1990s: modelos estadísticos, n -gramas

Word n -gram LMs: aplican la regla de la cadena y asumen que la probabilidad de una palabra solo depende de las $n - 1$ palabras previas (asunción Markoviana de orden $n - 1$)

$$p(w_{1:N}) = \prod_i p(w_i \mid w_{1:i-1}) \approx \prod_i p(w_i \mid w_{i-(n-1):i-1}), \quad w_i \in \mathcal{V} = \{1, \dots, V\} \text{ (vocabulario)}$$

Unigram LM: $n = 1$

$$p(w_{1:N}) \approx p(w_1) p(w_2) p(w_3) \cdots p(w_N)$$

Bigram LM: $n = 2$

$$p(w_{1:N}) \approx p(w_1) p(w_2 \mid w_1) p(w_3 \mid w_2) \cdots p(w_N \mid w_{N-1})$$

Trigram LM: $n = 3$

$$p(w_{1:N}) \approx p(w_1) p(w_2 \mid w_1) p(w_3 \mid w_1, w_2) \cdots p(w_N \mid w_{N-2}, w_{N-1})$$

Elementos del vocabulario: también pueden ser símbolos, tokens, etc.

Símbolo inicio: asumimos $n - 1$ $\langle s \rangle$ iniciales para modelizar probabilidades de las $n - 1$ primeras palabras de frase

Símbolo fin: asumimos un $\langle /s \rangle$ final; alternatively, sin $\langle /s \rangle$ final, tenemos un LM por cada longitud de frase

Ejemplo: probabilidad de **aab** según el LM bigrama

$$p(\mathbf{a} \mid \langle s \rangle) = 1 \quad p(\mathbf{a} \mid \mathbf{a}) = 0.9, p(\mathbf{b} \mid \mathbf{a}) = 0.1 \quad p(\mathbf{b} \mid \mathbf{b}) = 0.9, p(\langle /s \rangle \mid \mathbf{b}) = 0.1$$

$$p(\mathbf{aab}) = p(\mathbf{a} \mid \langle s \rangle) p(\mathbf{a} \mid \mathbf{a}) p(\mathbf{b} \mid \mathbf{a}) p(\langle /s \rangle \mid \mathbf{b}) = 1 \cdot 0.9 \cdot 0.1 \cdot 0.1 = 0.009$$

Perplejidad (a nivel de palabra): inversa de la media geométrica de la probabilidad por palabra

$$\text{Perplexity}(w_{1:N}) = p(w_{1:N})^{-\frac{1}{N}} = \left(\prod_i p(w_i \mid w_{i-(n-1):i-1}) \right)^{-\frac{1}{N}} \geq 1$$

Interpretación: factor de ramificación promedio, esto es, promedio de posibles palabras que pueden seguir a otra

Entropía (a nivel de palabra): media aritmética de la neg-log-probabilidad por palabra; igual a la log-perplejidad

$$\mathbb{H}(w_{1:N}) = -\frac{1}{N} \sum_i \log p(w_i \mid w_{i-(n-1):i-1}) \rightarrow \mathbb{H}(w_{1:N}) = \log \text{Perplexity}(w_{1:N})$$

Ejemplo (cont.): perplejidad y entropía del LM bigrama dado a partir de **aab**

$$\text{Perplexity}(\mathbf{aab}) = p(\mathbf{aab})^{-1/4} = 0.009^{-1/4} = 3.25$$

$$\mathbb{H}(\mathbf{aab}) = 1.70 \text{ bits}$$

Perplejidad y entropía de un modelo mínimamente perplejo:

- Solo una única palabra puede seguir a $n - 1$ previas (con probabilidad 1)
- El modelo concentra toda la probabilidad en una sola frase; sea $w_{1:N}$ dicha frase
- La perplejidad del modelo es 1 y su entropía 0; $\text{Perplexity}(w_{1:N}) = (1^N)^{-1/N} = 1$
- El modelo no necesita información extra sobre la siguiente palabra para determinarla

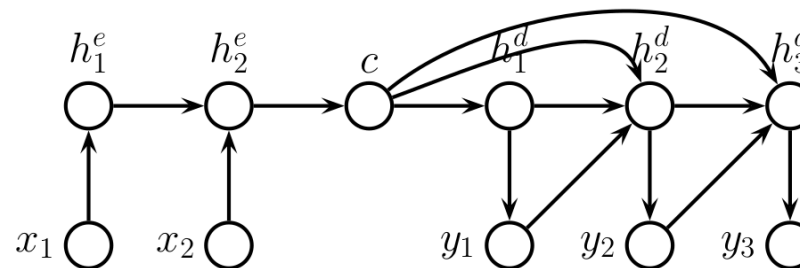
Perplejidad y entropía de un modelo máximamente perplejo:

- Cualquier palabra (salvo $\langle s \rangle$) puede seguir a $n - 1$ previas con probabilidad $1/V$ (V no incluye $\langle s \rangle$)
- La probabilidad asignada a una frase $w_{1:N}$ arbitraria, $N \geq 0$, es $1/(V - 1)^N \cdot 1/V \cong V^{-N}$
- La perplejidad del modelo es V (aprox.) y su entropía $\log V$; $\text{Perplexity}(w_{1:N}) \cong (V^{-N})^{-\frac{1}{N}} = V$
- El modelo necesita $\log V$ nats (bits con \log_2) de información extra sobre la siguiente palabra para determinarla

Limitaciones de los modelos estadísticos:

- **Curse of dimensionality:** demasiados parámetros (p.ej. 10^{50} con $V = 100K$ y $n = 10$), por lo que resulta difícil manejar contextos más allá de las dos palabras previas (trigramas)
- No tienen en cuenta la similitud (semántica) entre palabras; la palabra k -ésima del vocabulario puede verse como un vector disperso $\text{onehot}(k) \in \{0, 1\}^V$ a la misma distancia (semántica) de cualquier otra

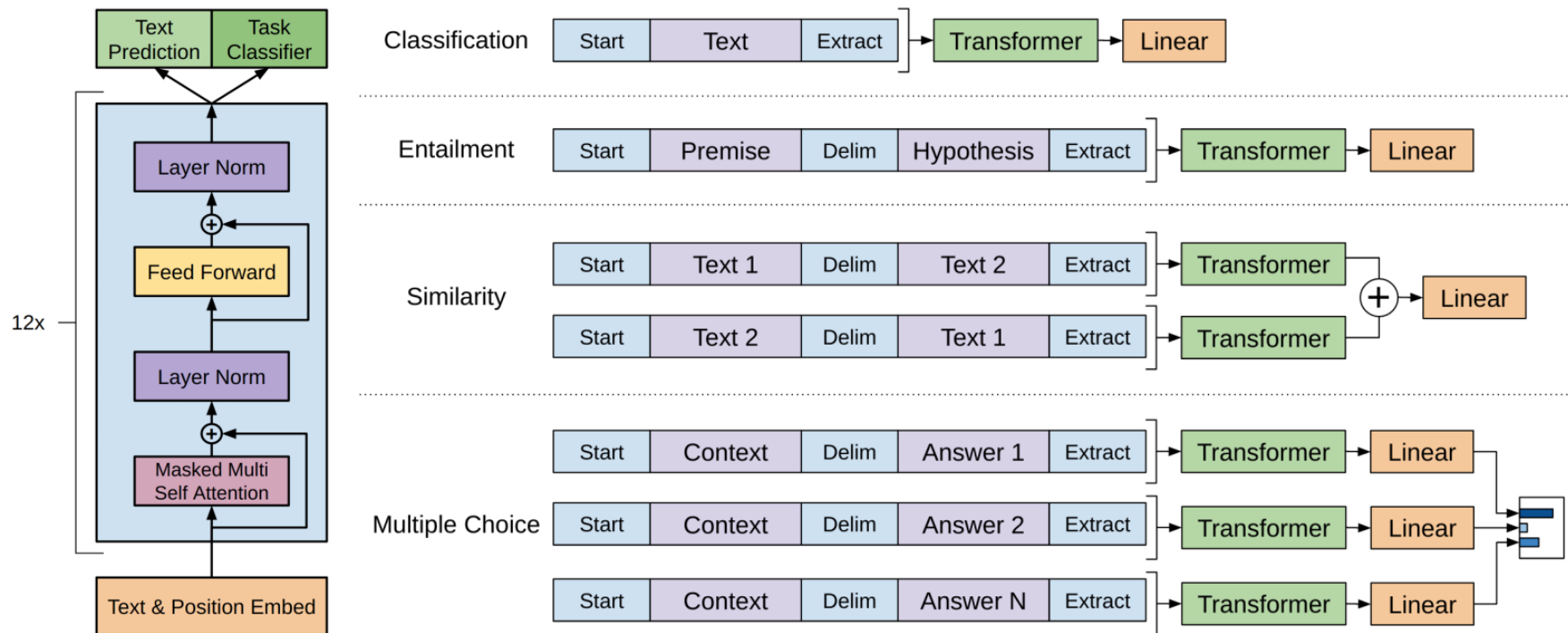
2000s y 2010s: redes, sobre todo recurrentes; p. ej. arquitectura encoder-decoder precursora de Transformer



3 GPT, BERT, T5 y ChatGPT

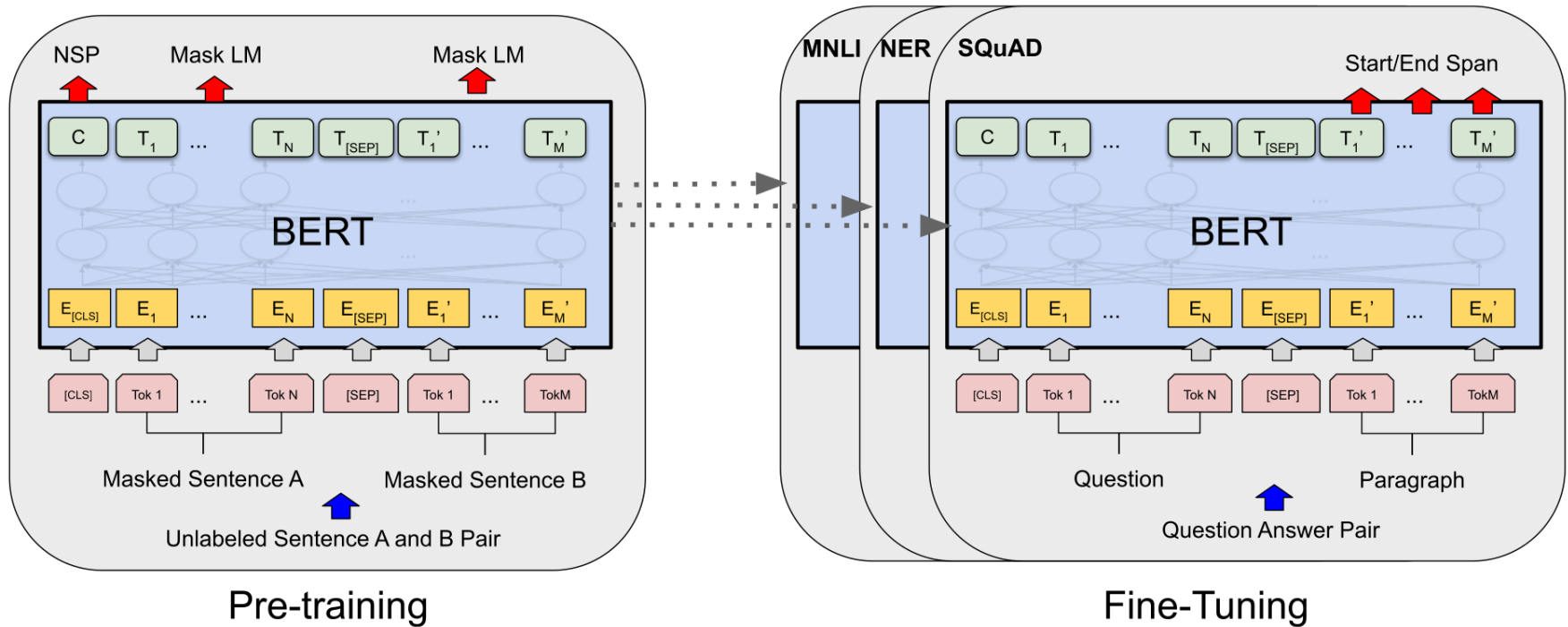
Generative Pre-Training (GPT-1, OpenAI, 2018): decoder-only transformer

- **Pre-entrenamiento:** con secuencias de texto largas para poder aprender de información lejana
- **Tareas de fine-tuning:** clasificación, entailment, similarity y multiple choice



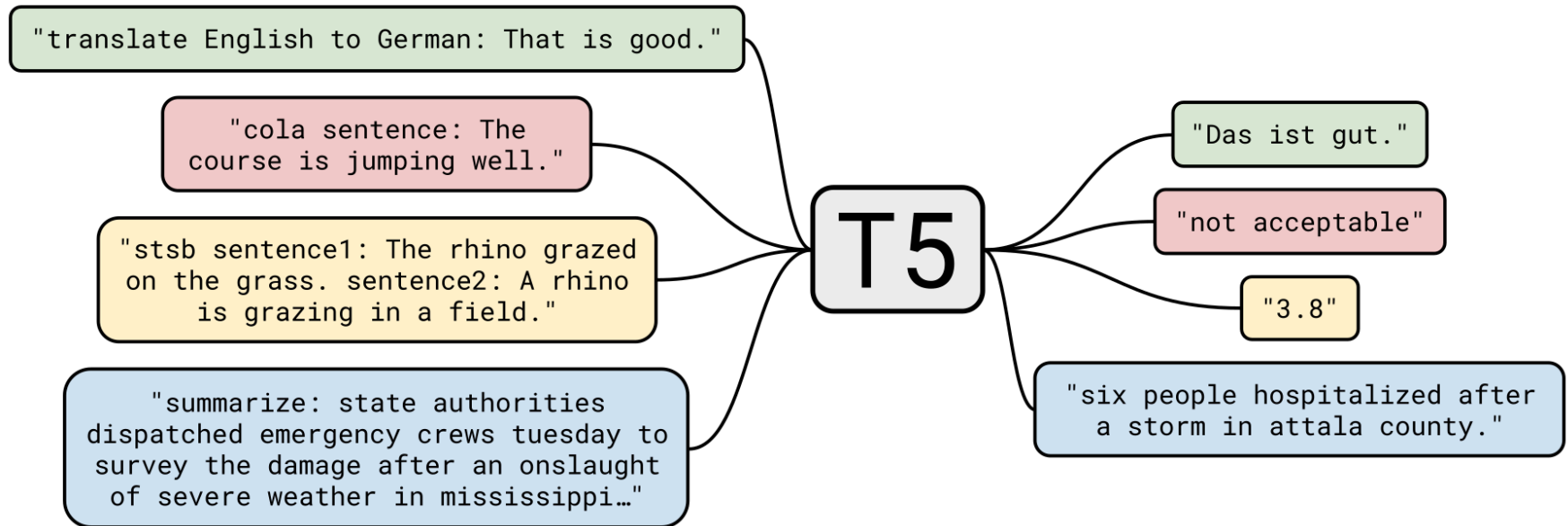
Bidirectional encoder representations from transformers (BERT, Google, 2018): encoder-only transformer

- **Tareas de pre-entrenamiento:** masked LM (MLM) y next sentence prediction (NSP)
- **Tareas de fine-tuning:** multi-genre natural language inference (MNLI), named entity recognition (NER), Stanford question answering dataset (SQuAD)

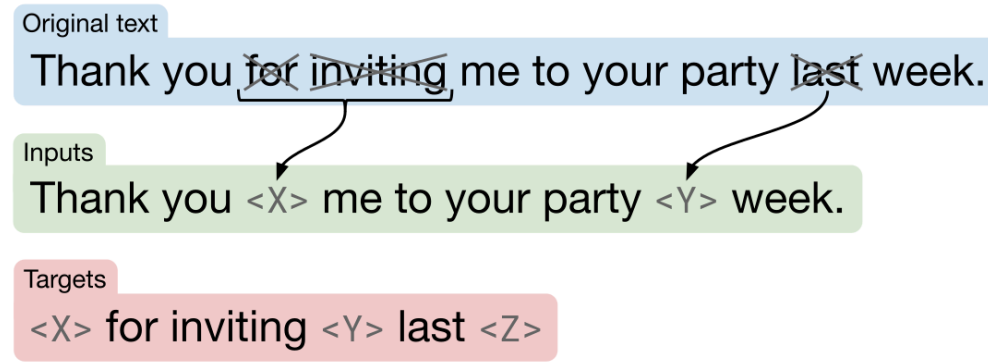


Text-to-text transfer transformer (T5, Google, 2019): encoder-decoder transformer

- **Tareas:** MT, corpus of linguistic acceptability (CoLA), semantic textual similarity benchmark (STSB), resúmenes, etc.



- **Entrenamiento:** pre-entrenamiento con masked LM seguido de fine-tuning text-to-text según tarea específica



GPT-2 (OpenAI, 2019): decoder-only transformer con **zero-shot task transfer**

- **WebText:** dataset de entrenamiento con 40GB de texto de 45M de enlaces Reddit hasta 2017
- **Resultados zero-shot en múltiples tareas:** sin entrenamiento o fine-tuning en ninguna de ellas

| | LAMBADA (PPL) | LAMBADA (ACC) | CBT-CN (ACC) | CBT-NE (ACC) | WikiText2 (PPL) | PTB (PPL) | enwik8 (BPB) | text8 (BPC) | WikiText103 (PPL) | 1BW (PPL) |
|-------|------------------|------------------|-----------------|-----------------|--------------------|--------------|-----------------|----------------|----------------------|--------------|
| SOTA | 99.8 | 59.23 | 85.7 | 82.3 | 39.14 | 46.54 | 0.99 | 1.08 | 18.3 | 21.8 |
| 117M | 35.13 | 45.99 | 87.65 | 83.4 | 29.41 | 65.85 | 1.16 | 1.17 | 37.50 | 75.20 |
| 345M | 15.60 | 55.48 | 92.35 | 87.1 | 22.76 | 47.33 | 1.01 | 1.06 | 26.37 | 55.72 |
| 762M | 10.87 | 60.12 | 93.45 | 88.0 | 19.93 | 40.31 | 0.97 | 1.02 | 22.05 | 44.575 |
| 1542M | 8.63 | 63.24 | 93.30 | 89.05 | 18.34 | 35.76 | 0.93 | 0.98 | 17.48 | 42.16 |

- **Language Modeling Broadened to Account for Discourse Aspect (LAMBADA):** [The LAMBADA dataset ...](#)
 - Tarea: predecir la palabra final de una frase que requiere al menos 50 tokens de contexto para predecirla

Context: Both its sun-speckled shade and the cool grass beneath were a welcome respite after the stifling kitchen, and I was glad to relax against the tree's rough, brittle bark and begin my breakfast of buttery, toasted bread and fresh fruit. Even the water was tasty, it was so clean and cold.

Target sentence: It almost made up for the lack of _____

Target word: coffee

- **Children's Book Test (CBT):** [The Goldilocks Principle: Reading Children's Books with Explicit Memory Repres.](#)
 - Ejemplo: izq.) extracto en azul; dcha.) contexto S , pregunta q , respuestas candidatas C y correcta a
 - Evaluación: de cada candidata para escoger la más probable
 - CBT-CN (ACC) y CBT-NE (ACC): precisión en la predicción de nombre comunes y entidades con nombre

"Well, Miss Maxwell, I think it only fair to tell you that you may have trouble with those boys when they do come. Forewarned is forearmed, you know. Mr. Cropper was opposed to our hiring you. Not, of course, that he had any personal objection to you, but he is set against female teachers, and when a Cropper is set there is nothing on earth can change him. He says female teachers can't keep order. He 's started in with a spite at you on general principles, and the boys know it. They know he'll back them up in secret, no matter what they do, just to prove his opinions. Cropper is sly and slippery, and it is hard to corner him."

"Are the boys big ?" queried Esther anxiously.

"Yes. Thirteen and fourteen and big for their age. You can't whip 'em -- that is the trouble. A man might, but they'd twist you around their fingers. You'll have your hands full, I'm afraid. But maybe they'll behave all right after all."

Mr. Baxter privately had no hope that they would, but Esther hoped for the best. She could not believe that Mr. Cropper would carry his prejudices into a personal application. This conviction was strengthened when he overtook her walking from school the next day and drove her home. He was a big, handsome man with a very suave, polite manner. He asked interestedly about her school and her work, hoped she was getting on well, and said he had two young rascals of his own to send soon. Esther felt relieved. She thought that Mr. Baxter had exaggerated matters a little.

S : 1 Mr. Cropper was opposed to our hiring you .
 2 Not , of course , that he had any personal objection to you , but he is set against female teachers , and when a Cropper is set there is nothing on earth can change him .
 3 He says female teachers ca n't keep order .
 4 He 's started in with a spite at you on general principles , and the boys know it .
 5 They know he 'll back them up in secret , no matter what they do , just to prove his opinions .
 6 Cropper is sly and slippery , and it is hard to corner him . ''
 7 `` Are the boys big ? ''
 8 queried Esther anxiously .
 9 `` Yes .
 10 Thirteen and fourteen and big for their age .
 11 You ca n't whip 'em -- that is the trouble .
 12 A man might , but they 'd twist you around their fingers .
 13 You 'll have your hands full , I 'm afraid .
 14 But maybe they 'll behave all right after all . ''
 15 Mr. Baxter privately had no hope that they would , but Esther hoped for the best.
 16 She could not believe that Mr. Cropper would carry his prejudices into a personal application .
 17 This conviction was strengthened when he overtook her walking from school the next day and drove her home .
 18 He was a big , handsome man with a very suave , polite manner .
 19 He asked interestedly about her school and her work , hoped she was getting on well , and said he had two young rascals of his own to send soon .
 20 Esther felt relieved .
 Q : She thought that Mr. _____ had exaggerated matters a little .
 C : Baxter, Cropper, Esther, course, fingers, manner, objection, opinion, right, spite.
 a : Baxter

GPT-3 (OpenAI, Jul20): decoder-only transformer con **in-context learning**

- **Aproximación:** zero-shot, one-shot y few-shot en lugar de fine-tuning

The three settings we explore for in-context learning

Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 cheese => ..... ← prompt
```

One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← example
3 cheese => ..... ← prompt
```

Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← examples
3 peppermint => menthe poivrée ←
4 plush girafe => girafe peluche ←
5 cheese => ..... ← prompt
```

Traditional fine-tuning (not used for GPT-3)

Fine-tuning

The model is trained via repeated gradient updates using a large corpus of example tasks.



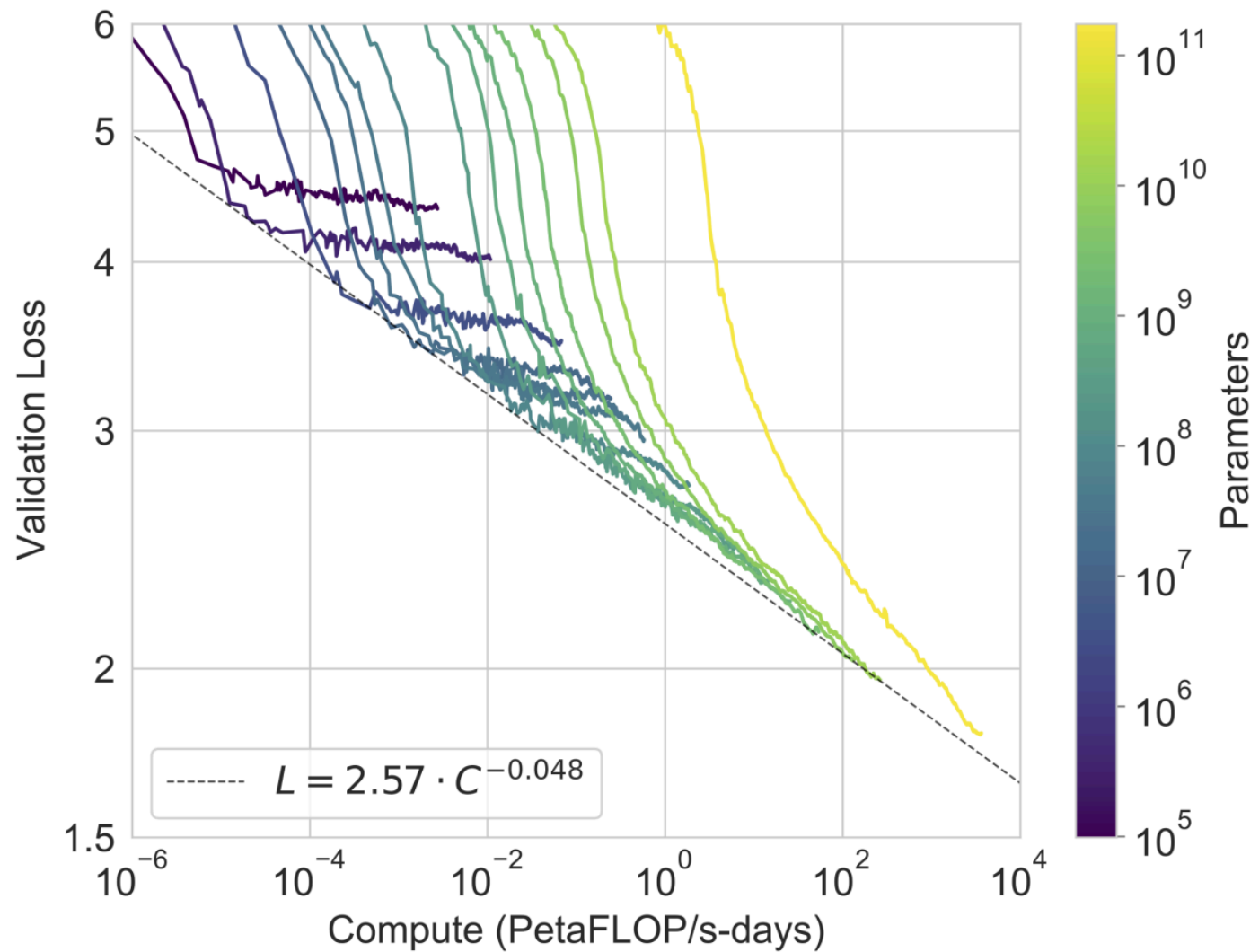
- **Modelos:** tamaños, arquitecturas e hiperparámetros de entrenamiento

| Model Name | n_{params} | n_{layers} | d_{model} | n_{heads} | d_{head} | Batch Size | Learning Rate |
|-----------------------|---------------------|---------------------|--------------------|--------------------|-------------------|------------|----------------------|
| GPT-3 Small | 125M | 12 | 768 | 12 | 64 | 0.5M | 6.0×10^{-4} |
| GPT-3 Medium | 350M | 24 | 1024 | 16 | 64 | 0.5M | 3.0×10^{-4} |
| GPT-3 Large | 760M | 24 | 1536 | 16 | 96 | 0.5M | 2.5×10^{-4} |
| GPT-3 XL | 1.3B | 24 | 2048 | 24 | 128 | 1M | 2.0×10^{-4} |
| GPT-3 2.7B | 2.7B | 32 | 2560 | 32 | 80 | 1M | 1.6×10^{-4} |
| GPT-3 6.7B | 6.7B | 32 | 4096 | 32 | 128 | 2M | 1.2×10^{-4} |
| GPT-3 13B | 13.0B | 40 | 5140 | 40 | 128 | 2M | 1.0×10^{-4} |
| GPT-3 175B or “GPT-3” | 175.0B | 96 | 12288 | 96 | 128 | 3.2M | 0.6×10^{-4} |

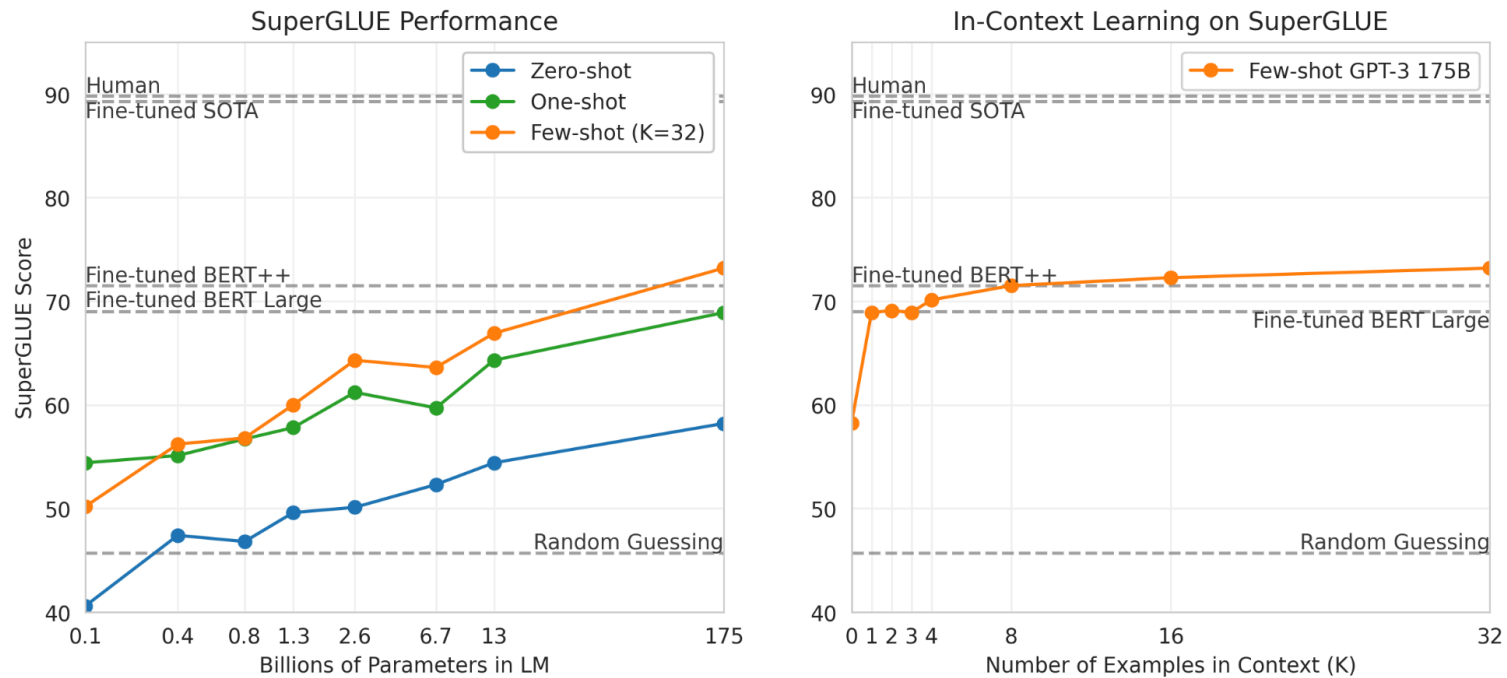
- **Coste de entrenamiento:** con 300B tokens de Common Crawl (filtrado), WebText2, Books1, Books2 y Wikipedia

| Model | Total train compute (PF-days) | Total train compute (flops) | Params (M) | Training tokens (billions) | Flops per param per token | Mult for bwd pass | Fwd-pass flops per active param per token | Frac of params active for each token |
|---------------|-------------------------------|-----------------------------|------------|----------------------------|---------------------------|-------------------|---|--------------------------------------|
| T5-Small | 2.08E+00 | 1.80E+20 | 60 | 1,000 | 3 | 3 | 1 | 0.5 |
| T5-Base | 7.64E+00 | 6.60E+20 | 220 | 1,000 | 3 | 3 | 1 | 0.5 |
| T5-Large | 2.67E+01 | 2.31E+21 | 770 | 1,000 | 3 | 3 | 1 | 0.5 |
| T5-3B | 1.04E+02 | 9.00E+21 | 3,000 | 1,000 | 3 | 3 | 1 | 0.5 |
| T5-11B | 3.82E+02 | 3.30E+22 | 11,000 | 1,000 | 3 | 3 | 1 | 0.5 |
| BERT-Base | 1.89E+00 | 1.64E+20 | 109 | 250 | 6 | 3 | 2 | 1.0 |
| BERT-Large | 6.16E+00 | 5.33E+20 | 355 | 250 | 6 | 3 | 2 | 1.0 |
| RoBERTa-Base | 1.74E+01 | 1.50E+21 | 125 | 2,000 | 6 | 3 | 2 | 1.0 |
| RoBERTa-Large | 4.93E+01 | 4.26E+21 | 355 | 2,000 | 6 | 3 | 2 | 1.0 |
| GPT-3 Small | 2.60E+00 | 2.25E+20 | 125 | 300 | 6 | 3 | 2 | 1.0 |
| GPT-3 Medium | 7.42E+00 | 6.41E+20 | 356 | 300 | 6 | 3 | 2 | 1.0 |
| GPT-3 Large | 1.58E+01 | 1.37E+21 | 760 | 300 | 6 | 3 | 2 | 1.0 |
| GPT-3 XL | 2.75E+01 | 2.38E+21 | 1,320 | 300 | 6 | 3 | 2 | 1.0 |
| GPT-3 2.7B | 5.52E+01 | 4.77E+21 | 2,650 | 300 | 6 | 3 | 2 | 1.0 |
| GPT-3 6.7B | 1.39E+02 | 1.20E+22 | 6,660 | 300 | 6 | 3 | 2 | 1.0 |
| GPT-3 13B | 2.68E+02 | 2.31E+22 | 12,850 | 300 | 6 | 3 | 2 | 1.0 |
| GPT-3 175B | 3.64E+03 | 3.14E+23 | 174,600 | 300 | 6 | 3 | 2 | 1.0 |

- Pérdida en validación en función de computación: ley de escala potencial $\mathcal{L} = 2.57 \cdot C^{-0.048}$



- **Rendimiento en SuperGLUE:** aumenta con el tamaño del modelo y número de ejemplos en contexto



- **SuperGLUE Score:** score medio de scores (medios) de las tareas

| Corpus | Train | Dev | Test | Task | Metrics | Text Sources |
|---------|-------|------|------|--------|---------------------|---------------------------------|
| BoolQ | 9427 | 3270 | 3245 | QA | acc. | Google queries, Wikipedia |
| CB | 250 | 57 | 250 | NLI | acc./F1 | various |
| COPA | 400 | 100 | 500 | QA | acc. | blogs, photography encyclopedia |
| MultiRC | 5100 | 953 | 1800 | QA | F1 _a /EM | various |
| ReCoRD | 101k | 10k | 10k | QA | F1/EM | news (CNN, Daily Mail) |
| RTE | 2500 | 278 | 300 | NLI | acc. | news, Wikipedia |
| WiC | 6000 | 638 | 1400 | WSD | acc. | WordNet, VerbNet, Wiktionary |
| WSC | 554 | 104 | 146 | coref. | acc. | fiction books |

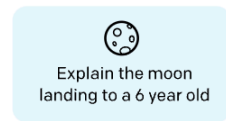
InstructGPT (OpenAI, Mar22): alineamiento de GPT-3 con la intención del usuario

- **Pasos:** fine-tuning supervisado (SFT), modelo de recompensa (RM) y aprendizaje por refuerzo (RL) con PPO(-ptx)

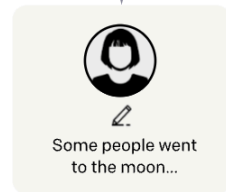
Step 1

Collect demonstration data, and train a supervised policy.

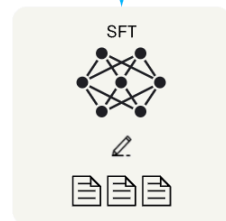
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



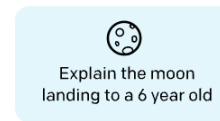
This data is used to fine-tune GPT-3 with supervised learning.



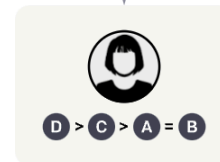
Step 2

Collect comparison data, and train a reward model.

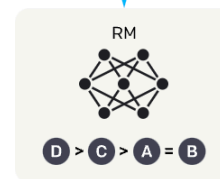
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.

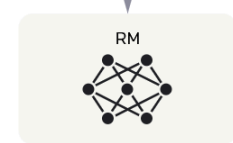


The policy generates an output.



Once upon a time...

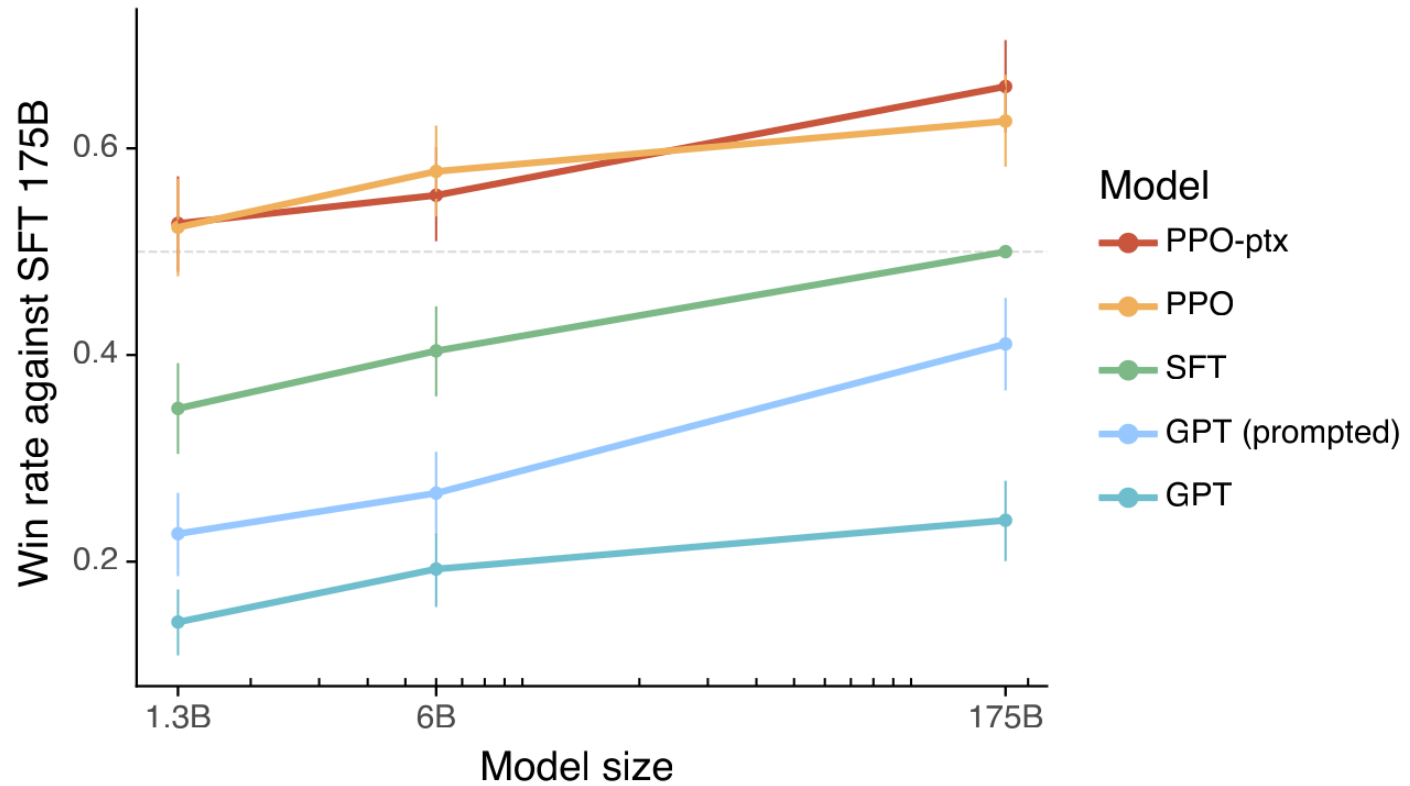
The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



- **Evaluaciones humanas:** preferencia por las salidas de cada modelo con respecto al modelo SFT 175B

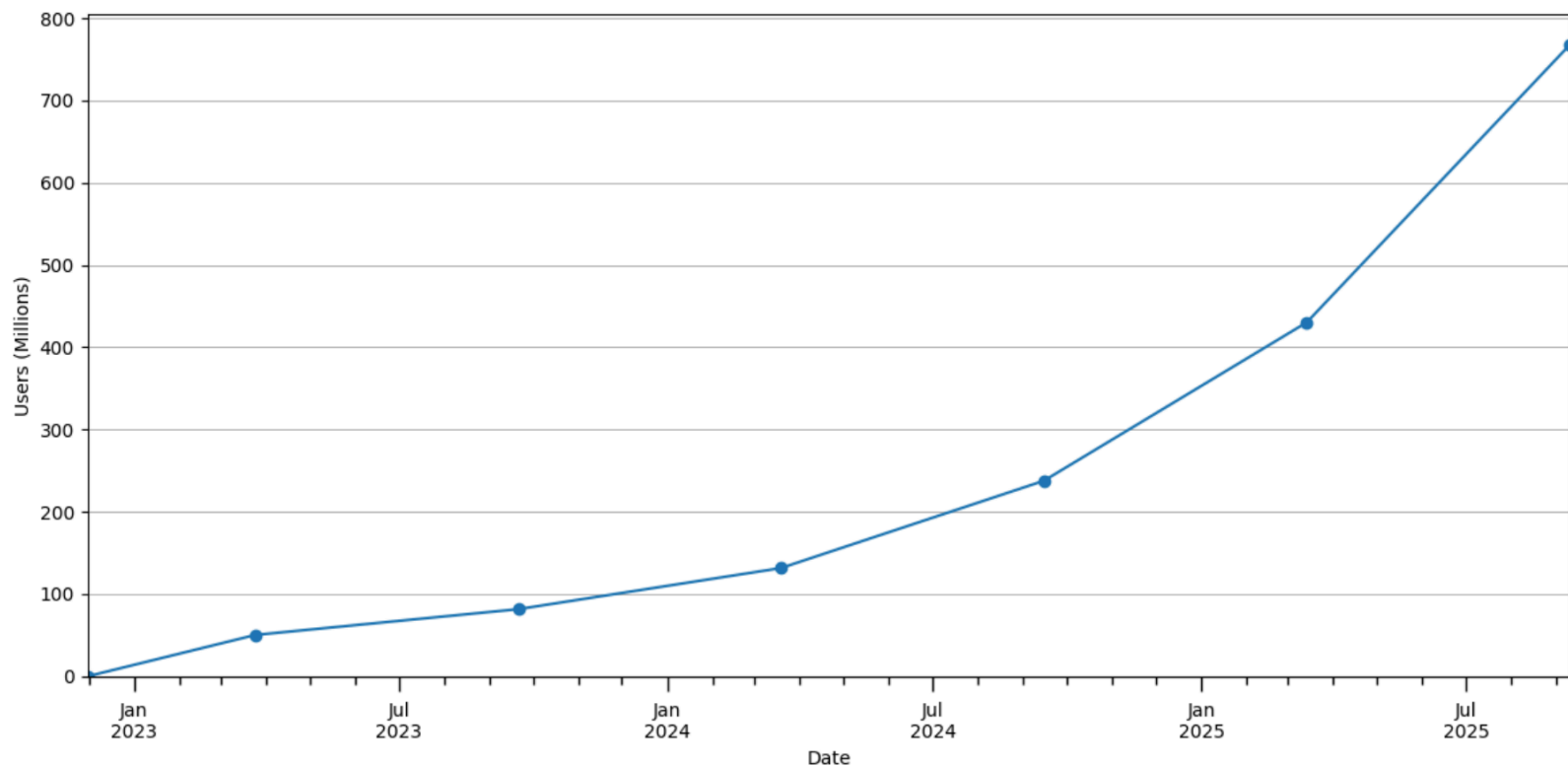


ChatGPT (OpenAI, Nov22): versión refinada de InstructGPT

- [The inside story of how ChatGPT was built from the people who made it \(MIT Technology Review, Mar23\)](#)

Liam Fedus: The ChatGPT model is fine-tuned from the same language model as InstructGPT, and we used a similar methodology for fine-tuning it. We had added some conversational data and tuned the training process a bit. So we didn't want to oversell it as a big fundamental advance. As it turned out, the conversational data had a big positive impact on ChatGPT. Microsoft and OpenAI evolve partnership to drive the next phase of AI

- [How people are using ChatGPT \(Sep25\):](#) usuarios activos semanalmente



4 Pre-entrenamiento

Objetivo: pre-entrenar un modelo base para predecir el siguiente token tras un contexto dado

Preproceso preliminar: adquisición y filtrado (limpieza) de texto (p. ej. duplicado y tóxico)

Preproceso básico: tokenización de texto con base en un vocabulario de tokens de tamaño adecuado

Construcción del vocabulario: con técnicas incrementales estándar, hasta alcanzar un tamaño dado

Byte-pair encoding (BPE): tokenizador estándar a nivel de carácter UTF8 (byte)

1. El vocabulario base se forma con (todos los) caracteres individuales
2. Regla de fusión: el par de tokens adyacentes más frecuente se fusiona en un nuevo token
3. El paso anterior se repite hasta alcanzar un vocabulario de tamaño dado

Ejemplo: a partir de `aaabdaaabac` hasta 5 tokens

1. `a a a b d a a a b a c` con vocabulario base `a b c d`
2. `aa a b d aa a b a c` tras fusionar `aa = a a`

Tokenización BPE: se aplican las reglas de fusión en el orden aprendido

Ejemplo (cont.): tokenización de `aabaa` con `a b c d aa`

1. `a a b a a`
2. `aa b aa`

WordPiece: variante de BPE que favorece la fusión de tokens adyacentes **relativamente** frecuentes

1. Cada palabra se divide añadiendo un prefijo a los caracteres interiores; p. ej. `word` → `w ##o ##r ##d`
2. El vocabulario base se forma con (todos los) tokens individuales, sin y con el prefijo de token interior
3. La frecuencia de cada par de tokens adyacentes, (i, j) , se evalúa en relación a la de sus tokens por separado, $\frac{F_{i,j}}{F_i F_j}$
4. Regla de fusión: el par de tokens adyacentes de mayor puntuación se fusiona en un nuevo token sin `##`
5. Los dos pasos anteriores se repiten hasta alcanzar un vocabulario de tamaño dado

Ejemplo: a partir de `aaabb` hasta 5 tokens

1. `a ##a ##a ##b ##b` con vocabulario base `a b ##a ##b`
2. `aa ##a ##b ##b` tras fusionar `aa = a ##a`

Tokenización WordPiece: no se aplican las reglas de fusión aprendidas; solo el vocabulario final

1. El proceso de tokenización se aplica palabra a palabra
2. Dada una palabra, se halla su prefijo más largo en el vocabulario y se divide de acuerdo con el mismo
3. Si queda sufijo, se procesa como una nueva palabra volviendo al paso anterior

Ejemplo (cont.): tokenización de `aaa` con `a b ##a ##b aa`

1. `aa ##a`

Ejemplo de dataset de pre-entrenamiento: [FineWeb2](#), 20TB de documentos en más de 1000 lenguas

```
In [ ]: from datasets import load_dataset
fw2 = load_dataset("HuggingFaceFW/fineweb-2", name="spa_Latn", split="train", streaming=True)
sample = next(iter(fw2)); sample
```

```
Resolving data files: 0%|          | 0/146 [00:00<?, ?it/s]
```

```
Out[ ]: {'text': 'En fín, Cruz .... Porque después de muchas negociaciones maritales, mi panadero, como yo le llamo, ha termi
nado convenciéndome de que no es buen momento para comprar la thermomix, y era algo de lo que yo tenía muchas ganas;
así es la vida, toca esperar.\nY Cara .... porque como casi todo en la vida, tiene un lado bueno; mi panadero para in
tentar subirme el ánimo, hasta que pueda comprarme la maquinita, pues me ha regalado una panificadora.\nHace unos dí
as, Cris en su blog "Kanela y Limón", uno de los mejores blogs de cocina que he visitado en mi vida, publicaba que p
onían a la venta en Lidl la panificadora Bifinett; y Alberto decidió comprármela.\nMe la regaló el mismo sábado que
se ponía a la venta, pero no pude abrirla y verla en todo el día, pues comíamos en casa de mis suegros. Se me hizo u
n día muy largo, porque moría de ganas de llegar a casa y abrirla, jejejeje. En cuanto llegamos a casa, ya de noche,
casi no me dio tiempo de soltar el bolso y ya estaba abriendo la caja. Seguí las instrucciones de "antes del primer
uso" y así quedó.\nEl Domingo, con los ojos aún medio cerrados y cambio de hora incluido, me fui corriendo a la coci
na, busqué la primera receta de la que tuviera los ingredientes, los pesé, y a la cubeta. Pan en el programa 2 "Espo
njoso".\nA las 2 en punto salía de mi nueva panificadora un hermoso pan cuadrado!!! No daba crédito a lo que veía. L
o saqué de la cubeta, lo puse en una bandeja y se lo llevé a mi suegra (que repetíamos comiendo allí). Menos mal, qu
e se me ocurrió hacerle una foto antes de irme, porque no quedó ni la muestra del corte. Sencillamente buenísimo!!!
Tan bueno, que mi suegra, que ya tiene una panificadora Princess, el lunes a primera hora fue al Lidl y se compró un
a como la mía.\nUn besazo a todas y todos!!',
'id': '<urn:uuid:85e1aa56-6833-4cb4-876f-95709ed428ae>',
'dump': 'CC-MAIN-2013-20',
'url': 'http://aguasanta-micocina.blogspot.com/2010_03_01_archive.html',
'date': '2013-05-19T07:11:09Z',
'file_path': 's3://commoncrawl/crawl-data/CC-MAIN-2013-20/segments/1368696384213/warc/CC-MAIN-20130516092624-00033-
ip-10-60-113-184.ec2.internal.warc.gz',
'language': 'spa',
'language_score': 0.9976795315742493,
'language_script': 'Latn',
'minhash_cluster_size': 8,
'top_langs': '{}'}

```

```
In [ ]: from datasets import load_dataset
fw2 = load_dataset("HuggingFaceFW/fineweb-2", name="cat_Latn", split="train", streaming=True)
sample = next(iter(fw2)); sample
```

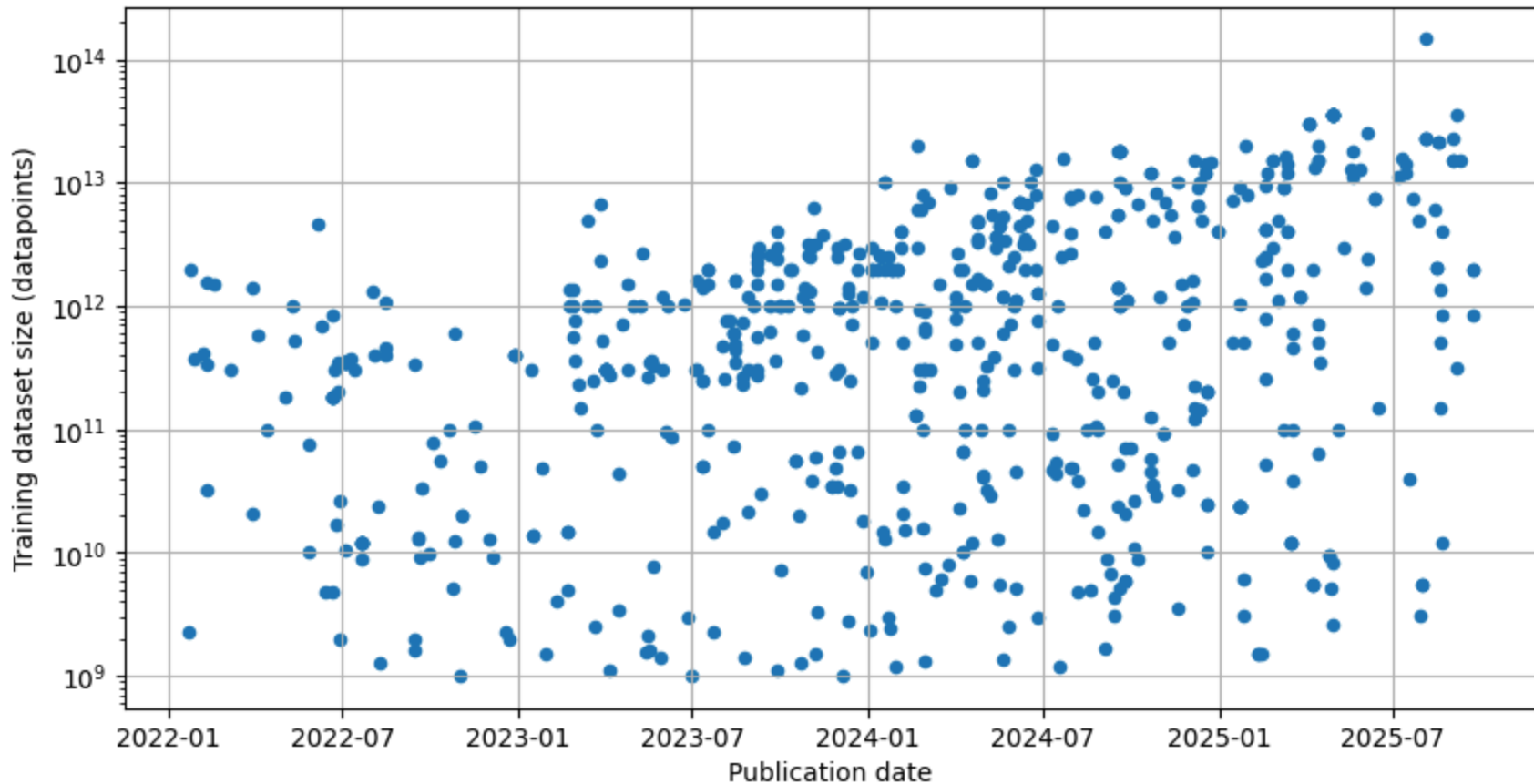
```
Out[ ]: {'text': 'Nosaltres som en Martí Palma, Roger Navarro,Irene Berdún, Judith Consentino, Daniel Martín i Víctor Gallar
do. Som alumnes de l'1.2 de batxillerat de l'INS Isaac Albéniz de Badalona. Nosaltres us presentarem una sèrie de re
creacions situades al Port d'Òstia i que narren, des de diferents punts de vista, la història de Clàudia, la Vesta
l.\nPort of Ostia during a Tempest - 1740-1750 de Leonardo Coccorante – Lowe Art Museum\nTothom parlava d'ella aquel
ls dies. Tots l'assenyalaven i la miraven malament. A la seva esquena li deien de tot, res de bo. La culpaven de l'e
ncallament del vaixell al riu, la culpaven de la guerra; fins i tot, la culpaven de mals que ella no havia sentit ma
i a parlar, per principalment la culpaven de la guerra. Ja que si aquell vaixell no arribava a port l'estàtua no hi
arribaria i la guerra no acabaria, tal com havia dit aquell vell oracle. Jo sóc Lucius, i sóc el que ningú coneix, p
erò tothom considera el seu "amant". Clàudia i jo no érem amants. Ella és una vestal, i no pot trencar ni trencarà m
ai el seu vot de castedat. Simplement ens vam enamorar. Clàudia era una veïna meva, vivia a la domus del costat. La
vaig veure créixer, els seus pares, patricis els dos, van morir a la guerra, i ella, sola com estava, va donar la se
va virginitat a Vesta. Poc temps després que fes el seu vot, la vaig veure un dia amb aquell quitó que li havia rega
lat. Estava molt guapa. Ens vam mirar i, poc temps després, passàvem tardes junts caminant i xerrant pel fòrum. Algu
ns dies visitàvem l'àgora, ens encantava anar per allà. Un dia li van dir que havia d'acompanyar l'estàtua de Cíbele
cap al port de Roma i ens vam acomiadar. Jo per sorprendre-la, la vaig seguir amb la meva gal·lera. Aquell viatge va
durar uns quants dies, però finalment vam arribar al riu. Allà el vaixell va encallar i ella, que anava en un carret
ó, era la principal culpable. Una vestal companya seva ens va veure per l'àgora ben junts, sense fer res que els déu
s no volguessin, i aquesta va anar a dir-ho per tot arreu. En poc temps la van culpar de la història del vaixell i d
e molts actes de la guerra pel fet de trencar el seu vot de castedat. Jo sabia que era fals, però és clar, si deia a
lguna cosa, sabrien que jo existia i potser l'acabarien enterrant viva igualment. Mentre tots parlaven d'ella, jo
l'observava des de la distància i veia el que feia a totes hores. Un bon dia pel matí, vaig seguir-la fins al port,
volia dir-li que jo estava amb ella, i que tots dos junts podríem fugir per sempre. I així evitar que l'enterressin
viva... Ella però, no em va veure. Va agafar una barca i es va dirigir al vaixell. Des de l'altura del port, veia com
amb molt esforç es lligava al cinturó quelcom que semblava una corda i com, per art dels déus, movia aquest vaixell.
Ella sola, sense ajuda de ningú, només dels déus, que així van provar que ni ella ni jo érem culpables, que ella no
havia trencat el seu vot de castedat. Així doncs, en veure que tot el poble l'aclamava, me'n vaig anar, i per decisi
ó dels déus la vaig deixar en pau, i vaig continuar amb la meva patètica vida sense ella.\nMARTÍ PALMA\n1.2 BATXILLE
RAT\nINS ISAAC ALBÉNIZ',
'id': '<urn:uuid:270b92bc-8619-4c38-b48c-0645684322b0>',
'dump': 'CC-MAIN-2013-20',
'url': 'http://blocs.xtec.cat/aracnefilaifila/tag/llati/',
'date': '2013-05-19T16:26:05Z',
'file_path': 's3://commoncrawl/crawl-data/CC-MAIN-2013-20/segments/1368697772439/warc/CC-MAIN-20130516094932-00041-
ip-10-60-113-184.ec2.internal.warc.gz',
'language': 'cat',
'language_score': 0.9999985694885254,
'language_script': 'Latn',
'minhash_cluster_size': 2,
'top_langs': '{}'}

```

Atlas de modelos públicos: [We Should Chart an Atlas of All the World's Models](#)

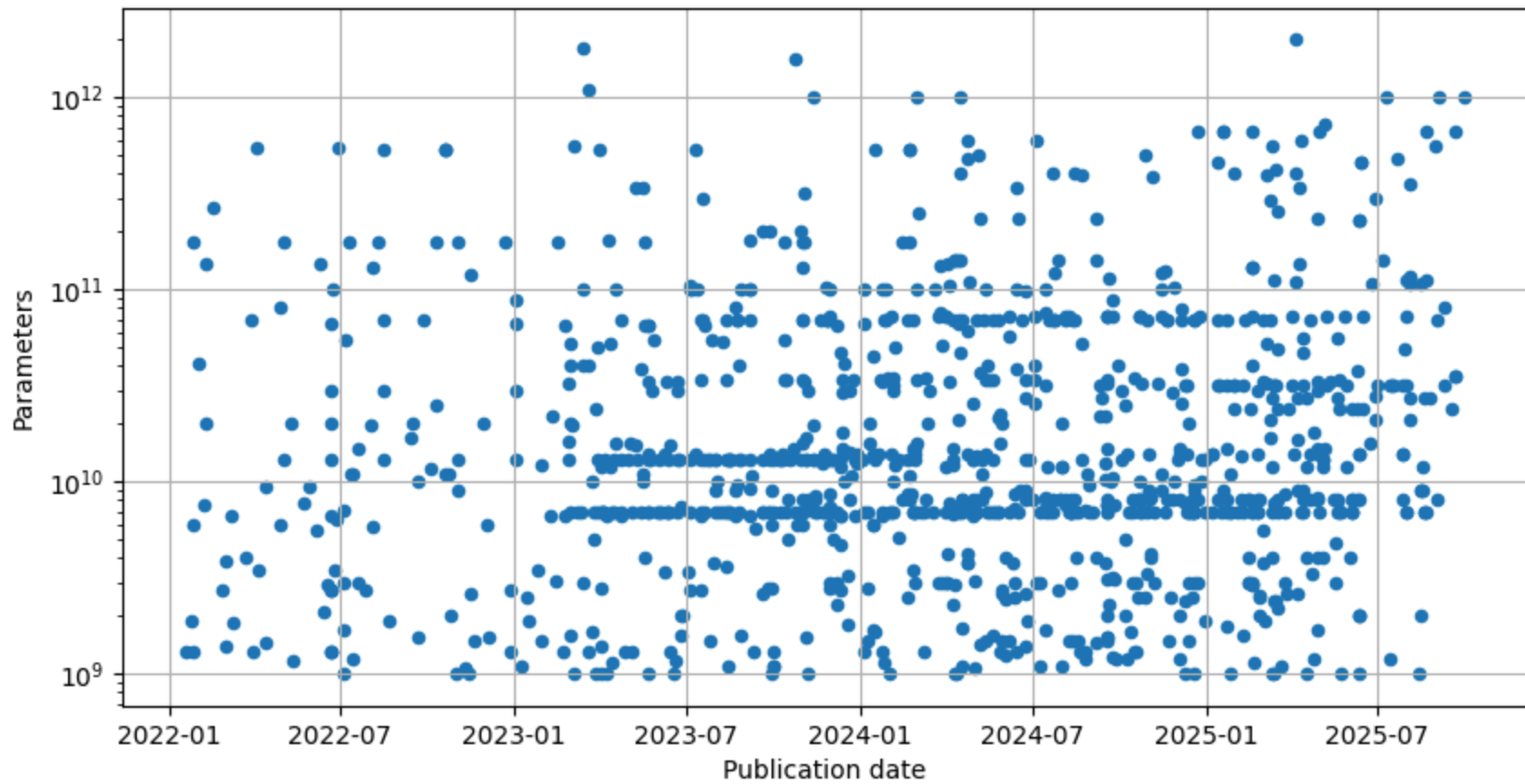
Datos de pre-entrenamiento de modelos recientes: decenas de trillones de datos (trillón = 10^{12}); tendencia creciente

```
In [ ]: import pandas as pd; file = "epoch.ai/all_ai_models.csv"
x = 'Publication date'; y = 'Training dataset size (datapoints)'
df = pd.read_csv(file, usecols=[x, y], parse_dates=[x])
df = df[(df[x] > "2022-01-01") & (df[y] > 9.9e8)]
plot = df.plot.scatter(x, y, logy=True, grid=True, figsize=(10, 5))
```



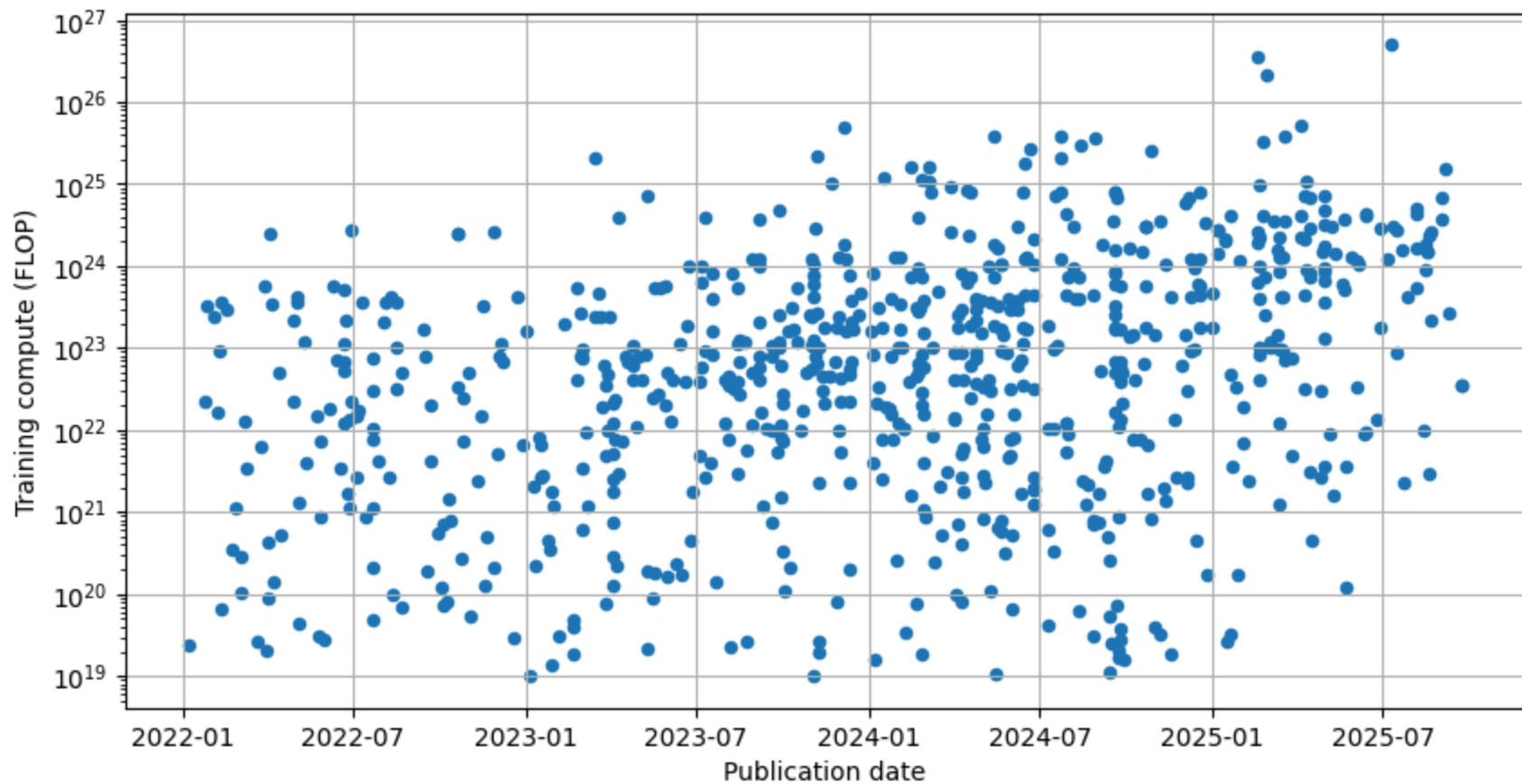
Tamaño de modelos recientes: entre el billón y trillón de parámetros; estable

```
In [ ]: import pandas as pd; file = "epoch.ai/all_ai_models.csv"
x = 'Publication date'; y = 'Parameters'
df = pd.read_csv(file, usecols=[x, y], parse_dates=[x])
df = df[(df[x]>"2022-01-01") & (df[y]>9.9e8)]
plot = df.plot.scatter(x,y,logy=True,grid=True,figsize=(10,5))
```



Coste de pre-entrenamiento de modelos recientes: entre 10^{19} y 10^{27} FLOPs; tendencia creciente

```
In [ ]: import pandas as pd; file = "epoch.ai/all_ai_models.csv"
x = 'Publication date'; y = 'Training compute (FLOP)'
df = pd.read_csv(file, usecols=[x, y], parse_dates=[x])
df = df[(df[x]>"2022-01-01") & (df[y]>9.9e18)]
plot = df.plot.scatter(x,y,logy=True,grid=True,figsize=(10,5))
```

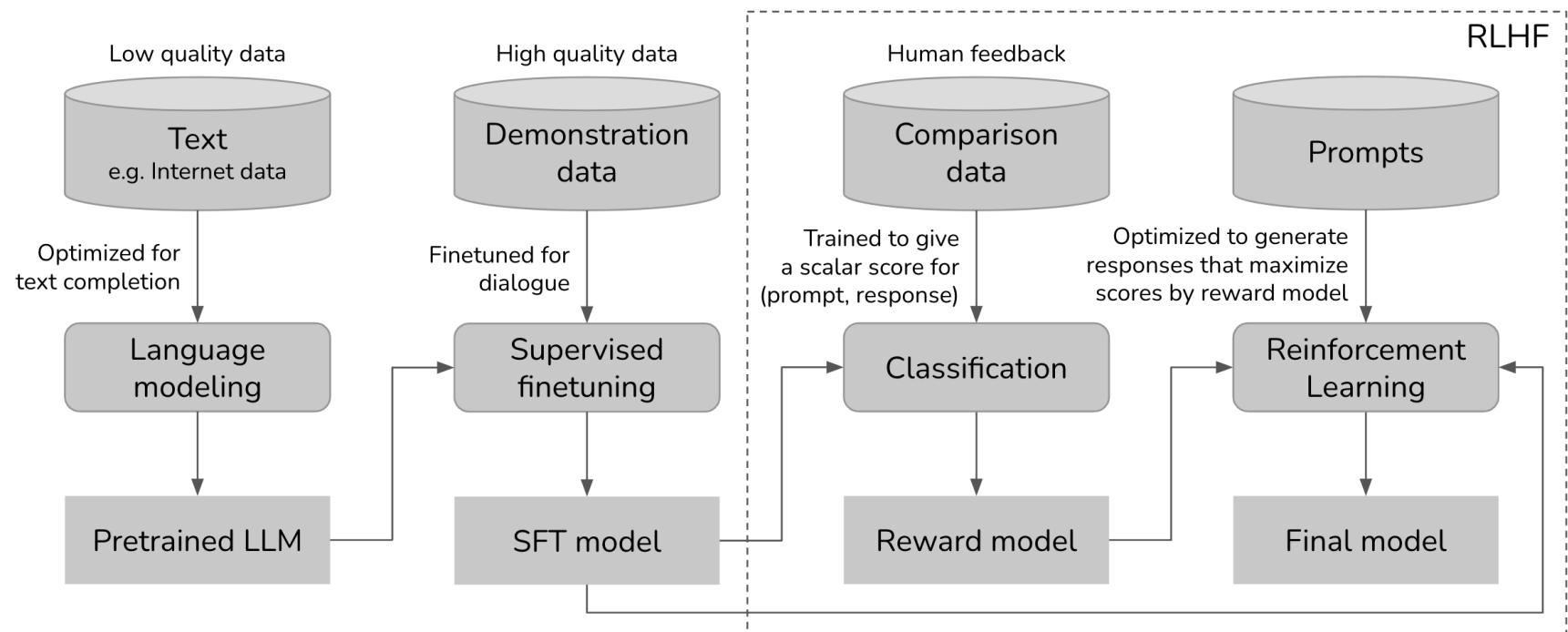


5 Post-entrenamiento

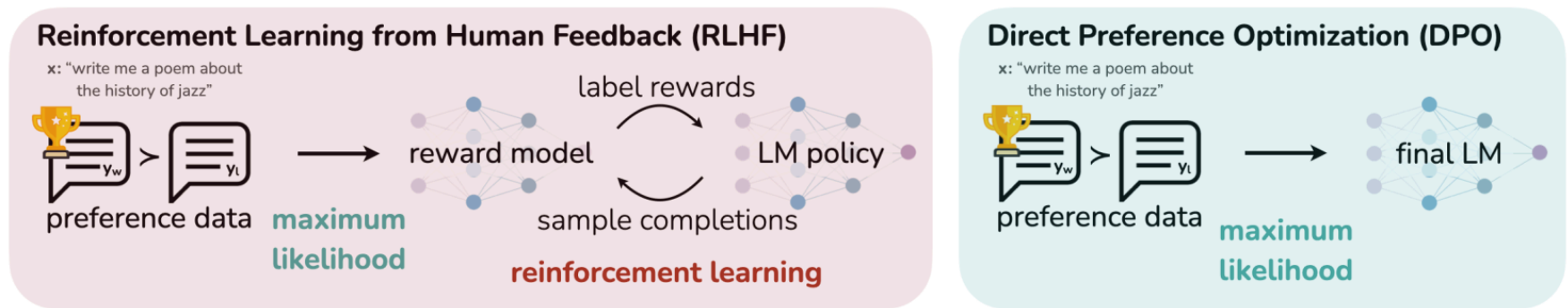
Objetivo: fine-tuning de un modelo pre-entrenado para adaptarlo según convenga

Supervised fine-tuning (SFT): entrenamiento muy similar al pre-entrenamiento; para conversar en un dominio específico

Reinforcement learning from human feedback (RLHF): aprendizaje por refuerzo basado en un modelo de recompensa entrenado a partir de preferencias humanas



Direct preference optimisation (DPO): versión simplificada de RLHF



Parameter-efficient fine-tuning (PEFT): entrenamiento eficiente de parámetros; unos "pocos" en lugar de todos

Retrieval-augmented generation (RAG): condicionamiento en inferencia con información externa relevante

Test-time prompting & reasoning: condicionamiento en inferencia con prompts elaborados; p. ej. chain-of-thought (CoT)

Compresión, cuantización, poda y knowledge distillation: reducción del modelo; necesario con recursos modestos

Continual Learning: actualización continua del modelo con nuevos datos

Seguridad: reducción de comportamientos indeseados (toxicidad, sesgos, fugas de datos, etc.)

6 Evaluación

Epoch AI: instituto de investigación en IA; SOTA, tendencias y predicciones

Epoch AI Publications & Commentary:

- **Papers & Reports:** artículos e informes detallados
- **Newsletter:** comentarios semanales breves y subjetivos
- **Podcast:** conversaciones informales sobre el presente y futuro de la IA

Epoch AI Datasets & Resources:

- **Overview:** datos sobre modelos ML desde 1950
- **AI Benchmarking:** datos de evaluaciones internas y externas sobre tareas actuales
- **AI Models:** datos sobre más de 3100 modelos
- **ML Hardware:** datos sobre más de 170 aceleradores IA (GPUs y TPUs)
- **GPU Clusters:** datos sobre más de 500 clústers GPU y supercomputadoras
- **AI Companies:** datos sobre empresas en IA frontera
- **AI Trends & Statistics:** cifras, gráficas y titulares sobre SOTA y tendencias
- **Data Insights:** análisis de datos y conclusiones

Epoch AI Projects:

- **FrontierMath:** problemas de matemáticas avanzados
- **GATE Playground:** modelo para evaluar el impacto económico de la IA
- **Distributed Training** simulador de entrenamiento distribuido
- **Model Counts:** proyección del número de modelos que superarán ciertos umbrales de computación

Artificial Analysis: consultora en IA

- **Models:** models, providers for a model, providers, open source models, evaluations, capabilities, etc.
- **Speech, Image, Video:** image generation, speech to text, text to speech, speech to speech, video generation
- **Hardware:** AI Hardware Benchmarking & Performance Analysis
- **Leaderboards:** [models](#), [providers](#), [text2image](#), [image-edit](#), [text2video](#), [image2video](#), [speech2text](#), [text2speech](#), [music](#)
- **AI Trends:** AI progress, efficiency, country analysis, open source models, model architecture, training analysis
- **MicroEvals:** evaluación de un mismo prompt con múltiples modelos
- **Arena:** [image arena](#), [video arena](#), [speech arena](#), [music arena](#)
- **Articles:** artículos e informes detallados

Modelos mejor valorados en Artificial Analysis (Oct25):

1. GPT-5 Codex (high)
2. GPT-5 (high)
3. GPT-5 (medium)

LMarena: plataforma popular para la evaluación colectiva y abierta de LLMs mediante preferencias

- **Overview:** resumen de cada arena seguido de un ranking global
- **Text Arena:** versatilidad, precisión lingüística y contexto cultural de texto (ver overall desplegado)
- **WebDev Arena:** tareas de desarrollo web con HTML, CSS y JavaScript
- **Vision Arena:** comprensión y procesamiento de entradas visuales
- **Text-to-Image Arena:** generación de imágenes a partir de descripciones textuales
- **Image Edit Arena:** generación y edición de imágenes
- **Search Arena:** búsqueda en tiempo real de información, conocimiento externo y citas reales
- **Text-to-Video Arena:** generación de vídeos a partir de un prompt dado
- **Image-to-Video Arena:** generación de vídeos a partir de imágenes dadas
- **Copilot Arena:** generación de código en varios lenguajes de programación y tareas

Modelos mejor valorados en LMarena (Oct25):

1. `claude-opus-4-1-20250805-thinking-16k`
2. `claude-sonnet-4-5-20250929-thinking-32k`
3. `gpt-4.5-preview-2025-02-27`

OpenCompass: plataforma china de evaluación de LLMs

- **Large Model:** Large Model Evaluation System
- **Embodied Intelligence:** Embodied Intelligence Evaluation System
- **Safety:** Security Evaluation System
- **AI Computing System:** AI Computing System Evaluation System
- **Industry Verticals:** AI x Finance Evaluation Framework

OpenCompass Large Model:

- **Introduction:** breve descripción de la plataforma
- **CompassHub:** benchmarks por categorías
- **CompassRank:** rankings por categorías
- **Rules:** página principal
- **CompassKit:** repo github
- **Docs:** documentación
- **AI4S Program:** registro en el programa de evaluación
- **Compass Arena:** evaluación mediante preferencias

Modelos mejor valorados en OpenCompass (Oct25):

1. GPT-5-2025-08-07
2. o3-high-2025-04-16
3. Doubao-Seed-1.6-thinking-250715