Introducción al aprendizaje de máquina

Aprendizaje de Máquina Aplicado

César Leandro Higuita Pérez
clhiguitap@eafit.edu.co

2023

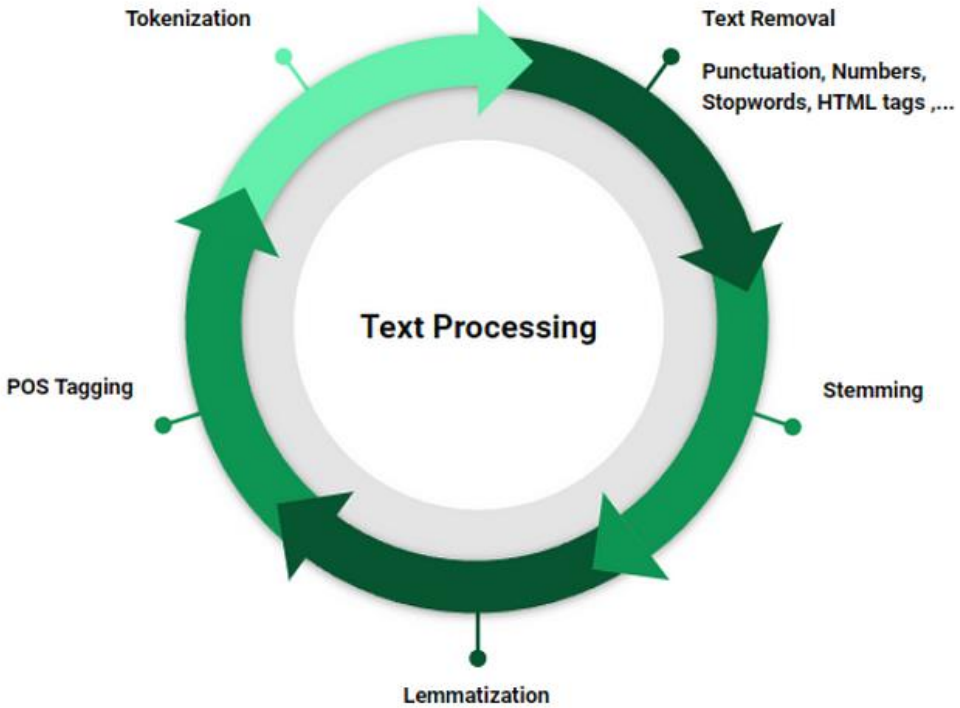
David Felipe Luna
dflunan@eafit.edu.co

UNIVERSIDAD EAFIT

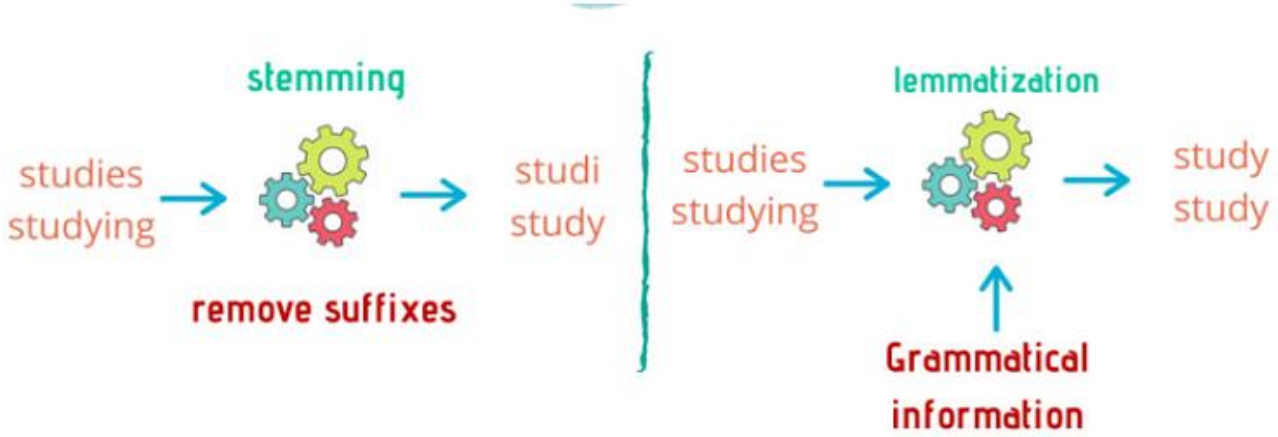
Agenda

- Introducción a NLP
- Aplicaciones de NLP
- NLP con transformers
- Ejemplo práctico de Q&A usando Hugging Face.

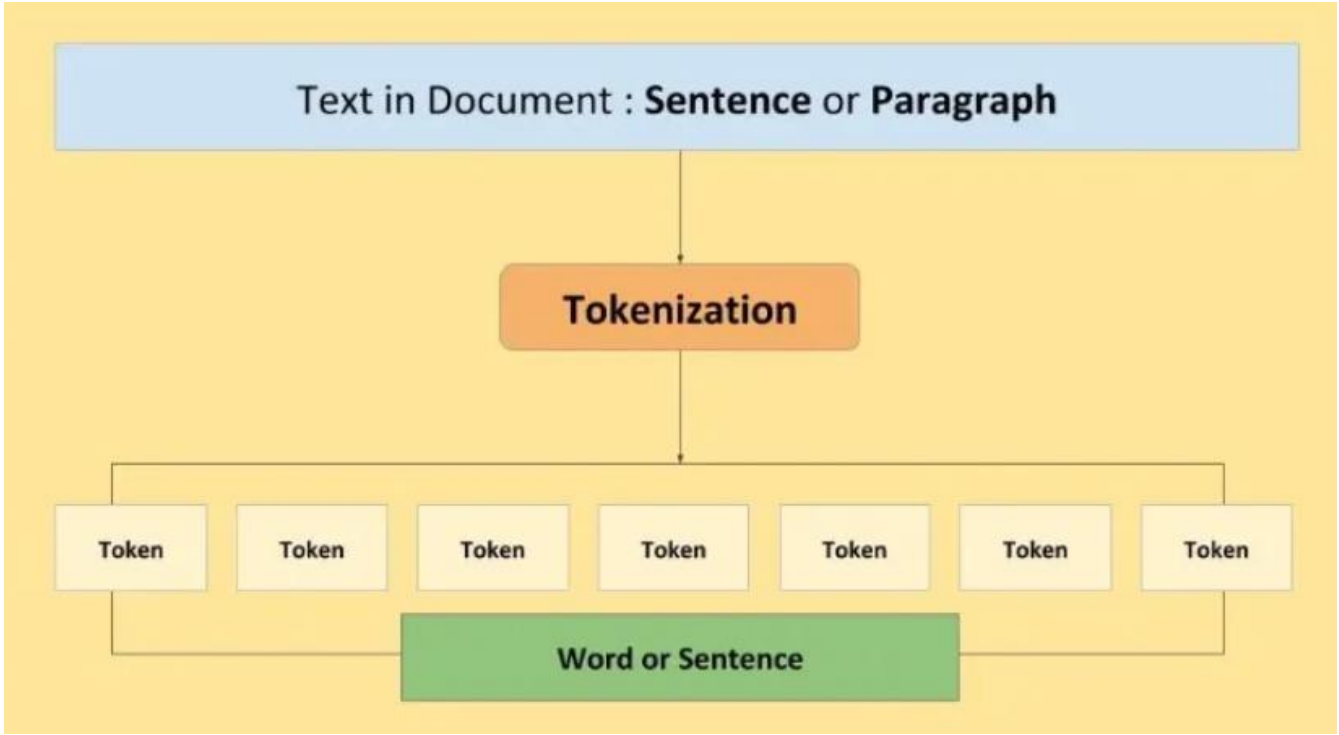
Introducción a NLP



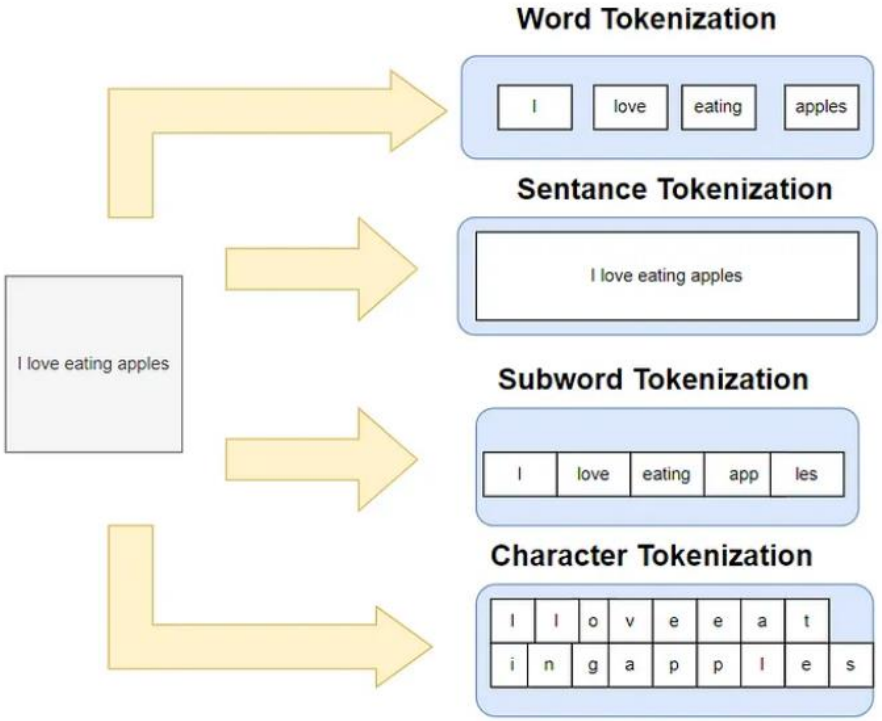
Stemming and Lemmatization



Tokenization

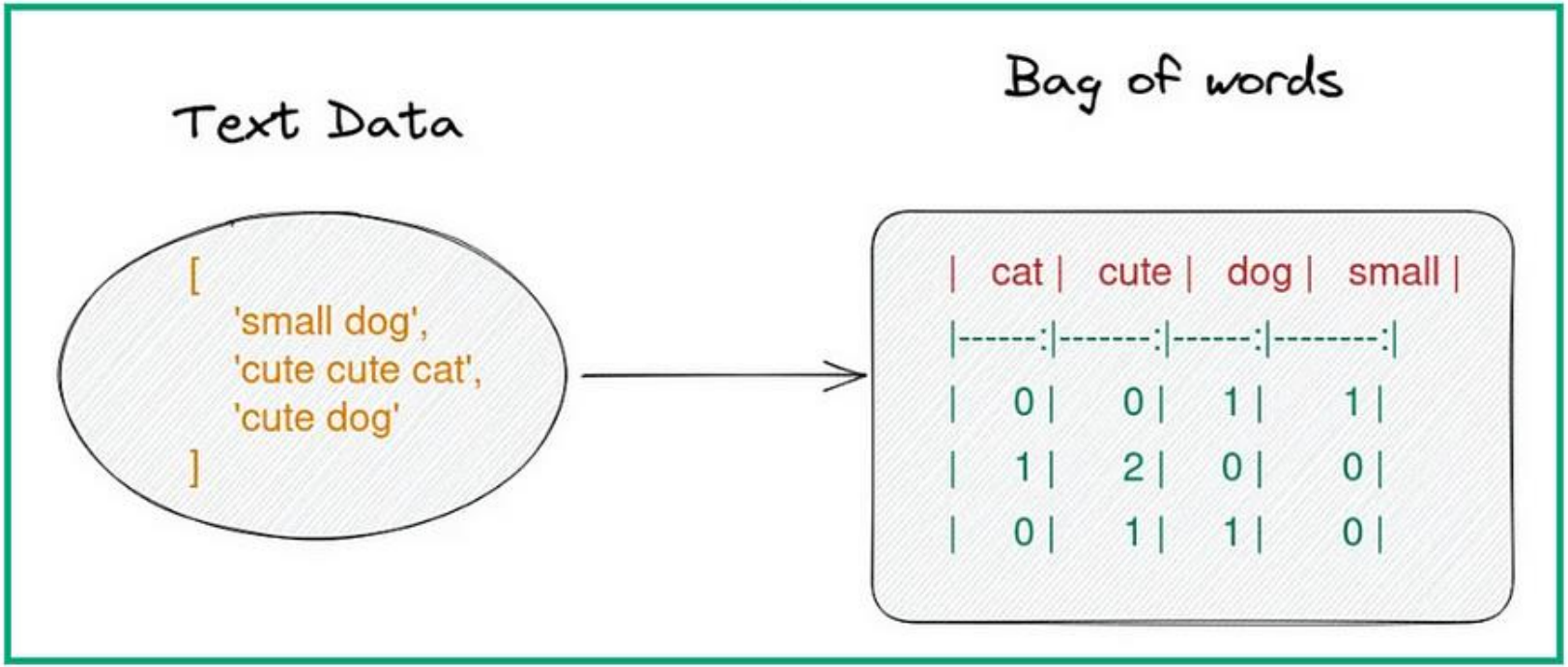


Tokenization

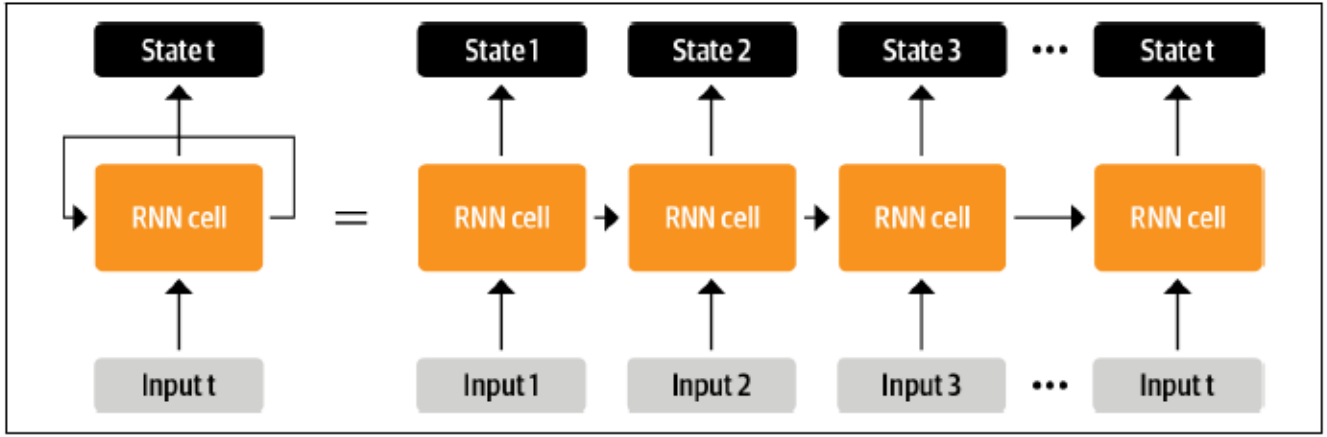


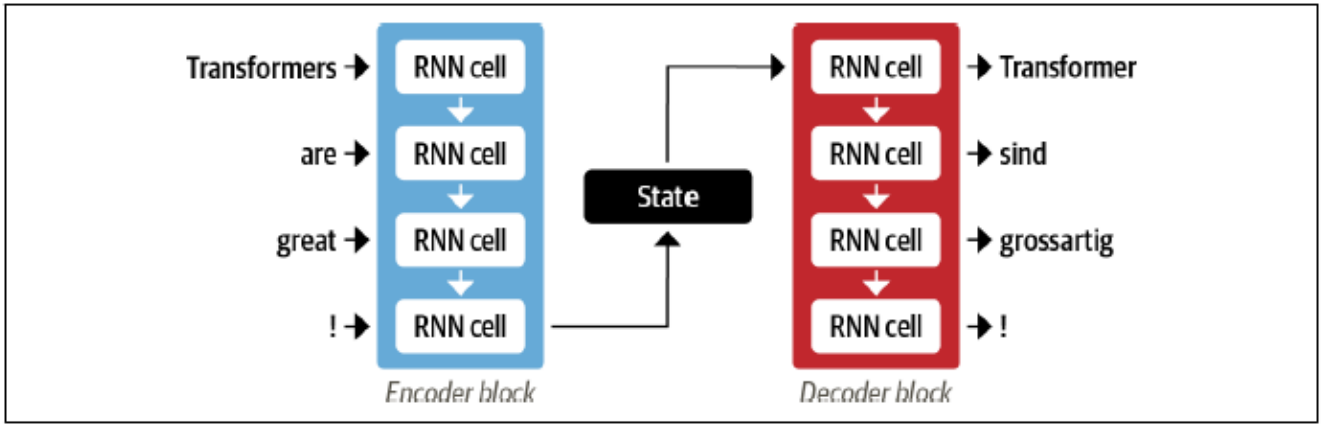
Tokenization techniques

Bag of words


Text Data
Bag of words
[
'small dog',
'cute cute cat',
'cute dog'
]
cat	cute	dog	small
0	0	1	1
1	2	0	0
0	1	1	0

RNN


State t
State 1
State 2
State 3
State t
RNN cell
RNN cell
RNN cell
RNN cell
RNN cell
Input t
Input 1
Input 2
Input 3
Input t


Transformers
are
great
!
RNN cell
RNN cell
RNN cell
RNN cell
State
RNN cell
RNN cell
RNN cell
RNN cell
Transformer
sind
grossartig
!
Encoder block
Decoder block

Hello Transformers

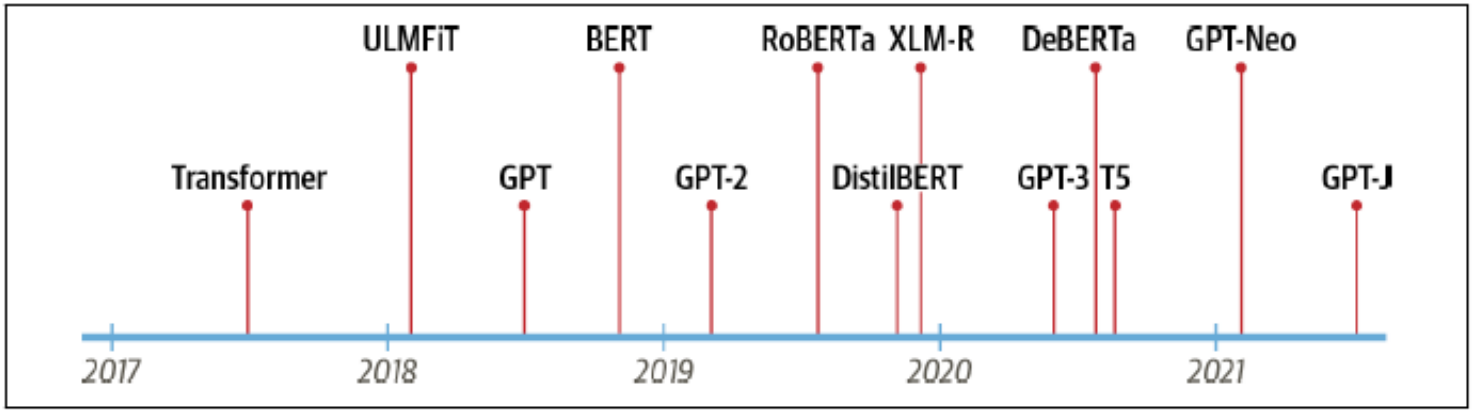
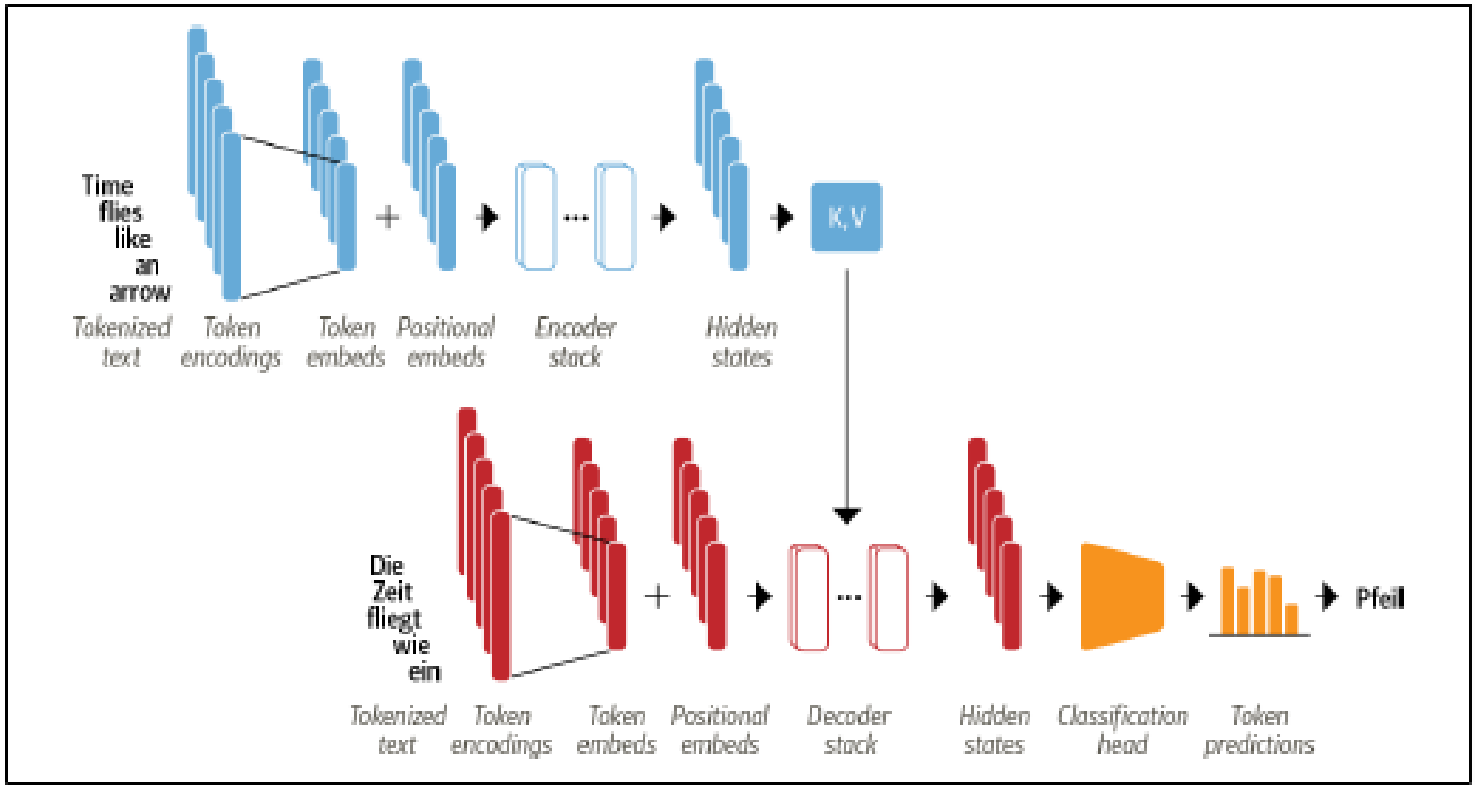
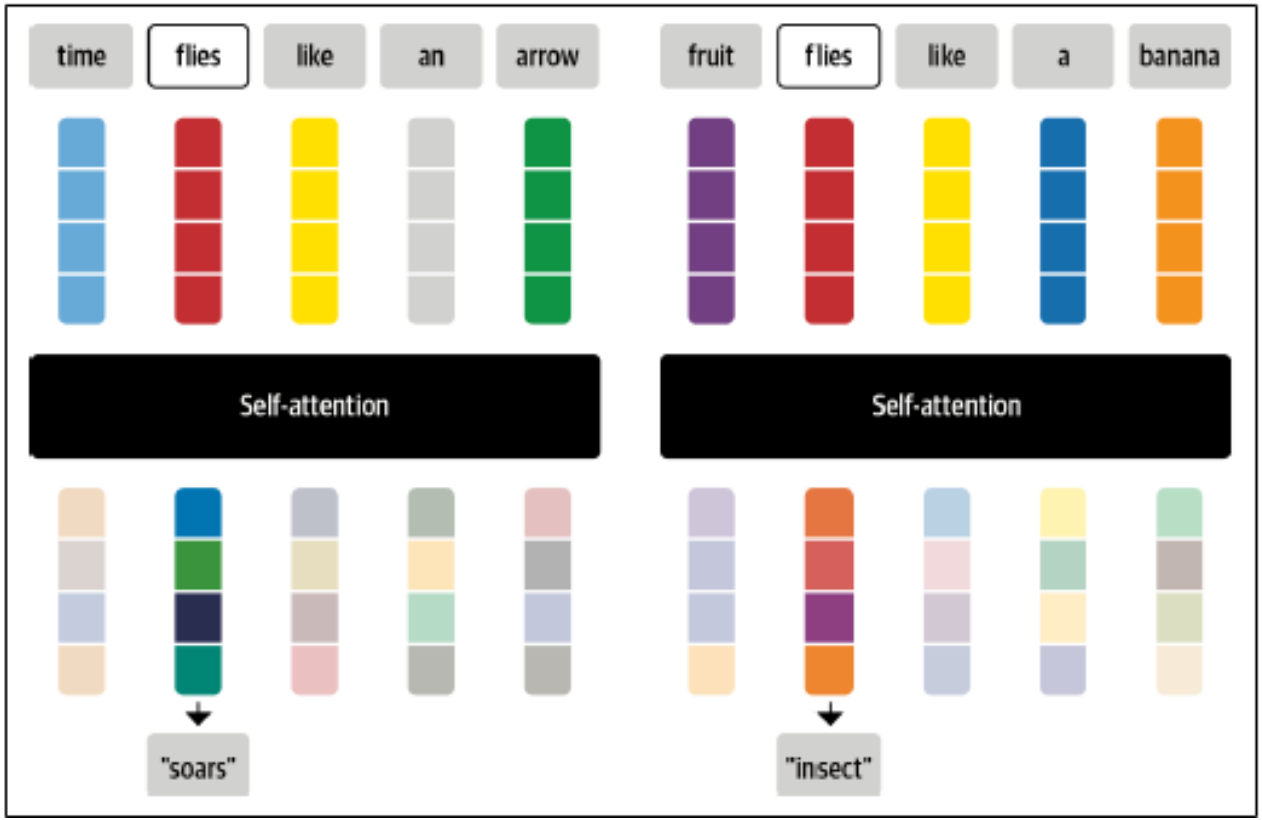


Figure 1-1. The transformers timeline

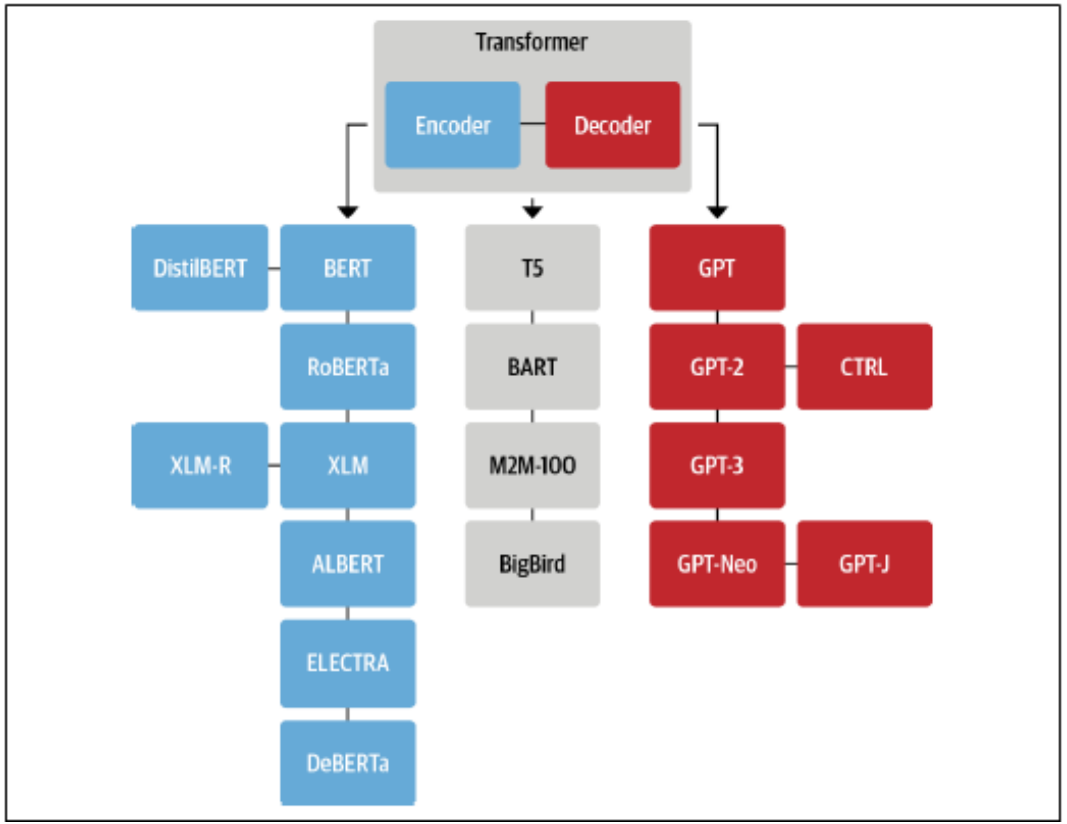
Transformer Arquitecture



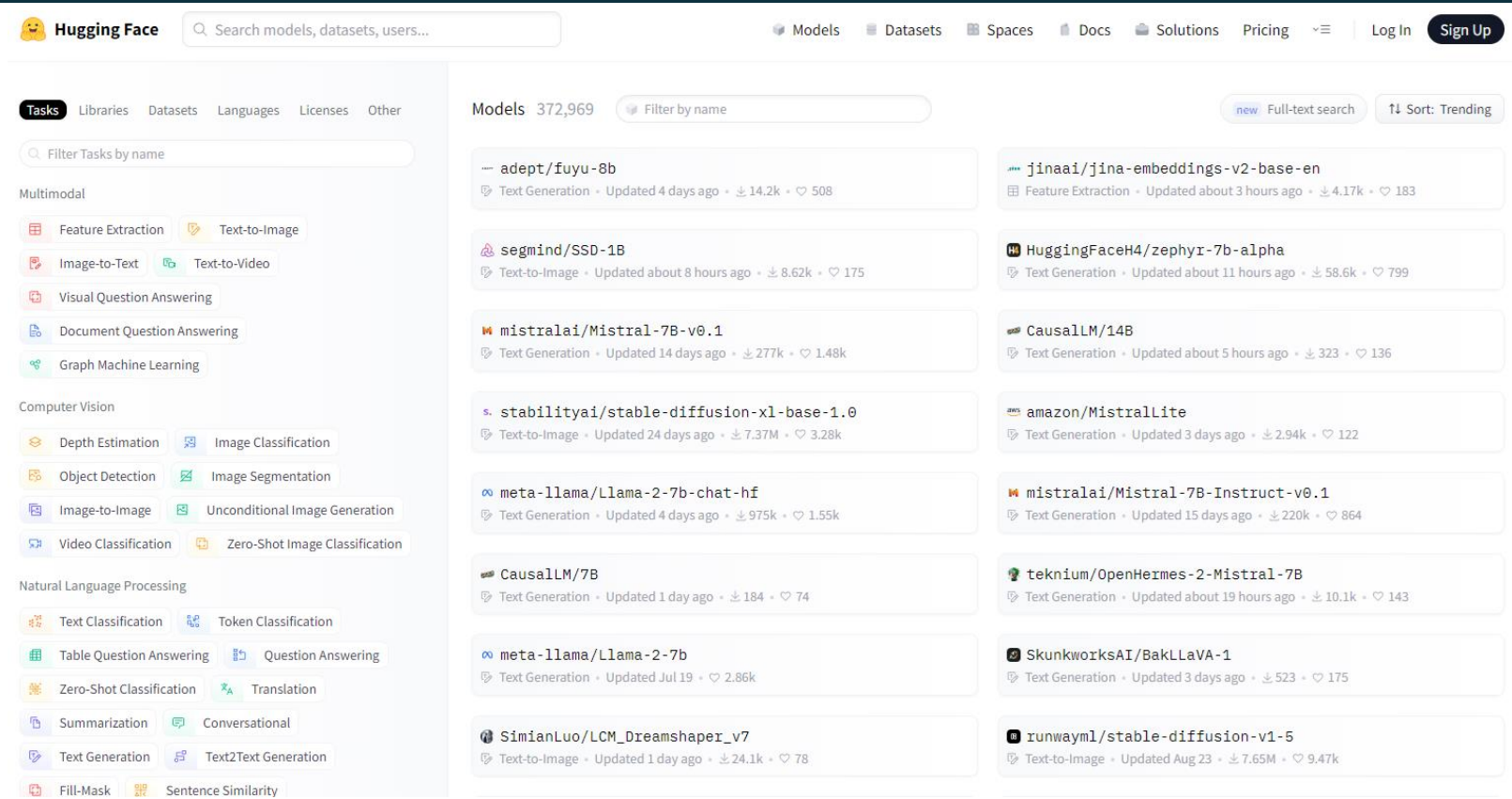
Self-Attention



Algunas Arquitecturas de Transformers



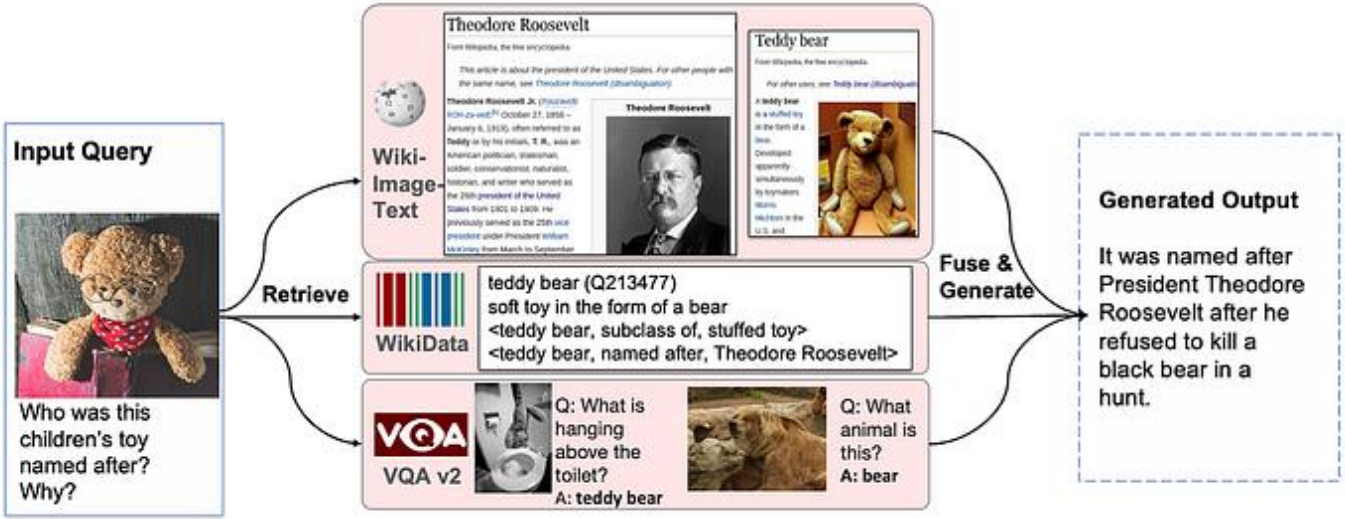
Hugging Face Hub



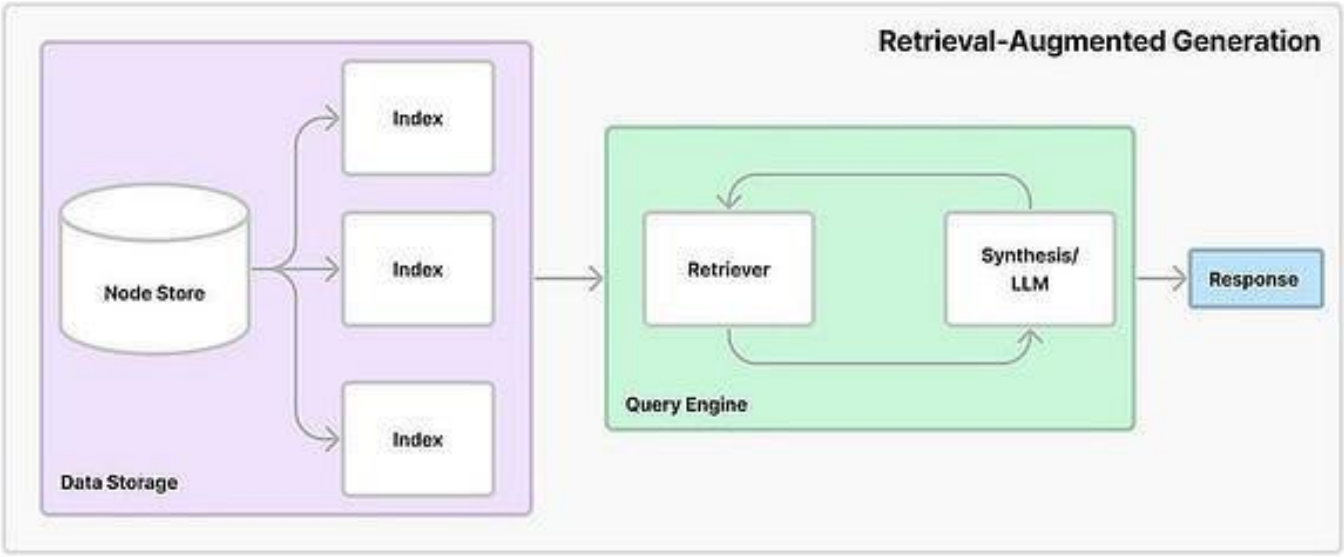
RAG

Retrieval-Augmented Generation (RAG)

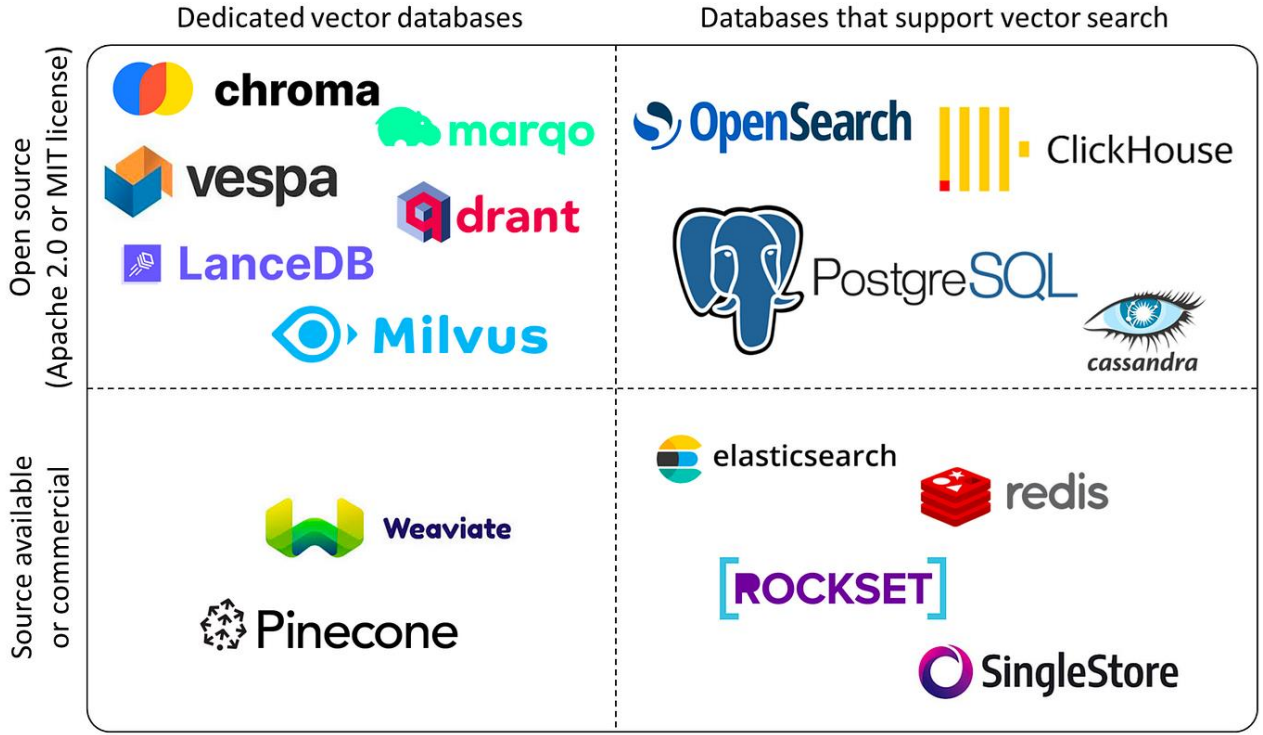
Retrieval-augmented generation (RAG) is a natural language processing (NLP) approach that combines the benefits of both retrieval-based and generation-based methods for content generation tasks. It aims to improve the quality and controllability of the generation tasks by leveraging a pre-trained language model in conjunction with a retrieval mechanism.



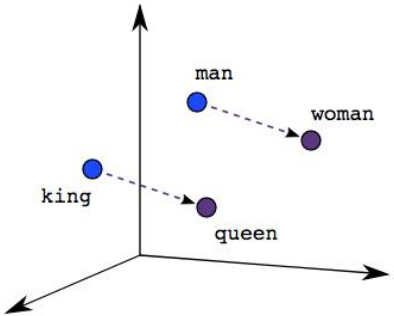
RAG: Componentes



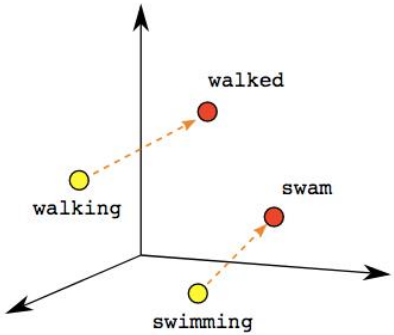
RAG: Database



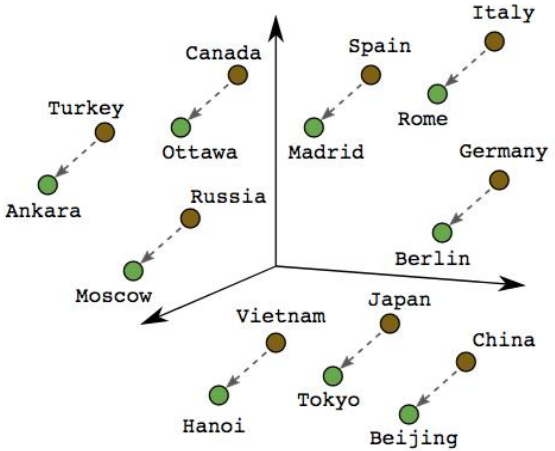
RAG: Embeddings



Male-Female



Verb Tense



Country-Capital

RAG: LLM

Hugging Face

The task illustrated in this tutorial is supported by the following model architectures:

ALBERT, BART, BERT, BigBird, BigBird-Pegasus, BLOOM, CamemBERT, CANINE, ConvBERT, Data2VecText, DeBERTa, DeBERTa-v2, DistilBERT, ELECTRA, ERNIE, ErnieM, Falcon, FlauBERT, FNet, Funnel Transformer, OpenAI GPT-2, GPT Neo, GPT NeoX, GPT-J, I-BERT, LayoutLMv2, LayoutLMv3, LED, LiLT, Longformer, LUKE, LXMERT, MarkupLM, mBART, MEGA, Megatron-BERT, MobileBERT, MPNet, MPT, MRA, MT5, MVP, Nezha, Nyströmformer, OPT, QDQBert, Reformer, RemBERT, RoBERTa, RoBERTa-PreLayerNorm, RoCBert, RoFormer, Splinter, SqueezeBERT, T5, UMT5, XLM, XLM-RoBERTa, XLM-RoBERTa-XL, XLNet, X-MOD, YOSO