



Curso de Especialización de Inteligencia Artificial y Big Data (IABD)



Modelos de Inteligencia Artificial

UD05. Áreas de estudio y desarrollo de la
Inteligencia Artificial.
Resumen.

JUAN ANTONIO GARCIA MUELAS

Modelos de Inteligencia Artificial UD05. Áreas de estudio y desarrollo de la IA

La **visión artificial automatiza** la extracción, el análisis, la clasificación y la comprensión de la **información útil** a partir de los datos **de las imágenes**. Los datos de la imagen adoptan muchas formas, como las siguientes:

- ✓ Imágenes individuales
- ✓ Secuencias de video
- ✓ Visualizaciones de varias cámaras
- ✓ Datos tridimensionales

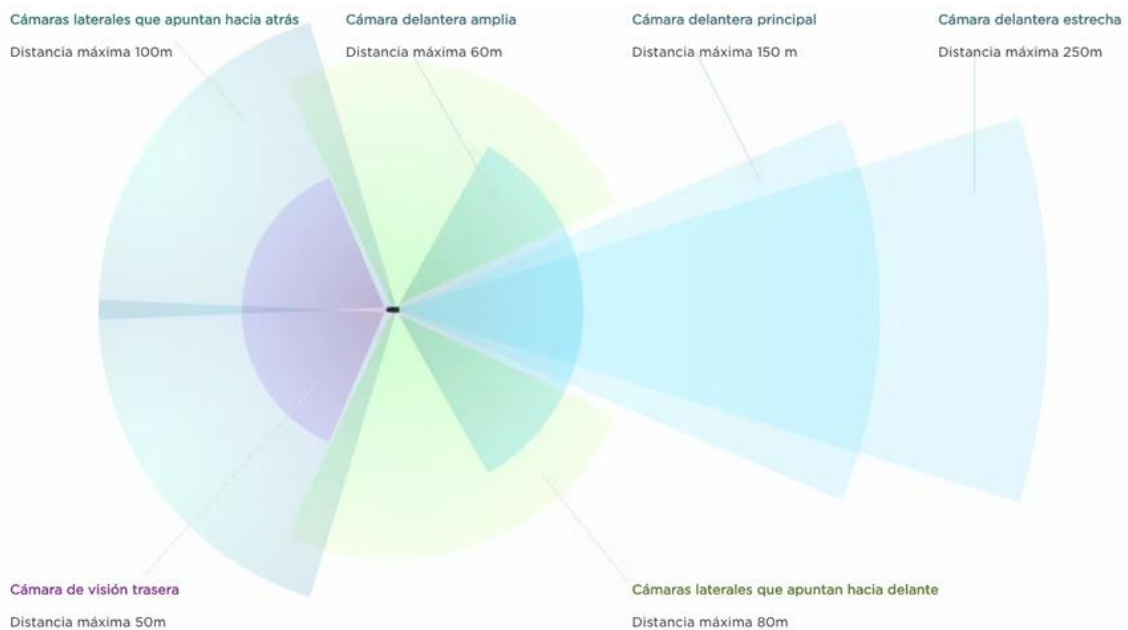
Un **analizador facial** es un software que identifica o confirma la identidad de una persona a partir del rostro. Funciona **mediante la identificación y medición de los rasgos faciales** (frontales o laterales) en una imagen.

La tecnología de reconocimiento facial por lo general busca lo siguiente:

- ✓ **Distancia entre los ojos**, de la **frente a la barbilla** o entre la **nariz y la boca**.
- ✓ **Profundidad** de las **cuecas** oculares.
- ✓ Forma de los **pómulos**.
- ✓ **Contorno** de los labios, las orejas y la barbilla.

El sistema **convierte los datos** de reconocimiento facial en una cadena de números o puntos denominada **huella facial**.

El **sistema de conducción autónoma** de vehículos implica varias **tareas y subsistemas**, pero uno de los más importantes, es el de visión artificial.



Para saber más

Si quieres conocer mejor la clasificación de niveles de conducción autónoma y cómo se relaciona el ámbito de la visión artificial con ellos, te recomendamos leer el artículo ["Autonomous Vehicles Are Driving Computer Vision Into the Future"](#) o, en una clave más técnica, el artículo ["Visual Perception for Self-Driving Cars"](#).

Los **robots** son **sistemas complejos** que **suelen ejecutar una serie de tareas** en el mundo físico **en base a una secuencia programada**.

Modelos de Inteligencia Artificial UD05. Áreas de estudio y desarrollo de la IA

Un sistema robótico tiene tres partes fundamentales:

- ✓ Sensores o entradas.
- ✓ Sistema de control.
- ✓ Actuadores.

El sistema de visión artificial de un robot le permite detectar objetos y posicionarse a sí mismo o a objetos que transporta en función de lo que está viendo.

En el campo del procesamiento y la interpretación de imágenes para el diagnóstico, la IA ofrece algoritmos que mejoran la calidad y la precisión del diagnóstico.

Uno de los grandes e inesperados avances de la computación de la década pasada, ha sido el de **los modelos generativos: las redes GAN para el campo de la imagen y los modelos de generación de texto basados en Transformers.**

Deep Dream

Creado en **2014**, se publica en **2015** DeepDream, un modelo de **generación de imágenes** creado por Google. El software Deep Dream fue desarrollado para el **imageNet (liderado por Fei-Fei Li)** large scale visual recognition challenge (ILSVRC) **capaz de reconocer cualquier imagen y clasificarla.**

Gaugan

GauGAN es una herramienta con la que se pueden **crear paisajes falsos** partiendo **de un boceto**. Este software **de Nvidia**, hace uso de una red de **confrontación generativa (GAN)**, basado en una técnica denominada "**normalización espacialmente adaptativa**" capaz de generar imágenes realistas a partir de un determinado diseño semántico.

DALL·E

Una de las más recientes incorporaciones es el modelo de generación de imágenes **Diffusion de openAI**. Se trata de una **implementación multimodal de GPT3**. El algoritmo interpreta una descripción escrita que se le proporciona a través de su interfaz, y **genera la imagen** correspondiente en base a lo que sus **12 mil millones de parámetros** del modelo GPT3 han interpretado de la entrada de texto dada.

El **Procesamiento del Lenguaje Natural** es el campo de conocimiento de la Inteligencia Artificial que se ocupa de la investigar la manera de comunicar las máquinas con las personas mediante el **uso de lenguas naturales**.

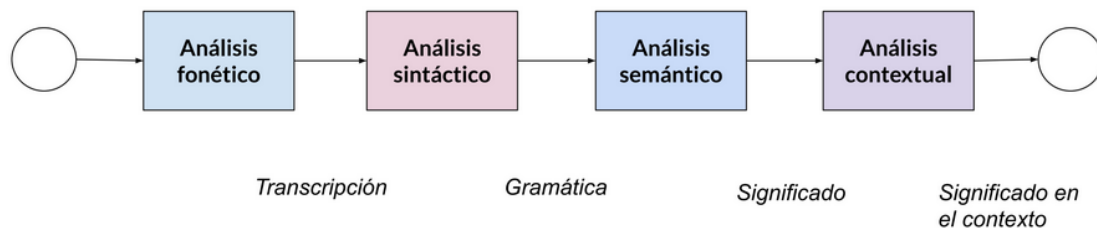
Tratar computacionalmente una lengua **implica un proceso de modelización matemática**.

Éstos son algunos de los **componentes del procesamiento del lenguaje natural**.

- ✓ **Análisis morfológico o léxico.** Consiste en el análisis interno de las palabras que forman oraciones para extraer lemas, rasgos flexivos, unidades léxica compuestas. Es esencial para la información básica: categoría sintáctica y significado léxico.
- ✓ **Análisis sintáctico.** Consiste en el análisis de la estructura de las oraciones de acuerdo con el modelo gramatical empleado (lógico o estadístico).
- ✓ **Análisis semántico.** Proporciona la interpretación de las oraciones, una vez eliminadas las ambigüedades morfosintácticas.
- ✓ **Análisis pragmático.** Incorpora el análisis del contexto de uso a la interpretación final. Aquí se incluye el tratamiento del lenguaje figurado (metáfora e ironía) como el conocimiento del mundo específico necesario para entender un texto especializado.

Un análisis morfológico, sintáctico, semántico o pragmático se aplicará dependiendo del objetivo de la aplicación. Por ejemplo, un conversor de texto a voz no necesita el análisis semántico o pragmático. Pero un sistema conversacional requiere información muy detallada del contexto y del dominio temático.

EL PROCESO DE RECONOCIMIENTO DEL LENGUAJE



¿Cómo comprender realmente el significado de un texto? ¿cómo intuir neologismos, ironías, chistes o poesía?

Las **técnicas** que permiten adaptar los **datos desestructurados** del lenguaje para utilizarlos con modelos de machine learning, son:

- ✓ **Tokenización:** separar **palabras del texto** en entidades llamadas tokens, con las que trabajaremos luego. Debemos pensar si utilizaremos los signos de puntuación como token, si daremos importancia o no a las mayúsculas y si unificamos palabras similares en un mismo token.
- ✓ **Bag of words:** es una manera de representar el vocabulario que utilizaremos en nuestro modelo y consiste en crear una **matriz en la que cada columna es un token** y se contabilizará la cantidad de veces que aparece ese token en cada oración (representadas en cada fila).
- ✓ **Word2vec:** Es una técnica que **aprende de leer enormes cantidades de textos** y memorizar qué palabras parecen ser similares en diversos contextos. Luego de entrenar suficientes datos, se **generan vectores** de 300 dimensiones para cada palabra conformando un nuevo vocabulario en donde las palabras “similares” se ubican cercanas unas de otras. Utilizando vectores pre-entrenados, logramos tener muchísima riqueza de información para comprender el significado semántico de los textos.

Harry walked down four blocks to pick up ice cream.



Tokenization



Harry	walked	down	four	blocks	to	pick	up	ice	cream	.
PROPN	VERB	ADV	NUM	NOUN	PART	VERB	ADP	NOUN	NOUN	PUNCT

Modelos de Inteligencia Artificial UD05. Áreas de estudio y desarrollo de la IA

Para la **generación de textos** tenemos:

BERT Creado por **Google en 2018**, integrado en su algoritmo de búsqueda, pero sin interfaz para pruebas.

GPT3 Generative Pre-trained Transformer 3, conocida por sus siglas (GPT-3), es un modelo de **lenguaje autorregresivo** que emplea aprendizaje profundo para producir textos que simulan la redacción humana.

Bloom El modelo BigScience Large Open-science Open-access Multilingual Language Model, más conocido como 'BLOOM' es un desarrollo de IA que comenzó en 2021, con el respaldo humano y financiero (100 millones de dólares) de la startup de machine learning Hugging Face, aunque también han colaborado Nvidia, Microsoft y el CNRS (el 'CSIC francés'). BLOOM está entrenado para generar texto un total de 59 idiomas.

Para saber más

Puedes consultar ejemplos de aplicaciones comerciales de generación de texto en este artículo: ["Los 15 mejores programas de escritura de IA para potenciar su redacción"](#)

Existen varios asistentes virtuales y conversacionales, que no son simples chatbots, pues no solo pueden mantener una conversación fluida con el usuario.

Google Assistant y Google Home

Google Home, es un altavoz inteligente desarrollado por Google. La aplicación Google Home está disponible en teléfonos y tablets Android, así como en dispositivos iPhone y iPad.

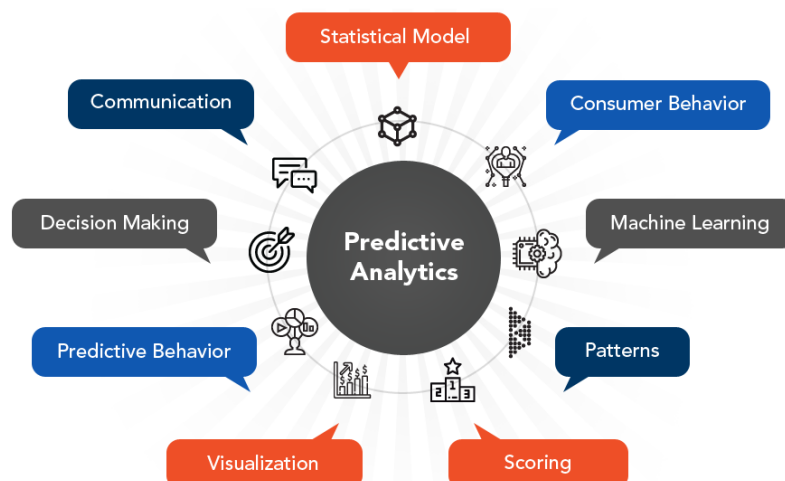
Alexa y Eco

Alexa es el asistente virtual controlado por voz creado por Amazon, y lanzado en noviembre de 2014 junto a su línea de altavoces inteligentes Echo. Su nombre fue elegido por tener una consonante fuerte al principio e incluir una x, algo que haría que el asistente reconociese más fácil su nombre, y también en honor a la Biblioteca de Alejandría.

Otros asistentes virtuales

Apple ha desarrollado **Siri** y sigue implementando funcionalidades de inteligencia artificial para mejorar su alcance y experiencia de usuario. También Microsoft continúa trabajando sobre su asistente **Cortana** para su sistema operativo Windows.

Los **Modelos Predictivos** son un grupo de técnicas que, mediante los campos del aprendizaje automático, la recolección de datos históricos, el Big Data y el reconocimiento de patrones, pretende dar una predicción de resultados futuros.



Modelos de Inteligencia Artificial UD05. Áreas de estudio y desarrollo de la IA

Los **sistemas robóticos** actuales cubren una gran cantidad de campos de aplicación del entorno. Destacan el caso de **Davinci, el robot cirujano, o agro-bot, el robot que recoge fresas** en su punto óptimo.

Hay otros casos representativos como:

Robots sociales.

Este tipo de robots tienen **muy desarrollados los módulos sensoriales**, es decir, el de reconocimiento de imagen, procesamiento de lenguaje natral, y su sistema de control es, básicamente, un asistente virtual que tiene ciertas opciones de movilidad y acciones remotas o conectadas, como encender la luz o hacer una llamada de emergencia.

Casas y ciudades inteligentes.

Estos sistemas robóticos cuentan con un elaborado sistema de sensores, y un hardware extendido por diversas localizaciones.

Conducción autónoma.

La unidad de control de un vehículo autónomo es el **paradigma** de las técnicas más avanzadas en **aprendizaje automático**. Se trata de **aprendizaje por refuerzo**, y se **entrena en simuladores virtuales**.

