

Semántica del lenguaje natural en informática

Cómo los modelos de lenguaje representan el texto



Resumen

1. Introducción
2. Cómo funciona
3. Tareas populares
4. Conclusión



1. Introducción

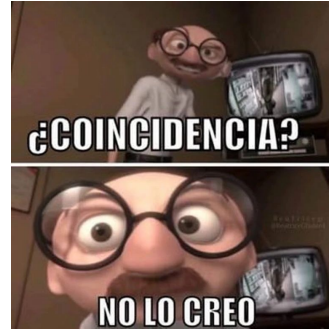
Traducción automática entre diferentes idiomas (~1950 / guerra fría)

Primer computadoras (~1950)

Estructuración de datos informáticos a gran escala

Comprender los comandos del lenguaje humano

/ Robot conversacional como ChatGPT y compañía



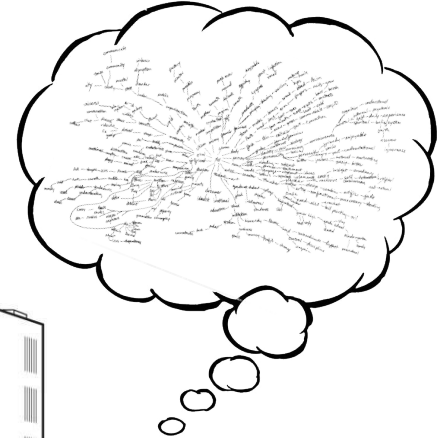
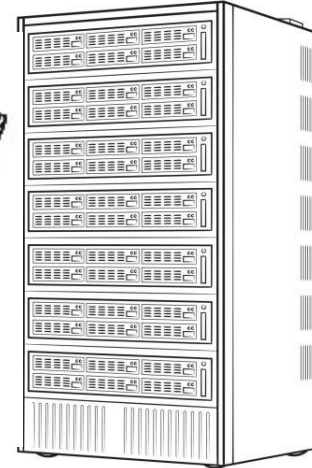
2. Cómo funciona

palabra / frases



(0.23 1.12 -0.15 0.71, ...)

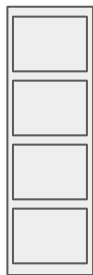
Los modelos se entrenan con grandes cantidades de datos.



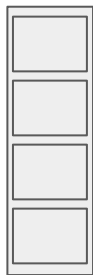
2. Cómo funciona

calcular distancias

frase 1



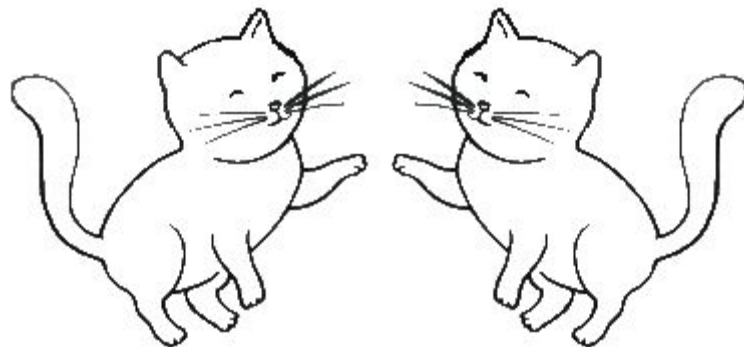
frase 2



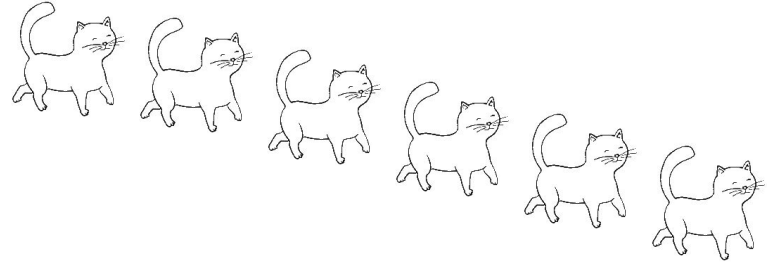
—



12.34



2. Cómo funciona



Actualmente : las frases se dividen en tokens

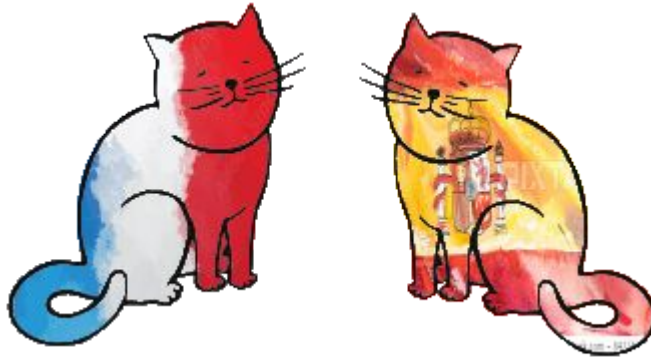
Procesamiento del lenguaje

España está al lado de Francia

3. Tareas populares

Traducción:

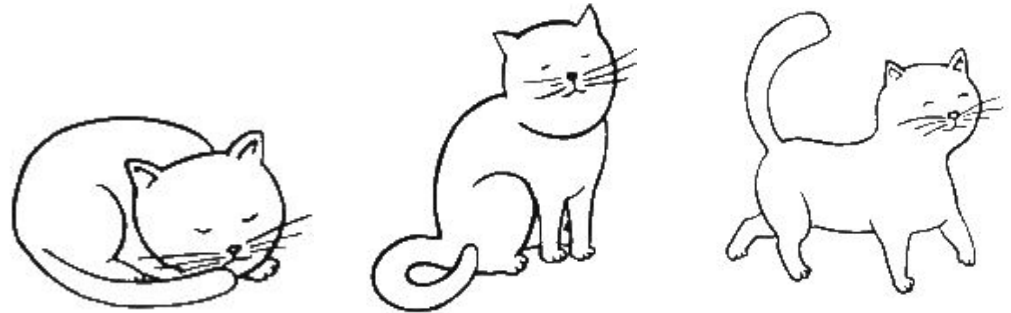
- “Bonjour, comment ca va ?”  <=>  “¿Hola, cómo estás?”



3. Tareas populares

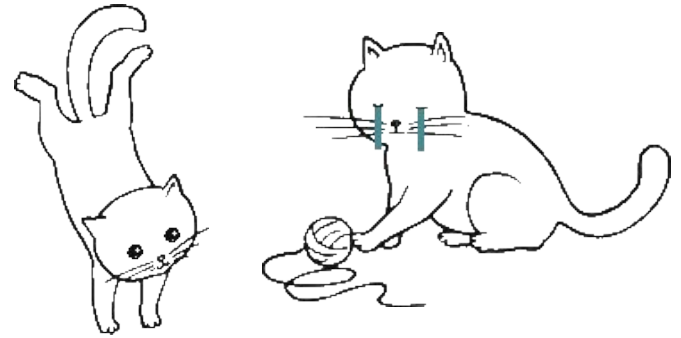
Predicción de la siguiente palabra:

- “¿Hola, cómo ” => “estás”
- “¿Hola, cómo estás” => “?”
- “¿Hola, cómo estás?” => .



3. Tareas populares

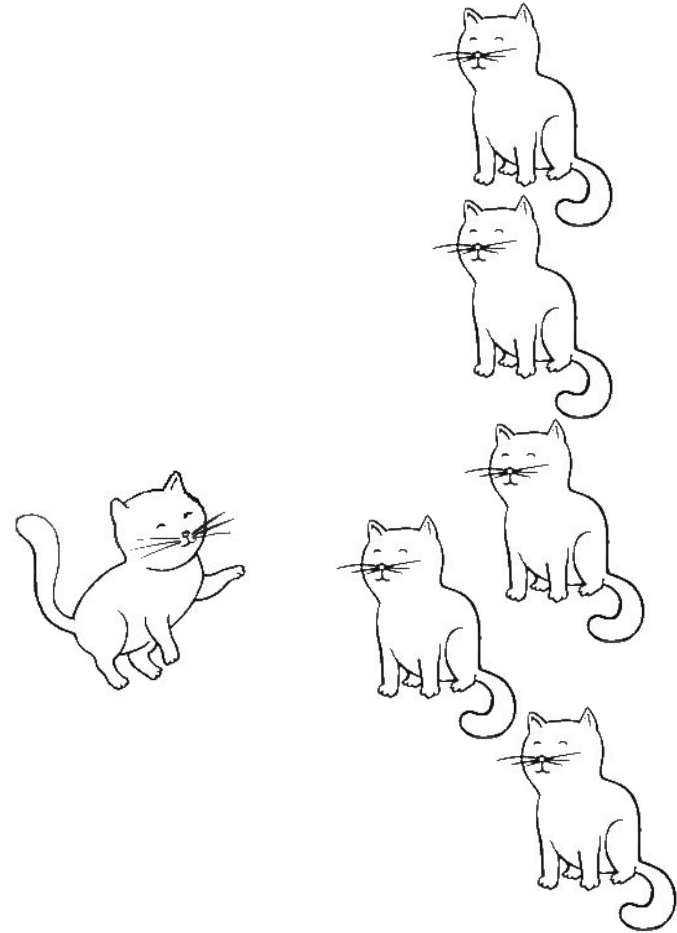
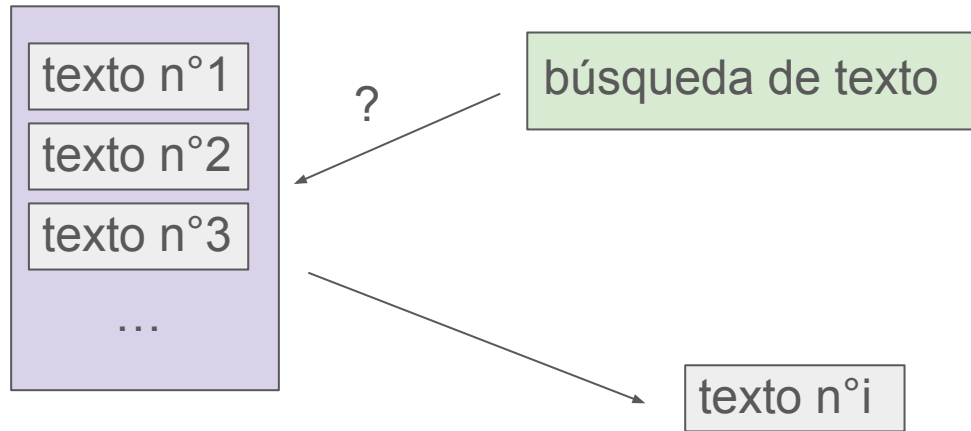
Clasificación semántica:



- “Este concierto fue fantástico, ¡la atmósfera era eléctrica!” -> **POSITIVO**
- “Me lo he pasado genial en esta conferencia.” -> **POSITIVO**
- “Estoy furiosa, cancelaron mi vuelo sin previo aviso.” -> **NEGATIVO**
- “Este producto es una estafa, no funciona en absoluto.” -> **NEGATIVO**
- “El tiempo está nublado hoy.” -> **NEUTRAL**
- “Estoy leyendo un artículo sobre política.” -> **NEUTRAL**

3. Tareas populares

Búsqueda semántica:



4. Conclusión

- Tema muy amplio y técnico
- Vectores para representar semántica
- Es útil saber cómo funciona -> vamos a ver más y más

