

# Todo sobre el procesamiento del lenguaje natural en un post del blog, resumen

Todo sobre el procesamiento del lenguaje natural en un post del blog, resumen. Muchos no creen que la semántica o la semiótica tengan que ver con el procesamiento del lenguaje natural o los motores de búsqueda, pero la historia de estas herramientas tecnológicas así lo demuestran. Vamos con nuestro compilado.

El procesamiento del lenguaje natural, también llamado procesamiento del lenguaje natural o PNL en inglés, es una tecnología que permite a las máquinas comprender el lenguaje humano a través de la inteligencia artificial. Descubra todo lo que necesita saber al respecto.

Con el auge anunciado en inteligencia artificial , pronto trataremos con robots y otras entidades digitales a diario. Sin embargo, para poder cooperar o incluso coexistir con las máquinas, es necesario poder comunicarse con ellas . Esta es la razón de la tecnología de procesamiento del lenguaje natural.

## ¿Qué es el procesamiento del lenguaje natural? ¿Cómo lo definiríamos?

El procesamiento del lenguaje natural (PNL) es una tecnología de inteligencia artificial diseñada para permitir que las computadoras entiendan el lenguaje humano .

El objetivo de esta tecnología es permitir que las máquinas lean, descifren, comprendan y den significado al lenguaje

humano. Es claro que se ha logrado un progreso significativo en esta área en los últimos años , y el procesamiento del lenguaje natural ahora se utiliza para una amplia variedad de casos de uso ...

## ¿En dónde se emplea o se aplica?

Puede que no lo sepa, pero muchas de las aplicaciones que usa todos los días dependen del procesamiento del lenguaje natural. Como ejemplo, podemos citar aplicaciones de traducción como Google Translate o asistentes personales como Apple Siri, Microsoft Cortana, Amazon Alexa o Microsoft Cortana. Lo mismo ocurre con todos los chatbots .

Del mismo modo, los procesadores de texto como Microsoft Word y Grammarly usan NLP para verificar la gramática y la ortografía de los textos. Finalmente, las aplicaciones de Respuesta de voz interactiva (IVR) utilizadas por los centros de llamadas permiten que ciertas solicitudes se procesen automáticamente.

Estos son solo algunos ejemplos concretos utilizados por la gente común. Sin embargo, hay una gran cantidad de aplicaciones para esta tecnología. En general , todos los programas basados en la comprensión automática del lenguaje se basan en PNL.

## ¿Cómo funciona el procesamiento del lenguaje natural?

La mayoría de las técnicas de procesamiento del lenguaje natural se basan en el aprendizaje profundo . Los algoritmos de inteligencia artificial se basan en datos para analizar el lenguaje humano y aprender a encontrar patrones y correlaciones.

El papel de los algoritmos es identificar y extraer las reglas del lenguaje natural, para convertir los datos del lenguaje no

estructurado, en una forma que las computadoras y servidores web puedan entender .

En el pasado, los viejos enfoques de procesamiento del lenguaje natural se basaban en un enfoque basado en reglas. Los algoritmos de Machine Learning de la época recibieron instrucciones de buscar palabras y oraciones en un texto y dieron respuestas específicas en consecuencia. Sin embargo, el aprendizaje profundo permite un enfoque más flexible e intuitivo, más cercano al lenguaje natural y la forma en que los humanos lo aprenden en la infancia.

Típicamente, una interacción entre humanos y máquinas a través de PNL se basa en el siguiente proceso : primero, los humanos hablan con la máquina. La máquina captura el sonido y lo convierte en texto. Los datos de texto se procesan y luego se vuelven a convertir en datos de audio. La máquina reproduce el archivo de audio para responder al interlocutor humano.

## ¿Cuáles son las diferentes técnicas de PNL?

Las dos técnicas principales utilizadas para el procesamiento del lenguaje natural son el análisis sintáctico y el análisis semántico .

El análisis sintáctico consiste en identificar las reglas gramaticales en una oración para descifrar su significado.

Existen varias técnicas de análisis semántico . El análisis consiste en analizar la gramática de una oración. La segmentación de palabras consiste en dividir un texto en unidades, mientras que la segmentación morfológica divide las palabras en grupos.

Mientras tanto, el análisis semántico consiste en descifrar directamente el significado de un texto utilizando algoritmos para analizar palabras y la estructura de las oraciones. Los

algoritmos pueden basarse notablemente en el contexto o comparar textos con bases de datos para comprender su significado. Sin embargo, es un enfoque complejo y todavía no existe un algoritmo realmente capaz de comprender el significado de un texto de esta manera ...

## ¿Cuáles son las dificultades asociadas con el procesamiento del lenguaje natural?

E  
l  
p  
r  
o  
c  
e  
s  
a  
m  
i  
e  
n  
t  
o  
d  
e  
e



l lenguaje natural está lejos de ser una tarea fácil. Causar el lenguaje humano es intrínsecamente complejo y diferentes reglas son difíciles de entender para una computadora . (Y dejémonos de cuentos, el inglés es más simple que las lenguas romances como el español, sobre todo para este tipo de ejercicios).

Algunas de estas reglas pueden ser muy abstractas. Por ejemplo, cuando una persona usa un comentario sarcástico para

transmitir un mensaje sutil. Es casi imposible para una máquina actual percibir tales matices .

Del mismo modo, el uso de la letra «s» para significar la pluralidad es complejo de asimilar para una máquina. Para comprender verdaderamente el lenguaje humano, es necesario comprender ambas palabras y cómo se conectan los conceptos para transmitir un mensaje.

Si los humanos pueden dominar fácilmente un idioma, la ambigüedad y las características imprecisas del lenguaje lo hacen mucho más complicado para las máquinas que no están acostumbradas a «pensar» de esa manera. Los humanos generalmente «hablan» a las computadoras usando lenguajes de programación precisos, concretos y altamente estructurados.

Sin embargo, las inversiones y la investigación en el campo del procesamiento del lenguaje natural son cada vez más importantes. En los próximos años, esta tecnología se desarrollará a un ritmo rápido y las máquinas pronto podrán comunicarse de forma completamente natural con los humanos ...

## **¿Sabía que las bases de datos públicas son una mina de información útil?**

Nadamos en información, tenemos múltiples herramientas de procesamiento de datos y servidores en línea trabajando para nosotros. Pero seamos honestos, muchos no sabemos qué hacer con la data, no tenemos idea de su practicidad en nuestras profesiones, no logramos armar el conjunto visionar la totalidad de sus implicaciones.

Si usted es un negocio o un individuo, los datos pueden ser invaluable. El análisis de Big Data puede, por ejemplo, aumentar la precisión de la investigación de mercado y permitir el desarrollo de nuevos productos y servicios que

satisfagan la demanda de los consumidores.

Lamentablemente, para las PYME y las personas, acceder a los datos puede ser difícil . A diferencia de gigantes como Google o Facebook, que generan grandes volúmenes de datos a través de redes sociales u objetos conectados, una pequeña empresa solo tiene una visión limitada sobre el perfil de su objetivo.

Afortunadamente, hay datos públicos accesibles para todos y recopilados en bases de datos públicas. Estas herramientas están disponibles para todos y le permiten beneficiarse de las muchas ventajas de Big Data .

Una base de datos pública agrupa los datos disponibles al público para el interés general. Las razones pueden ser múltiples: puede ser información del servicio público o datos ambientales.

Los datos pueden ser documentos administrativos, decisiones, documentos de la ONU, fondos digitales o archivos. Como regla general, estos datos son proporcionados por el Estado o por una autoridad local .

Para citar un ejemplo relevantes. En la Unión Europea, los artículos 41 y 42 de la Carta de los Derechos Fundamentales de la UE garantizan el derecho de todos los ciudadanos de la UE a acceder a documentos de la Comisión Europea, el Parlamento y el Consejo. Todos los documentos de organizaciones del sector público también se pueden ver y reutilizar .

Sin embargo, en aras de la confidencialidad, los datos relacionados con la privacidad, la seguridad nacional o los derechos de propiedad intelectual son una excepción. Además, las bases de datos públicas no deben confundirse con las bases de datos libres de regalías.

Al aprovechar las bases de datos públicas, las empresas pueden lograr sus objetivos comerciales . Del mismo modo, los investigadores pueden usar esta información para realizar su investigación.

## Cómo usar bases de datos públicas para su proyecto

Existen todo tipo de bases de datos públicas para todos los sectores de actividad. Algunos recopilarán datos de salud y serán utilizados por los investigadores. Otros enumeran datos sobre el mercado inmobiliario y pueden ser utilizados por agencias o personas que buscan vivienda.

Como ejemplo, podemos citar la base de datos Datadock que enumera las organizaciones de capacitación que respetan 21 indicadores de calidad precisos.

Por lo tanto, esta base de datos pública puede ser muy útil si planea crear una organización de capacitación o si desea convertirse en un capacitador. También puede consultar el informe educativo y financiero de una organización, y su número de declaración de actividad.

En todos los sectores de actividad , las bases de datos públicas pueden ser muy útiles. Se pueden usar para investigadores, compañías para sus estudios de mercado o incluso individuos para apoyar su investigación en Internet. Por lo tanto, es otro ejemplo de las oportunidades que ofrece Big Data ...

Leer también: [¿Cómo está mejorando la inteligencia artificial la industria del alojamiento web?](#) ; [Por qué es tan importante el análisis Big Data](#) ; [Deep learning, aprendizaje profundo para construir centros de datos inteligentes](#)