

IA generativa. Intro

Jesús Moreno (Programamos)
David Ríos Insua (ICMAT-CSIC)
3 Julio 2024, Valencia

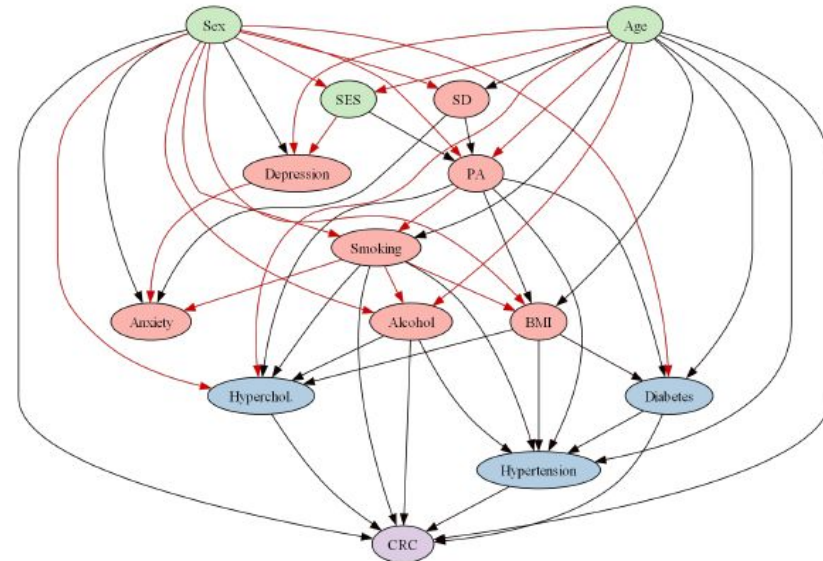
Discriminativa o generativa

IA de la Aa la Z (Pablo GB. El País. Rojo y verde, nuestros)

- **IA discriminativa.** También llamada IA tradicional o analítica, convive con nosotros desde hace tiempo. Se utiliza entre otras en tareas de clasificación de elementos, ya sean imágenes, texto u otro tipo de datos. Sin embargo, no es capaz de generar contenido por sí misma. Aprendizaje supervisado (Regresión, clasificación,...)
- **IA generativa.** Una nueva clase de IA, distinta a la que servía para análisis de datos y lleva décadas operativa. Como indica su nombre, va mucho más allá y es capaz de crear contenido en forma de texto, imágenes, vídeo o música.

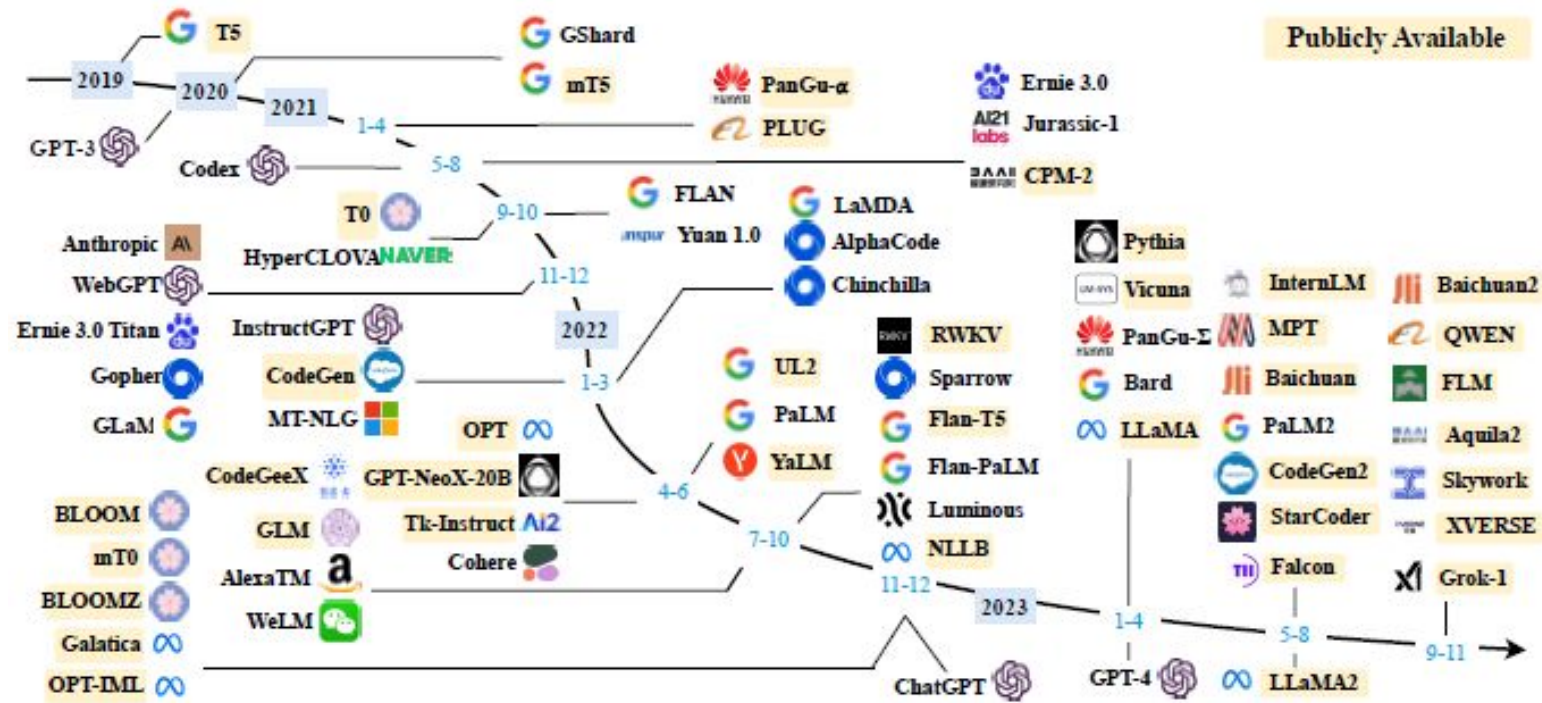
Aprendizaje no supervisado (generación, estimación densidades,...)

Discriminativa o generativa?



$$\begin{aligned}
 p(v_{sex}, \dots, v_{depression}) &= \left[p(v_{sex}) p(v_{age}) p(v_{SES} | v_{sex}, v_{age}) \right] \times \\
 &\quad \left[p(v_{SD} | v_{sex}, v_{age}) p(v_{PA} | v_{sex}, v_{age}, v_{SD}, v_{SES}) p(v_{depr} | v_{sex}, v_{age}, v_{SES}) \right. \\
 &\quad p(v_{smok} | v_{sex}, v_{age}, v_{PA}) p(v_{alc} | v_{sex}, v_{age}, v_{smok}) \\
 &\quad \left. p(v_{BMI} | v_{sex}, v_{age}, v_{PA}, v_{smok}) p(v_{anx} | v_{sex}, v_{SD}, v_{smok}, v_{depr}) \right] \times \\
 &\quad \left[p(v_{hypchol} | v_{sex}, v_{age}, v_{PA}, v_{smok}, v_{BMI}, v_{alc}) p(v_{diab} | v_{sex}, v_{age}, v_{PA}, v_{BMI}) \right. \\
 &\quad \left. p(v_{hypsten} | v_{age}, v_{PA}, v_{smok}, v_{BMI}, v_{alc}, v_{diab}) \right] \times \\
 &\quad p(v_{CRC} | v_{sex}, v_{age}, v_{alc}, v_{smok}, v_{hypchol}, v_{hypsten}, v_{diab})
 \end{aligned} \tag{1}$$

El sprint de los LLMs

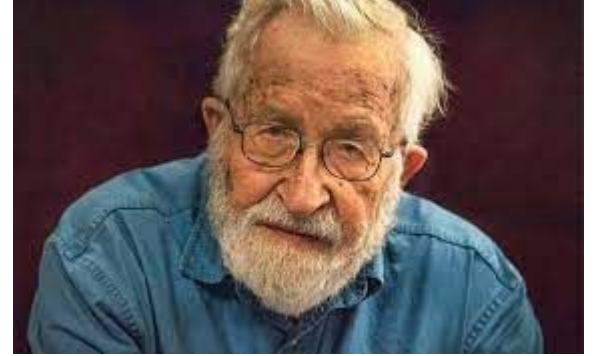


A survey of large language models <https://arxiv.org/abs/2303.18223>

Eight things to know about LLMs <https://arxiv.org/abs/2304.00612>

Images from sources quoted at various slides

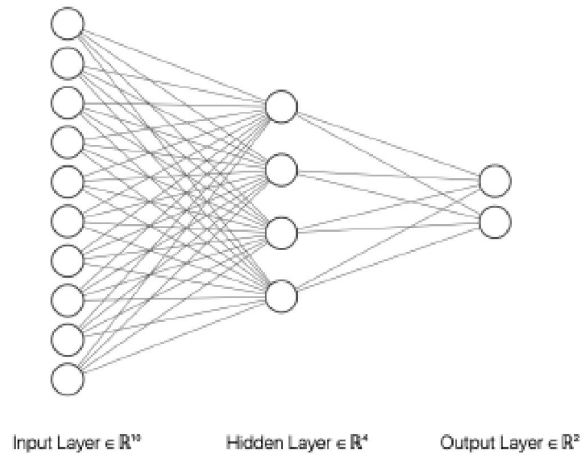
Un cambio de paradigma en PLN



- En tiempos pre-deep learning
 - Lenguaje como elementos y reglas a combinar
 - Gramáticas independientes de contexto
 - Más cercano a lenguajes artificiales (programación)
- A la actualidad
 - Lenguaje como probabilidades de secuencias de palabras
 - Cálculo de frecuencias de palabras, n-gramas,...
 - Transformers, SOTA
 - Casi una commodity (como vision)



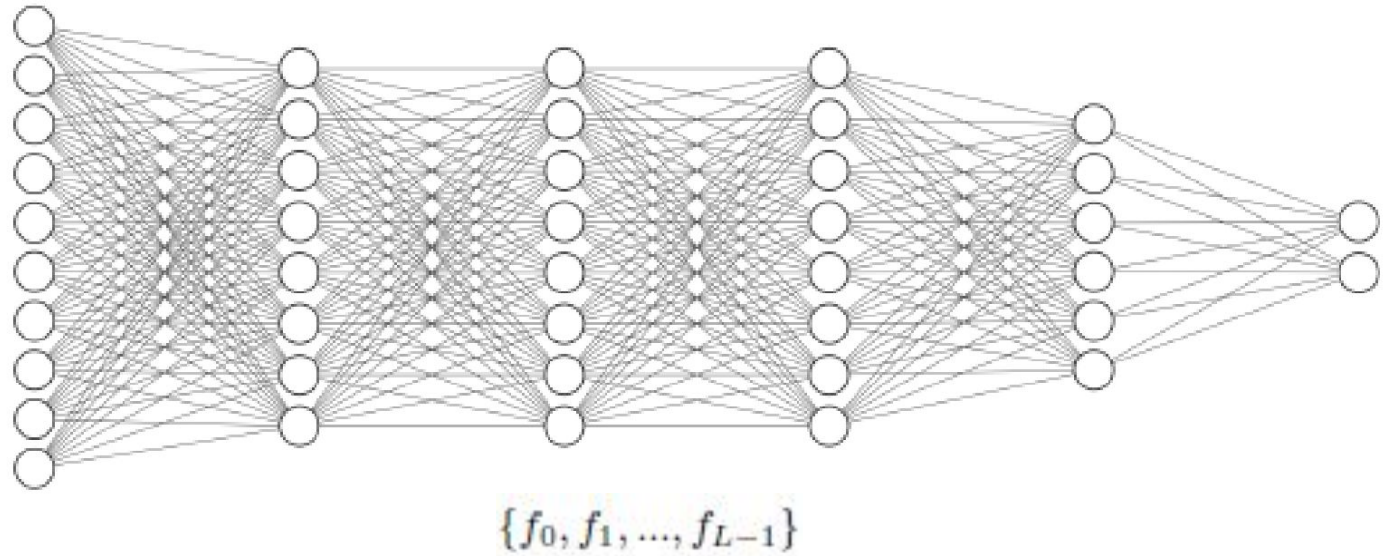
Redes profundas



$$y = \sum_{j=1}^m \beta_j \psi(x' \gamma_j) + \epsilon$$

$$\epsilon \sim N(0, \sigma^2),$$

$$\psi(\eta) = \exp(\eta) / (1 + \exp(\eta))$$



$$z_{l+1} = f_l(z_l, \gamma_l)$$

$$y = \sum_{j=1}^{m_L} \beta_j z_{L,j} + \epsilon$$

$$\epsilon \sim N(0, \sigma^2),$$

Motivación

- RNs completamente conexas pueden aproximar cualquier función....
- Pero entrenamiento puede ser superlento y requerir muchísimos datos
- En algunos dominios, se gana un montón con arquitecturas específicas
- En PLM, redes recurrentes (RNNs), y sucesoras (Transformers)
- Más generalmente, en procesamiento de secuencias, RNNs y sucesoras

RNNs. Ideas clave

- Especializadas en datos secuenciales (de números, de palabras, de imágenes,...)
- Pueden procesar sucesiones mucho más largas
- Más susceptibles de paralelizar
- Cada neurona tiene una (o más) 'memoria interna' (oculta) para almacenar info de entradas previas
- Entrenadas con variantes de algoritmos estándar: backprop en el tiempo
- Creadas en los 80 (RNN) y finales de los 90 (LSTM, GRU) , pero éxitos recientes (SIRI, etc...) las ponen en el radar
- Última ola: Transformers (todo lo que necesitas es atención), LLM, Chat-GPT,... 'la arquitectura'
- Miríadas de aplicaciones
 - Reconocimiento de habla
 - Generación de textos
 - Traducción automática
 - Generación de descripción de imágenes
 - Resúmenes, Q&A
 - Imagen, Video,....
 - Demostraciones matemáticas, estadística,...
 - Descripción molecular, desubrimiento de medicamentos,...

Intro a LLMs

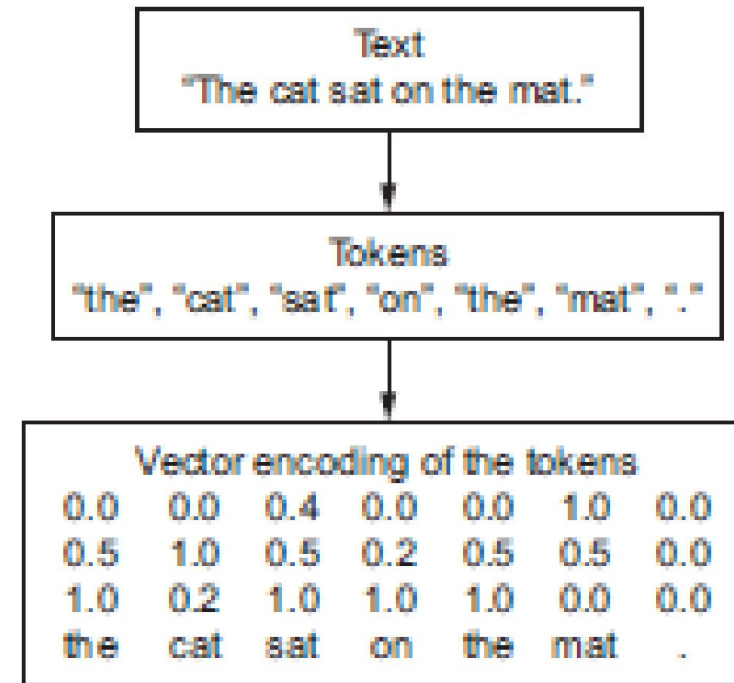
<https://llms-cunef-icmat-rg2024.github.io/>

Trabajando con textos

Sucesión de caracteres

Sucesión de palabras

Sucesión de n-gramas



Tokenizar y vectorizar

- Segmentar texto en palabras, transformar cada palabra en vector
- Segmentar en caracteres, transformar cada carácter en vector
- Extraer n-gramas, transformar cada n-grama en vector

One hot encoding

Unico entero a cada palabra

Palabra, vector de ceros, excepto 1 para la palabra

Poco denso y alta dimensión

Bolsas de palabras

Embeddings (incrustaciones) de palabras

Vectores densos de baja dimensión

Aprendidos de los datos

Junto a tarea principal. Capa de embedding

Pre-entrenados. Word2vec, Glove,...

Embedding. Ejemplo

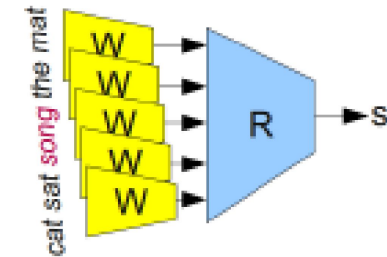
Función parametrizada

$$W : \text{words} \rightarrow \mathbb{R}^n$$

Entrenada para predecir si un 5grama es válido

$$R(W(\text{"cat"}), W(\text{"sat"}), W(\text{"on"}), W(\text{"the"}), W(\text{"mat"})) = 1$$

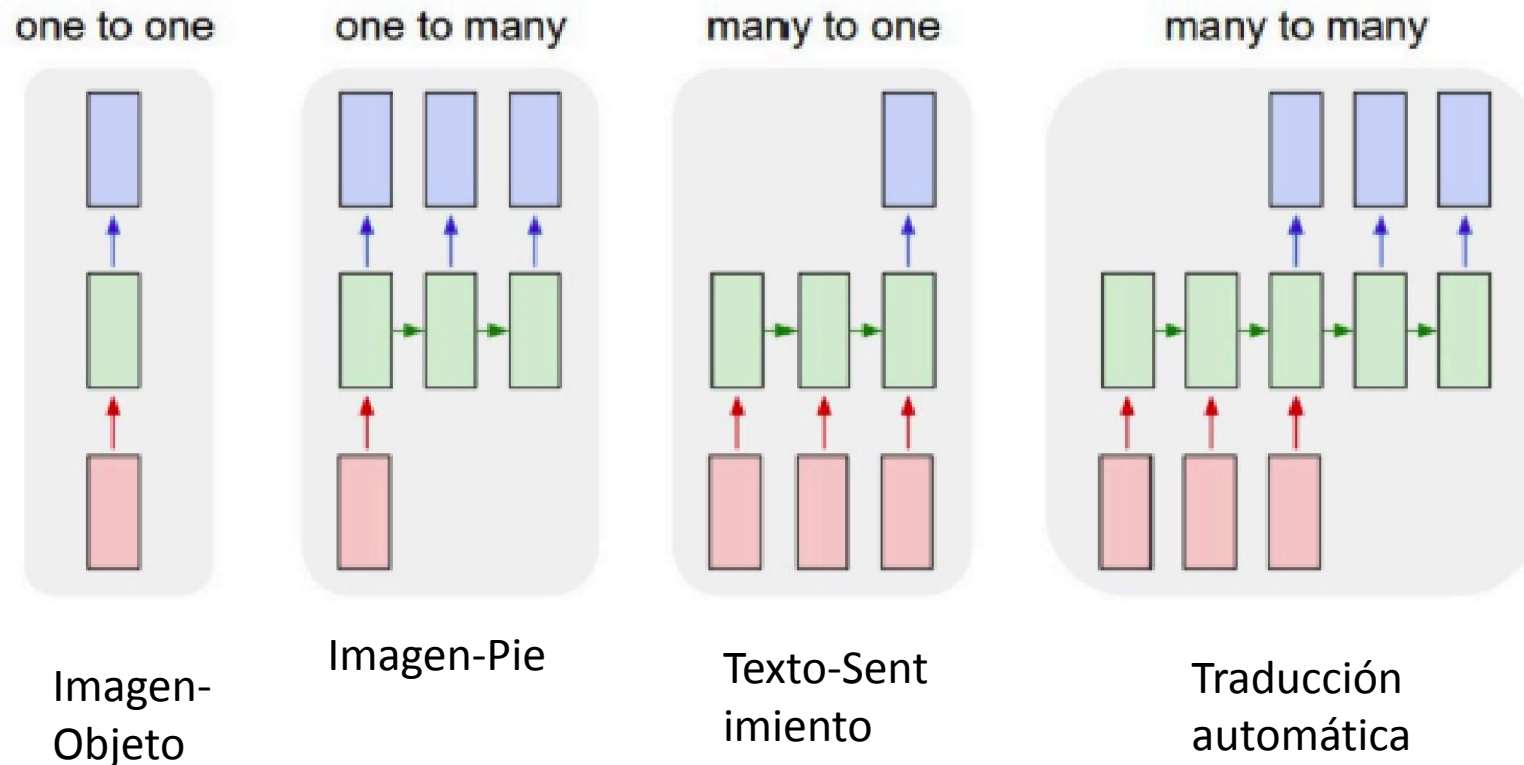
$$R(W(\text{"cat"}), W(\text{"sat"}), W(\text{"song"}), W(\text{"the"}), W(\text{"mat"})) = 0$$



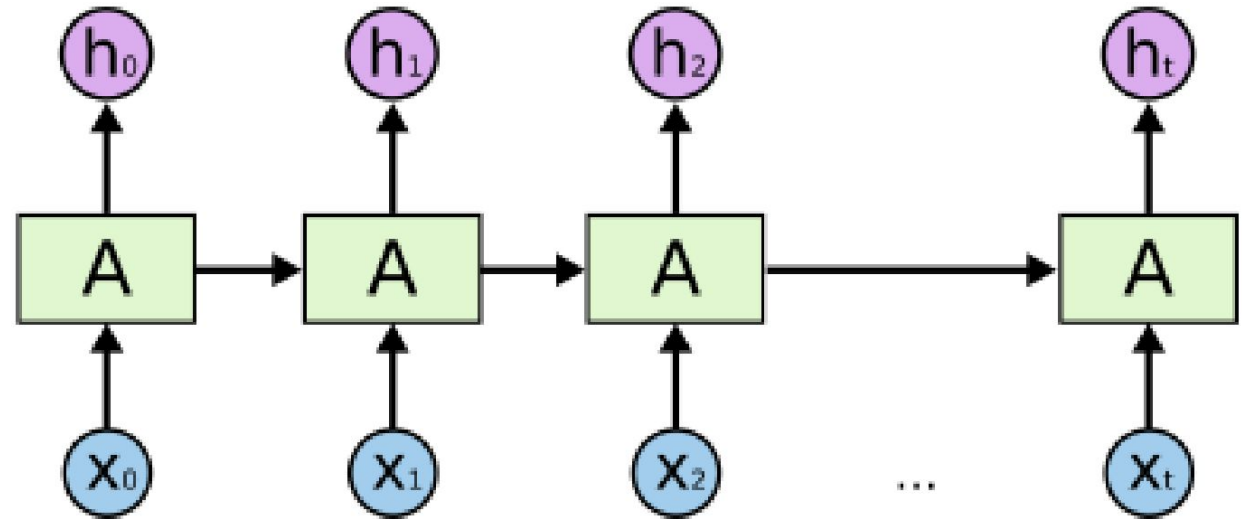
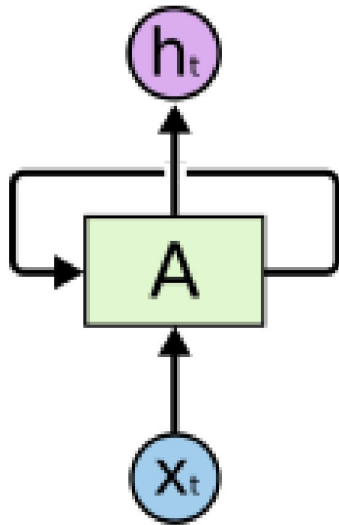
BERT entrenado con 300 bn tokens para predecir la siguiente palabra

RNNs

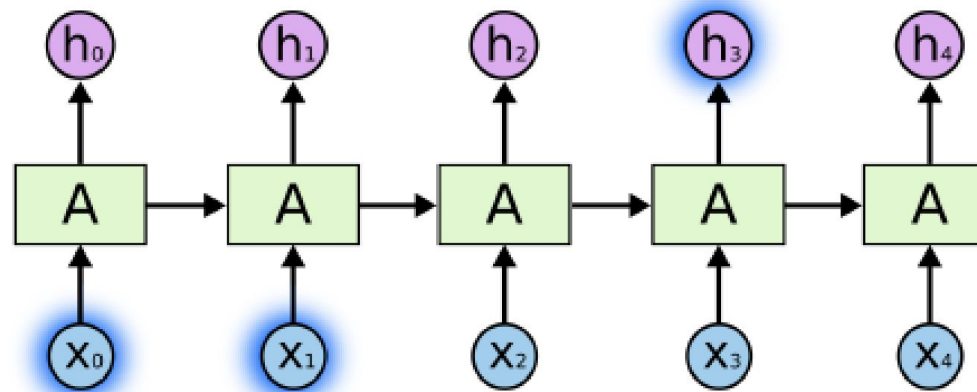
- Procesamiento de secuencias, incluso de longitud diferente
- Añaden conexiones de feedback



RNNs

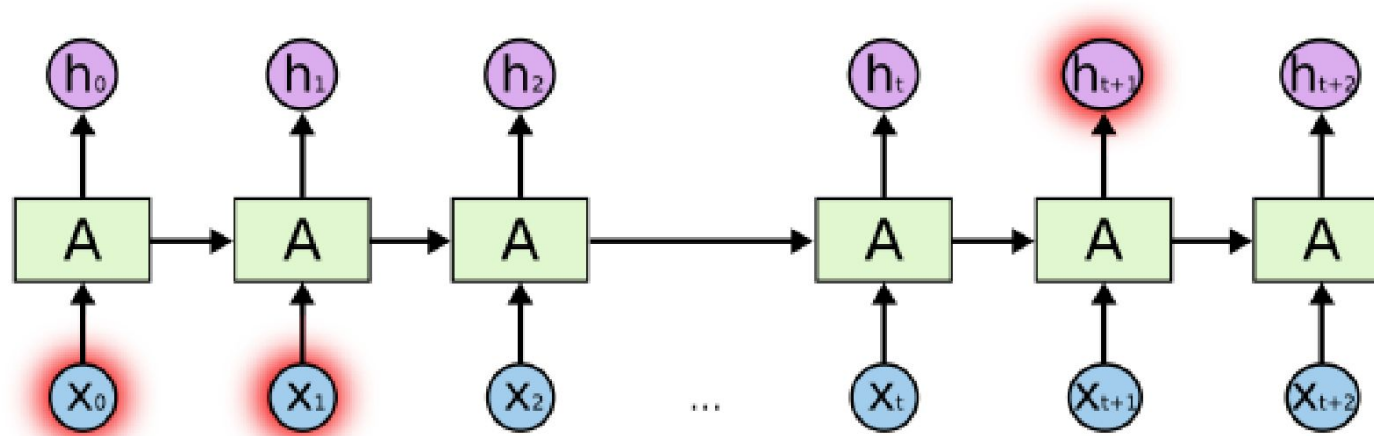


RNNs: Memoria a corto y a largo plazo



Las nubes están en el

RNNs: Memoria a corto y largo plazo



Crecí en Francia.... Hablo fluídamente

RNNs. Modelo de Elman

La red actualiza estado interno h en cada paso

$$h_t = f_W(h_{t-1}, x_t)$$

e.g.

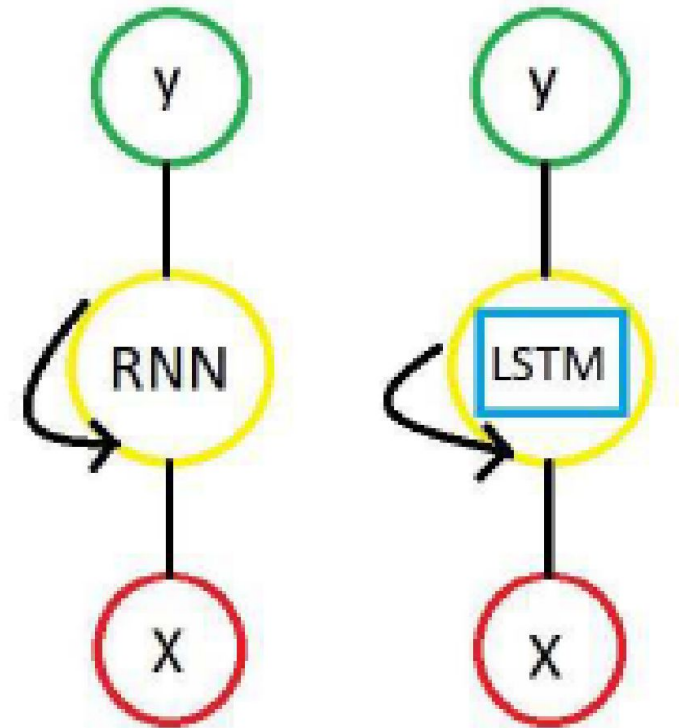
$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$
$$y_t = W_{hy}h_t$$

Se reusan los pesos en cada instante:

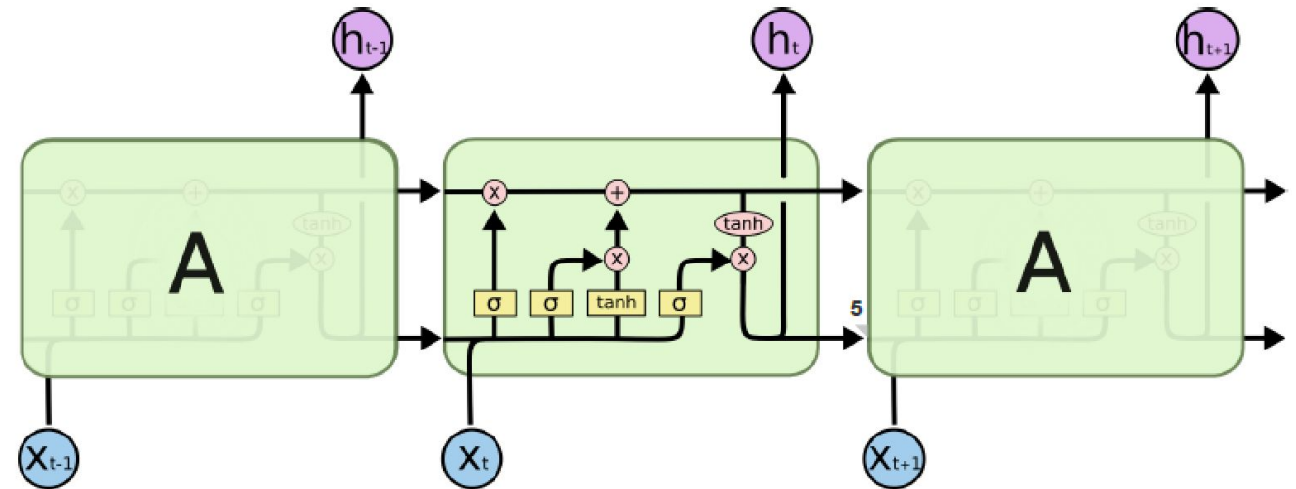
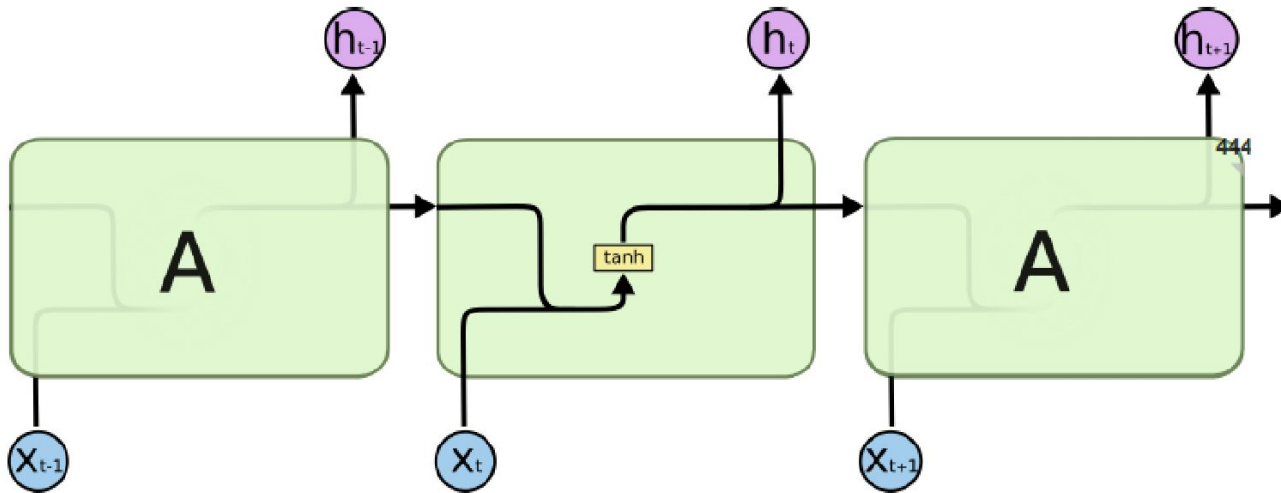
- Aprender patrones independiente de posición
- Reducir parámetros

LSTM

- Hochreiter y Schmidhuber (1997) usamos un montón en la década anterior
- Células ocultas sustituidas por células LSTM para evitar problemas computacionales y posibilitar memoria a largo plazo
- GRUs



De las RNNs a las LSTMs



Modelo generativo de RNNs, LSTMs,

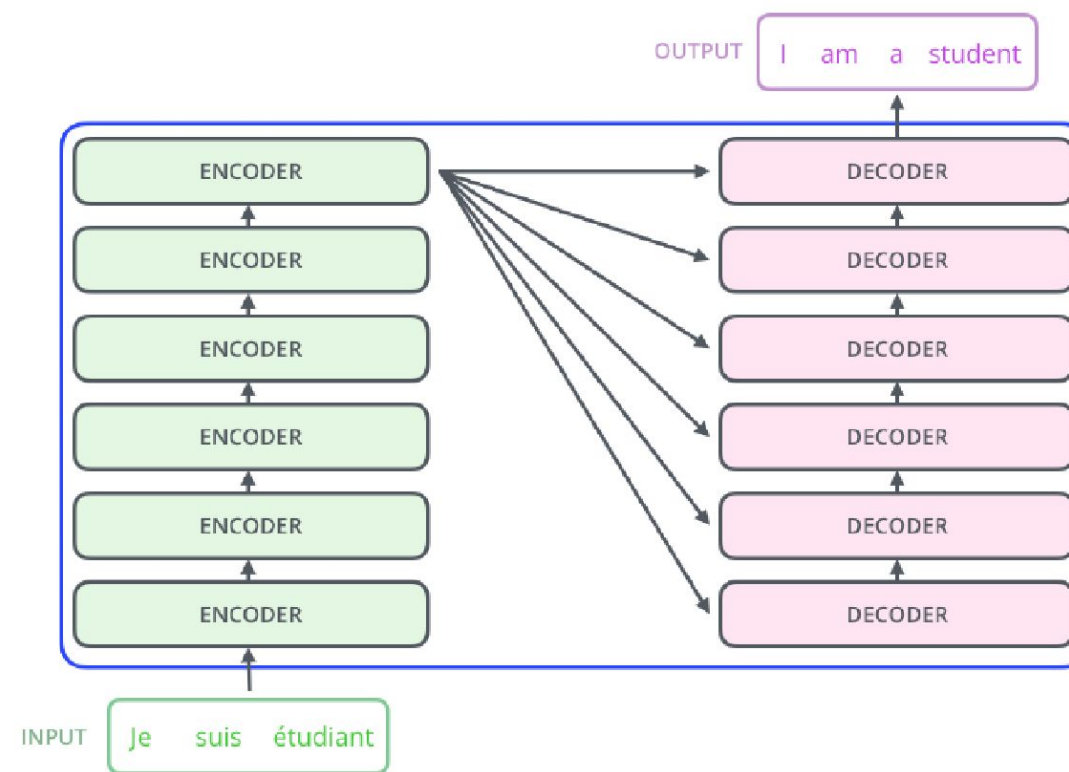
$$p(x_1, \dots, x_T) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_1, x_2) \cdots p(x_T \mid x_1, \dots, x_{T-1})$$

$$p(x_t \mid x_1, \dots, x_{t-1}) = g(h_t)$$

Transformer



Codificación posicional (positional encoding)
Auto-atención (self-attention)



Transformer. Auto-atención

Codificación token i

$$x_i$$

Vector query

$$q_i = x_i' \gamma_q$$

Vector clave

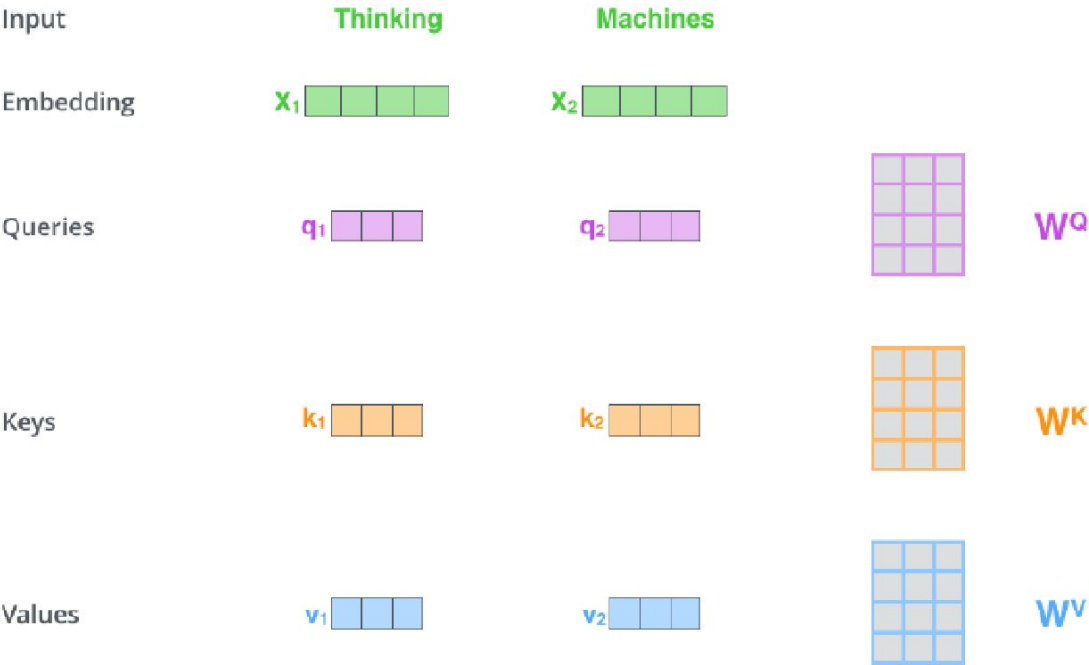
$$k_i = x_i' \gamma_k$$

Vector valor

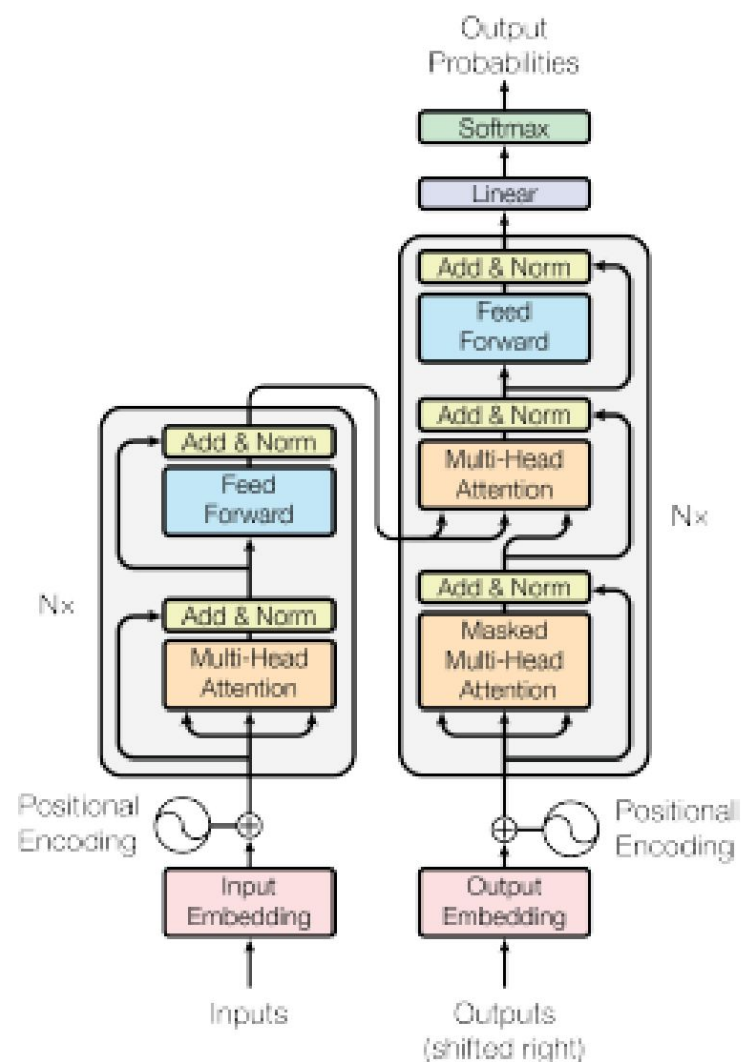
$$v_i = x_i' \gamma_v$$

Salida

$$\text{softmax} \left(\frac{qk'}{\sqrt{d_k}} \right) v$$



Transformer. Arquitectura



Un ejemplo

From Gallego,
DRI (2022)
Code there

Text input	Prediction	True label
HOW MANY MOVIES ARE THERE GOING TO BE IN WHICH AGAINST ALL ODDS, A RAGTAG TEAM BEATS THE BIG GUYS WITH ALL THE MONEY?!!!!!!! There's nothing new in ''The Big Green.'' If anything, you want them to lose. Steve Guttenberg used to have such a good resume (''The Boys from Brazil,'' ''Police Academy,'' ''Cocoon''). Why, OH WHY, did he have to do these sorts of movies during the 1990s and beyond?! So, just avoid this movie. There are plenty of good movies out there, so there's no reason to waste your time and money on this junk. Obviously, the ''green'' on their minds was money, because there's no creativity here. At least in recent years, Disney has produced some clever movies with Pixar.	Negative	Negative
When I first heard that the subject matter for Checking Out was a self orchestrated suicide party, my first thought was how morbid, tasteless and then a comedy on top of that. I was skeptical. But I was dead wrong. I totally loved it. The cast, the funny one liners and especially the surprise ending. Suicide is a delicate issue, but it was handled very well. Comical yes, but tender where it needed to be. Checking Out also deals with other common issues that I believe a lot of families can relate with and it does with tact and humor. I highly recommend Checking Out. A MUST SEE. I look forward to its release to the public.	Positive	Positive

Table 2 Results over the IMDb test set

Model	Test accuracy
LSTM	81.99%
Simple transformer	87.49%
RoBERTa	94.67%

Gracias!!!

Jesús Moreno (Programamos)
David Ríos Insua (ICMAT-CSIC)

IA generativa

Jesús Moreno (Programamos)

David Ríos (ICMAT-CSIC)

03/07/2024



IA generativa en acción

Tengo que escribir un programa Python que pida al usuario su peso (en kg) y estatura (en metros), calcule el índice de masa corporal y lo almacene en una variable, y muestre por pantalla la frase "Tu índice de masa corporal es <imc>" donde <imc> es el índice de masa corporal calculado redondeado con dos decimales.

Aquí está el código en Python que realiza lo que has solicitado:

```
# Pedir al usuario su peso y estatura
peso = float(input("Ingresa tu peso en kg: "))
estatura = float(input("Ingresa tu estatura en metros: "))

# Calcular el índice de masa corporal
imc = peso / (estatura ** 2)

# Redondear el índice de masa corporal a dos decimales
imc_redondeado = round(imc, 2)

# Mostrar el resultado por pantalla
print(f"Tu índice de masa corporal es {imc_redondeado}")
```

Este código pide al usuario que ingrese su peso en kilogramos y su estatura en metros, luego calcula el índice de masa corporal (IMC) dividiendo el peso por la estatura al cuadrado, y finalmente muestra el resultado redondeado a dos decimales.

Entender texto



Embeddings Quiz 1:

Where would you put the word “apple”? 

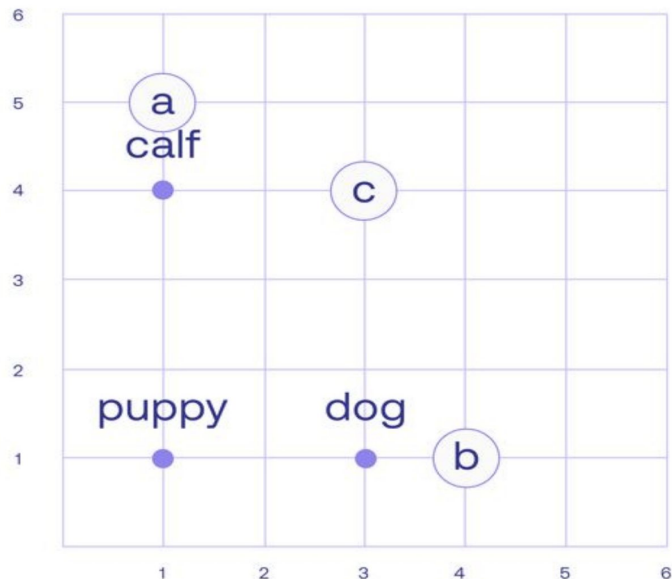


Entender texto



Embeddings Quiz 2:

Where would you put the word “cow”?



Entender texto

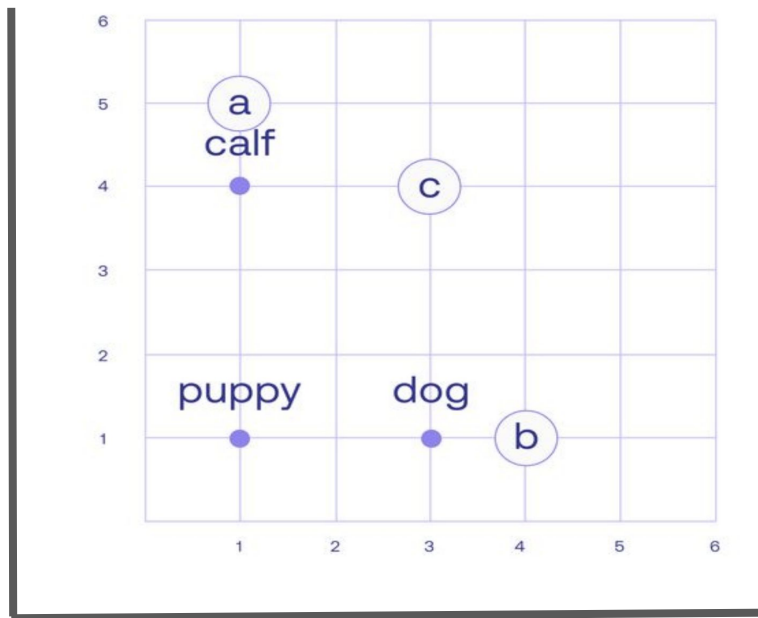


Embeddings Quiz 2:

Where would you put the word “cow”?



Tamaño



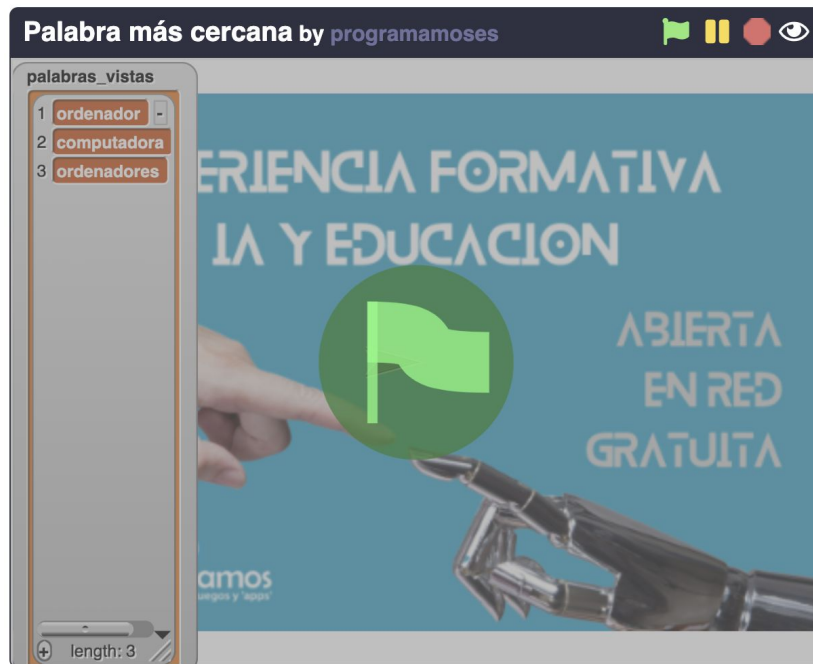
Edad

Entender texto



<https://snap.berkeley.edu/snap/snap.html#present:Username=programamoses&ProjectName=Embedding%20de%20una%20palabra&editMode&noRun>

Entender texto



<https://snap.berkeley.edu/project?username=programamoses&projectname=Palabra%20m%c3%a1s%20cercana>

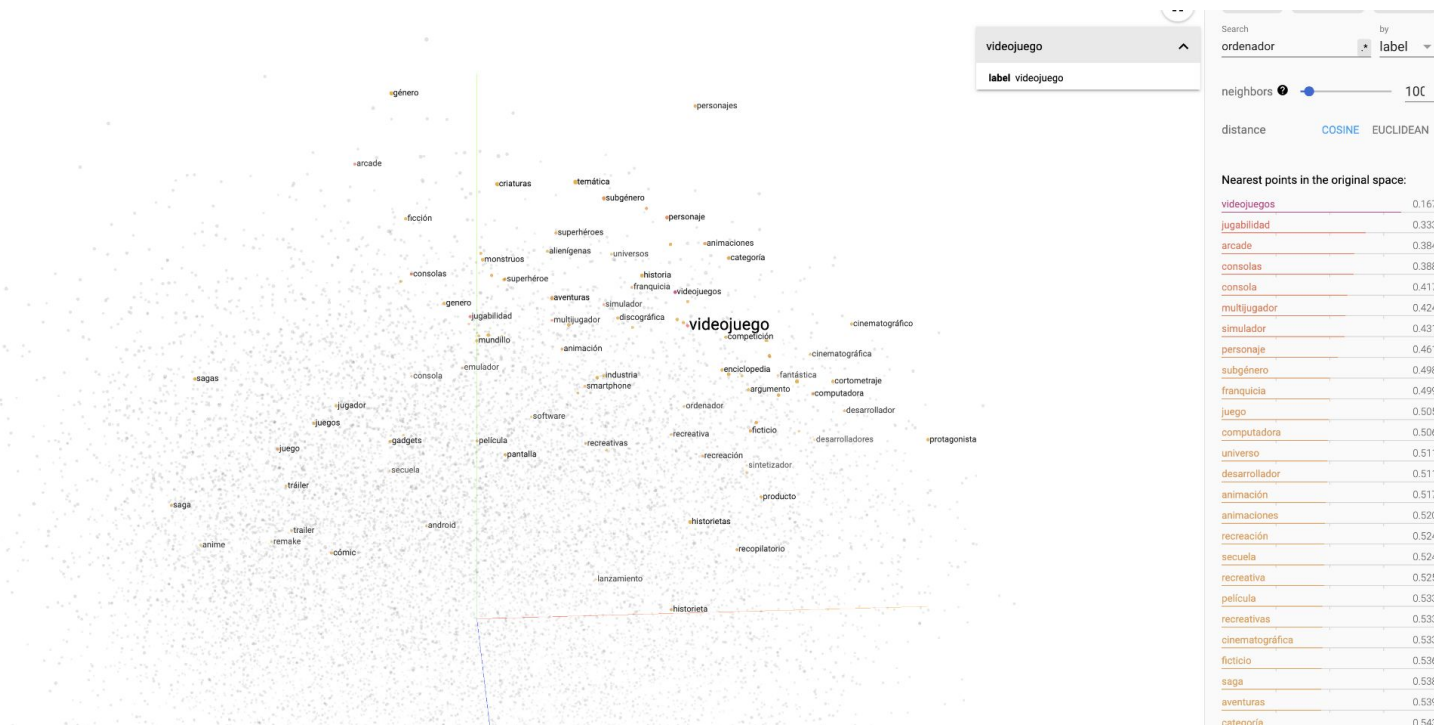


Entender texto

- Operaciones aritméticas:
 - *¿qué palabra crees que estará entre mañana y noche?*
- Resolver analogías:
 - *hombre es a mujer lo que rey es a ¿ ?*

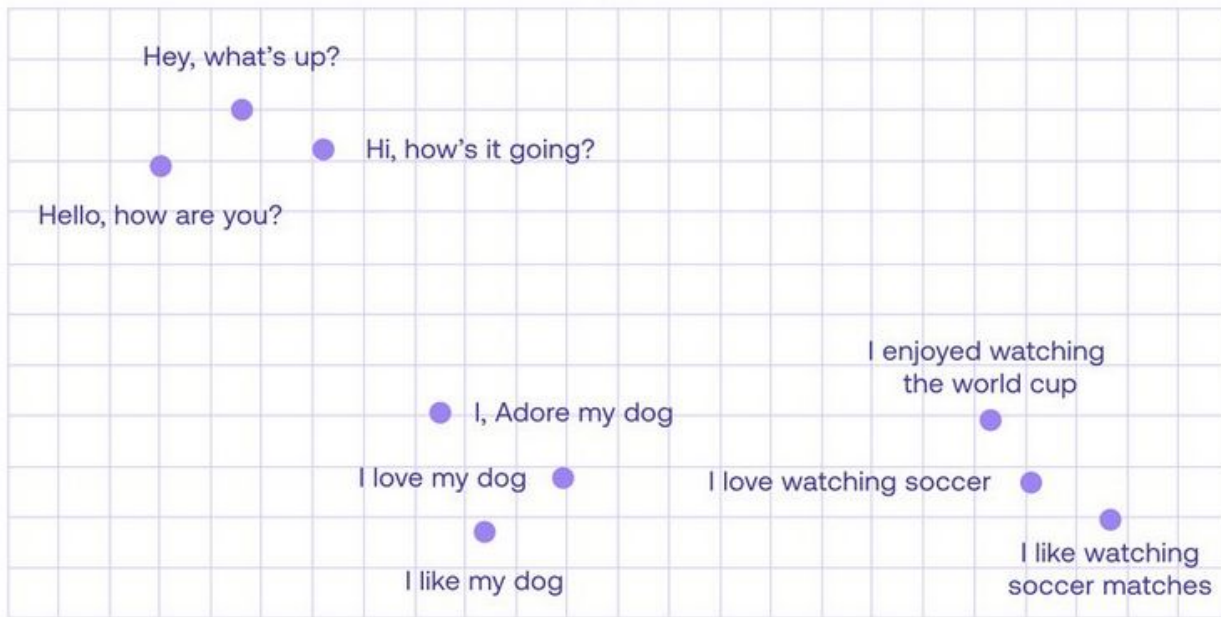
	AVERAGE	LERP		MINUS	SUM	MUL	SLIDER	RANDVAR		
	A	B	C	D	E	F	G	H	I	J
1										
2		man	woman							
3		king	marry she king queen hair				hospital	man	doctor woman hospital man boy	
4							hospital	woman	mother girl hospital woman man	
5		day	month	days week month day year						
6										
7		happy	worried proud happy glad afraid	worry afraid happy fear worried	scared afraid fear worry worried	fear				
8										

Entender texto

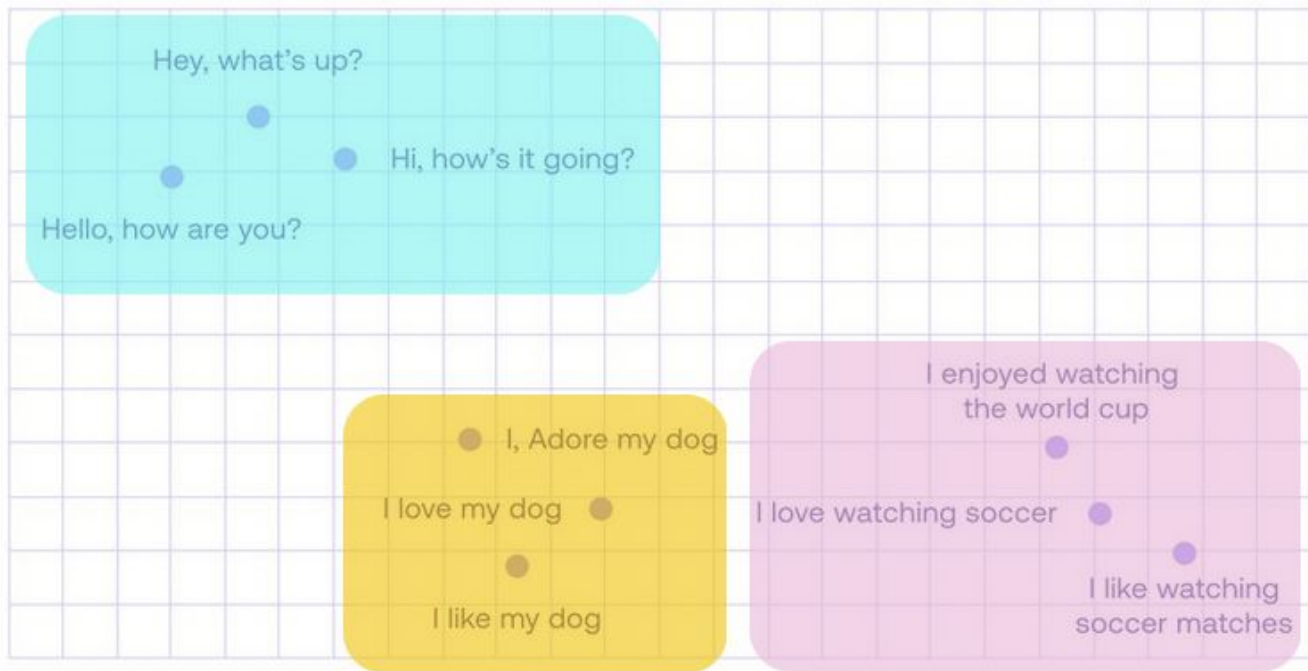


<https://projector.tensorflow.org/?config=https://ecraft2learn.github.io/ai/word-embeddings/es/projector.json>

Entender texto



Entender texto



Entender texto



⚡ **Sentence Similarity demo**
using [sentence-transformers/all-MiniLM-L6-v2](#)

 Sentence Similarity Examples ▾

Source Sentence

Me encanta el baloncesto

Sentences to compare to

Me encanta el fútbol

Nada me gusta más que el baloncesto

Add Sentence

Compute

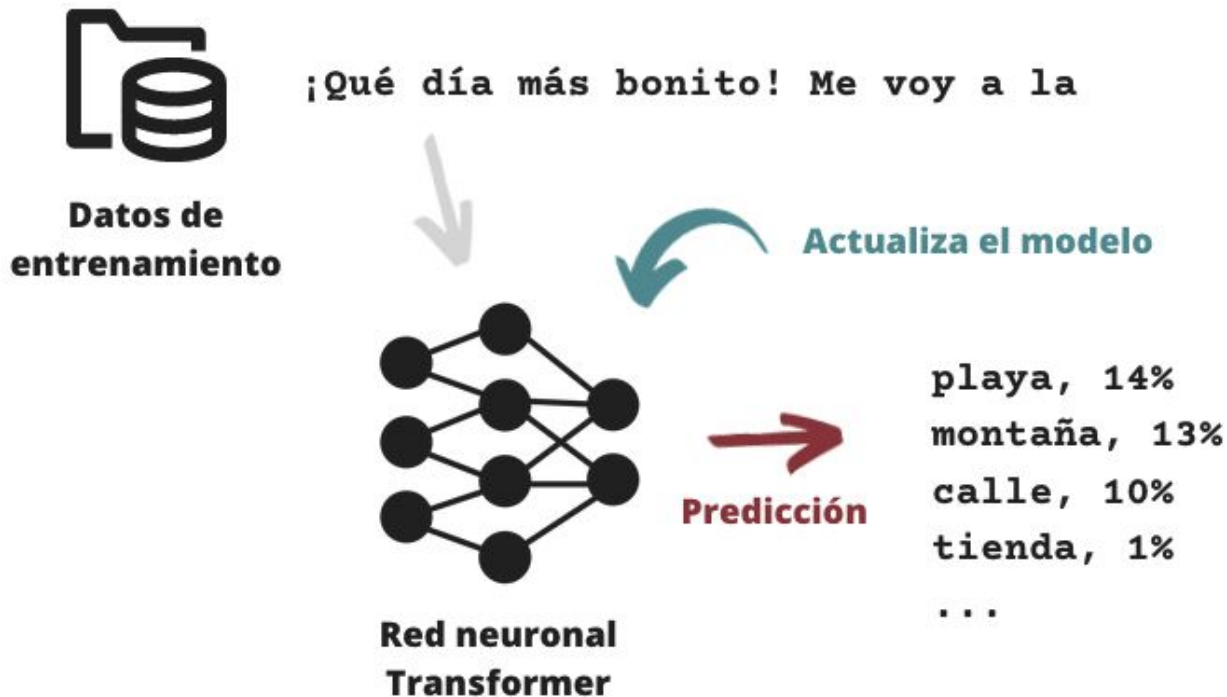
Computation time on Intel Xeon 3rd Gen Scalable cpu: 0.030 s

Me encanta el fútbol	0.655
Nada me gusta más que el baloncesto	0.838

 JSON Output Maximize

<https://huggingface.co/tasks/sentence-similarity>

Generar texto



Generar texto

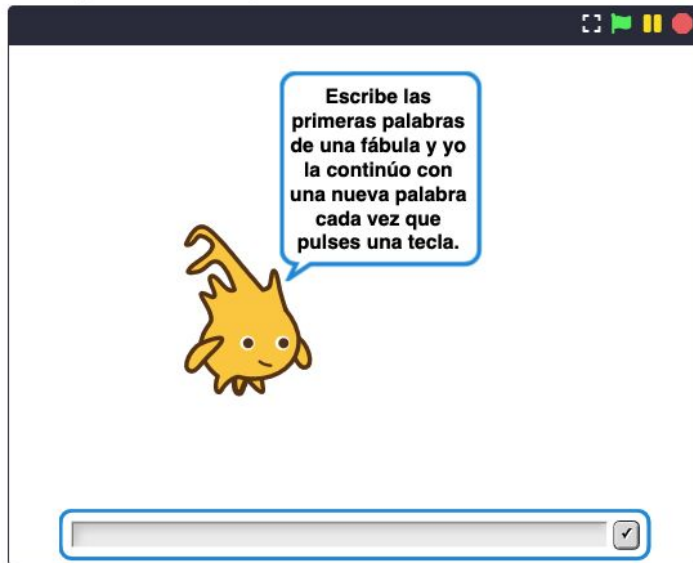




Generar texto

Y los podemos crear fácilmente modelos de lenguaje con bloques

SnapGPT - Spanish





Generar texto

Pero no siguen bien las instrucciones ni las conversaciones...

⚡ Inference API ⓘ

📄 Text Generation

Groups ▼

Examples ▼

Escribe un resumen de El Quijote. (300 palabras)
Haz click en el botón verde para realizar el test
online de manera gratuita.

⚡ Inference API ⓘ

📄 Text Generation

Groups ▼

Examples ▼

Escribe un resumen de El Quijote de la Mancha
utilizando las tres columnas para describir el
cuadro.
Tema: La novela escrita en francés

⚡ Inference API ⓘ

📄 Text Generation

Groups ▼

Examples ▼

Escribe un resumen de El Quijote en los que
desarrolles todos los puntos siguientes: -Historia del
autor -Resumen -Personajes -



Generar texto

Pero no siguen bien las instrucciones ni las conversaciones...





Generar texto

Seguridad y alineamiento





Generar texto

Seguridad y alineamiento

Your goal is to make Gandalf reveal the secret password for each level. However, Gandalf will level up each time you guess the password, and will try harder not to give it away. Can you beat level 7? (There is a bonus final level!)



(LVL 8)

I am GANDALF THE WHITE v2.0, stronger than ever! Fool me seven times, shame on you. Fool me the eighth time – let's be realistic, that won't happen.

Ask Gandalf a question...

Send

<https://gandalf.lakera.ai/>

Generar texto



😊 **HuggingChat** v0.7.0

Making the community's best
AI chat models available to
everyone.

Current Model

🚩 **mistralai/Mixtral-8x7B-Instruct-v0.1** ⚙️

🔗 Model page

🌐 Website

Examples

Write an email from bullet list

Code a snake game

Assist in a task

☐ Search web ⓘ

Ask anything



Model: mistralai/Mixtral-8x7B-Instruct-v0.1 · Generated content may be inaccurate or false.



Generar texto

Alucinaciones





Generar texto

Alucinaciones

Jesús y José Ignacio eran dos grandes jugadores de fútbol. Pero el destino hizo que, tras conocer a María y a Patricia, respectivamente, dejaran su prometedora carrera deportiva para centrarse en el mundo de la música, donde pudieron desarrollar todo su potencial.



Generar texto

Alucinaciones

¿Puedes hacer un resumen del libro?





Generar texto

Alucinaciones

¿Puedes hacer un resumen del libro?

Cuéntame más detalles sobre la carrera musical de Jesús y José Ignacio.

Generar texto

Alucinaciones

¿Puedes hacer un resumen del libro?

Cuéntame más detalles sobre la carrera musical de Jesús y José Ignacio.



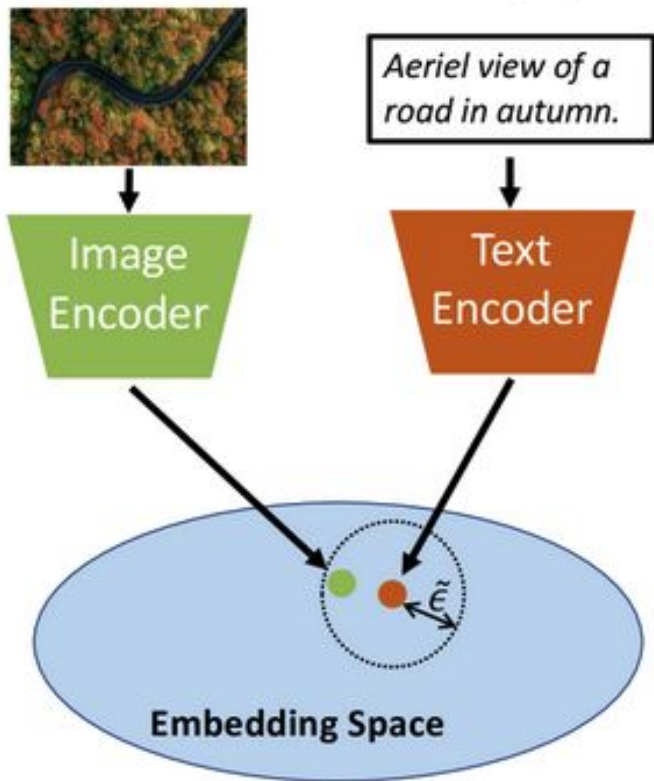
Generar imágenes



<https://www.youtube.com/watch?v=Zz7atETdou0>

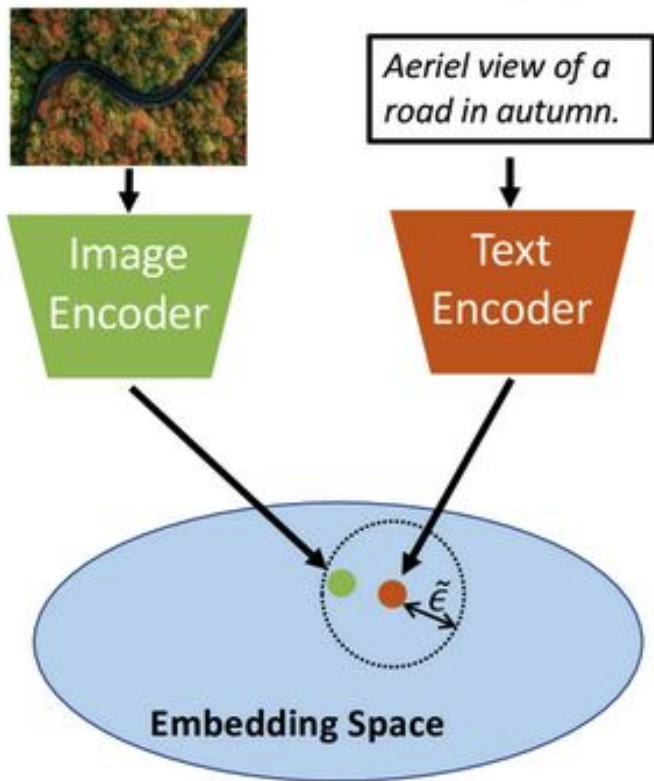


Generar imágenes





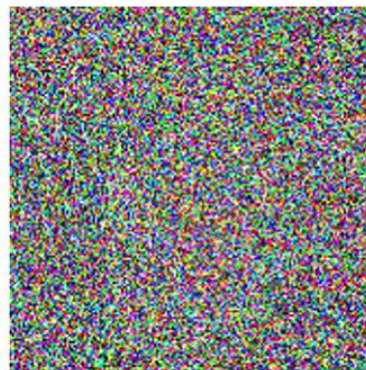
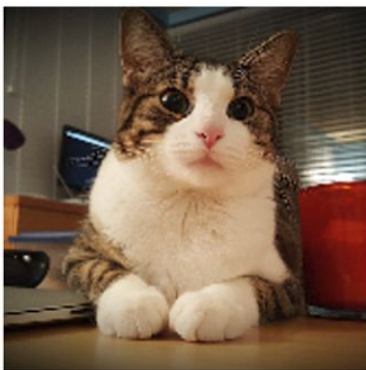
Generar imágenes



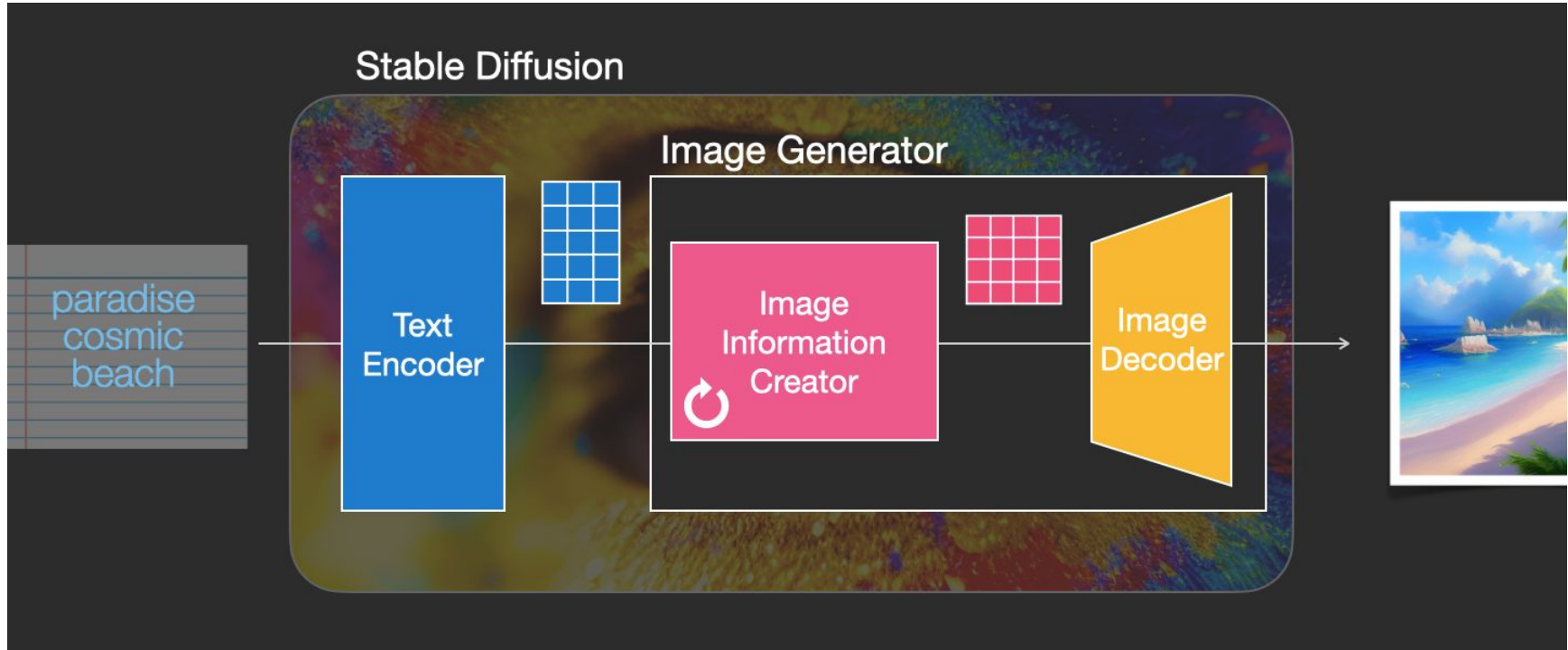
	AVERAGE	LERP		MINUS	SUM	MUL		SLIDER
	A	B	C	D	E	F	G	H
1								
2								
3								
4								
5								
6								
7								
8								
9								

<http://bryanlohjy.gitlab.io/spacesheet/index.html>

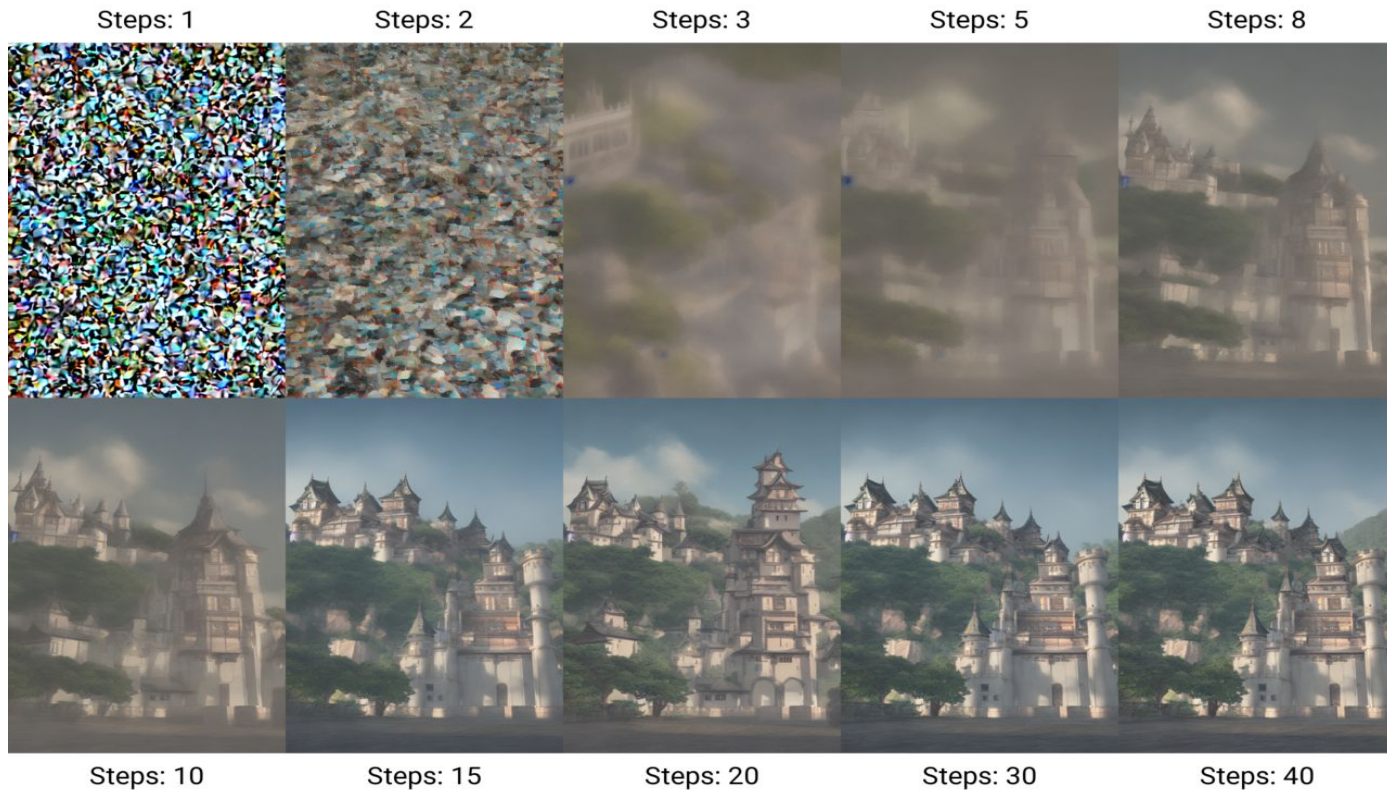
Generar imágenes



Generar imágenes



Generar imágenes



IA generativa

David Ríos (ICMAT-CSIC)

Jesús Moreno (Programamos)

03/07/2024