***Unidad 1.3***: ***Áreas de aplicación***

***Visión por computadora:*** en el campo de la inteligencia artificial se incluye la Visión por Computadora (CV), un área especializada que tiene como objetivo hacer que las máquinas comprendan datos visuales a partir de imágenes y videos. ‘Para ver’ Con computadoras, robots y sistemas automatizados, la tecnología CV busca equiparlos con la capacidad de capturar la entrada de la cámara o del sensor, procesarla y analizarla tal como la perciben los humanos. La idea es enseñar a estos sistemas a reconocer señales visuales y tomar decisiones basadas en datos, interpretando su entorno para realizar tareas que requieren comprensión visual.

La visión por computadora le brinda muchas capacidades a las máquinas, lo que permite la automatización en muchas operaciones como las siguientes

* Clasificación de objetos: este proceso implica categorizar objetos en una imagen según etiquetas o grupos. Por ejemplo, con la clasificación de objetos, una computadora puede saber qué personas hay en una imagen y cuántas personas hay en la imagen.
* Identificación de objetos: este paso es más que una clasificación; identifica objetos específicos en una imagen o vídeo. Por ejemplo, el sistema no sólo puede identificar objetos en una fotografía, sino que también puede examinar las características de las personas en la fotografía para inferir identidad o rasgos.
* Seguimiento de objetos: en este caso, el sistema monitorea la posición de un objeto en movimiento a lo largo del tiempo en una secuencia de video. Por ejemplo, en un estacionamiento, las cámaras de vigilancia pueden registrar la ubicación y el movimiento de un vehículo tiempo.
* Reconocimiento óptico de caracteres (OCR): el proceso de detectar letras y números en imágenes y convertirlos en texto legible por máquina que otras aplicaciones pueden usar o modificar.

Algunos de los usos de la visión por computadora son los siguientes:

* Organización de contenidos: es el proceso de analizar y categorizar medios visuales. Puede ayudarle a ordenar, escanear y etiquetar imágenes y vídeos en busca de cantidades masivas de medios de una manera ordenada y fácil de encontrar.
* Extracción de texto: extraer texto de imágenes y documentos escaneados y utiliza reconocimiento óptico de caracteres (OCR) para convertirlo en texto digital editable. Le permite digitalizar documentos en papel, automatizar la entrada de datos y hacer que el texto se pueda buscar.
* Realidad aumentada (RA): permite anclar objetos digitales al mundo real asociándolos con superficies y objetos. Esto hace que las experiencias de RA sean más inmersivas porque los usuarios pueden interactuar con elementos virtuales en un entorno en tiempo real, lo que resulta útil para aplicaciones de juegos, venta minorista y simulación de entrenamiento.
* Agricultura: las imágenes satelitales y de drones ayudan a la visión por computadora a mejorar las áreas agrícolas al monitorear la salud de los cultivos, detectar malezas y analizar la calidad del suelo. Usando lo visual análisis de datos los agricultores no sólo pueden maximizar el rendimiento de sus cultivos sino también minimizar el desperdicio de recursos, así como identificar problemas para poder ayudar a impulsar prácticas agrícolas más sostenibles.
* Salud: en este ámbito la visión artificial se utiliza para ayudar a analizar imágenes médicas (patrones de rayos X, resonancias magnéticas, etc.) para detectar anomalía, como tumores o fracturas. Ayuda a acelerar el diagnóstico y el tratamiento preciso, incluso con datos visuales en tiempo real disponibles para los cirujanos durante los procedimientos, constituyendo una valiosa herramienta de diagnóstico y tratamiento médico.
* Deportes: la visión por computadora se utiliza en los deportes para el análisis y la planificación del rendimiento. Los jugadores pueden usarlo para monitorear sus movimientos, desarrollar tácticas en el juego e incluso dibujar posiciones y opciones de los jugadores.
* Fabricación: la visión por computadora se utiliza en la fabricación para ayudar con el control de calidad observando los productos, inspeccionándolos y viendo si están libres de defectos y cumplen con algunas especificaciones determinadas. Además, admite la automatización, que es necesaria para máquinas con capacidades de visión, que pueden monitorear flujos de trabajo, verificar el estado del equipo y desempeñar un papel en el aumento de la eficiencia general.
* Reconocimiento facial: el reconocimiento facial es un proceso mediante el cual visión por ordenador confirma identidades utilizando rasgos faciales. Esta tecnología se ha utilizado ampliamente en sistemas de seguridad y autenticación y permite la verificación sin contacto en muchas situaciones, incluida la verificación del acceso a teléfonos inteligentes, la entrada segura a las instalaciones y la experiencia personalizada del cliente.

***Procesamiento del Lenguaje Natural (NLP):*** El procesamiento de lenguaje natural o PLN es un subcampo de la informática y la inteligencia artificia que emplea el machine learning para permitir que las computadoras comprendan y se comuniquen con el lenguaje humano. Permite que los ordenadores procesen, analicen y generen lenguaje humano, lo que facilita interacciones intuitivas y fluidas. Los modelos de PNL se esfuerzan por interpretar con precisión las complejidades del lenguaje humano, incluidos los datos de texto y voz.

La investigación de PLN ha ayudado a habilitar la era de la IA generativa, desde las habilidades de comunicación de los modelos de lenguaje de gran tamaño (LLM) hasta la capacidad de los modelos de generación de imágenes para comprender las solicitudes. El PLN ya forma parte de la vida cotidiana de muchos, accionando los motores de búsqueda, impulsando los chatbots de atención al cliente con instrucciones habladas, los sistemas GPS operados por voz y los asistentes digitales de preguntas frecuentes en teléfonos inteligentes como Alexa de Amazon, Siri de Apple y Cortana de Microsoft.

El PLN combina el poder de la lingüística computacional con algoritmos de machine learning y aprendizaje profundo. La lingüística computacional emplea la ciencia de datos para analizar el lenguaje y el habla. Incluye dos tipos principales de análisis: análisis sintáctico y análisis semántico. El análisis sintáctico determina el significado de una palabra, frase u oración al analizar la sintaxis de las palabras y aplicar reglas gramaticales preprogramadas. El análisis semántico emplea la salida sintáctica para extraer significado de las palabras e interpretar su significado dentro de la estructura de la oración.

El análisis de palabras puede tomar una de dos formas. El análisis de dependencias analiza las relaciones entre palabras, como la identificación de sustantivos y verbos, mientras que el análisis de constituyentes construye un árbol de análisis (o árbol de sintaxis): una representación enraizada y ordenada de la estructura sintáctica de la oración o cadena de palabras. Los árboles de análisis resultantes subyacen a las funciones de los language translators y el reconocimiento de voz. Idealmente, este análisis hace que el resultado, ya sea texto o voz, sea comprensible tanto para los modelos de PLN como para las personas.

El PLN funciona combinando varias técnicas computacionales para analizar, comprender y generar lenguaje humano de una manera que las máquinas puedan procesarlo, esas técnicas son las siguientes:

* Preprocesamiento de texto: El preprocesamiento de texto del PLN prepara el texto sin procesar para su análisis transformándolo en un formato que las máquinas puedan entender más fácilmente. Comienza con la tokenización, que implica dividir el texto en unidades más pequeñas, como palabras, oraciones o frases. Esto ayuda a desglosar el texto complejo en partes manejables. A continuación, se estandariza el texto al convertir todos los caracteres a minúsculas, lo que garantiza que palabras como “Apple” y “apple” se traten de la misma manera. La eliminación de palabras de detención es otro paso común, en el cual las palabras de uso frecuente como “es” o “el” se filtran porque no agregan un significado importante al texto. Derivación o lematización reducen las palabras a su forma raíz (por ejemplo, “corriendo” se convierte en “correr”), lo que facilita el análisis del lenguaje agrupando diferentes formas de la misma palabra. Además, la limpieza de texto elimina elementos no deseados, como puntuación, caracteres especiales y números que pueden saturar el análisis. Luego del preprocesamiento, el texto está limpio, estandarizado y listo para que los modelos de aprendizaje automático lo interpreten de manera efectiva.
* Extracción de características: La extracción de características es el proceso de convertir texto sin procesar en representaciones numéricas que las máquinas pueden analizar e interpretar. Esto implica transformar texto en datos estructurados mediante el uso de técnicas de PLN como Bag of Words y TF-IDF, que cuantifican la presencia e importancia de las palabras en un documento. Los métodos más avanzados incluyen incrustaciones de palabras como Word2Vec o GLOve, que representan palabras como vectores densos en un espacio continuo, capturando las relaciones semánticas entre palabras. Las incrustaciones contextuales mejoran aún más esto al considerar el contexto en el que aparecen las palabras, lo que permite tener mejores representaciones y más matizadas.
* Análisis de textos: El análisis de textos consiste en interpretar y extraer información significativa de los datos textuales mediante diversas técnicas computacionales. Este proceso incluye tareas como el etiquetado de partes del discurso (POS), que identifica las funciones gramaticales de las palabras, y el reconocimiento de entidades nombradas (NER), que detecta entidades específicas como nombres, lugares y fechas. El análisis sintáctico de dependencias analiza las relaciones gramaticales entre las palabras para comprender la estructura de las frases, mientras que el análisis de sentimientos determina el tono emocional del texto, evaluando si es positivo, negativo o neutro. El modelado de temas identifica los temas subyacentes en un texto o en un corpus de documentos. La comprensión del lenguaje natural (NLU) es un subconjunto de NLP que se centra en analizar el significado detrás de las oraciones. La NLU permite al software encontrar significados similares en distintas frases o procesar palabras que tienen significados diferentes. Mediante estas técnicas, el análisis de textos de PLN transforma el texto no estructurado en insights.
* Entrenamiento de modelos: a continuación, los datos procesados se emplean para entrenar modelos de machine learning, que aprenden patrones y relaciones dentro de los datos. Durante el entrenamiento, el modelo ajusta sus parámetros para reducir los errores y mejorar su rendimiento. Una vez entrenado, el modelo puede emplearse para hacer predicciones o generar resultados con datos nuevos y desconocidos. La eficacia de los modelos de PLN se perfecciona continuamente mediante la evaluación, la validación y el ajuste para mejorar la precisión y la pertinencia en las aplicaciones del mundo real. Diferentes entornos de software son útiles a lo largo de dichos procesos. Por ejemplo, el Natural Language Toolkit (NLTK) es una suite de bibliotecas y programas para inglés escrito en el lenguaje de programación Python. Admite funciones de clasificación de texto, tokenización, derivación, etiquetado, análisis y razonamiento semántico. TensorFlow es una biblioteca de software gratuita y de código abierto para machine learning e IA que se puede emplear para entrenar modelos para aplicaciones de PLN. Existen numerosos tutoriales y certificaciones para aquellos interesados en familiarizarse con estas herramientas.

***Robótica:*** La robótica es una rama de la ingeniería y la informática que implica la concepción, diseño, fabricación y operación de robots. El objetivo del campo de la robótica es crear máquinas inteligentes que puedan ayudar a los humanos de diversas maneras.

La robótica puede adoptar diversas formas. Un robot podría parecerse a un humano o tener la forma de una aplicación robótica, como por ejemplo automatización robótica de procesos, que simula cómo los humanos interactúan con el software para realizar tareas repetitivas basadas en reglas.

Si bien el campo de la robótica y la exploración de los usos potenciales y la funcionalidad de los robots han crecido sustancialmente en el siglo XXI, la idea ciertamente no es nueva.

Los robots y la inteligencia artificial van de la mano, debido a que los robots existen gracias a un software desarrollado en IA.

Equipar a los robots con herramientas de IA, les permite aprender y tomar decisiones de manera autónoma y en tiempo real gracias al uso de algoritmos y técnicas que les hace procesar información de los sensores que les conecta con su entorno. Así, pueden comprender lo que les rodea para moverse y actuar de forma correcta.

La integración de ambas disciplinas ha impulsado el desarrollo de sistemas robóticos cada vez más avanzados y autónomos, ya que su ámbito de colaboración será cada vez mayor.

La robótica también puede contribuir al desarrollo de la IA, proporcionando datos y situaciones del mundo real para entrenar y mejorar los algoritmos de aprendizaje automático. Además, los robots pueden ser utilizados como plataformas para probar nuevas técnicas de IA y aprendizaje por refuerzo.