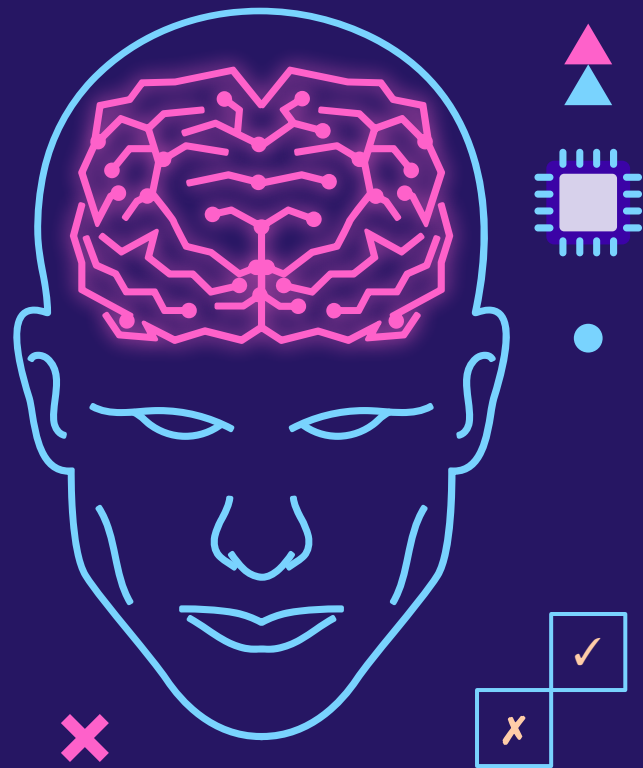
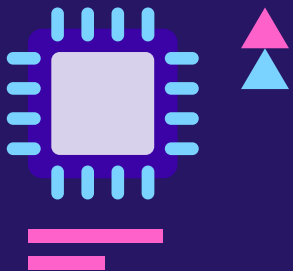




TP4 - Parser





01

Procesamiento de Lenguaje Natural



PLN

El Procesamiento de Lenguaje Natural, o NLP por sus siglas en inglés (Natural Language Processing), es un campo de la inteligencia artificial (IA) que se centra en la interacción entre las computadoras y el lenguaje humano. En esencia, NLP se ocupa de cómo las máquinas pueden comprender, interpretar y generar texto y habla en un formato que sea comprensible para los seres humanos.



Aplicaciones del PLN

- Asistentes Virtuales y Chatbots
- Traducción Automática
- Resúmenes de Texto
- Análisis de Sentimientos
- Reconocimiento de Voz



Nuestro gran amigo que usa PLN...



Tu, eres un caso de uso de NLP? Como funcionas?



Sí, soy un caso de uso de Procesamiento de Lenguaje Natural (NLP). Funciono utilizando un modelo de lenguaje desarrollado por OpenAI llamado GPT-3.5, que está basado en la arquitectura GPT-3. Este modelo ha sido entrenado en una gran cantidad de texto en varios idiomas y puede comprender y generar lenguaje humano de manera efectiva.



Mi funcionamiento se basa en procesar y generar texto en respuesta a las consultas y solicitudes de los usuarios. Cuando un usuario me hace una pregunta o proporciona información, utilizo algoritmos de NLP para comprender el contexto, identificar patrones y generar respuestas coherentes y relevantes.

Mi capacidad para responder preguntas, generar texto, proporcionar información y mantener conversaciones es posible gracias a las técnicas avanzadas de NLP y el aprendizaje automático. Los modelos de lenguaje como el mío son ampliamente utilizados en aplicaciones que van desde asistentes virtuales y motores de búsqueda hasta sistemas de recomendación y análisis de texto.

En resumen, soy un ejemplo de cómo el Procesamiento de Lenguaje Natural se utiliza para permitir la comunicación efectiva entre humanos y sistemas de inteligencia artificial.



Conceptos fundamentales

Conceptos fundamentales



Tokenization

El proceso de dividir un texto en unidades más pequeñas llamadas tokens. Estos tokens pueden ser palabras individuales, frases o incluso caracteres, dependiendo de los requisitos de procesamiento.



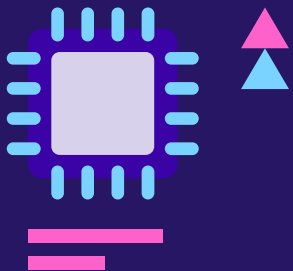
Análisis morfológico

Se centra en el estudio de la estructura y las propiedades de las palabras. Esto incluye la identificación de la forma base de una palabra (lema), sus afijos (como prefijos y sufijos) y su categoría gramatical (sustantivo, verbo, adjetivo, etc.).



Análisis sintáctico

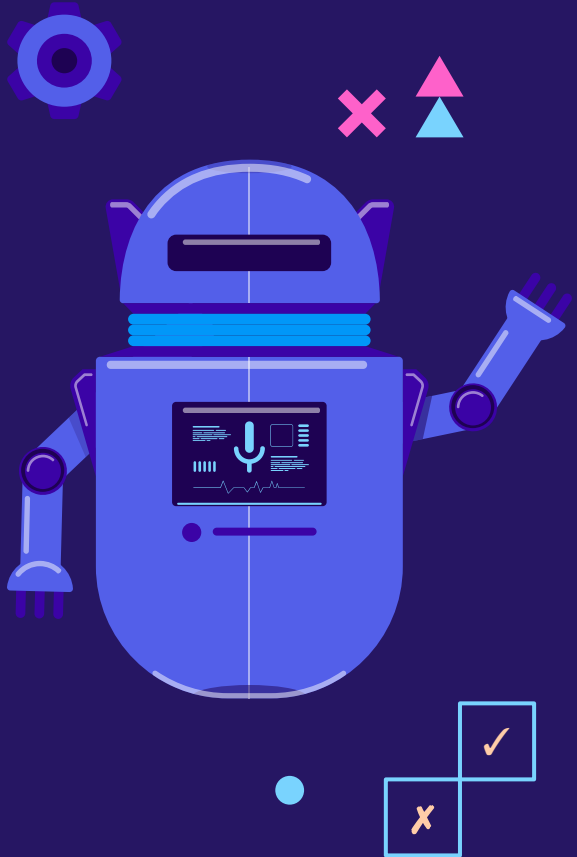
Se enfoca en la estructura de las oraciones y cómo las palabras se combinan gramaticalmente para transmitir significado. Esto implica identificar las relaciones sintácticas entre las palabras y cómo se organizan en frases y cláusulas.



Recursos y herramientas

Biblioteca NLTK





La biblioteca Natural Language Toolkit (NLTK) es una plataforma en Python que brinda acceso a datos lingüísticos, recursos léxicos y algoritmos especializados en NLP.

— **NLTK**

NLTK

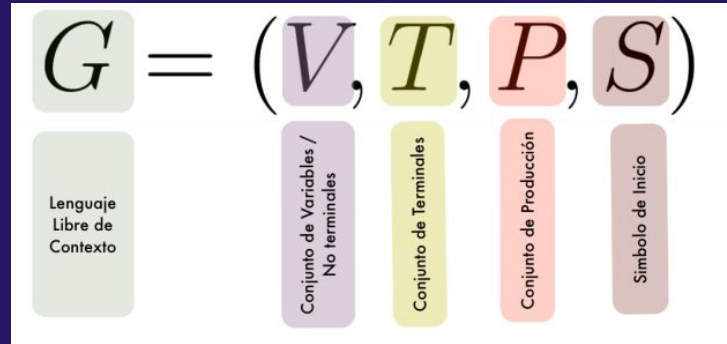
El código proporcionado utiliza NLTK para definir gramáticas y realizar análisis sintácticos en un texto dado.

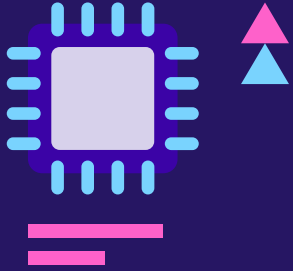
1. Definición de Gramáticas
2. Análisis Sintáctico
3. Extracción de Información

Gramática libre de contexto.

Una gramática libre de contexto se caracteriza por ser capaz de describir la estructura de un lenguaje basado en reglas generales que son independientes del contexto en el que se encuentren las palabras o símbolos en una oración.

1. **Terminales (T)**
2. **No Terminales (N)**
3. **Reglas de Producción (P)**
4. **Símbolo Inicial (S)**





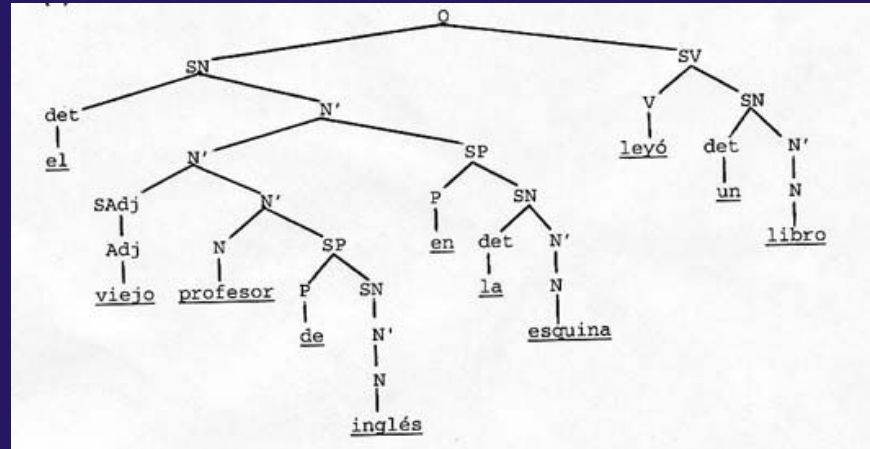
Explicacion problema

Parser



Parser

El objetivo del problema es desarrollar un analizador que pueda analizar oraciones en inglés y determinar su estructura utilizando gramática libre de contexto. Esto implica aplicar reglas de reescritura para transformar símbolos no terminales en otros símbolos y, finalmente, generar una oración completa de símbolos terminales (palabras).





Vamos al codigo...





Desafíos y Futuro del PLN

Desafíos y futuro del PLN

Desafíos

1. Ambigüedad
2. Idiomas no estándar
3. Lenguaje coloquial
4. Contexto y coherencia

Futuro

1. Modelos de lenguaje pre-entrenados
2. Análisis de dependencias
3. Multilingüismo
4. Procesamiento de lenguaje coloquial
5. Procesamiento de contexto

MUCHAS
GRACIAS
