

Deep Learning: Natural Language Processing (NLP), Modelos Generativos y Transformers

Inteligencia Artificial e Ingeniería del Conocimiento

Constantino Antonio García Martínez

Universidad San Pablo CEU

arquitectura de

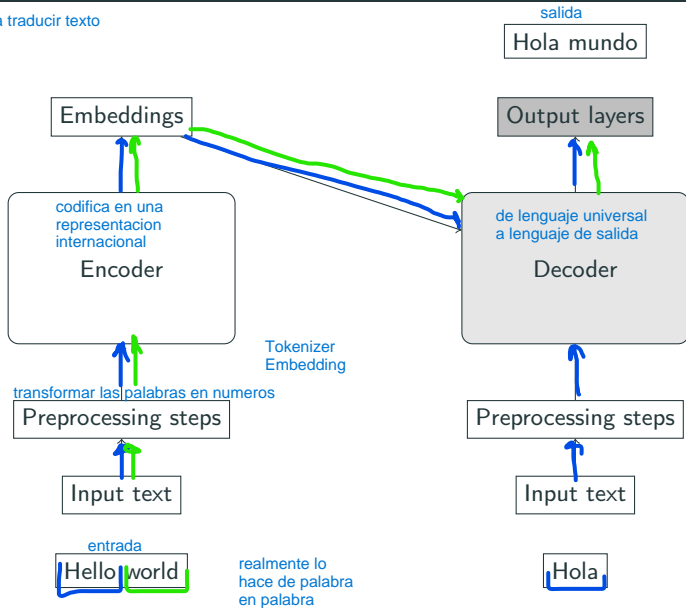
Transformers y NLP

Transformers: modelos estado del arte en NLP

- Introducidos en "Attention is all you need" (Vaswani et al., 2017).
- Inicialmente diseñados para traducción neural, pero desde entonces se han aplicado a muchas tareas de Natural Language Processing (NLP).
- La arquitectura tiene un encoder y un decoder.
- Ambos se basan en la idea del mecanismo de self-attention.

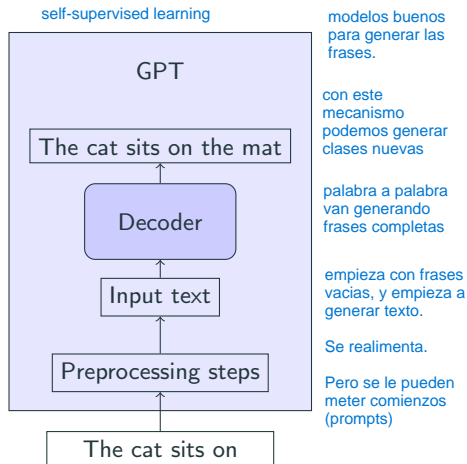
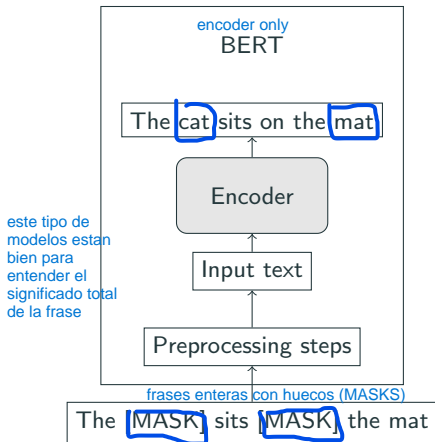
Transformers: modelos estado del arte en NLP

para traducir texto



Modelos Encoder-only Vs. Decoder-only

- Los modelos modernos de NLP suelen basarse en **modelos encoder-only** o **decoder-only**.
- Los modelos modernos se entrenan con conjuntos de datos masivos de lenguaje, generalmente de internet. Por ejemplo, el dataset Common Crawl ≈ 1 petabyte



¿Cuándo usar Encoder-only Vs. Decoder-only?

Vienen bien para comprensión global.

Para tareas tipo clasificación vienen bien porque entienden bien el contexto de una frase

- **Encoder-only (como BERT):**

- Mejor para tareas de comprensión: clasificación, análisis de sentimiento, reconocimiento de entidades
- Comprensión de contexto bidireccional
- Cuando necesitas analizar o extraer información del texto

- **Decoder-only (como GPT):**

- Mejor para tareas generativas: generación de texto, completado, traducción
- Generación de texto auto-regresiva
- Cuando necesitas generar texto coherente o continuar secuencias

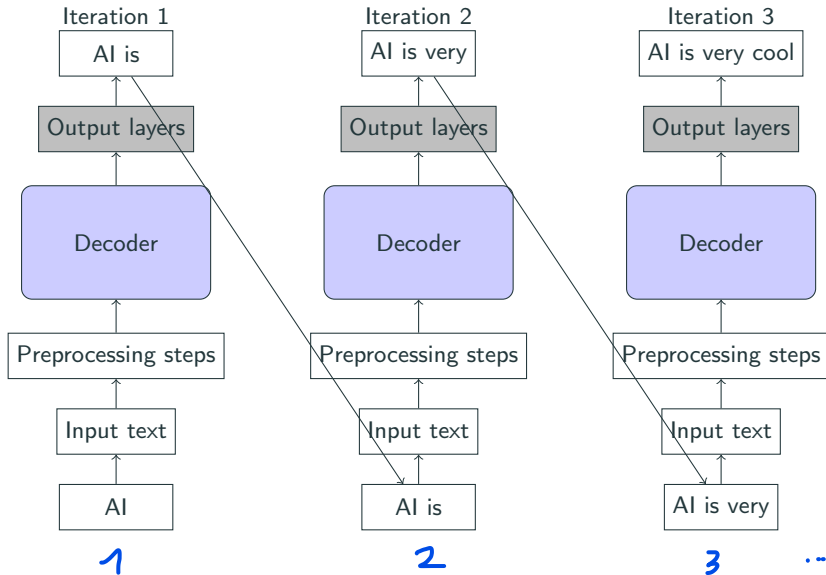
Para tareas generativas, por eso los chats usan GPTs

GPT lo hace paso a paso: autoregresivo

Modelos GPT

- GPT significa Generative Pre-trained Transformer.

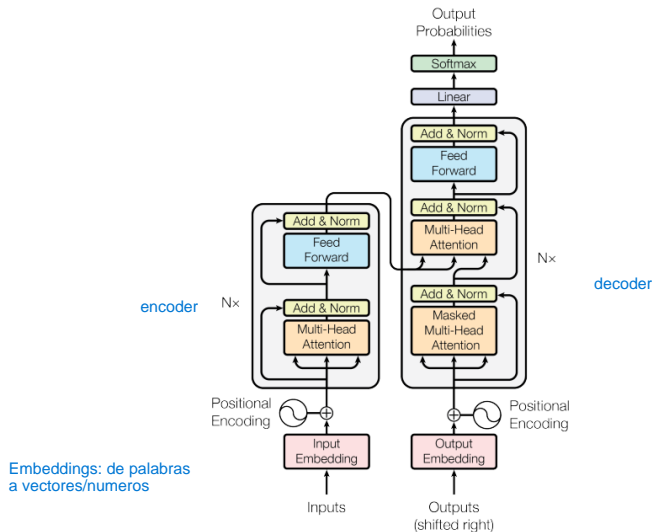
Paso a paso completando hasta que decide cuando parar



Transformers y NLP

Detalles de la Arquitectura Transformer

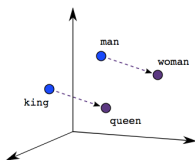
La Arquitectura Transformer



Del artículo "Attention is all you need" de Vaswani et al.

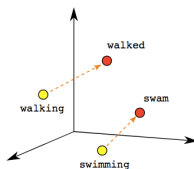
Detalle: **Embeddings**

- Los embeddings transforman “palabras” en vectores comprensibles por una red neuronal.
- Estas transformaciones se aprenden al entrenar el Transformer de principio a fin.
- ¡Los **Transformers** aprenden a **asociar direcciones del espacio de embeddings** con significado semántico!



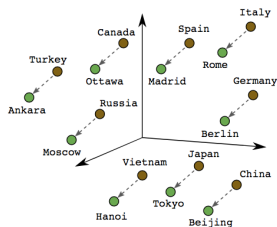
asocia una direccion
con el sexo

Male-Female



con conjugaciones

Verb Tense

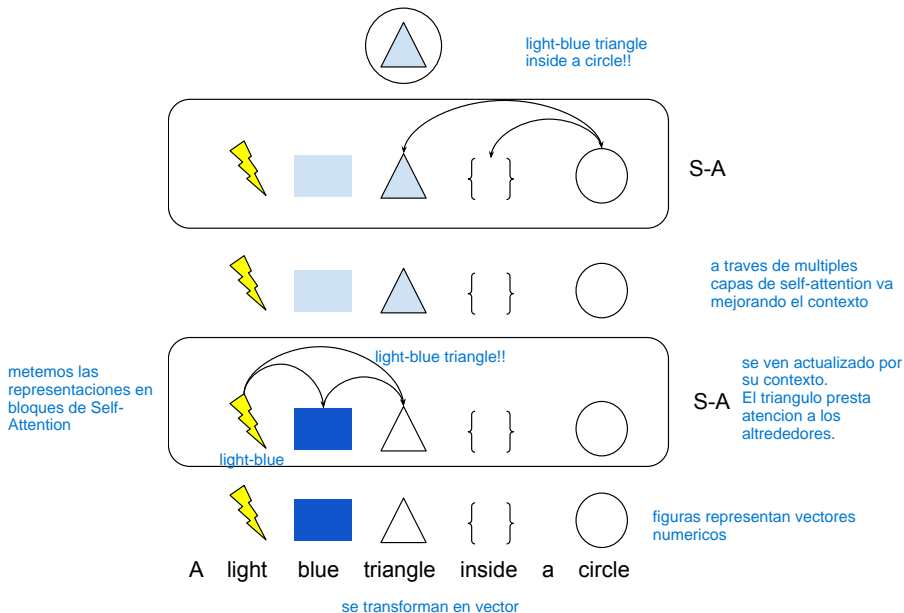


con relaciones
capital-pais

Country-Capital

pone las palabras en un espacio vectorial, y asocia direcciones con atributos.

Detalle: Mecanismo de Self-Attention



Retrieval Augmented Generation (RAG)

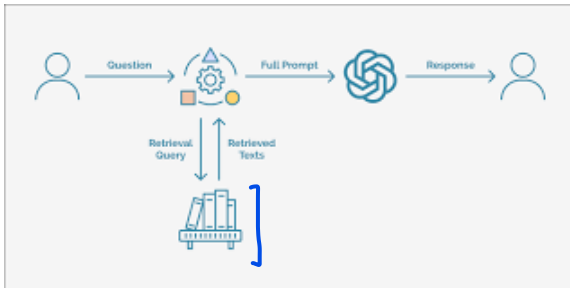
Retrieval Augmented Generation (RAG)

- Combina modelos de lenguaje con recuperación de conocimiento externo
- Beneficios:
 - Mayor precisión y factualidad
 - Reducción de alucinaciones
 - Acceso a información actualizada
- Componentes clave:
 - Base de conocimiento/Base de datos vectorial
 - Recuperador (encuentra documentos relevantes)
 - Generador (crea respuestas usando el contexto recuperado)

no puedo hacer fine-tuning con chatGPT con datos clasificados

puedo hacer una red local con datos internos clasificados.

con un modelo de lenguaje natural que genera clases con esos datos



¿Cómo se entrena ChatGPT?

¿Cómo se entrena ChatGPT?

Saca datos de sitios como reddit lol
Cuidao

Reinformer Learning: Metemos a gente en proceso de
realimentación

datos de preguntas y respuestas

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-3.5
with supervised
learning.

Modelo que genera res a
preguntas

personas
generando
respuestas

esto hace un
fine-tuning

Step 2

Collect comparison data and
train a reward model.

A prompt and
several model
outputs are
sampled.

pedimos a gente
que elija su fav

A labeler ranks the
outputs from best
to worst.

esto entrena un
neural net

This data is used
to train our
reward model.

aprende a imitar a humanos

Hay sesgo de respuestas
peligrosas.
Salen cosas 'educadas', como
"crucial"

Cuando se equivoca, hostia.
Cuando acaban, premio

Step 3

Optimize a policy against the
reward model using the PPO
reinforcement learning algorithm.

A new prompt is
sampled from
the dataset.

The PPO model is
initialized from the
supervised policy.

The policy generates
an output.

The reward model
calculates a reward
for the output.

The reward is used
to update the
policy using PPO.

simulaciones que van repitiendo hasta que
da respuestas que le gusten a humanos

Existe Jail-Breaking