

1.- Lenguaje Natural

martes, 13 de febrero de 2024 05:23 p. m.

Espacio que conforma a todos --> Ciencias de datos

Conjunto que contiene a los demás conjuntos --> Inteligencia Artificial

Conjuntos dentro del conjunto de IA dentro del espacio de Ciencia de D.:

- 1.- Machine learning
- 2.- Procesamiento del lenguaje natural
- 3.- Deep Learning

1.- Los algoritmos ya existían desde antes de que fueran posibles sus aplicaciones.

2.- La evolución de los transistores nos ha brindado la mejor capacidad de procesamiento a diferencia de la utilidad que tuvieron inicialmente los bulbos de vacío.

La cantidad de transistores hasta hace poco cumplía la Ley de Moore.

3.- Incremento de volumen en los datos. Debido a la innovación de web 2.0

La inteligencia artificial trata de replicar el comportamiento humano.

El machine learning son las reglas de operación del sistema no codificado explícitamente, si no que el sistema es capaz de "actuar" en condiciones "nuevas".

El procesamiento del lenguaje natural es una especialización en el machine learning.

***Conocimiento:**

- 1.- El conocimiento es la acción y efecto de conocer, es decir, de adquirir información valiosa para comprender la realidad por medio de la razón, el entendimiento y la inteligencia. Se refiere, pues, a lo que resulta de un proceso de aprendizaje
- 2.- En PLN, el conocimiento se refiere a la comprensión y el análisis profundo del lenguaje humano que se obtiene a partir de los datos y la información. Esto puede incluir la capacidad de responder preguntas, traducir entre idiomas, resumir texto, entre otras tareas.

3.- En términos generales, se puede decir que existen dos grandes tipos de conocimiento: el conocimiento *a priori* y el conocimiento *a posteriori*.

- **Conocimiento *a priori*:** el conocimiento puede ser *a priori* cuando se basa en el proceso de la razón personal o la introspección para formularse, sin verificarse en la experiencia.
- **Conocimiento *a posteriori*:** se habla de conocimiento *a posteriori* cuando surge a partir de una experiencia, y esa misma experiencia se vuelve en una validación del aprendizaje.

***Predicción:**

- 1.- La predicción en el contexto científico es una declaración precisa de lo que ocurrirá en determinadas condiciones especificadas.

Investigar.

***Datos:**

- 1.- Es la entrada sin procesar de la cual se produce la información.
- 2.- Los datos se refieren a cualquier tipo de texto o discurso que se utiliza para entrenar modelos de procesamiento del lenguaje natural. Estos datos pueden incluir libros, artículos de noticias, conversaciones, correos electrónicos, tweets, etc.
- 3.- Los datos son el objeto de estudio de la informática, y son información en forma formalizada (codificada en una forma digital comprensible para un ordenador, en un código binario).

***Información:**

- 1.- Información se refiere a los datos que han sido procesados y comunicados de tal manera que pueden ser entendidos e interpretados por el receptor.
- 2.- La información es un conjunto organizado de datos relevantes para uno o más sujetos que extraen de él un conocimiento. Es decir, es una serie de conocimientos comunicados, compartidos o transmitidos y que constituyen por lo tanto algún tipo de mensaje.
- 3.- La información se refiere al conocimiento extraído de los datos mediante técnicas de PLN. Esto puede incluir la identificación de entidades nombradas (como nombres de personas o lugares), el análisis de sentimientos, la extracción de información relevante de textos, entre otros.

***Agrupar:**

- 1.- Un agrupamiento es una colección de grupos con características similares.
- 2.- En el contexto del PLN, agrupar se refiere a la tarea de organizar documentos de texto en grupos o clústeres basados en similitudes lingüísticas. Esto puede ser útil para descubrir temas comunes o tendencias dentro de un conjunto de documentos.
- 3.- El clustering o agrupamiento en machine learning es una técnica de aprendizaje no supervisado en la que se da un algoritmo de agrupamiento de conjuntos de datos que presentan características similares.

***Clasificar:**

- 1.- Organización de datos en categorías predefinidas con ciertas características y atributos.
- 2.- La clasificación en PLN implica asignar una etiqueta o categoría predefinida a un documento de texto en función de su contenido. Por ejemplo, clasificar correos electrónicos como spam o no spam, clasificar noticias en diferentes categorías como deportes, política, entretenimiento, etc.
- 3.- Un clasificador es un algoritmo que, recibiendo como entrada cierta información de un objeto, es capaz de indicar la categoría o clase a que pertenece de entre un número acotado de clases posibles.

2.- En PLN, la predicción implica predecir algún aspecto del lenguaje humano en función de datos históricos. Esto puede incluir predecir la próxima palabra en una oración, predecir el siguiente elemento en una secuencia de texto o predecir la intención detrás de una declaración.

3.- La capacidad de un programa informático de utilizar el análisis estadístico para identificar pautas, anticipar comportamientos y prever los acontecimientos futuros.