



Procesamiento de Lenguaje Natural

Santiago Pineda Quintero

Grupo de Control y Procesamiento Digital de Señales (GCPDS)
Dep. de Ing. Eléctrica Electrónica y Computación (DIEEC)
Facultad de Ingeniería y Arquitectura (FIA)
Universidad Nacional de Colombia sede Manizales

Contenido



UNIVERSIDAD
NACIONAL
DE COLOMBIA

- 1 Introducción
- 2 Modelos Tokenizadores
- 3 Tareas Clásicas De Procesamiento De Lenguaje Natural
- 4 Conclusiones

Contenido



1 Introducción

2 Modelos Tokenizadores

3 Tareas Clásicas De Procesamiento De Lenguaje Natural

4 Conclusiones



¿Qué es inteligencia artificial?



Inteligencia Artificial

1. Subdisciplina del campo de la Informática, que busca la creación de máquinas que puedan imitar comportamientos inteligentes.

Usado por primera vez: 1955

CAMPOS - IA



UNIVERSIDAD
NACIONAL
DE COLOMBIA





¿Qué es PLN?

Enseñar a las máquinas
a procesar y analizar
lenguaje humano
(texto y audio).





"El Lenguaje es la joya de la corona de la cognición"
Steven Pinker 1994

¿Qué es un modelo de lenguaje amplio (LLM)?



ChatGPT

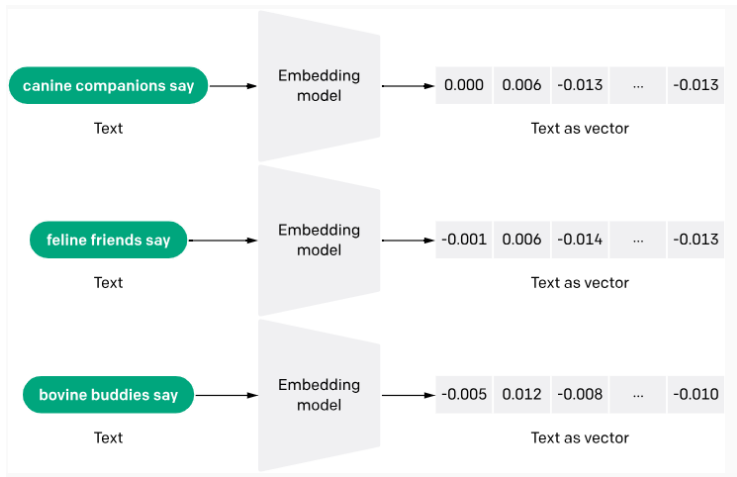
Contenido



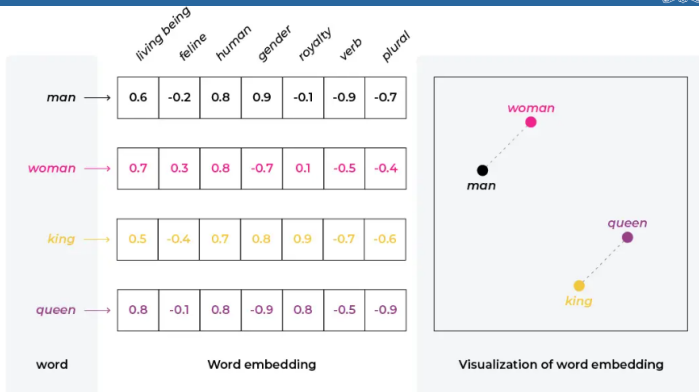
UNIVERSIDAD
NACIONAL
DE COLOMBIA

- 1 Introducción
- 2 Modelos Tokenizadores
- 3 Tareas Clásicas De Procesamiento De Lenguaje Natural
- 4 Conclusiones

Embedding



Espacio Dimensional

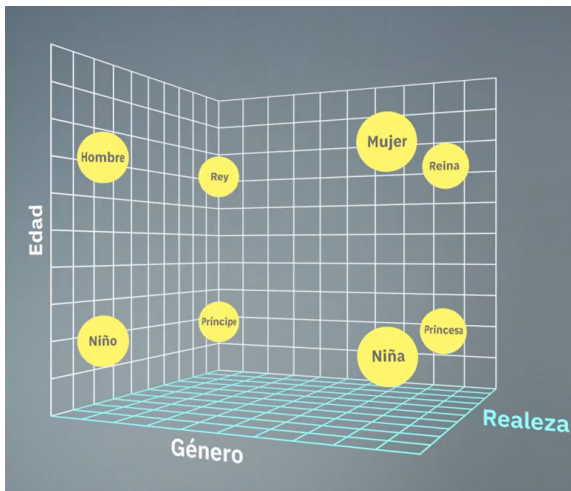


$$\mathbb{X} \in [0, 1]^d \quad (1)$$

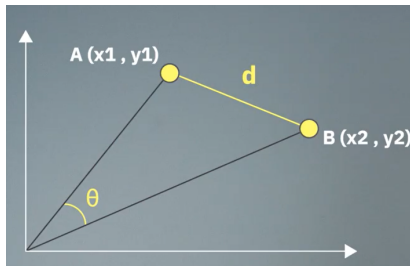
Espacio Dimensional



UNIVERSIDAD
NACIONAL
DE COLOMBIA



Medidas de Similitud Semántica



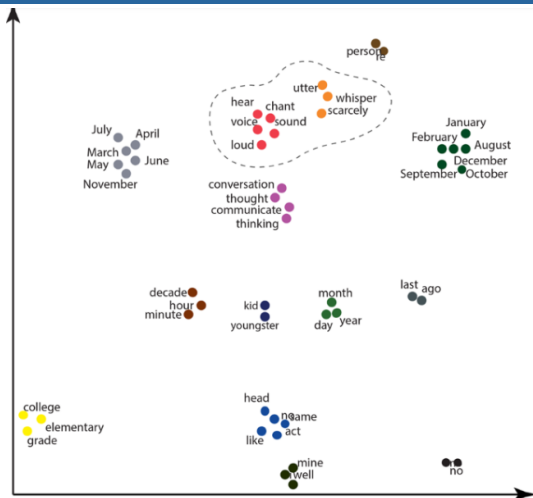
$$d = \sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2} \quad (2)$$

$$\theta = \cos^{-1} \left(\frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} \right) \quad (3)$$

Clusterización de Palabras



UNIVERSIDAD
NACIONAL
DE COLOMBIA



Contenido



UNIVERSIDAD
NACIONAL
DE COLOMBIA

- 1 Introducción
- 2 Modelos Tokenizadores
- 3 Tareas Clásicas De Procesamiento De Lenguaje Natural**
- 4 Conclusiones

Tareas clásicas de NLP con LangChain



UNIVERSIDAD
NACIONAL
DE COLOMBIA



Contenido



- 1 Introducción
- 2 Modelos Tokenizadores
- 3 Tareas Clásicas De Procesamiento De Lenguaje Natural
- 4 Conclusiones**

Conclusiones



- El procesamiento de lenguaje natural es clave para transformar datos textuales en conocimiento accionable mediante modelos avanzados como LLMs.información relevante.
- Técnicas como embeddings y arquitecturas RAG facilitan la integración de recuperación y generación de texto para soluciones más precisas e inteligentes.
- Los agentes conversacionales y las tareas clásicas de PLN tienen aplicaciones prácticas en sectores como servicio al cliente, análisis de datos y automatización.



Gracias!

Santiago Pineda Quintero

email: `spinedaq@unal.edu.co`