



Programa académico: Ingeniería en Inteligencia Artificial

Unidad de aprendizaje : Aplicaciones del Lenguaje Natural

Nombre del Alumno: _____

Fecha de entrega: _____

Calificación: _____

Practica de laboratorio 6: Introducción desambiguación de las palabras

A) Marco teórico

En el procesamiento del lenguaje natural, la desambiguación del sentido de las palabras es el problema que consiste en determinar qué "sentido" (significado) de una palabra se activa al utilizarla en un contexto determinado, un proceso que parece ser en gran medida inconsciente en las personas. La WSD es un problema de clasificación natural: dada una palabra y sus posibles sentidos, definidos por un diccionario, clasifique una aparición de la palabra en el contexto en una o más de sus clases de sentido. Las características del contexto (como las palabras vecinas) proporcionan las pruebas para la clasificación.

Existen diferentes enfoques para la desambiguación de palabras en NLP, y algunos de ellos incluyen:

1. **Contexto local:** Analizar las palabras circundantes en la misma oración para determinar el significado correcto de la palabra ambigua.
2. **Contexto más amplio:** Considerar un contexto más extenso, como el párrafo o incluso el documento completo, para entender mejor el significado de la palabra en cuestión.
3. **Uso de características lingüísticas:** Examinar características lingüísticas, como la gramática y la sintaxis, para inferir el significado correcto.
4. **Uso de información externa:** Utilizar recursos externos, como diccionarios, tesauros o bases de conocimiento, para obtener información adicional sobre los posibles significados de una palabra.
5. **Aprendizaje automático:** Aplicar técnicas de aprendizaje automático, como modelos de clasificación, para entrenar al sistema en la tarea de desambiguación, utilizando conjuntos de datos etiquetados que indican el significado correcto en diferentes contextos.

La desambiguación de palabras es un componente esencial en muchas aplicaciones de procesamiento del lenguaje natural, como la traducción automática, la recuperación de información y los sistemas de respuesta a preguntas, ya que contribuye significativamente a mejorar la comprensión del lenguaje natural por parte de las máquinas.

B) Procedimiento:

Realizar el análisis de los códigos proporcionados y resaltar las operaciones dentro de estos que son relevantes a tomar en cuenta. Genere ejemplos propuestos por usted para llevar a cabo la desambiguación de las palabras.

Por ejemplo, lo primero que se debe analizar o investigar es ¿cómo funciona el algoritmo de LESK?

¿Qué es t-SNE, o t-distributed Stochastic Neighbor Embedding?

Procesar la información con: 'firefox.txt', 'grail.txt', 'overheard.txt', 'pirates.txt', 'singles.txt', 'wine.txt', implica observar el número de frases convertidas a partir del fichero seleccionado (además de firefox).

Como idea o ejercicio a realizar con los códigos proporcionados, fusionar ambos procesamientos en uno solo para obtener un rendimiento significativo (mejora) de la desambiguación de las palabras.

C) Resultados

Proporcionar los resultados obtenidos al ejecutar cada código, de forma adicional realice y/o análisis como los mencionados en la etapa de procedimiento para mejorar el rendimiento de los códigos implementados de forma individual o fusionados.

D) Análisis de Resultados

E) Conclusiones

F) Referencias bibliográficas