

**Metodologías en el aprendizaje automático para el análisis de imágenes: avances,
aplicaciones y perspectivas**

Ana María Arias Ramírez

Asesor

Jorge Enrique Arboleda Puerta

Universidad Nacional Abierta y a Distancia – UNAD

Escuela de Ciencias, Tecnología e Ingenierías ECBTI

Ingeniería Electrónica

2025

Resumen

La monografía aborda el análisis de imágenes mediante aprendizaje automático, explorando tanto modelos clásicos como avanzados de aprendizaje profundo. Se analizan los fundamentos de la visión por computadora y su aplicación en áreas como clasificación, segmentación y reconocimiento de objetos. Se destacan las metodologías tradicionales que cimentaron el aprendizaje automático, y cómo el aprendizaje profundo, especialmente las redes neuronales, ha mejorado la capacidad de interpretar imágenes. Además, se examinan los desafíos éticos y técnicos, como el sesgo de los modelos, la comprensión de los resultados y la seguridad de los datos.

Las conclusiones resaltan la importancia de evaluar los desafíos en este campo, reconociendo que, aunque se han logrado avances significativos, aún existen áreas por mejorar. Se enfatiza la necesidad de una formación sólida en matemáticas y programación, así como una comprensión ética en la implementación de estas tecnologías, subrayando la evolución constante y el potencial prometedor del campo.

Palabras Clave: Aprendizaje automático, digital, inteligencia artificial.

Abstract

The monograph addresses image analysis through machine learning, exploring both classical and advanced deep learning models. It examines the fundamentals of computer vision and its application in areas such as classification, segmentation, and object recognition. The traditional methodologies that laid the foundation for machine learning are highlighted, as well as how deep learning, especially neural networks, has improved the ability to interpret images. Additionally, ethical and technical challenges are explored, such as model bias, the understanding of results, and data security.

The conclusions emphasize the importance of evaluating the challenges in this field, acknowledging that, although significant progress has been made, there are still areas that need improvement. The need for a strong foundation in mathematics and programming, as well as an ethical understanding in the implementation of these technologies, is stressed, underscoring the constant evolution and promising potential of the field.

Keywords: Machine learning, digital, artificial intelligence.

Tabla de Contenido

Introducción.....	8
Justificación	10
Objetivos	11
Objetivo General.....	11
Objetivos Específicos.....	11
Estado Del Arte.....	12
Fundamentos del Aprendizaje Automático.....	14
Tipos de Aprendizaje.....	16
<i>Aprendizaje Supervisado</i>	<i>16</i>
<i>Aprendizaje No Supervisado</i>	<i>19</i>
<i>Aprendizaje Semi-Supervisado</i>	<i>20</i>
<i>Aprendizaje por Refuerzo</i>	<i>21</i>
<i>Transfer Learning</i>	<i>21</i>
Algoritmos de Aprendizaje Automático	25
<i>Regresión Lineal.....</i>	<i>25</i>
<i>Regresión Logística.....</i>	<i>28</i>
<i>Máquinas de Soporte Vectorial (SVM)</i>	<i>32</i>
<i>K-Vecinos más Cercanos (KNN).....</i>	<i>36</i>
<i>Árboles de Decisión.....</i>	<i>40</i>
<i>Random Forest.....</i>	<i>42</i>
<i>AdaBoost.....</i>	<i>45</i>

<i>Q-Learning</i>	47
<i>Actor-Critic Methods</i>	48
<i>Métodos de Etiquetado Propagado (Label Propagation)</i>	50
<i>Aprendizaje Profundo</i>	51
<i>Redes neuronales convolucionales (CNN)</i>	54
<i>Redes Generativas Adversarias (GAN)</i>	59
<i>Análisis De Imágenes Por Aprendizaje Automático</i>	60
<i>Segmentación de imágenes</i>	60
<i>Clasificación de imágenes</i>	61
Aplicaciones Prácticas	64
Casos de estudio y ejemplos reales	65
<i>Aplicaciones Practicas</i>	65
<i>Medicina</i>	66
<i>Seguridad</i>	69
<i>Automoción</i>	71
<i>Agricultura</i>	72
<i>Desafíos y Limitaciones</i>	74
Discusión	76
Conclusiones y Recomendaciones	79
Referencias Bibliográficas	82

Lista de Tablas

Tabla 1 <i>Algorithm 2: Regresión Logística</i>	30
Tabla 2 <i>Algorithm 3: SVM Simplificado</i>	34
Tabla 3 <i>Algorithm 4: K-Vecinos más Cercanos (KNN)</i>	39

Lista de Figuras

Figura 1 <i>Modelo de Aprendizaje Supervisado</i>	18
Figura 2 <i>Ilustración del Transfer Learning</i>	22
Figura 3 <i>Ejemplo de Regresión Lineal</i>	27
Figura 4 <i>Regresión Lineal vs. Regresión Logística</i>	29
Figura 5 <i>Plano Generado por una SVM</i>	33
Figura 6 <i>Ilustración Sobre el Funcionamiento de KNN</i>	37
Figura 7 <i>Ilustración de un Árbol de Decisión</i>	41
Figura 8 <i>Ejemplo de Random Forest</i>	43
Figura 9 <i>Red Neuronal Profunda</i>	51
Figura 10 <i>Modelo de Neurona Artificial</i>	53
Figura 11 <i>Ejemplo Red Neuronal Convolutacional</i>	55
Figura 12 <i>Modelo de Red Neuronal Recurrente</i>	57
Figura 13 <i>Análisis y Mapa de características del TIL (Linfocitos Infiltrantes de Tumor)</i>	67
Figura 14 <i>Reconocimiento Facial en Zonas Publicas</i>	70
Figura 15 <i>Dos Enfoques para el Módulo de Percepción en Vehículos Autónomos</i>	72

Introducción

Uno de los campos que ha tomado mayor relevancia en los últimos años es la inteligencia artificial y sus diferentes aplicaciones. Entre ellas, el aprendizaje automático aplicado al análisis de imágenes destaca como uno de los más llamativos. Es importante distinguir entre el procesamiento de imágenes y la visión por computadora: el procesamiento de imágenes se enfoca en manipular y mejorar imágenes digitales, ajustando parámetros como el brillo, el contraste o eliminando el ruido, sin interpretar su contenido. En cambio, la visión por computadora va un paso más allá, permitiendo que las máquinas comprendan e interpreten automáticamente el contenido visual, realizando tareas como la detección de objetos, el reconocimiento de patrones y la segmentación de escenas. Así, la visión por computadora representa una convergencia entre la informática y la interpretación automática de imágenes, desafiando los límites tradicionales y abriendo nuevas posibilidades para la comprensión del contenido visual

En la actualidad, el crecimiento exponencial de los datos visuales provenientes de diversas fuentes, como cámaras digitales, sensores y plataformas de redes sociales, ha impulsado la necesidad de herramientas avanzadas capaces de procesar y extraer información significativa de estas imágenes de manera eficiente (Goodfellow, Bengio & Courville, 2016). En este contexto, el empleo de algoritmos de inteligencia artificial, en particular técnicas de aprendizaje automático, se ha consolidado como un enfoque central para abordar los desafíos asociados con el análisis y la interpretación de imágenes a gran escala (Russell & Norvig, 2021). El principal problema de investigación radica en comprender cómo las técnicas de aprendizaje automático y semi-automático pueden optimizar el análisis de imágenes, lo que implica abordar cuestiones clave como la precisión en la identificación de patrones visuales, la capacidad de adaptación a diversas condiciones

y la eficiencia en el procesamiento de grandes volúmenes de datos visuales (Zhou et al., 2021). Además, se busca explorar la aplicación de estas tecnologías en áreas específicas como la medicina, la vigilancia y la agricultura, entre otras. Por ello, el objetivo principal es analizar las metodologías de aprendizaje automático más relevantes para el análisis de imágenes, investigando sus aplicaciones actuales y potenciales en campos específicos relacionados con el procesamiento y la interpretación de datos visuales. Finalmente, se pretende determinar el rendimiento y la eficacia de los modelos de aprendizaje automático en comparación con enfoques convencionales para el análisis de imágenes, mediante el uso de métricas y técnicas adecuadas que permitan una comparación objetiva y rigurosa.

La creciente importancia del análisis de imágenes impulsado por inteligencia artificial en diversos sectores ha generado un notable interés en individuos y empresas, que lo consideran un campo emergente con potencial para desarrollar nuevos modelos económicos (IMF, 2024). La capacidad de las máquinas para interpretar y extraer conocimiento de datos visuales no solo tiene implicaciones significativas en términos de eficiencia y precisión, sino que también puede impactar positivamente la toma de decisiones, el diagnóstico médico, la seguridad y otros aspectos fundamentales de la vida cotidiana (Russell & Norvig, 2021). A través de este trabajo se busca contribuir a la comprensión y aplicación efectiva de estas tecnologías en el análisis de imágenes, explorando su potencial transformador en la sociedad contemporánea.

Justificación

La inteligencia artificial, y en particular su aplicación al procesamiento de imágenes, ha revolucionado múltiples sectores, generando un impacto significativo en áreas como la medicina, donde facilita el diagnóstico por imágenes; la seguridad, mediante sistemas avanzados de reconocimiento facial; y la automoción, con el desarrollo de vehículos autónomos, entre otros. Estas aplicaciones no solo mejoran la eficiencia y precisión de los procesos, sino que también abren nuevas posibilidades para la innovación tecnológica y el desarrollo económico.

A pesar de estos avances, el campo enfrenta desafíos importantes que limitan su pleno potencial. Entre ellos se encuentran la necesidad de grandes volúmenes de datos etiquetados para entrenar modelos precisos, lo que implica un alto costo y esfuerzo en la recolección y anotación de datos. Además, el elevado consumo computacional de los algoritmos de aprendizaje automático representa una barrera para su implementación en entornos con recursos limitados. Otro reto crucial es la presencia de sesgos en los modelos, que puede afectar la equidad y la confiabilidad de los sistemas, generando resultados discriminatorios o erróneos. Finalmente, los problemas relacionados con la privacidad y la protección de datos personales son cada vez más relevantes, dado el uso masivo de imágenes que pueden contener información sensible.

Abordar estos desafíos es fundamental para avanzar en el desarrollo de soluciones más robustas, eficientes y éticas en el procesamiento de imágenes mediante inteligencia artificial. Superar estas barreras permitirá no solo mejorar la calidad y aplicabilidad de los sistemas actuales, sino también ampliar su impacto en sectores clave, contribuyendo a un uso responsable y beneficioso de estas tecnologías en la sociedad.

Objetivos

Objetivo General

Analizar las principales metodologías de aprendizaje automático aplicadas al análisis de imágenes, destacando sus avances, aplicaciones y perspectivas futuras.

Objetivos Específicos

Describir los principales algoritmos de aprendizaje automático utilizados en el procesamiento y análisis de imágenes.

Examinar el impacto de estos algoritmos en áreas clave como la medicina, la seguridad y la automatización.

Identificar los desafíos actuales y las oportunidades de mejora en el uso de aprendizaje automático para el análisis de imágenes.

Comparar el rendimiento y la eficacia de los modelos de aprendizaje automático con enfoques convencionales para el análisis de imágenes, mediante métricas y métodos objetivos.

Estado Del Arte

En los últimos años, el campo del análisis de imágenes mediante el aprendizaje automático ha experimentado avances significativos, impulsados principalmente por el desarrollo de nuevas arquitecturas y algoritmos. Los enfoques basados en el aprendizaje profundo han revolucionado el procesamiento de imágenes, permitiendo a los modelos aprender representaciones complejas directamente a partir de los datos crudos, eliminando la necesidad de preprocesamiento manual intensivo (StarTechUp, 2023).

Una de las principales metodologías utilizadas para el análisis de imágenes son las redes neuronales convolucionales (CNNs, por sus siglas en inglés). Las CNNs han demostrado ser altamente efectivas en tareas como clasificación de imágenes, detección de objetos y segmentación. Modelos como AlexNet, VGG, ResNet y EfficientNet han evolucionado en términos de profundidad y eficiencia computacional, mejorando la capacidad de generalización en grandes conjuntos de datos como ImageNet (FocalX, 2025).

Otra metodología importante y relevante son las denominadas Redes Generativas Antagónicas (GANs). Las GANs han abierto nuevas fronteras en la síntesis de imágenes y la creación de modelos realistas. Aplicaciones como la generación de imágenes sintéticas, la mejora de la resolución de imágenes (superresolución) y la conversión de imágenes entre dominios (por ejemplo, de fotos a bocetos o de verano a invierno) son algunas de las contribuciones clave de las GANs en la visión por computadora (FocalX, 2025).

Recientemente, los Transformers, que han tenido un gran impacto en el procesamiento del lenguaje natural, se han adaptado con éxito al análisis de imágenes. Los Transformers de Visión (ViT) han demostrado resultados competitivos con CNNs tradicionales en tareas de clasificación de imágenes, logrando un rendimiento notable sin necesidad de operaciones convolucionales, lo que apunta a un cambio de paradigma en la

arquitectura de los modelos (StarTechUp, 2023).

Una de las principales aplicaciones prácticas del análisis de imágenes mediante el uso de aprendizaje automático proviene del campo de la medicina, particularmente en radiología y en la patología digital, que implica la obtención, gestión e interpretación de imágenes de tejidos para el diagnóstico (Litjens et al., 2017). Modelos de aprendizaje profundo se utilizan para detectar enfermedades como el cáncer en imágenes de resonancia magnética, tomografías computarizadas y radiografías.

Estos sistemas mejoran la precisión del diagnóstico, reducen los errores humanos y aceleran el tiempo de evaluación (Choy et al., 2018). En radiología, el aprendizaje automático permite la detección automatizada de hallazgos como nódulos pulmonares, fracturas y tumores, facilitando la interpretación de imágenes complejas y apoyando la toma de decisiones clínicas (McBee et al., 2018). En patología digital, los algoritmos ayudan a identificar patrones en imágenes histológicas que pueden ser indicativos de distintas enfermedades, mejorando la eficiencia y precisión diagnóstica (Litjens et al., 2017).

El aprendizaje automático ha emergido como una disciplina clave dentro de la inteligencia artificial, proporcionando modelos capaces de aprender y mejorar su rendimiento a partir de datos mediante procesos matemáticos y programación compleja desarrollada por expertos, más que por una simple configuración de datos por parte del usuario. Una de las áreas de aplicación más destacadas es el procesamiento de imágenes, donde los algoritmos permiten analizar, clasificar y extraer información relevante con un nivel de precisión sin precedentes.

El procesamiento de imágenes mediante aprendizaje automático ha sido aplicado en numerosos campos, como la detección temprana de cáncer. Redes profundas como U-Net han sido utilizadas para segmentar imágenes de tejidos con alta precisión, especialmente en

imágenes de resonancia magnética, lo que ha contribuido a reducir significativamente los tiempos de diagnóstico en cuanto a la segmentación y localización de tumores malignos (Quiles-Fernández et al., 2023). Otro sector que se ha beneficiado es el agrícola, donde el análisis de imágenes capturadas por drones y satélites permite monitorear el estado de los cultivos y detectar enfermedades en plantas. Estudios recientes han aplicado aprendizaje automático para analizar índices de vegetación como el NDVI, logrando identificar estrés hídrico y deficiencias nutricionales en tiempo real (Kamilaris & Prenafeta-Boldú, 2018). En la industria manufacturera, las redes neuronales convolucionales (CNN) se utilizan para inspeccionar productos y detectar defectos, mejorando la calidad y eficiencia de los procesos productivos.

Fundamentos del Aprendizaje Automático

El Aprendizaje Automático (también conocido como Machine Learning) ha emergido como un campo fundamental dentro de la inteligencia artificial, revolucionando la manera en que los sistemas computacionales procesan y toman decisiones basadas en datos complejos que requieren análisis. En este capítulo se explorarán los fundamentos esenciales del Aprendizaje Automático.

El Aprendizaje Automático (también conocido como Machine Learning) es un campo fundamental dentro de la inteligencia artificial que se centra en el desarrollo de modelos matemáticos y estadísticos capaces de identificar patrones en datos complejos y tomar decisiones basadas en esos patrones, sin que el sistema "comprenda" en sentido humano, sino mediante cálculos y optimizaciones (Goodfellow, Bengio & Courville, 2016). En este capítulo se explorarán los fundamentos esenciales del Aprendizaje Automático.

Para comprender su funcionamiento, es importante destacar que existen diversas

técnicas y metodologías. Los modelos son representaciones matemáticas que capturan patrones y relaciones dentro de los datos. Estos modelos pueden ser supervisados o no supervisados, según la disponibilidad de etiquetas en los datos de entrenamiento. El aprendizaje supervisado implica entrenar un modelo con datos etiquetados para predecir salidas en nuevos datos, siendo común en tareas de clasificación y regresión. Por otro lado, el aprendizaje no supervisado trabaja con datos no etiquetados, descubriendo patrones ocultos, con aplicaciones como el clustering y la reducción de dimensionalidad (Bishop, 2006).

Los algoritmos son conjuntos de instrucciones o procedimientos utilizados para entrenar estos modelos. Aunque se dice que un sistema "aprende", en realidad es el algoritmo el que ajusta los parámetros del modelo mediante procesos matemáticos y estadísticos para minimizar errores en las predicciones (Cormen, Leiserson, Rivest & Stein, 2022). La efectividad de un modelo se mide a través de métricas específicas y técnicas como la validación cruzada y matrices de confusión, que permiten evaluar su rendimiento y capacidad de generalización (Kohavi, 1995). La optimización del modelo se realiza ajustando sus parámetros para minimizar funciones de pérdida que cuantifican la discrepancia entre las predicciones y los valores reales (Bottou, Curtis & Nocedal, 2018).

De forma general, el proceso de aprendizaje automático se puede dividir en dos etapas principales: entrenamiento y evaluación. El entrenamiento implica exponer al modelo a datos de entrada y ajustar sus parámetros para mejorar su rendimiento en la tarea específica. Posteriormente, la evaluación verifica la capacidad del modelo para generalizar y funcionar correctamente con datos no vistos durante el entrenamiento. Es importante considerar en esta etapa los fenómenos de *overfitting* y *underfitting*: el *overfitting* ocurre

cuando un modelo se ajusta demasiado a los datos de entrenamiento, perdiendo capacidad de generalización a nuevos datos, mientras que el *underfitting* se refiere a modelos demasiado simples que no logran capturar patrones complejos presentes en los datos (Géron, 2022).

Cabe destacar que el aprendizaje automático, a diferencia del aprendizaje humano, no implica comprensión ni razonamiento consciente. Más bien, se trata de un proceso matemático y estadístico que permite a los sistemas computacionales tomar decisiones basadas en patrones extraídos de los datos. Aunque se utiliza el término "aprendizaje", este proceso es fundamentalmente distinto al aprendizaje de un ser vivo, ya que no implica conciencia ni experiencia subjetiva, sino la optimización de funciones matemáticas para mejorar el desempeño en tareas específicas.

Tipos de Aprendizaje

El análisis de imágenes con ayuda de herramientas de inteligencia artificial abarca diversas metodologías de aprendizaje que permiten a los sistemas computacionales extraer patrones, características y conocimientos relevantes de datos visuales. Para analizar cualquier tipo de conjunto de datos de los cuales se requiere extraer información, generalmente se aplican métodos ya desarrollados y de efectividad comprobada, a continuación, se muestran los más conocidos y extendidos.

Aprendizaje Supervisado

En el aprendizaje supervisado, el modelo se entrena utilizando un conjunto de datos que contiene ejemplos de entrada emparejados con salidas deseadas; es decir, para unas entradas conocidas, se conoce la respuesta correcta para cada caso. En el análisis de imágenes, esto implica proporcionar algoritmos con imágenes etiquetadas, donde cada

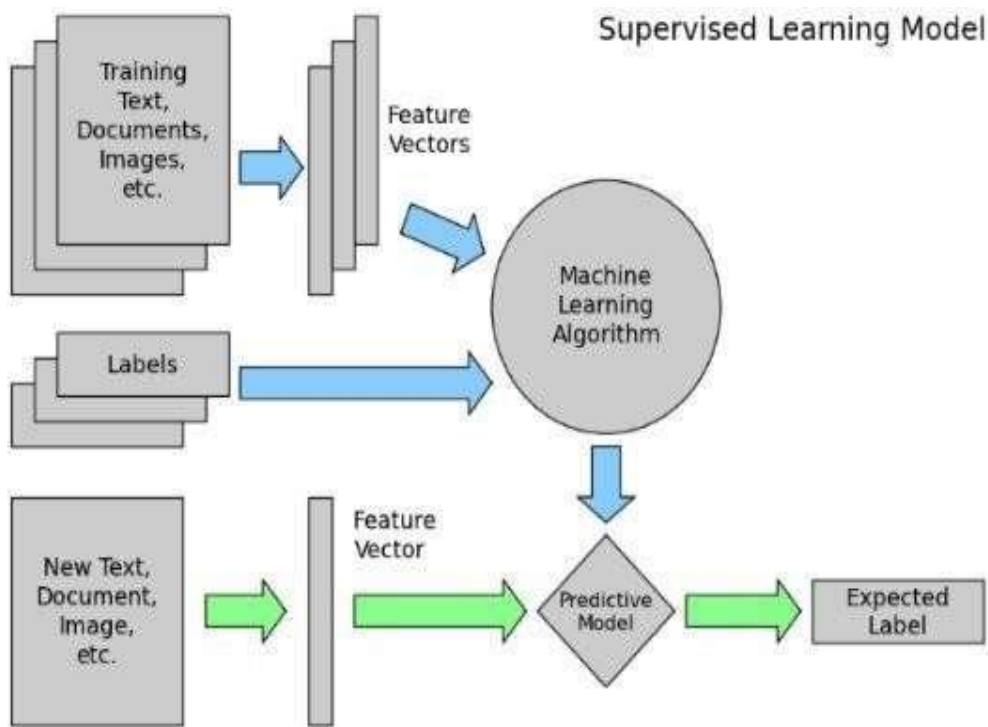
imagen tiene asociada una categoría o clase específica.

Este enfoque es esencial para tareas de clasificación y regresión, donde el objetivo es asignar una etiqueta predefinida o predecir un valor numérico para nuevas imágenes no vistas durante el entrenamiento (Johnston & Mathur, 2019, p. 45).

Entre los modelos más utilizados en aprendizaje supervisado para análisis de imágenes se encuentran:

- Máquinas de soporte vectorial (SVM), que encuentran el hiperplano óptimo para separar clases en espacios de alta dimensión.
- K-vecinos más cercanos (KNN), que clasifican nuevas instancias según la mayoría de las clases de sus vecinos más cercanos.
- Redes neuronales convolucionales (CNN), ampliamente usadas para clasificación, detección y segmentación de imágenes.
- Modelos basados en campos aleatorios de Markov (MRF), que modelan relaciones espaciales en imágenes para tareas como segmentación.
- Estimación de volumen parcial (**Partial Volume Estimation**, PVE), utilizada para mejorar la precisión en la segmentación de imágenes médicas.

Estos modelos, junto con otros algoritmos supervisados, permiten construir sistemas capaces de reconocer patrones complejos en imágenes, facilitando aplicaciones en medicina, seguridad, agricultura y otros campos (Bishop, 2006, pp. 320-325).

Figura 1*Modelo de Aprendizaje Supervisado*

Nota. Gráfico modelo de aprendizaje supervisado. Tomado de. Nasteski. (2017)

La Figura 1 muestra el proceso general para el funcionamiento del aprendizaje supervisado; de los datos de entrada se extrae un vector de características (*feature vector*) y se entregan sus respectivas etiquetas (*lables*), esto pasa por el algoritmo de aprendizaje máquina lo cual genera un modelo predictor. Posteriormente, se entrega un nuevo conjunto de datos de los cuales se extrae un vector de características y el modelo predictor es capaz de generar las etiquetas para realizar la clasificación de cada uno de los componentes entregados en la base de datos.

El aprendizaje supervisado se utiliza comúnmente en aplicaciones como la clasificación de correos electrónicos como spam o no spam, o en la predicción de

enfermedades a partir de datos médicos. Los sistemas de recomendación, como los que utilizan plataforma de *streaming* para sugerir películas o música, también son ejemplos de aplicaciones prácticas del aprendizaje supervisado, donde un algoritmo aprende a partir de datos etiquetados para realizar predicciones precisas.

Aprendizaje No Supervisado

El aprendizaje no supervisado se emplea cuando se busca descubrir patrones intrínsecos en conjuntos de datos de imágenes sin etiquetas previas. A diferencia del aprendizaje supervisado, aquí los modelos trabajan con datos sin clasificar y sin salidas conocidas, lo que obliga a los algoritmos a identificar estructuras subyacentes por sí mismos. Entre las aplicaciones más comunes en el análisis de imágenes se encuentran el *clustering* (agrupamiento) y la reducción de dimensionalidad.

Existen diversas técnicas de *clustering* que permiten organizar datos visuales en grupos significativos según sus similitudes. Entre las más importantes están:

- **K-means.** Divide los datos en un número predefinido de grupos, buscando minimizar la distancia dentro de cada clúster.
- **DBSCAN.** Agrupa datos basándose en la densidad, identificando regiones densas y etiquetando como ruido los puntos aislados.
- **Algoritmos jerárquicos.** Crean una estructura en forma de árbol que agrupa datos en niveles según sus características.

Además, técnicas probabilísticas y modelos basados en densidad también se emplean para descubrir patrones complejos.

En el análisis de imágenes, también se utilizan métodos no supervisados para la extracción automática de características, que facilitan la reducción de dimensionalidad y la

representación eficiente de los datos visuales. Estas técnicas permiten simplificar la información manteniendo la esencia de las imágenes, lo que es fundamental para tareas como la segmentación y clasificación sin supervisión.

El aprendizaje no supervisado es especialmente útil en escenarios donde no se dispone de datos etiquetados o cuando se busca explorar y entender la estructura interna de grandes volúmenes de datos visuales, como en imágenes médicas, satelitales o de vigilancia (IBM, 2024; PyPro, 2024; DataCamp, 2024; OBS Business, 2025).

El aprendizaje no supervisado es útil en problemas de agrupamiento, como la segmentación de clientes en *marketing*. Las empresas utilizan este enfoque para identificar patrones ocultos en los datos de clientes y ofrecer promociones personalizadas. Otro ejemplo es la detección de anomalías en seguridad cibernética, donde los algoritmos de agrupamiento pueden identificar comportamientos inusuales que podrían indicar una amenaza.

Aprendizaje Semi-Supervisado

El aprendizaje semi-supervisado combina elementos de los enfoques supervisado y no supervisado. Este método se beneficia de conjuntos de datos que contienen tanto imágenes etiquetadas como no etiquetadas. La inclusión de datos no etiquetados permite al modelo mejorar su capacidad de generalización y adaptación a nuevos patrones, especialmente cuando la obtención de datos etiquetados es costosa o limitada.

Este tipo de aprendizaje es útil cuando se dispone de grandes cantidades de datos no etiquetados y solo una pequeña porción etiquetada. Por ejemplo, en el reconocimiento facial, es costoso etiquetar manualmente cada imagen, por lo que los algoritmos semi-supervisado pueden aprender de un conjunto limitado de imágenes etiquetadas y mejorar su precisión utilizando datos no etiquetados.

Aprendizaje por Refuerzo

El aprendizaje por refuerzo encuentra aplicación en el análisis de imágenes cuando se abordan problemas de toma de decisiones secuenciales. En este contexto, el modelo aprende a través de la interacción con un entorno visual, recibiendo recompensas o penalizaciones según las acciones tomadas. Esto es crucial para tareas como la navegación autónoma de robots basada en visión, donde el sistema debe aprender a interpretar y reaccionar a su entorno.

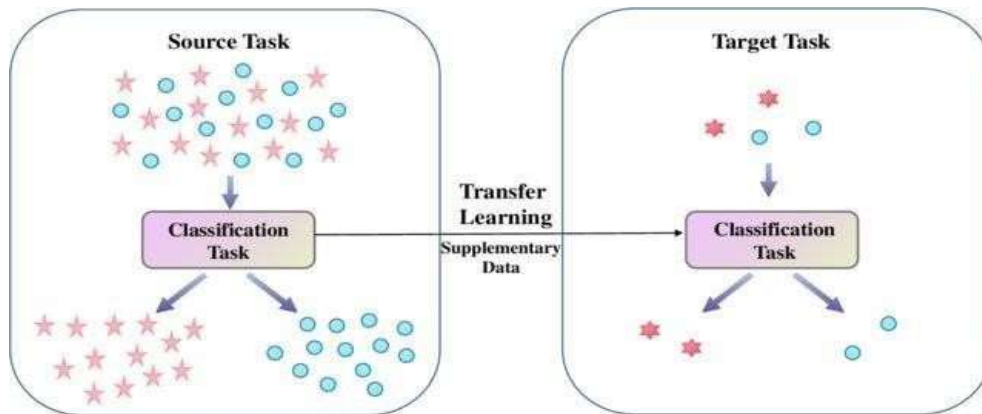
El aprendizaje por refuerzo se utiliza en la optimización de estrategias en video juegos y sistemas de conducción autónoma. Un ejemplo destacado es AlphaGo, un sistema que aprende a jugar mediante la interacción con el entorno. Además, en robots industriales, se emplea para que aprendan a realizar tareas complejas a través de prueba y error, mejorando continuamente sus acciones.

Transfer Learning

El *transfer learning* se ha vuelto esencial en el análisis de imágenes, permitiendo la reutilización de conocimientos aprendidos en un dominio para mejorar el rendimiento en otro. Este enfoque es particularmente valioso cuando se cuenta con conjuntos de datos limitados, ya que modelos pre-entrenados en grandes conjuntos de datos pueden transferir características útiles a tareas específicas de análisis de imágenes.

Figura 2

Ilustración del Transfer Learning



Nota. Ilustración que explica el proceso de Transfer Learning. Tomado de. Yang and Mei. (2021)

La figura 2 muestra un diagrama que ilustra el proceso que sigue el modelo *transfer learning* en la clasificación de tareas. La figura es un diagrama dividido en dos partes, etiquetadas como “*Source Task*” y “*Target Task*”; en “*Source Task*” hay dos grupos, estrellas y círculos. Hay una caja etiquetada como “*Classification Task*” conectada a los dos grupos mencionados; esta sección del diagrama hace referencia al modelo de clasificación utilizado. En el medio del diagrama, hay una etiqueta denominada “*Transfer Learning Supplementary Data*” que conecta la tarea fuente con la tarea objetivo.

El término “*Transfer Learning Supplementary Data*” se refiere a los datos adicionales que se utilizan para mejorar el rendimiento del modelo de aprendizaje automático en la tarea de clasificación. El aprendizaje por transferencia es una técnica de aprendizaje automático que utiliza conocimientos previos adquiridos en una tarea para mejorar el rendimiento en otra tarea relacionada; los datos adicionales se utilizan para mejorar la

capacidad del modelo para generalizar y mejorar el rendimiento en la tarea objetivo (Hosna et al., 2022). Finalmente, en “*Target Task*”, hay menos datos presentes con solo unos pocos íconos de estrellas y círculos visibles; en esta parte se entregan nuevos datos para que el clasificador aplique el aprendizaje obtenido en los pasos anteriores.

El análisis de imágenes en el aprendizaje automático involucra diversas estrategias, cada una con sus fortalezas y limitaciones. En esta evaluación (tabla 1), se examinan los principales enfoques del aprendizaje automático, aprendizaje supervisado (S), no supervisado (NS), semi-supervisado (SS), por refuerzo (R) y transfer learning (T); con el fin de distinguir cual ofrece las mejores prestaciones para su implementación en problemáticas que requieran el análisis de imágenes.

Tabla 1

Comparación entre Tipos de Aprendizaje Automático

Aspecto	S	NS	SS	R	T
Utilización de etiquetas	5	0	4	0	3
Adaptabilidad a estructuras	4	4	4	3	4
Manejo de grandes conjuntos	3	4	4	4	4
Sensibilidad a calidad de representación	2	4	3	3	4
Aplicabilidad en entornos dinámicos	3	3	3	5	3

Nota. Esta tabla explica la Comparación entre Tipos de Aprendizaje Automático

Los aspectos evaluados son:

Utilización De Etiquetas. Mide la dependencia de información etiquetada para el entrenamiento. El aprendizaje supervisado requiere datos completamente etiquetados, mientras que el no supervisado no utiliza etiquetas.

Adaptabilidad A Estructuras Complejas. Evalúa la capacidad del método para modelar patrones complejos presentes en los datos visuales, como texturas, formas y relaciones espaciales.

Manejo De Grandes Conjuntos De Datos. Considera la eficiencia computacional y la escalabilidad del método al procesar grandes volúmenes de datos. Actualmente, “grandes volúmenes” se refieren a conjuntos con miles o millones de imágenes o muestras, dependiendo del dominio y la capacidad computacional disponible (Google Cloud, 2024; Iberdrola, 2023).

Sensibilidad A Calidad De Representación. Evalúa la robustez del método ante variaciones en la calidad de las imágenes, como ruido, baja resolución o condiciones de iluminación variables. La sensibilidad se determina mediante pruebas empíricas en datasets con diferentes niveles de calidad y mediante métricas de desempeño (Zendesk, 2020; Lisdatasolutions, 2022).

Aplicabilidad en entornos dinámicos. Se refiere a la capacidad del método para adaptarse a cambios en el entorno o en la distribución de datos, como variaciones temporales o contextuales frecuentes.

Cada aspecto se evalúa en una escala de 0 a 5, donde 0 indica el rendimiento más bajo y 5 el rendimiento más alto. La tabla presenta una visión resumida de cómo cada enfoque aborda estos aspectos clave, proporcionando una guía para la selección según las

necesidades específicas del análisis de imágenes. La elección del modelo dependerá de la disponibilidad de etiquetas, la necesidad de adaptabilidad a entornos dinámicos y la sensibilidad a la calidad de la representación. La combinación de enfoques, como el uso de *transfer learning* en conjunto con el aprendizaje semisupervisado, podría ofrecer una solución equilibrada para diversas aplicaciones en el análisis de imágenes.

El aprendizaje por transferencia (*transfer learning*) es muy popular en la visión por computadora y el procesamiento del lenguaje natural. Por ejemplo, en la detección de objetos, un modelo pre-entrenado en un gran conjunto de datos puede adaptarse a un conjunto de imágenes más pequeño para identificar objetos específicos. Este enfoque acelera el entrenamiento y mejora la precisión, incluso con menos datos.

Algoritmos de Aprendizaje Automático

En el campo del aprendizaje automático, la importancia de los distintos algoritmos juega un papel fundamental, ya que son estos los que permiten realizar cada una de las tareas propuestas para cada enfoque o necesidad particular. Estos algoritmos, concebidos para la identificación de patrones, la realización de predicciones y la toma de decisiones autónomas, engloban una diversidad sustancial de enfoques y métodos.

Regresión Lineal

La Regresión Lineal es un método fundamental en estadística y aprendizaje automático utilizado para modelar la relación lineal entre una variable dependiente y una o más variables independientes. Su objetivo principal es establecer un modelo matemático que describa la tendencia general de los datos, permitiendo realizar predicciones y comprender la naturaleza de la relación entre las variables.

El modelo de regresión lineal parte de la ecuación de la recta, pero su forma general

incluye un término de error para capturar la variabilidad no explicada por las variables independientes. En su forma más simple, con una variable independiente (regresión lineal simple), el modelo se expresa como:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p. \quad (1)$$

donde y es el valor estimado de la variable dependiente, x es la variable independiente, $\beta_1 x_1$ es la pendiente que indica el cambio esperado en y por unidad de cambio en x , y β_0 es la intersección con el eje y , es decir, el valor estimado de y cuando $x=0$ (Montgomery, Peck & Vining, 2021).

Para casos con múltiples variables independientes (regresión lineal múltiple), el modelo se generaliza como:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n \quad (2)$$

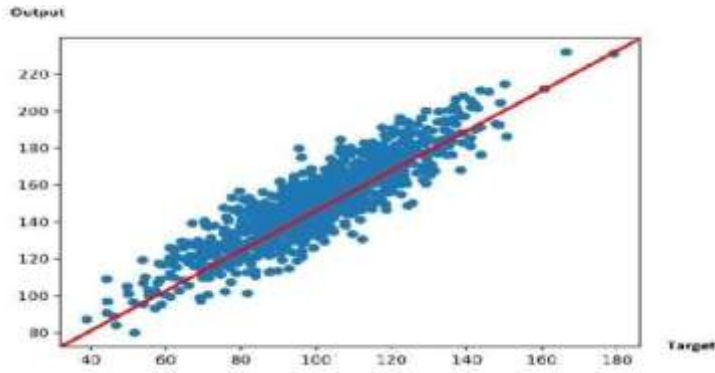
donde cada x_i representa una variable independiente y cada b_i su coeficiente asociado, que mide la influencia de esa variable en la variable dependiente.

El proceso de ajuste del modelo implica encontrar los valores óptimos para los coeficientes

b_i que minimizan la diferencia entre las predicciones del modelo y los valores observados, comúnmente mediante el método de mínimos cuadrados ordinarios. Este método busca minimizar la suma de los cuadrados de las diferencias entre los valores observados y los valores predichos por el modelo (DataCamp, 2024; AWS, 2024).

Figura 3

Ejemplo de Regresión Lineal



Algorithm 1: Regresión Logística

Data: Conjunto de datos: $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$

Result: Coeficientes de regresión: b_0, b_1

Inicialización: $b_0 \leftarrow 0;$
 $b_1 \leftarrow 0;$
 $\alpha \leftarrow$ Tasa de aprendizaje;

Entrenamiento: for $i \leftarrow 1$ to Iteraciones do

Calcular la $y_{pred} \leftarrow b_0 + b_1 x_i$; Calcular el error:

predicción: $error \leftarrow y_{pred} - y_i$; Actualizar los coeficientes: $b_0 \leftarrow b_0 - \alpha \cdot error;$
 $b_1 \leftarrow b_1 - \alpha \cdot error \cdot x_i;$

Nota. Expone de manera detallada los pasos esenciales requeridos para el modelo de regresión línea. Tomado de. Sultan Mohd. (2021)

El algoritmo referenciado como 1 expone de manera detallada cada uno de los pasos esenciales requeridos en el proceso de construcción de un modelo de regresión lineal, específicamente diseñado para un conjunto de datos dado. Cada paso dentro de este procedimiento desglosa las acciones fundamentales que se deben llevar a cabo para desarrollar un modelo que capture de manera efectiva la relación lineal entre las variables involucradas. Desde la inicialización de los parámetros hasta el entrenamiento del modelo mediante técnicas como el descenso de gradiente, este algoritmo proporciona una guía completa para implementar la regresión lineal. Cada instrucción está diseñada para maximizar la comprensión y eficacia en la construcción de un modelo que pueda realizar predicciones precisas basadas en el conjunto de datos proporcionado. Este enfoque paso a paso garantiza una implementación sistemática y rigurosa del proceso de regresión lineal, estableciendo así una base sólida para el análisis y la modelización de datos.

La regresión lineal es ampliamente utilizada en aplicaciones financieras para predecir tendencias de mercado, como el precio de las acciones o la demanda de productos. En el sector inmobiliario, este método ayuda a estimar el precio de una propiedad en función de variables como el tamaño, ubicación y número de habitaciones.

Regresión Logística

La regresión logística es un método de aprendizaje automático utilizado para modelar la probabilidad de que una variable categórica binaria asuma uno de dos valores posibles. A diferencia de la regresión lineal, que se emplea para predecir valores numéricos, la regresión logística se utiliza en problemas de clasificación binaria, donde la respuesta deseada es dicotómica, como sí/no, éxito/fracaso o positivo/negativo. Este modelo estima la probabilidad de que ocurra un evento en función de variables independientes, utilizando una

función logística para garantizar que las predicciones estén entre 0 y 1 (IBM, 2024).

La figura 4 muestra una comparación entre la regresión logística y la regresión lineal.

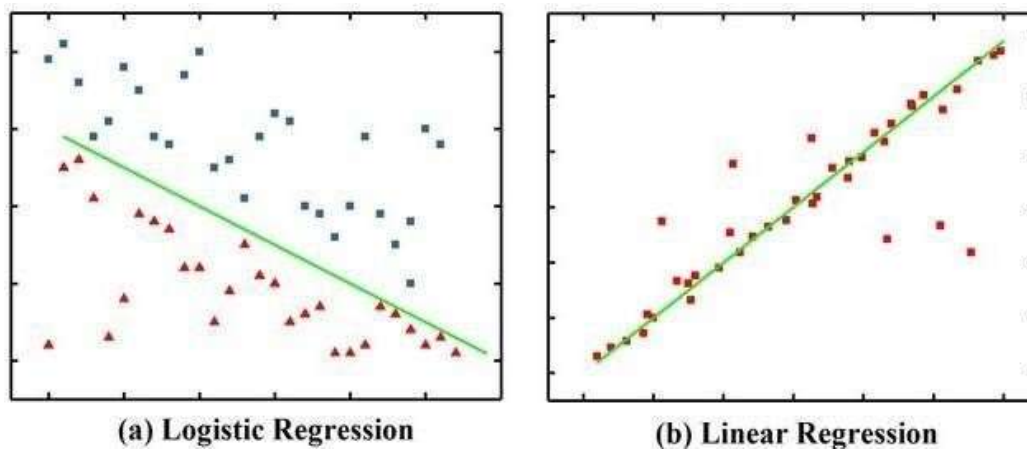
Es importante para mostrar los detalles de cada uno de estos modelos y como se deben aplicar a los diferentes casos en el análisis de datos.

El principio fundamental de la regresión logística radica en la transformación logística de la variable de respuesta. La función logística, también conocida como la función sigmoide, se define en la ecuación 3.

$$f(x) = \frac{1}{1 + e^{-(b_0 + b_1X_1 + \dots + b_nX_n)}}$$

Figura 4

Regresión Lineal vs. Regresión Logística



Nota. Tomado de. Fang. (2019)

En la ecuación 3, el término $P(Y = 1)$ es la probabilidad de que la variable de respuesta Y sea igual a 1. e es la base del logaritmo natural. b_0, b_1, \dots, b_n son los coeficientes del modelo. Finalmente, X_1, \dots, X_n son las variables predictoras. El lado derecho de la ecuación representa el log-odds¹ (logaritmo de la razón de probabilidades) de que Y sea igual

a 1; la función sigmoidea transforma este log-odds en una probabilidad que oscila entre 0 y 1. El proceso de entrenamiento de la regresión logística implica la estimación de los coeficientes del modelo de manera que maximice la verosimilitud de los datos observados. Esto generalmente se realiza mediante métodos como el descenso de gradiente; en el algoritmo 2 se muestra cada uno de los pasos generales que se deben seguir para establecer el modelo de regresión logística sobre un conjunto de datos.

La regresión logística presenta varias ventajas para el análisis de imágenes; su interpretación sencilla permite entender fácilmente el impacto de cada característica en la probabilidad de la clase, facilitando la comprensión de los resultados. Además, su eficiencia computacional la hace adecuada para conjuntos de datos grandes, y es particularmente efectiva en problemas de clasificación binaria, como identificación de condiciones

¹El logodds es el logaritmo de la razón de probabilidades y se usa en la regresión logística para

expresar de manera lineal la relación entre las variables predictoras y la probabilidad de un evento binario. Es fundamental para modelar la probabilidad en un formato que sea fácil de manipular matemáticamente.

Tabla 1

Algorithm 2: Regresión Logística

Algorithm 2: Regresión Logística	
Data:	Conjunto de datos: $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$
Resultado:	Coefficientes del modelo: b_0, b_1, \dots, b_n
Inicialización:	Iteraciones \leftarrow Número máximo de iteraciones; $\alpha \leftarrow$ Tasa de aprendizaje;

```

B0, B1, . . . , Bn ← Inicializar coeficientes;
Entrenamiento:;
Entrenamiento:   for i ← 1 to Iteraciones do
Probabilidad predicha:   $Y_i = B_0 + B_1 X_i + u_i$ 
                    Calcular el error: error ←  $P(Y = 1) - Y$  ; Actualizar
                    los coeficientes:  $B_0 \leftarrow B_0 - \alpha \cdot \text{error}$ ;  $B_1$ 
                    ←  $B_1 - \alpha \cdot \text{error} \cdot X_1$ ;
                     $B_n \leftarrow B_n - \alpha \cdot \text{error} \cdot X_n$ ;

```

Nota. Explicación regresiones médicas. Tomado de. Regresiones médicas en imágenes Hosmer Jr. (2013).

La regresión logística presenta ciertas limitaciones importantes. Su suposición de una relación lineal entre las variables predictoras y el logit de la variable dependiente puede ser restrictiva en problemas complejos donde se requieren modelos no lineales. Además, puede ser sensible a características irrelevantes o a la multicolinealidad, lo que demanda un cuidadoso proceso de selección de variables para evitar resultados inestables. En situaciones de desbalance de clases, la regresión logística puede enfrentar dificultades para aprender correctamente las características de la clase minoritaria. Asimismo, su capacidad para manejar imágenes de alta resolución es limitada, especialmente cuando se necesita capturar la complejidad de datos no lineales (Analytics Lane, 2018; ISGlobal, 2024).

Este método es útil para predecir resultados binarios. En el ámbito médico, la regresión logística se emplea para estimar la probabilidad de que un paciente desarrolle una

enfermedad, considerando factores como edad, hábitos de vida o antecedentes familiares. También es común en marketing para predecir la probabilidad de que un cliente realice una compra en línea (IBM, 2024)

²La colinealidad en estadística ocurre cuando dos o más variables predictoras están altamente correlacionadas en un modelo de regresión, afectando la precisión de las estimaciones de coeficientes. Se requieren técnicas específicas para abordar este problema y garantizar resultados más estables.

Máquinas de Soporte Vectorial (SVM)

Las Máquinas de Soporte Vectorial (*Support Vector Machine, SVM*) son un método muy utilizado en el aprendizaje automático utilizado para la clasificación y regresión. Su principal objetivo es encontrar un hiperplano en un espacio de características que mejor separe los puntos de datos de distintas clases. En el contexto de clasificación, este hiperplano sirve como una frontera de decisión óptima que maximiza el margen entre las clases.

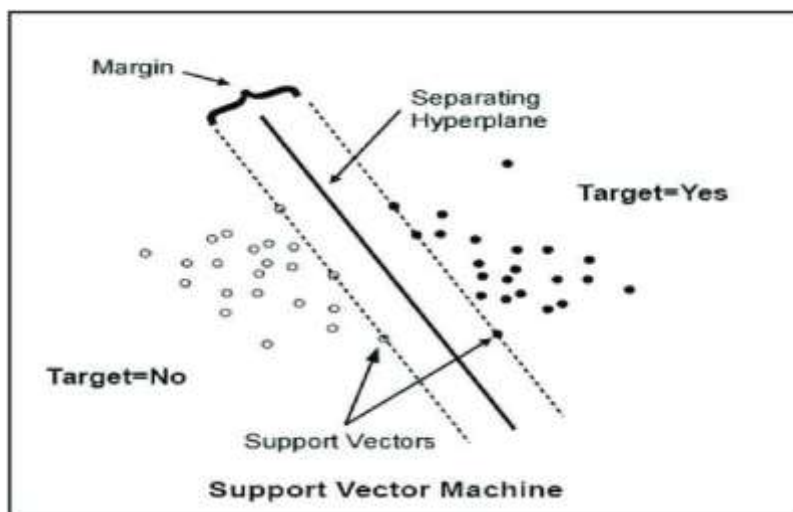
El espacio de características, en las Máquinas de Soporte Vectorial (SVM) y otros métodos de aprendizaje automático, se refiere al espacio matemático donde se representan y analizan los datos. Cada instancia o punto de datos en un conjunto de datos se describe mediante un conjunto de características, y el espacio de características es el espacio multidimensional que surge al considerar todas estas características como ejes.

La figura 5 muestra un diagrama de una Máquina de Soporte Vectorial donde se observan dos grupos de datos clasificados como “Target=Yes” y “Target=No”, separados por un hiperplano; también se indican los vectores de soporte y el margen. Hay dos grupos distintos de puntos, uno a la izquierda etiquetado como “Target=No” y otro a la derecha etiquetado como “Target=Yes”; un hiperplano separador, indicado por una línea sólida,

divide estos dos grupos. Las líneas punteadas representan los márgenes que tocan los vectores de soporte. Los vectores de soporte son puntos que están justo en el borde o dentro del margen. La palabra “*Margin*” está indicada con una flecha que apunta hacia las líneas punteadas que representan el margen. Para entender de manera más profunda las Maquinas de Soporte Vectorial, se deben describir algunos conceptos clave:

Figura 5

Plano Generado por una SVM



Nota. Tomado de. Rady and Anwar. (2019)

- **Hiperplano.** Un hiperplano es un subespacio de dimensiones inferiores al espacio de características completo. Se busca encontrar el hiperplano que maximiza el margen entre las clases.
- **Vectores de Soporte.** Son los puntos de datos más cercanos al hiperplano de separación. Estos puntos son críticos para la definición del hiperplano y, por ende, para la eficacia del modelo. Cambiar o eliminar un vector de soporte podría alterar el hiperplano y, en consecuencia, la capacidad del modelo para generalizar (Boswell, 2002).

- **El margen.** Es la distancia perpendicular entre el hiperplano separador y los vectores de soporte, que son los puntos de datos más cercanos al hiperplano. El objetivo de las Máquinas de Vectores de Soporte es maximizar este margen, ya que un margen más amplio conduce a un modelo más robusto y con mayor capacidad de generalización a nuevos datos. Un mayor margen está asociado con una mayor resistencia al ruido y una reducción del riesgo de sobreajuste, mejorando la precisión y estabilidad del clasificador (IBM, 2023; AprendeIA, 2019).
- **Función de Pérdida.** Para problemas de clasificación, las Maquinas de Soporte Vectorial utilizan una función de pérdida que penaliza la clasificación incorrecta y premia el mantenimiento del margen óptimo. La función de pérdida a menudo incluye términos de regularización para evitar el sobreajuste del modelo (Shalev-Shwartz and Ben-David, 2014).
- **Kernel Trick.** Las Maquinas de Soporte Vectorial pueden manejar de manera eficiente conjuntos de datos no lineales mediante el “*kernel trick*”. Este enfoque implica mapear los datos a un espacio de características de mayor dimensión donde la separación lineal es posible. Los *kernels* permiten realizar este mapeo de manera implícita, evitando la necesidad de calcular explícitamente las nuevas características (Burges, 1998).
- El algoritmo 3 presenta de forma simplificada cada uno de los pasos necesarios para implementar una SVM.

Tabla 2

Algorithm 3: SVM Simplificado

Algorithm 3: SVM Simplificado	
Data:	Conjunto de datos: $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$
Result:	Modelo SVM entrenado

Entrenamiento SVM: Seleccionar un kernel (lineal, RBF, polinómico, etc.); Iniciar parámetros del modelo;

for iteración ← 1 to Número máximo de iteraciones do

 Calcular la función de pérdida;

Nota. Explicación Algorithm 3: SVM Simplificado

– Calcular los gradientes o realizar optimización cuadrática para actualizar los parámetros; Las Máquinas de Soporte Vectorial ofrecen ventajas notables en el análisis de imágenes, su eficacia en espacios de alta dimensión, como los encontrados en imágenes pixeladas, las hace adecuadas para capturar patrones complejos. Además, las SVM generalizan bien a conjuntos de datos no vistos, lo que las convierte en una elección versátil. Pueden modelar eficientemente relaciones no lineales en datos de imágenes a través de funciones de kernel, ampliando su aplicabilidad (Bishop and Nasrabadi, 2006).

Sin embargo, las Máquinas de Vectores de Soporte (SVM) requieren un preprocesamiento cuidadoso y la elección adecuada de parámetros, especialmente del hiperparámetro de regularización C y de la función kernel, ya que estos influyen significativamente en su rendimiento. La identificación del hiperplano óptimo, que maximiza el margen entre clases, es un problema de optimización convexa que puede volverse complejo en la práctica. Además, las SVM fueron diseñadas inicialmente para problemas de clasificación binaria, y su extensión a problemas multiclase mediante estrategias como *one-versus-one* o *one-versus-rest* puede generar inconsistencias y complicaciones, como asignaciones múltiples o desbalance en el entrenamiento de clasificadores individuales.

En conjuntos de datos grandes, el tiempo computacional puede volverse prohibitivo debido a la complejidad del problema de optimización. La interpretación de los resultados también puede ser compleja, y la sensibilidad a datos ruidosos o con muchas características irrelevantes puede limitar su eficacia. Estos problemas están relacionados con la selección del hiperplano óptimo y la capacidad del modelo para generalizar, especialmente cuando los datos no son linealmente separables o presentan ruido significativo (MathWorks, s.f.;; DataCamp, 2024).

Las SVM son eficaces en aplicaciones de clasificación de imágenes y texto. Por ejemplo, en el reconocimiento de dígitos escritos a mano (como en sistemas de corrección de exámenes) o la detección de fraude en transacciones bancarias, donde se requiere separar categorías de datos de manera precisa.

K-Vecinos más Cercanos (KNN)

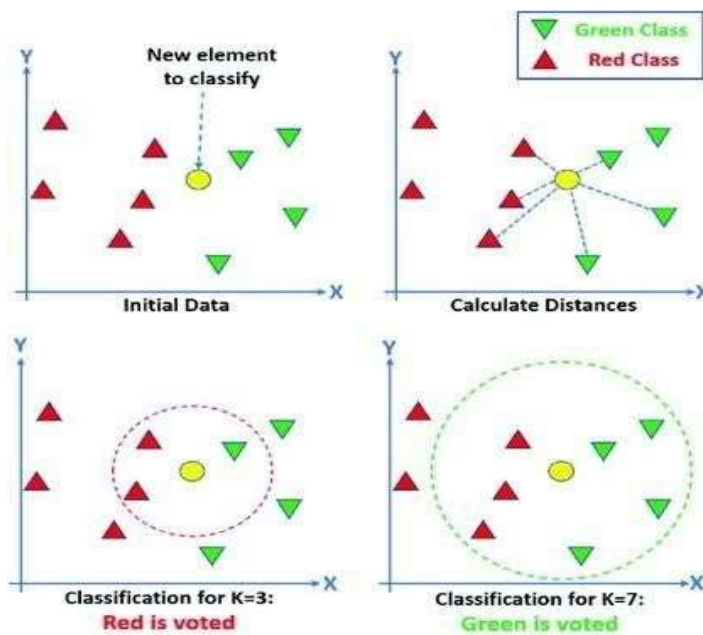
El método K-Vecinos más cercanos (KNN) es un algoritmo de aprendizaje supervisado que se utiliza para problemas de clasificación y regresión; la predicción para un nuevo punto de datos se basa en la mayoría de las etiquetas de sus vecinos más cercanos (en el caso de clasificación) o en la media de los valores de los vecinos (en el caso de regresión) (Cunningham and Delany, 2021, Müller and Guido, 2016).

La figura 6 es un diagrama que ilustra el proceso de clasificación en un sistema de k-vecinos más cercanos (K-NN) con diferentes valores de k. La imagen es un diagrama dividido en cuatro partes que ilustran los pasos para clasificar un nuevo elemento usando el algoritmo K-NN. En la parte superior izquierda, titulada “*Initial Data*”, muestra triángulos rojos y verdes en un plano cartesiano, representando dos clases diferentes; un círculo amarillo representa el nuevo elemento a clasificar. En la parte superior derecha, titulada “*Calculate*

Distances”, se han dibujado líneas desde el círculo amarillo hasta los elementos cercanos para calcular las distancias. En la parte inferior izquierda, “*Classification for $K = 3$* ”, se ha dibujado un círculo punteado alrededor del círculo amarillo y los tres vecinos más cercanos. El texto indica que la clase roja ha sido votada por- que hay dos triángulos rojos dentro del círculo punteado. En la parte inferior derecha, titulada “*Classification for $K = 7$* ”, se muestra otro círculo punteado más grande que incluye siete vecinos más cercanos. El texto indica que la clase verde ha sido votada porque hay cuatro triángulos verdes dentro del círculo punteado.

Figura 6

Ilustración Sobre el Funcionamiento de KNN



Nota. Tomado de. Zidi. (2023)

Los pasos importantes para generar el algoritmo KNN son:

- **Selección del Parámetro K.** KNN depende del parámetro K, que representa el número de vecinos más cercanos a considerar. La elección de K es crucial y puede afectar la

precisión del modelo. Un K pequeño puede ser sensible al ruido, mientras que un K grande puede hacer que el modelo pierda detalles importantes.

- **Medida de Distancia.** KNN utiliza una medida de distancia, comúnmente la distancia euclidiana, para determinar qué puntos son los vecinos más cercanos. Otros métodos de distancia, como la distancia de Manhattan o la distancia de Minkowski, también pueden ser utilizados según el problema específico (Alpaydin, 2020).
- **Búsqueda de Vecinos más Cercanos.** Para predecir la etiqueta o el valor para un nuevo punto, KNN encuentra los K vecinos más cercanos en función de la medida de distancia. Estos vecinos pueden provenir de un conjunto de datos de entrenamiento previamente etiquetado.
- **Clasificación.** En el caso de clasificación, el nuevo punto se asigna a la clase que predomina entre sus vecinos más cercanos (mayoría de votos). Por ejemplo, si la mayoría de los vecinos son de la clase “A”, el nuevo punto se clasifica como “A”.
- **Regresión.** En el caso de regresión, el algoritmo K -vecinos más cercanos (KNN) calcula la media o la mediana de los valores de los vecinos más cercanos para predecir el valor de un nuevo punto. KNN es un método estadístico no paramétrico, lo que significa que no hace suposiciones específicas sobre la forma funcional de los datos. Esta característica lo hace muy versátil y aplicable a diversos tipos de problemas, especialmente cuando la relación entre variables es compleja o no lineal.

Sin embargo, KNN es sensible a la escala de las características, por lo que es importante estandarizar o normalizar los datos antes de su aplicación para asegurar que todas las variables contribuyan equitativamente a la distancia calculada. Además, este algoritmo puede ser computacionalmente costoso, especialmente en conjuntos de datos

grandes, ya que implica calcular distancias entre el nuevo punto y todos los puntos del conjunto de entrenamiento.

Para mejorar la predicción, se puede ponderar la influencia de los vecinos en función de su distancia, dando mayor peso a los más cercanos y menor a los más alejados. Esto ayuda a reducir el impacto de puntos atípicos o ruido en los datos.

El modelo general de KNN para regresión se basa en los siguientes pasos: seleccionar un valor para k , identificar los k vecinos más cercanos al punto a predecir, y calcular la media ponderada o simple de sus valores para obtener la predicción final (IBM, 2021; DataCamp, 2024; KeepCoding, 2024).

Tabla 3

Algorithm 4: K-Vecinos más Cercanos (KNN)

Algorithm 4: K-Vecinos más Cercanos (KNN)	
Data:	Conjunto de entrenamiento: $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$; Nuevo punto de datos: X_{new}
Result:	Predicción para X_{new}
Paso 1:	Calcular distancias; for $i \leftarrow 1$ to n do Calcular la distancia entre X_{new} y X_i (por ejemplo, distancia euclidiana);
Paso 2:	Seleccionar los K vecinos más cercanos; Ordenar las distancias y seleccionar los K vecinos más cercanos;

Clasificación o regresión según el problema;

if Problema de clasificación then

Paso 3: Votar por la clase más común entre los K vecinos;

else if Problema de regresión then

Calcular la media de los valores de los K vecinos;

Nota. Explicación Algorithm 4: K-Vecinos más Cercanos (KNN)

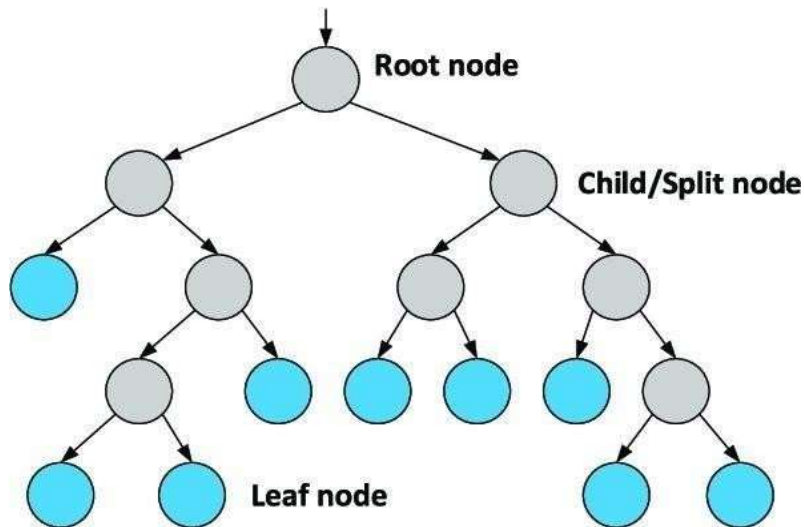
Similares en función de compras anteriores de los usuarios. También se aplica en la clasificación de enfermedades, donde se pueden identificar patrones similares en datos médicos para clasificar correctamente nuevos casos.

Árboles de Decisión

Los Árboles de Decisión son modelos de aprendizaje supervisado utilizados para tareas de clasificación y regresión. Representan decisiones y sus posibles consecuencias en forma de un árbol, donde cada nodo interno corresponde a una decisión basada en una característica, cada rama representa el resultado de esa decisión, y cada hoja indica la etiqueta de clasificación o el valor de regresión final (Safavian & Landgrebe, 1991).

Figura 7

Ilustración de un Árbol de Decisión



Nota. Gráfico explicación árbol de decisiones. Tomado de. Acosta et al., (2020)

Una de las principales ventajas de estos modelos es su facilidad para ser interpretados y visualizados; la lógica de toma de decisiones es intuitiva, lo que facilita su comprensión incluso para usuarios no técnicos. Además, pueden manejar tanto variables categóricas como numéricas sin requerir preprocesamiento exhaustivo, y no son sensibles a la escala de las características. También son capaces de capturar relaciones no lineales y patrones complejos en los datos (Loh, 2011; Rokach & Maimon, 2008).

Sin embargo, los árboles de decisión pueden ser propensos al sobreajuste, especialmente cuando se ajustan demasiado a los detalles del conjunto de entrenamiento, lo que afecta su capacidad de generalización. Son sensibles a pequeñas variaciones en los datos, lo que puede generar árboles muy diferentes con datos ligeramente distintos, mostrando inestabilidad. En problemas con clases desbalanceadas, pueden sesgarse hacia la clase mayoritaria. Además, la construcción del árbol utiliza criterios locales para maximizar

la pureza en cada nodo, pero esto no garantiza una optimización global del árbol. Una vez construido, el árbol no puede adaptarse a nuevos datos sin ser reconstruido. También presentan dificultades para resolver problemas con patrones XOR u otros problemas no lineales complejos debido a su estructura jerárquica (Safavian & Landgrebe, 1991; Rokach & Maimon, 2008).

Los árboles de decisión se aplican en diversos campos, como el sector financiero, para evaluar la aprobación de créditos basándose en criterios como ingresos, historial crediticio y edad, y en medicina para apoyar diagnósticos basados en síntomas clínicos (Loh, 2011).

Los árboles de decisión se emplean comúnmente en el sector financiero, por ejemplo, para decidir si se aprueba o no un crédito basándose en criterios como los ingresos del solicitante, historial crediticio y edad. También se utilizan en medicina para ayudar a los médicos a tomar decisiones de diagnóstico basadas en síntomas.

Random Forest

Random Forest es un algoritmo de aprendizaje automático que se utiliza tanto para tareas de clasificación como de regresión. Se basa en la idea de construir múltiples árboles de decisión y combinar sus resultados para obtener un modelo más robusto y preciso. Fue propuesto por Leo Breiman en 2001 ((Breiman, 2001)) y se ha convertido en una herramienta popular en el aprendizaje automático.

A continuación, se describen cada uno de los pasos necesarios para la implementación del algoritmo (Breiman, 2001, Hastie et al., 2009).

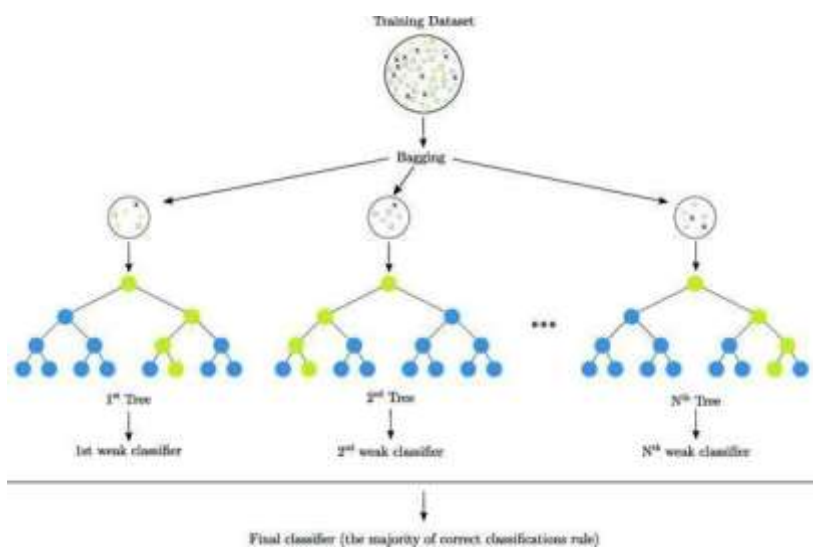
- **Construcción del Bosque Aleatorio.** Se inicia con la creación de múltiples árboles de decisión, que forman el “bosque aleatorio”. La cantidad de árboles se determina

previamente y puede ajustarse según los requisitos del problema.

- **Selección Aleatoria de Características.** Antes de construir cada árbol, se selecciona aleatoriamente un conjunto de características (columnas) del conjunto de datos. Esta selección aleatoria ayuda a introducir diversidad entre los árboles y evita la dependencia excesiva de ciertas características.

Figura 8

Ejemplo de Random Forest



Nota. Explicación Random Forest. Tomado de. *Sapountzoglou et al., (2020)*

- **Construcción de Árboles de Decisión.** Para cada árbol en el bosque, se utiliza un subconjunto aleatorio de datos y características para construir un árbol de decisión. Cada nodo en el árbol realiza divisiones basadas en las características seleccionadas y los criterios de división, como la ganancia de información o la reducción del error cuadrático medio.
- **Votación o Promedio.** Una vez que todos los árboles se han construido, se realiza la predicción combinada. Para problemas de clasificación, se realiza una votación

entre los árboles, y la clase con más votos se elige la predicción final. Para problemas de regresión, se promedian las predicciones de todos los árboles.

- **Manejo de Nuevos Datos.** Cuando se presenta un nuevo conjunto de datos para predecir, cada árbol del bosque realiza una predicción y se lleva a cabo la votación o el promedio, según la tarea. Este proceso garantiza robustez y generalización del modelo.
- **Evaluación de la Importancia de las Características.** Después de construir el bosque, se puede evaluar la importancia relativa de cada característica. Esto se logra al observar cuánto contribuye cada característica a la precisión del modelo a lo largo de todos los árboles.

En términos de ventajas, destaca por su alta precisión y robustez. La combinación de múltiples árboles de decisión reduce el riesgo de sobreajuste, ofreciendo resultados confiables en tareas complejas como el análisis de imágenes. Además, la capacidad de seleccionar aleatoriamente características en cada árbol facilita el manejo eficiente de conjuntos de datos con numerosas características, comunes en este tipo de análisis (Breiman, 2001, Hastie et al., 2009).

El algoritmo puede manejar conjuntos de datos desbalanceados, equilibrando la influencia de las clases minoritarias. La estimación de la importancia de características proporciona *insights* valiosos para comprender qué aspectos visuales son más relevantes en las decisiones del modelo (Kang et al., 2022).

Sin embargo, existen desventajas a considerar en el uso de *Random Forest*. La interpretación de las decisiones del modelo puede ser un desafío debido a la complejidad inherente del bosque aleatorio, lo que dificulta la transparencia y la explicación del proceso de toma de decisiones. Además, el algoritmo requiere un alto consumo de recursos

computacionales, especialmente para conjuntos de datos grandes y complejos, lo que puede limitar su aplicación en entornos con restricciones de *hardware*. En el caso de imágenes de alta resolución, *Random Forest* puede volverse computacionalmente costoso y menos eficiente. También es sensible a datos ruidosos y puede perder información espacial relevante, lo que afecta su desempeño en tareas que requieren preservar relaciones espaciales (Louppe, 2014; Fernández-Delgado et al., 2014; Cutler et al., 2007).

³Se refiere a percepciones o entendimientos profundos obtenidos al interpretar información o datos. En el análisis de datos, son observaciones clave que proporcionan comprensión estratégica o apoyo a la toma de decisiones (Tufte, 2001).

Random Forest se aplica en problemas de clasificación y regresión. En la predicción del riesgo de crédito, este método es usado para analizar grandes volúmenes de datos de clientes y evaluar la probabilidad de impago. Además, se emplea en la predicción de ventas, donde puede combinar muchos factores diferentes para hacer predicciones precisas.

AdaBoost

AdaBoost o “*Adaptive Boosting*”, es un algoritmo de aprendizaje automático desarrollado por Yoav Freund y Robert Schapire en 1996 (Freund et al., 1996). Su objetivo principal es mejorar el rendimiento de modelos más débiles mediante la combinación ponderada de múltiples clasificadores. El proceso de *AdaBoost* se inicia asignando pesos a cada instancia de entrenamiento. Inicialmente, todas las instancias tienen el mismo peso. Luego, se entrena un clasificador débil (un modelo ligeramente mejor que el azar) en el conjunto de datos ponderado. Se calcula el error ponderado del clasificador débil, dando más peso a las instancias mal clasificadas. El peso del clasificador débil se calcula en función de su rendimiento, y luego se actualizan los pesos de las instancias, otorgando más importancia a

las instancias mal clasificadas. Esto permite al algoritmo centrarse en corregir errores anteriores. El clasificador débil actual se combina con los clasificadores anteriores, asignando más peso a los más precisos. Estos pasos se repiten iterativamente, construyendo una secuencia de clasificadores débiles. Finalmente, los clasificadores débiles se combinan mediante una suma ponderada para formar un clasificador fuerte. Este clasificador final tiene en cuenta la contribución de cada clasificador débil de acuerdo con su rendimiento. AdaBoost es efectivo en conjuntos de datos desequilibrados y puede adaptarse a datos ruidosos, siendo utilizado con éxito en problemas de clasificación y detección de objetos, entre otros (*Schapire and Freund, 2013, Ferreira and Figueiredo, 2012*).

AdaBoost destaca por su capacidad para mejorar la precisión en tareas de clasificación, siendo especialmente efectivo en la detección de objetos y características visuales en imágenes. Aunque puede adaptarse a datos desequilibrados y manejar cierto nivel de ruido, es importante considerar que en el procesamiento de imágenes siempre se realiza un preprocesamiento previo, como la aplicación de filtros para eliminar ruido y mejorar la calidad de los datos de entrada. Esta etapa es crítica, ya que *AdaBoost*, aunque robusto, no reemplaza la necesidad de preparar adecuadamente los datos, especialmente en aplicaciones en línea donde el preprocesamiento puede ser limitado o difícil de implementar.

AdaBoost puede utilizarse con una variedad de clasificadores débiles, combinándolos iterativamente para formar un modelo fuerte. Sin embargo, su rendimiento puede verse afectado por datos atípicos o ruido residual en las imágenes. Además, requiere un cuidadoso ajuste de parámetros y selección de características para optimizar su desempeño. El proceso iterativo de entrenamiento puede volverse computacionalmente costoso en grandes conjuntos de datos, y existe el riesgo de sobreajuste si no se controla adecuadamente

AdaBoost se ha utilizado en el reconocimiento de rostros y en la clasificación de objetos en imágenes. Este método también es útil en la detección de fraudes en tarjetas de crédito, donde mejora la precisión de la identificación de transacciones fraudulentas al combinar varios modelos débiles.

Q-Learning

El algoritmo *Q-Learning* es una técnica de aprendizaje por refuerzo utilizada en el aprendizaje máquina para enseñar a un agente a tomar decisiones óptimas en un entorno desconocido y dinámico. En este algoritmo, el agente aprende a maximizar una función de recompensa a lo largo del tiempo al explorar y explotar las acciones disponibles en su entorno. La esencia de este método radica en la estimación de valores de calidad (*Q-values*) para cada par estado-acción posible en el entorno del agente. Estos valores representan la “calidad” de tomar una determinada acción en un estado dado, considerando las recompensas futuras esperadas. Durante el entrenamiento, el agente actualiza iterativamente estos valores de calidad utilizando la ecuación de Bellman (ver ecuación 4), que refleja la relación entre los *Q-values* actuales y los futuros (Sutton and Barto, 2018, Clifton and Laber, 2020, Kim and Yang, 2020).

$$Q_n(s,a) = (1-\alpha)Q_{p(s,a)} + \alpha Q_0(s,a);$$

$$\text{SIENDO, } Q_0(s,a) = r(s,a) + \gamma \max_{a^2} Q(s^2, a^2) \quad (4)$$

Para la ecuación 4, los términos:

- $Q_n(s,a)$ es el valor del estado actual s,a
- $r(s,a)$ es la recompensa inmediata al realizar una acción en el estado s,a
- γ es el factor de descuento que representa la importancia de las recompensas futuras

■ $\max Q(s^2, a^2)$ es el valor máximo de calidad (Q-value) para todas las acciones posibles a en el estado s

A medida que el agente explora el entorno y recibe recompensas, ajusta sus estimaciones de Q-values, aprendiendo así a tomar decisiones que maximicen las recompensas acumulativas. El proceso continúa hasta que el agente converge en una estrategia óptima para realizar acciones en diferentes estados del entorno.

En el procesamiento de imágenes, el algoritmo Q-Learning tiene ventajas, como su adaptabilidad a entornos dinámicos y capacidad para aprender sin supervisión. Sin embargo, enfrenta desafíos computacionales, especialmente en conjuntos de datos grandes, y requiere una representación efectiva del estado, lo cual puede ser crítico para su rendimiento en imágenes. La capacidad de adaptación y el aprendizaje no supervisado son puntos fuertes, pero la complejidad computacional y la necesidad de una representación de estado eficaz son desafíos a considerar (Kim and Yang, 2020, Clifton and Laber, 2020).

Q-Learning, una técnica clave en el aprendizaje por refuerzo, se utiliza en el control de robots y en el desarrollo de sistemas autónomos de toma de decisiones. Un ejemplo práctico es su uso en vehículos autónomos, donde el algoritmo aprende a optimizar rutas y tomar decisiones sobre maniobras en tiempo real.

Actor-Critic Methods

Los métodos *Actor-Critic* son una clase de algoritmos utilizados en el aprendizaje por refuerzo. Estos métodos combinan elementos de dos enfoques: el actor, que aprende una metodología para tomar decisiones, y el crítico, que evalúa la calidad de esas decisiones (Konda and Tsitsiklis, 1999). Los principales elementos de este método son:

- **Actor (Actor).** Este componente aprende una metodología que mapea los

estados a acciones. El “actor” es responsable de decidir qué acciones tomar en función de la observación actual del entorno.

- **Critic (Critic).** Este componente, llamado “crítico”, evalúa las acciones tomadas por el actor. Su función es estimar la función de valor, que indica la utilidad o calidad de una acción en un estado particular.

La interacción entre el actor y el crítico es lo que distingue a los métodos Actor-Critic.

El actor toma decisiones basadas en su metodología aprendida, mientras que el crítico evalúa esas decisiones y proporciona retroalimentación. Esta retroalimentación se utiliza para actualizar tanto la metodología del actor como la función de valor del crítico (Hong et al., 2023, Bhatnagar et al., 2023, Zahavy et al., 2020).

En el análisis de imágenes, los métodos Actor-Critic tienen ventajas en entornos visuales complejos y en tareas con recompensas no inmediatas. Sin embargo, su rendimiento puede depender de la calidad de la representación visual y puede requerir mayor complejidad de entrenamiento y ajuste de hiperparámetros en comparación con enfoques más simples. Los hiperparámetros son parámetros externos al modelo de aprendizaje automático que no se aprenden directamente del conjunto de datos durante el entrenamiento. Estos parámetros son configurados antes de iniciar el proceso de entrenamiento y afectan el rendimiento y comportamiento del modelo. A diferencia de los parámetros del modelo, que son ajustados automáticamente durante el entrenamiento, los hiperparámetros son establecidos por el diseñador del modelo antes de comenzar el proceso de entrenamiento (Géron, 2022).

Estos métodos se utilizan en la robótica y en la toma de decisiones en tiempo real, donde es necesario optimizar acciones continuas. En aplicaciones como robots humanoides que aprenden a caminar o drones que deben mantener una trayectoria estable, los métodos

actor-critic permiten mejorar el rendimiento a través de una política de acciones aprendidas.

Métodos de Etiquetado Propagado (Label Propagation)

Los métodos de etiquetado propagado, también conocidos como *Label Propagation*, son técnicas utilizadas en aprendizaje automático y minería de grafos para asignar etiquetas a nodos no etiquetados en función de la información de los nodos vecinos en un grafo. La idea principal es propagar las etiquetas conocidas a través de la red para asignar etiquetas a nodos sin etiquetar. (Zhur and Ghahramanirh, 2002). El proceso general de etiquetado propagado puede describirse de la siguiente manera:

- **Inicialización de Etiquetas.** Se asignan etiquetas a un conjunto inicial de nodos conocidos.
- **Propagación de Etiquetas.** Las etiquetas se propagan a través de las conexiones en el grafo. Nodos cercanos tienden a influirse mutuamente, y esta influencia se propaga a lo largo de las conexiones.
- **Actualización de Etiquetas.** Se actualizan las etiquetas de los nodos en función de las etiquetas propagadas y de la información local.

Este proceso se repite iterativamente hasta que las etiquetas convergen o se alcanza un criterio de parada. Los métodos de etiquetado propagado son útiles en situaciones donde la información de las etiquetas se propaga de manera natural a través de conexiones en un grafo, como en clasificación de nodos en redes sociales o recomendación de productos basada en la relación entre usuarios (Zhou et al., 2003).

En el análisis de imágenes, el método de etiquetado propagado tiene ventajas al utilizar información contextual de nodos vecinos y adaptarse a estructuras complejas. Sin embargo, puede ser sensible a la calidad del grafo, dependiendo de la representación

adecuada. Además, en conjuntos de datos grandes, el proceso iterativo puede volverse computacionalmente costoso.

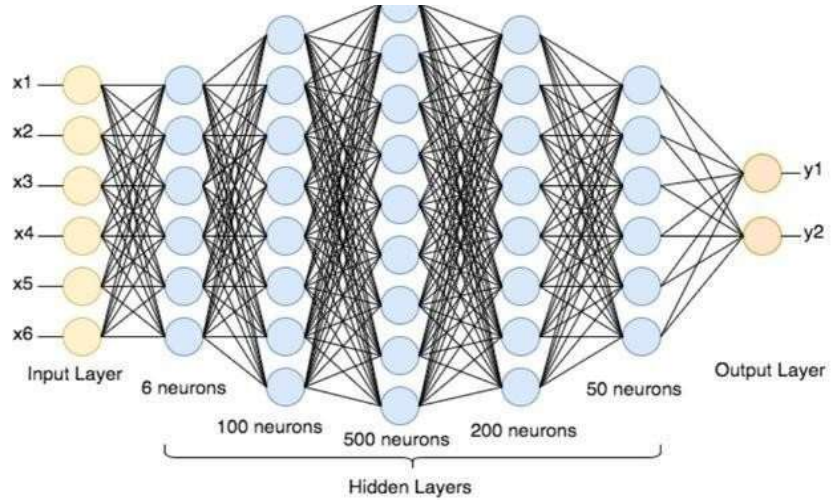
Los métodos de etiquetado propagado son utilizados en redes sociales para la clasificación de comunidades o la predicción de enlaces. Estos algoritmos permiten identificar grupos de usuarios con intereses comunes o relaciones potenciales basadas en pocos datos etiquetados, propagando la información a lo largo de la red de usuarios.

Aprendizaje Profundo

El aprendizaje profundo, también conocido como *Deep Learning*, es una rama avanzada del aprendizaje automático que ha revolucionado la capacidad de las máquinas para realizar tareas complejas. En esencia, el aprendizaje automático consiste en la toma de decisiones mediante métodos matemáticos y estadísticos basados en datos (Patterson and Gibson, 2017). A diferencia de los modelos de aprendizaje convencionales, que pueden requerir una definición explícita de características, el aprendizaje profundo utiliza redes neuronales artificiales compuestas por múltiples capas para aprender automáticamente representaciones jerárquicas de los datos. Cada capa procesa la información de manera incremental, permitiendo una abstracción cada vez más compleja y sofisticada a medida que se profundiza en la red (Aggarwal et al., 2018).

Figura 9

Red Neuronal Profunda

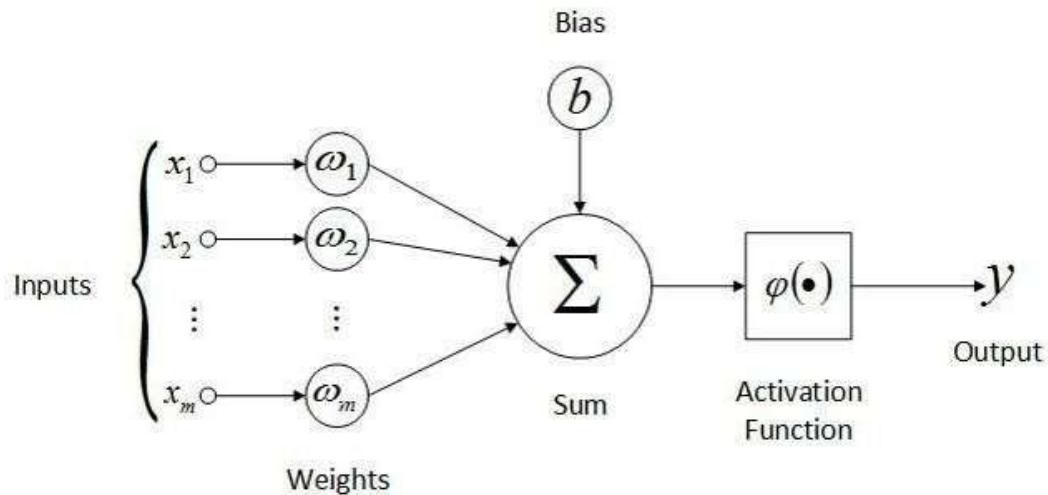


Nota. Explicación de una Red Neuronal Profunda. Tomado de. Bahi and Batouche, (2018)

El fundamento clave del aprendizaje profundo radica en los conceptos de neuronas artificiales y pesos sinápticos. Cada neurona artificial está conectada a otras (ver figura 9) mediante pesos que se ajustan durante el proceso de entrenamiento, permitiendo que la red aprenda patrones y relaciones en los datos. A través de iteraciones y optimizaciones, la red neuronal puede generalizar y realizar predicciones precisas en datos no vistos anteriormente.

Figura 10

Modelo de Neurona Artificial.



Nota. Explicación funcionamiento modelo de neurona Artificial. Tomado de. Oliveira et al., (2017)

Una característica distintiva del aprendizaje profundo es su capacidad para abordar problemas de gran escala y complejidad, como el reconocimiento de imágenes, el procesamiento del lenguaje natural y la toma de decisiones en entornos complejos. La adaptabilidad y autonomía inherentes a estas redes neuronales profundas han llevado a avances significativos en campos como la inteligencia artificial y la visión por computadora (Goodfellow et al., 2016).

Sin embargo, con sus capacidades también posee limitaciones; la necesidad de grandes conjuntos de datos para entrenar eficazmente las redes profundas, así como la complejidad de ajustar sus numerosos parámetros, son cuestiones que los investigadores y profesionales deben abordar. A pesar de estos desafíos, el aprendizaje profundo continúa transformando la forma en que interactuamos con la tecnología, y su potencial para

innovaciones futuras sigue siendo motivo de considerable interés y estudio (Kirkpatrick et al., 2017, Zhang et al., 2021).

El aprendizaje profundo es fundamental para el análisis de imágenes por diversas razones. En primer lugar, las redes neuronales profundas pueden aprender representaciones jerárquicas de características en imágenes, capturando información significativa a diferentes niveles de abstracción. Esta capacidad se traduce en la extracción automática de características, eliminando la necesidad de diseñar descriptores manualmente (Szeliski, 2022). Además, el aprendizaje profundo destaca por su eficaz manejo de grandes volúmenes de datos, siendo esencial en escenarios de *big data* en constante crecimiento. Su versatilidad le permite adaptarse a diversas tareas de análisis de imágenes, como clasificación, detección de objetos, segmentación y generación de imágenes. La capacidad de transferencia de conocimientos aprendidos a tareas relacionadas es crucial, especialmente cuando los conjuntos de datos son limitados (Venkatesan and Li, 2017).

La mejora continua con grandes conjuntos de datos y los avances en arquitecturas específicas, como las redes neuronales convolucionales (CNN) y las redes generativas adversarias (GAN), contribuyen a su efectividad. El aprendizaje profundo también se destaca en la aplicación a problemas complejos como el reconocimiento facial, la clasificación de imágenes médicas o la interpretación de escenas, abordando la complejidad y variabilidad inherentes a estos problemas (Goodfellow et al., 2016, Creswell et al., 2018).

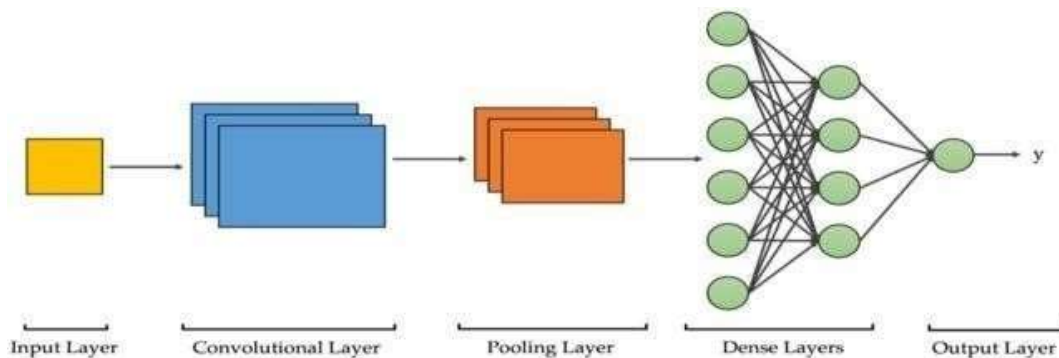
Redes neuronales convolucionales (CNN)

Las Redes Neuronales Convolucionales (CNN) son un tipo de arquitectura de redes neuronales diseñadas para procesar y analizar datos de tipo *grid*, como imágenes. Están compuestas por capas convolucionales que aplican operaciones de convolución para extraer

características específicas de las imágenes. Luego, se utilizan capas de *pooling* para reducir la dimensionalidad y mantener las características más importantes. Finalmente, se emplean capas completamente conectadas para realizar la clasificación (Goodfellow et al., 2016) (ver figura 11). El *pooling*, es una operación utilizada para reducir la dimensionalidad de las representaciones de características obtenidas tras las capas convolucionales. La capa de *pooling* toma pequeñas regiones de la entrada, llamadas ventanas, y reduce la información en cada región a un solo valor representativo. Las dos operaciones de *pooling* más comunes son el *max pooling*, que toma el valor máximo en cada región, y el *average pooling*, que toma el promedio (Charu, 2018). Esta reducción de dimensionalidad ayuda a disminuir la carga computacional y a preservar las características más relevantes (Venkatesan and Li, 2017).

Figura 11

Ejemplo Red Neuronal Convolutiva



Nota. Ejemplo Red Neuronal Convolutiva. Tomado de. Sit et al., (2020)

Las Redes Neuronales Convolucionales (CNN) son especialmente eficaces para el análisis de imágenes debido a varias consideraciones clave:

- **Extracción Jerárquica de Características.** Las capas convolucionales permiten la extracción automática de características en diferentes niveles de abstracción, desde bordes

simples hasta características más complejas. Esto facilita la captura de información significativa en imágenes.

- **Preservación de Estructura Espacial.** Las operaciones de convolución conservan la estructura espacial de las imágenes. Esto es crucial para el análisis visual, ya que las relaciones espaciales entre píxeles desempeñan un papel importante en la interpretación de la información visual.
- **Parámetros Compartidos.** Las CNN utilizan parámetros compartidos en las capas convolucionales, lo que reduce significativamente la cantidad de parámetros en comparación con redes completamente conectadas. Esto facilita el entrenamiento eficiente de modelos, especialmente en conjuntos de datos grandes.
- **Invariancia a Traslaciones Locales.** Las capas de *pooling* proporcionan invariancia a pequeñas traslaciones locales, lo que significa que la red puede reconocer patrones independientemente de su ubicación exacta en la imagen.
- **Adaptabilidad a Tareas Diversas.** La arquitectura modular de las CNN permite adaptarse a diversas tareas de análisis de imágenes, como clasificación, detección de objetos, segmentación y generación de imágenes, haciendo que sean versátiles y aplicables en diferentes dominios.

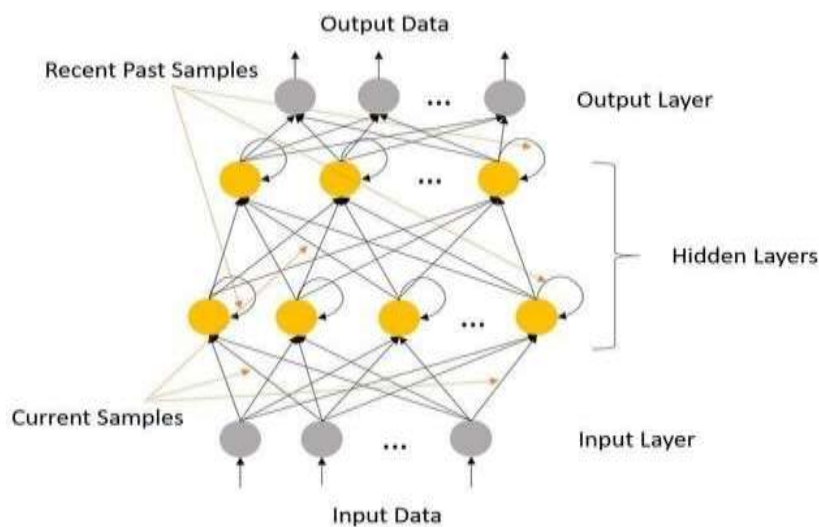
Redes Neuronales Recurrentes (RNN)

Las Redes Neuronales Recurrentes (RNN) son un tipo de red neuronal artificial que se utiliza para procesar datos secuenciales o de series temporales. Estos algoritmos de aprendizaje profundo se utilizan comúnmente para problemas temporales como la traducción de idiomas, el procesamiento del lenguaje natural (NLP), el reconocimiento de voz y los subtítulos de imágenes (Goodfellow et al., 2016). A diferencia de las redes neuronales

profundas tradicionales, las RNN tienen una “memoria” que les permite obtener información de entradas anteriores para influir en la entrada y salida actuales (ver figura 12); esto significa que la salida de las RNN depende de los elementos anteriores dentro de la secuencia (Caterini et al., 2018).

Figura 12

Modelo de Red Neuronal Recurrente



Nota. Explicación del Modelo de Red Neuronal Recurrente. Tomado de. Modelo de Red Neuronal Recurrente

Las Redes Neuronales Recurrentes (RNN) desempeñan un papel crucial en el análisis de imágenes debido a su capacidad para modelar dependencias temporales y secuenciales. Sus ventajas son:

- **Manejo de Secuencias.** Las imágenes a menudo tienen información estructurada y dependencias temporales. Las RNN pueden procesar secuencias de datos, como vídeos o series temporales de imágenes, capturando la relación entre los elementos anteriores y los actuales.

- **Contexto a Largo Plazo.** La “memoria” de las RNN permite mantener y utilizar información relevante de entradas anteriores, lo que es esencial para comprender contextos a largo plazo en una secuencia de imágenes.
- **Adaptabilidad a Variaciones Temporales.** En situaciones donde la información visual cambia con el tiempo, como en la detección de objetos en un vídeo, las RNN pueden adaptarse para tener en cuenta la evolución temporal de las características.
- **Generación de Descripciones.** Las RNN también se utilizan en la generación de descripciones de imágenes, donde la salida depende de la comprensión acumulativa de la información visual a lo largo de la secuencia.

Implementar RNN para el análisis de imágenes requiere considerar varios aspectos clave:

- **Tamaño de Secuencia.** El tamaño de secuencia determina la longitud de la secuencia que la Red Neuronal Recurrente (RNN) debe procesar. Ajustar este parámetro es crucial para manejar la información temporal en las imágenes de manera efectiva, ya que permite capturar dependencias relevantes sin sobrecargar el modelo. Además, un tamaño de secuencia adecuado contribuye a disminuir el costo computacional, ya que el modelo debe hacer menos predicciones y procesar menos datos en cada paso, lo que reduce el tiempo de entrenamiento y la memoria requerida. Por el contrario, secuencias demasiado largas pueden provocar problemas como el desvanecimiento o explosión del gradiente, dificultando el aprendizaje efectivo
- **Capacidad de Memoria.** La capacidad de la RNN para recordar depende de su arquitectura y de cómo se gestionan las conexiones recurrentes. Asegurarse de que la red pueda retener la información relevante es esencial.

- **Preprocesamiento de Datos.** Las imágenes deben transformarse en secuencias de datos manejables para las RNN. Esto podría implicar dividir las imágenes en cuadros o aplicar técnicas específicas de preprocesamiento.
- **Arquitectura de la Red.** Seleccionar la arquitectura adecuada, como LSTM o GRU, que se adapte mejor al problema y a la naturaleza de las imágenes es esencial. Esto afectará la capacidad de la red para modelar dependencias temporales.
- **Ajuste de Hiperparámetros.** Los parámetros como la tasa de aprendizaje, el tamaño de lote y otros deben ajustarse cuidadosamente para lograr un equilibrio entre el rendimiento y la eficiencia del entrenamiento.
- **Consideraciones Computacionales.** Implementar RNN para análisis de imágenes puede ser computacionalmente intensivo. Asegurarse de tener el poder de cómputo necesario es fundamental, especialmente para conjuntos de datos grandes.
- **Regularización.** Dado que las RNN pueden ser propensas al sobreajuste, la aplicación de técnicas de regularización, como *dropout*, puede ser crucial para mejorar la generalización.
- **Evaluación del Rendimiento.** Seleccionar métricas apropiadas para evaluar el rendimiento de la RNN en tareas específicas de análisis de imágenes, como la clasificación o la generación de descripciones.

Redes Generativas Adversarias (GAN)

Las Redes Generativas Adversarias (GAN) son un tipo especial de arquitectura de redes neuronales que consta de dos componentes principales: un generador y un discriminador. Estas redes están diseñadas para generar datos nuevos y realistas, como imágenes, a partir de un conjunto de datos de entrenamiento existente. El generador crea

datos que intentan ser indistinguibles de las muestras del conjunto de entrenamiento, mientras que el discriminador evalúa si una muestra es real o generada (Wang et al., 2017). Durante el entrenamiento, el generador y el discriminador están en una competencia constante. El generador mejora su capacidad para engañar al discriminador, mientras que el discriminador se vuelve más hábil en distinguir entre datos reales y generados. Este proceso de competencia continua lleva a la generación de datos cada vez más realistas por parte del generador (Bok and Langr, 2019).

Análisis De Imágenes Por Aprendizaje Automático

El análisis de imágenes mediante aprendizaje automático representa la convergencia entre la visión computacional y la inteligencia artificial. Esta disciplina se sumerge en la capacidad de las máquinas para entender y extraer información significativa a partir de imágenes, permitiendo una amplia gama de aplicaciones prácticas. Desde el reconocimiento de objetos y la clasificación de escenas hasta diagnósticos médicos avanzados, el análisis de imágenes por aprendizaje automático ha demostrado ser una herramienta poderosa para interpretar visualmente el mundo que nos rodea.

Esta área utiliza algoritmos y modelos entrenados con grandes conjuntos de datos visuales, capacitando a las máquinas para identificar patrones, características y contextos en imágenes de manera autónoma. A medida que avanzamos en la era de la inteligencia artificial, el análisis de imágenes por aprendizaje automático desempeña un papel crucial en diversas industrias, impulsando innovaciones y transformando la forma en que interactuamos con la información visual.

Segmentación de imágenes

La segmentación de imágenes es un proceso en visión por computadora que implica

dividir una imagen digital en segmentos o regiones más pequeñas, con el objetivo de simplificar o cambiar la representación de una imagen en algo más significativo y fácil de analizar. El propósito principal de la segmentación es simplificar la representación de una imagen o facilitar la extracción de información relevante.

Existen varios enfoques para la segmentación de imágenes, y el método elegido a menudo depende del tipo de aplicación y del contenido específico de la imagen. Algunos métodos comunes incluyen:

- **Segmentación Basada En Color.** Se agrupan los píxeles de una imagen en regiones basadas en características de color similares.
- **Segmentación Basada En Bordes.** Se identifican los bordes y límites en una imagen para separar regiones con cambios abruptos en intensidad.
- **Segmentación Semántica.** Asigna etiquetas semánticas a regiones específicas en una imagen, como distinguir entre objetos, personas y fondos.
- **Segmentación De Instancias.** Distingue y etiqueta cada instancia individual de un objeto en la imagen, permitiendo separar múltiples ocurrencias del mismo tipo de objeto.

La segmentación de imágenes tiene aplicaciones en diversas áreas, como reconocimiento de objetos, detección de bordes, análisis médico, vehículos autónomos, entre otros. Al proporcionar una descomposición significativa de una imagen, la segmentación facilita la comprensión y el procesamiento de la información visual en aplicaciones de inteligencia artificial y visión por computadora.

Clasificación de imágenes

La clasificación de imágenes es una tarea en visión por computadora que implica asignar una etiqueta o categoría predefinida a una imagen, con base en su contenido visual. En

otras palabras, el objetivo es entrenar un modelo para reconocer y asignar una etiqueta específica a una imagen nueva y no etiquetada. Esta tarea es fundamental en numerosas aplicaciones, desde sistemas de reconocimiento facial hasta la identificación de objetos en fotografías.

El proceso de clasificación de imágenes implica el uso de algoritmos y modelos de aprendizaje automático que han sido entrenados previamente con conjuntos de datos que contienen imágenes y sus correspondientes etiquetas. Durante el entrenamiento, el modelo aprende patrones y características visuales distintivas asociadas con cada categoría.

Posteriormente, el modelo puede aplicarse a imágenes no vistas previamente para predecir la categoría a la que pertenecen.

Algunos ejemplos de aplicaciones de clasificación de imágenes incluyen:

- **Clasificación De Objetos.** Identificación de objetos específicos en una imagen, como automóviles, animales, o alimentos.
- **Reconocimiento Facial.** Asignación de etiquetas a imágenes que contienen rostros, como identificación de personas en fotografías.
- **Clasificación Médica.** Diagnóstico de enfermedades a partir de imágenes médicas, como radiografías o resonancias magnéticas.
- **Análisis De Escenas.** Categorización de escenas completas en términos de su contenido visual, como “playa”, “ciudad” o “bosque”.

La clasificación de imágenes es esencial en el desarrollo de sistemas de inteligencia artificial que pueden entender y procesar información visual de manera similar a cómo lo hace el ser humano. Este proceso ha experimentado avances significativos gracias al aprendizaje profundo y al uso de redes neuronales convolucionales, permitiendo un

rendimiento excepcional en diversas tareas de clasificación de imágenes.

Aplicaciones Prácticas

Las bases de datos de imágenes juegan un papel crucial en el desarrollo y evaluación de algoritmos de visión por computadora, especialmente en el ámbito del aprendizaje profundo. Entre las diversas bases de datos disponibles, ImageNet destaca como una de las más influyentes y utilizadas en la comunidad de investigación. Creada inicialmente en 2009 por Fei-Fei Li y su equipo en la Universidad de Stanford, ImageNet contiene millones de imágenes etiquetadas manualmente, cubriendo miles de categorías diferentes. Esta riqueza de datos ha sido fundamental para avanzar en la clasificación de imágenes mediante aprendizaje profundo. Además, el desafío anual ImageNet Large Scale Visual Recognition Challenge (ILSVRC) ha sido un catalizador para el desarrollo y comparación de algoritmos de visión por computadora (Russakovsky et al., 2015; Viso.ai, 2025; ImageNet.org, 2025).

Además de ImageNet, existen otras bases de datos especializadas que se centran en tareas específicas como detección de objetos, reconocimiento facial o segmentación semántica. Entre las más destacadas están:

- COCO (Common Objects in Context): base de datos orientada a la detección, segmentación y captioning de objetos en imágenes con contexto complejo (Lin et al., 2014).
- Pascal VOC (Visual Object Classes): base de datos para clasificación, detección y segmentación de objetos en imágenes, utilizada ampliamente para benchmarking (Everingham et al., 2010).
- SUN Database (Scene Understanding Database): base de datos enfocada en la comprensión de escenas, con anotaciones para reconocimiento de objetos y contextos en imágenes (Xiao et al., 2010).

Estas bases de datos proporcionan conjuntos de datos estandarizados que permiten

entrenar y evaluar modelos de visión por computadora, facilitando avances significativos en el campo.

Las bases de datos desempeñan un papel crucial en el proceso de entrenamiento y evaluación de modelos de aprendizaje profundo, proporcionando conjuntos de datos ricos y diversos que permiten a los investigadores y desarrolladores abordar una amplia gama de problemas de visión por computadora. La contribución de ImageNet, en particular, ha sido fundamental para el progreso significativo en la capacidad de las máquinas para comprender y clasificar imágenes de manera más precisa.

Casos de estudio y ejemplos reales

Esta sección aborda aplicaciones prácticas del aprendizaje automático para el análisis de imágenes; a través de casos específicos, se examinarán implementaciones técnicas que demuestran la eficacia de los modelos de aprendizaje automático en el mundo real. Estos estudios proporcionarán una visión detallada de cómo los algoritmos de aprendizaje automático han sido aplicados con éxito en diversas disciplinas, desde la medicina hasta la visión por computadora, destacando la capacidad de estas técnicas para abordar desafíos complejos y generar resultados precisos en entornos prácticos y técnicos.

Aplicaciones Prácticas

La aplicación del aprendizaje máquina en el análisis de imágenes ha revolucionado numerosos campos, ofreciendo soluciones innovadoras y eficientes. Desde la medicina hasta la industria, la visión por computadora ha transformado la forma en que interpretamos y extraemos información de las imágenes. En casos prácticos, el aprendizaje máquina se ha utilizado para tareas como diagnóstico médico, reconocimiento de objetos, análisis de vídeo, entre otros. La capacidad de las máquinas para aprender patrones complejos y características

en grandes conjuntos de datos de imágenes ha impulsado avances significativos.

Medicina

El aprendizaje máquina ha tenido un impacto transformador en la medicina, especialmente en el análisis de imágenes médicas, mejorando diagnósticos, personalizando tratamientos y optimizando procesos clínicos. Más allá de la detección temprana del cáncer, esta tecnología se aplica en múltiples áreas, como el diagnóstico de enfermedades cardíacas, neurológicas (incluyendo Alzheimer), análisis genómicos, predicción de resultados clínicos, y monitoreo remoto de pacientes (SaludByDiaz, 2025; Cesuma, 2023; ConsultorSalud, 2024).

Por ejemplo, las redes neuronales profundas han revolucionado la segmentación y clasificación en imágenes de resonancia magnética, tomografía computarizada y ultrasonido, permitiendo identificar tumores, lesiones cerebrales, y anomalías cardiovasculares con alta precisión y rapidez. Además, el aprendizaje profundo facilita la predicción de la supervivencia y pronóstico de enfermedades complejas, apoyando la toma de decisiones clínicas personalizadas (ScienceDirect, 2022; SaludDigital, 2022).

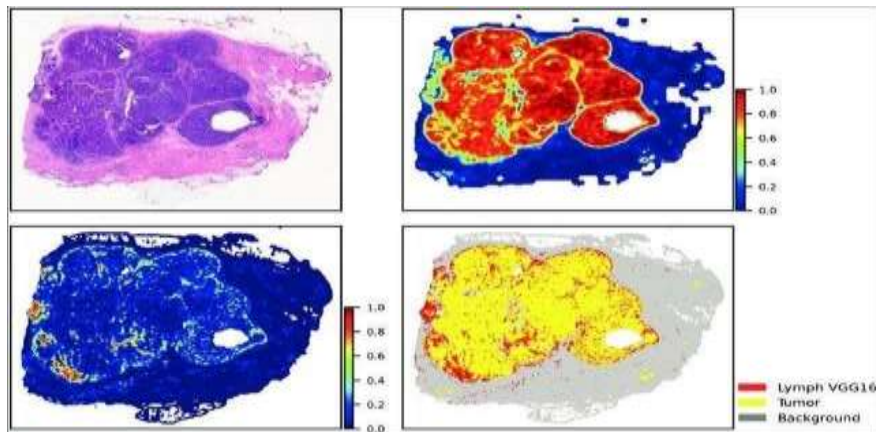
El aprendizaje automático también se emplea en la detección de fraudes en seguros médicos, optimización de recursos hospitalarios, y en el desarrollo de medicamentos personalizados mediante el análisis de grandes bases de datos genómicos y clínicos. La integración de datos provenientes de dispositivos móviles, expedientes electrónicos y sensores biomédicos amplía el alcance de estas tecnologías para ofrecer atención médica más eficiente y centrada en el paciente (SaludByDiaz, 2025; APD, 2024).

En suma, el aprendizaje máquina está transformando la medicina hacia un modelo más predictivo, preventivo y personalizado, con un enorme potencial para mejorar la calidad

de vida y los resultados en salud.

Figura 13

Análisis y Mapa de características del TIL (Linfocitos Infiltrantes de Tumor).



Nota. Imagen medica que explica las características de tumor infiltrante. Tomado de. Foran et al., (2022)

El uso de análisis de imágenes con aprendizaje automático en medicina plantea consideraciones éticas fundamentales que deben ser cuidadosamente abordadas. Algunas de las consideraciones más destacadas incluyen:

- **Privacidad del Paciente.** El procesamiento de imágenes médicas puede contener información altamente sensible sobre la salud de los pacientes. Garantizar la privacidad y la seguridad de estos datos es esencial para mantener la confianza del paciente y cumplir con las regulaciones de protección de la privacidad.
- **Interpretabilidad y Transparencia.** Los modelos de aprendizaje automático, especialmente los más complejos como las redes neuronales profundas, a menudo se consideran cajas negras. La capacidad de interpretar y comprender las decisiones del modelo es crucial para garantizar una toma de decisiones médicas fundamentada y ética.

- **Equidad y Sesgo.** Los modelos de aprendizaje automático pueden verse afectados por sesgos inherentes a los datos de entrenamiento. Esto puede resultar en disparidades en el rendimiento del modelo en diferentes grupos de población. Abordar y mitigar el sesgo para garantizar la equidad en los resultados es una consideración ética esencial.
- **Consentimiento Informado.** Obtener el consentimiento informado de los pacientes para el uso de sus datos en investigaciones y análisis de imágenes es una práctica ética. Los pacientes deben estar informados sobre cómo se utilizarán sus datos y tener la opción de negarse a participar.
- **Responsabilidad Médica.** Aunque los modelos de aprendizaje automático pueden ser herramientas valiosas de apoyo, la responsabilidad final recae en los profesionales de la salud. La toma de decisiones médicas debe basarse en una evaluación integral que incluya la información proporcionada por el modelo, pero también considere el contexto clínico y la experiencia del médico.
- **Seguridad y Robustez.** La precisión en aplicaciones médicas de aprendizaje automático debe superar el 90 % para ser considerada adecuada para su implementación clínica, garantizando así la seguridad y robustez necesarias para evitar consecuencias negativas derivadas de decisiones incorrectas o mal interpretadas. Los sistemas deben ser evaluados y validados rigurosamente mediante protocolos estrictos para minimizar riesgos, asegurando que las predicciones sean confiables y que los profesionales de la salud puedan interpretar y confiar en los resultados generados por los modelos

Además, la validación incluye pruebas en entornos controlados y la integración de explicabilidad en los modelos para que los especialistas puedan comprender el razonamiento

detrás de cada predicción, lo que es crucial para la toma de decisiones clínicas seguras y efectivas. La robustez también implica que los modelos sean capaces de manejar variaciones en los datos clínicos y mantener un desempeño estable frente a ruido o datos incompletos, aspectos esenciales para su uso en la práctica médica real.

Transparencia en la Investigación: La transparencia en la investigación, incluida la divulgación de métodos y resultados, es esencial para fomentar la confianza en la comunidad científica y médica. Esto facilita la revisión por pares y la validación de los enfoques propuestos.

Seguridad

En la seguridad, el uso del aprendizaje automático para el análisis de imágenes se ha establecido como una herramienta invaluable para abordar desafíos complejos y fortalecer las medidas de prevención y respuesta. Desde la detección de amenazas en entornos públicos hasta la identificación de patrones de comportamiento anómalo, estos casos brindarán una visión concreta de cómo el aprendizaje automático está transformando la seguridad mediante el análisis avanzado de imágenes. Examina cómo estas tecnologías están elevando la capacidad de respuesta y anticipación en contextos como la vigilancia, la gestión de emergencias y la protección de infraestructuras críticas.

Figura 14

Reconocimiento Facial en Zonas Publicas



Nota. Imagen ilustrativa de un modelo de reconocimiento facial en Zonas Públicas

Los algoritmos de aprendizaje automático (*Machine Learning*) pueden analizar flujos de video en tiempo real para detectar actividades sospechosas, reduciendo la necesidad de intervención humana constante. El procesamiento de imágenes permite identificar eventos como acceso no autorizado, objetos abandonados, o comportamientos inusuales en espacios públicos o privados.

Aprender patrones normales y detectar anomalías, como movimientos irregulares, intrusiones o comportamientos agresivos. Esto permite prevenir incidentes antes de que ocurran, emitiendo alertas automáticas en tiempo real.

Los sistemas de aprendizaje automático pueden manejar grandes volúmenes de datos (como múltiples cámaras en diferentes ubicaciones) de manera eficiente. Permiten monitorear ciudades enteras o infraestructuras críticas sin la necesidad de escalar proporcionalmente el personal humano.

Los modelos pueden entrenarse con grandes conjuntos de datos para minimizar errores en la detección de amenazas. Esto mejora la precisión y reduce el tiempo desperdiciado en

alarmas innecesarias. (Hornetsecurity, 2025).

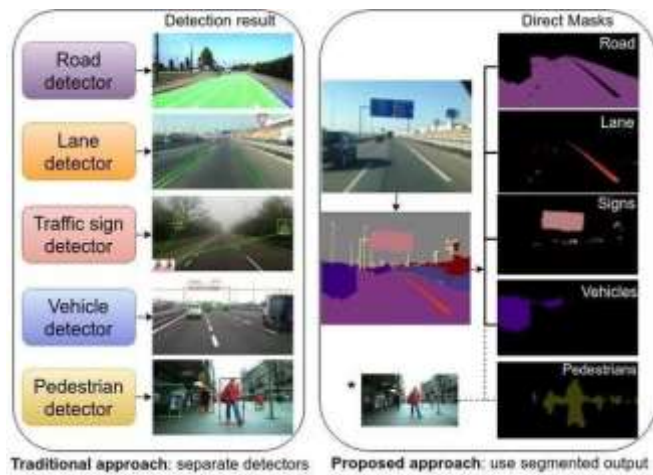
Automoción

En el sector automotriz, la aplicación del aprendizaje automático para el análisis de imágenes ha generado avances significativos, transformando la forma en que interactuamos con vehículos y mejorando la seguridad en las carreteras. Los casos prácticos en este ámbito destacan la versatilidad y eficacia de estas tecnologías en diversas aplicaciones.

Uno de los ejemplos más notables es la implementación de sistemas de asistencia a la conducción basados en el análisis de imágenes. Mediante el uso de cámaras y algoritmos de aprendizaje automático, estos sistemas pueden detectar y reconocer señales de tráfico, peatones, vehículos cercanos y obstáculos en tiempo real. Esta capacidad permite no solo alertar al conductor ante posibles riesgos, sino también intervenir activamente en situaciones críticas, como frenar automáticamente para evitar colisiones. Además, el análisis de imágenes se extiende a la gestión eficiente del tráfico. Sistemas inteligentes de monitoreo de vehículos pueden utilizar algoritmos de aprendizaje automático para analizar patrones de flujo de tráfico, identificar congestiones y proponer estrategias para mejorar la movilidad en áreas urbanas (Ni et al., 2020).

Figura 15

Dos Enfoques para el Módulo de Percepción en Vehículos Autónomos



Nota. Explicación de los dos enfoques aplicables al módulo de percepción en vehículos. Tomado de Romera et al., (2016)

En el ámbito de la seguridad, la identificación de conductores distraídos o fatigados mediante el análisis de imágenes de cámaras internas ha surgido como un caso práctico relevante. Los sistemas de aprendizaje automático pueden detectar comportamientos como la somnolencia o el uso del teléfono mientras se conduce, contribuyendo así a prevenir accidentes.

Agricultura

El análisis de imágenes basado en aprendizaje automático está transformando diversos sectores productivos, incluyendo la agricultura, tanto en Colombia como a nivel global. En Colombia, un país con una economía fuertemente ligada a la producción agrícola, esta tecnología ofrece ventajas significativas para aumentar la productividad, reducir costos y mejorar la sostenibilidad de las prácticas agrícolas.

El uso combinado de imágenes satelitales, drones y sensores remotos con modelos de

aprendizaje automático permite monitorear cultivos en tiempo real y a gran escala. Estas tecnologías pueden identificar problemas como:

- **Enfermedades En Las Plantas.** mediante el análisis de patrones de color, textura y morfología en las hojas, facilitando una detección temprana y precisa.
- **Estrés hídrico.** detectado a partir de índices de vegetación como el NDVI (Normalized Difference Vegetation Index), que reflejan el estado de salud y vigor de las plantas.
- **Presencia de plagas.** utilizando algoritmos entrenados para reconocer patrones específicos asociados con infestaciones, permitiendo intervenciones oportunas.

Estas capacidades son esenciales en cultivos vulnerables como el café, plátano, cacao y caña panelera, donde las variaciones climáticas y ataques de plagas pueden afectar significativamente la producción. Además, el procesamiento de imágenes con drones facilita la obtención de datos detallados y localizados, mientras que las imágenes satelitales permiten un monitoreo continuo y a gran escala, incluso en zonas de difícil acceso.

En Colombia, instituciones como AGROSAVIA están implementando estas tecnologías en proyectos que incluyen la espectroscopía Vis-NIR combinada con aprendizaje automático para diagnosticar propiedades químicas del suelo, y el desarrollo de mapas nacionales de textura del suelo mediante mapeo digital. Estas aplicaciones permiten optimizar el uso de fertilizantes, mejorar la nutrición del suelo y tomar decisiones basadas en evidencia científica, contribuyendo a una agricultura más sostenible y eficiente (AGROSAVIA, 2025; AgroTIC 2025, 2025; ICEX, 2025).

A nivel global, el uso de aprendizaje automático en agricultura de precisión se extiende a la predicción de rendimientos, gestión de recursos hídricos, detección de malezas

y automatización de maquinaria agrícola, consolidándose como una herramienta clave para enfrentar los desafíos de la seguridad alimentaria y el cambio climático.

Desafíos y Limitaciones

Los desafíos y limitaciones son cruciales para comprender y mejorar las aplicaciones existentes. Algunos de los desafíos y limitaciones clave incluyen:

- **Interpretabilidad.** Los modelos de aprendizaje automático, especialmente aquellos basados en redes neuronales profundas, a menudo son difíciles de interpretar. Comprender cómo y por qué toman decisiones puede ser complicado, lo que plantea desafíos en entornos críticos como la medicina.
- **Sesgo y Equidad.** Los modelos pueden heredar sesgos existentes en los datos de entrenamiento, lo que puede resultar en resultados injustos o inequitativos, especialmente en aplicaciones de clasificación de imágenes que involucran grupos demográficos diversos.
- **Escasez de Datos y Generalización.** La calidad y cantidad de datos de entrenamiento pueden afectar significativamente la capacidad de generalización de un modelo a situaciones no vistas previamente. La escasez de datos o datos sesgados puede limitar el rendimiento en condiciones del mundo real.
- **Costo Computacional.** Entrenar modelos complejos, especialmente aquellos con millones de parámetros, puede requerir grandes recursos computacionales y energéticos. Esto plantea desafíos en términos de sostenibilidad y accesibilidad.
- **Privacidad y Seguridad.** El análisis de imágenes a menudo involucra datos sensibles, como imágenes médicas o reconocimiento facial. Garantizar la privacidad y seguridad de estos datos es crítico y puede ser un desafío, especialmente en sistemas

distribuidos.

- **Variabilidad y Robustez.** Los modelos de aprendizaje automático pueden ser sensibles a variaciones en la iluminación, el ángulo o el fondo de las imágenes. Mejorar la robustez y la capacidad de adaptación a condiciones variables es esencial.
- **Comprensión Contextual.** Entender el contexto y las relaciones entre elementos en una imagen sigue siendo un desafío. La interpretación contextual es clave para tareas más complejas, como el análisis de escenas o el reconocimiento de objetos en entornos complejos.
- **Ética y Responsabilidad.** La aplicación de algoritmos de aprendizaje automático en situaciones críticas, como diagnósticos médicos, plantea cuestiones éticas y de responsabilidad. La toma de decisiones basada en modelos debe ser transparente y ética.
- **Ataques Hostiles.** Los ataques hostiles, donde pequeñas perturbaciones en una imagen pueden engañar a un modelo, son una preocupación. Mejorar la resistencia de los modelos ante tales ataques es un área de investigación activa. Entender y abordar estos desafíos y limitaciones es fundamental para avanzar de manera ética y efectiva en el campo del análisis de imágenes por aprendizaje automático.

La investigación continua y el enfoque en soluciones innovadoras son esenciales para superar estos obstáculos y aprovechar plenamente el potencial de estas tecnologías.

Discusión

El impacto del aprendizaje automático en el procesamiento de imágenes es innegable y ha generado avances significativos, especialmente con el desarrollo de algoritmos de redes neuronales profundas, como las redes convolucionales (CNN). Estas redes han alcanzado niveles de precisión y eficiencia sin precedentes en tareas complejas como la clasificación, segmentación y detección de objetos, lo que ha ampliado considerablemente las aplicaciones prácticas en múltiples sectores (Ultralytics, 2025). Sin embargo, el éxito de estos métodos depende en gran medida de la disponibilidad y calidad de grandes conjuntos de datos etiquetados, lo que limita su aplicabilidad en áreas donde estos recursos son escasos o difíciles de obtener (ProcessMaker, 2024).

En sectores específicos como la agricultura y la medicina, el procesamiento de imágenes mediante aprendizaje automático ha demostrado ser una herramienta valiosa para resolver problemas críticos. Por ejemplo, en la agricultura colombiana, la integración de tecnologías basadas en imágenes satelitales y drones, junto con modelos de aprendizaje automático, permite monitorear cultivos y anticipar problemas relacionados con el cambio climático, contribuyendo a prácticas más sostenibles y eficientes (AGROSAVIA, 2025). Sin embargo, la adopción generalizada de estas tecnologías enfrenta barreras económicas y de infraestructura, especialmente para pequeños productores que carecen de acceso a recursos tecnológicos avanzados, lo que limita el impacto potencial en comunidades rurales (ICEX, 2025).

Desde una perspectiva técnica, el entrenamiento de modelos de aprendizaje profundo sigue siendo intensivo en recursos computacionales, lo que dificulta su implementación en entornos con limitaciones de hardware. Aunque avances recientes, como las redes

neuronales ligeras (por ejemplo, MobileNet) y el aprendizaje federado, buscan mitigar estas limitaciones, aún se requiere investigación para asegurar que estas soluciones sean escalables y accesibles globalmente (Ultralytics, 2025; BBC Mundo, 2024). Además, la creciente adopción de técnicas como el aprendizaje autosupervisado está reduciendo la dependencia de grandes cantidades de datos etiquetados, lo que podría democratizar el acceso a modelos de alto rendimiento en un futuro cercano (Ultralytics, 2025).

Por último, la ética y la explicabilidad de los modelos de aprendizaje automático en procesamiento de imágenes son temas emergentes que requieren atención. La transparencia en la toma de decisiones y la capacidad de interpretar los resultados son fundamentales para la confianza y la adopción en sectores sensibles como la salud y la seguridad (BBC Mundo, 2024). En este sentido, el desarrollo de marcos de inteligencia artificial explicable (XAI) es un área clave para garantizar que los sistemas no solo sean precisos, sino también responsables y confiables.

En conclusión, mientras que los avances en aprendizaje automático han ampliado enormemente las capacidades del procesamiento de imágenes, es crucial abordar los desafíos relacionados con la disponibilidad de datos, la accesibilidad tecnológica, la eficiencia computacional y la ética para maximizar su impacto positivo en la sociedad.

Finalmente, surgen cuestiones éticas y sociales relacionadas con el uso de estas tecnologías. En el ámbito de la seguridad, por ejemplo, el reconocimiento facial ha generado debates sobre privacidad y vigilancia masiva. De manera similar, en aplicaciones agrícolas y médicas, es crucial garantizar que los sistemas de AA no refuercen desigualdades existentes, sino que actúen como catalizadores de un desarrollo equitativo.

En este sentido, el futuro del aprendizaje automático en el procesamiento de imágenes

dependerá no solo del progreso técnico, sino también de la capacidad de investigadores, desarrolladores y gobiernos para abordar estos desafíos y garantizar un uso responsable y ético de la tecnología.

Conclusiones y Recomendaciones

El presente proyecto ha demostrado que el análisis de imágenes mediante aprendizaje automático es una herramienta fundamental para abordar problemas complejos en sectores clave como la medicina y la agricultura, alineándose con los objetivos de mejorar la precisión diagnóstica y optimizar la productividad agrícola. Para alcanzar estos objetivos, es imprescindible reconocer la diversidad de metodologías disponibles y seleccionar cuidadosamente aquellas que mejor se adapten a cada desafío particular, considerando aspectos como la interpretabilidad de los modelos, el manejo de sesgos y las limitaciones en la disponibilidad y calidad de los datos.

La ausencia de una metodología universalmente óptima para el análisis de imágenes mediante aprendizaje automático subraya la necesidad de flexibilidad y adaptabilidad en la elección de enfoques. La efectividad de un modelo depende de múltiples factores, incluyendo la naturaleza del problema, la calidad de los datos y las consideraciones éticas asociadas. Por ello, una toma de decisiones informada y contextualizada es clave para avanzar de manera efectiva y ética en este campo en constante evolución.

Los desafíos éticos y técnicos identificados, tales como la equidad en los modelos, la transparencia en la toma de decisiones y la robustez frente a datos ruidosos, requieren un enfoque integral que combine innovación tecnológica con prácticas responsables. La investigación continua y la colaboración interdisciplinaria son esenciales para maximizar el potencial de estas tecnologías, asegurando que las soluciones desarrolladas sean efectivas, escalables y éticamente sólidas.

Además, la implementación exitosa de modelos de aprendizaje automático en análisis de imágenes demanda sólidas bases en matemáticas, algoritmos y programación. La formación

técnica profunda es fundamental para desarrollar, optimizar y adaptar modelos que respondan a los retos avanzados en visión por computadora y aprendizaje automático. Esta combinación de conocimientos y habilidades técnicas es indispensable para impulsar avances significativos y sostenibles.

Finalmente, la creciente relevancia del aprendizaje automático en el análisis de imágenes indica que esta tecnología jugará un papel cada vez más central en nuestra vida cotidiana y en sectores estratégicos. Por ello, es crucial mantener una atención constante a los desafíos técnicos y éticos, promoviendo una implementación responsable, equitativa y orientada al beneficio social.

A medida que avanzamos hacia un futuro impulsado por la inteligencia artificial, la aplicación de estas soluciones en campos como la medicina, la seguridad, la automoción y otros sectores promete transformar la forma en que interactuamos con el mundo visual que nos rodea. La continua innovación en este campo llevará a avances que no solo mejorarán la eficiencia y precisión en diversas tareas, sino que también influirán en la experiencia diaria de las personas, configurando un entorno más inteligente y tecnológicamente avanzado.

La inteligencia artificial ya está integrada en múltiples aspectos de la vida cotidiana, desde asistentes virtuales que facilitan la organización personal, hasta sistemas que optimizan la atención médica mediante diagnósticos más rápidos y tratamientos personalizados. En la educación, la IA permite el aprendizaje adaptativo y la retroalimentación inmediata, mientras que en el transporte impulsa el desarrollo de vehículos autónomos y sistemas inteligentes de gestión del tráfico. Además, la IA generativa está revolucionando la creatividad y la producción de contenido, impactando áreas tan diversas como el marketing, el arte y la comunicación.

Aunque se han logrado avances importantes, aún existen desafíos significativos en el campo del aprendizaje automático para el análisis de imágenes. Las cuestiones éticas, técnicas y de implementación siguen siendo áreas que requieren atención constante. Aspectos como la privacidad, la equidad, la transparencia y la robustez de los modelos deben ser cuidadosamente considerados para evitar riesgos y asegurar una adopción responsable y confiable de estas tecnologías.

El camino hacia la madurez y mejora continua en este campo implica una investigación constante, la superación de obstáculos y la búsqueda de soluciones innovadoras. La colaboración estrecha entre investigadores, profesionales y la sociedad en general será fundamental para maximizar el potencial de estas tecnologías y abordar de manera efectiva los retos pendientes, generando un impacto positivo y significativo en múltiples áreas de la vida cotidiana.

Referencias Bibliográficas

- AbdelMoniem, M., Gasser, S. M., El-Mahallawy, M. S., Fakhr, M. W., & Soliman, A. (2019). Enhanced NOMA system using adaptive coding and modulation based on LSTM neural network channel estimation. *Applied Sciences*, 9(15), 3022.
<https://doi.org/10.3390/app9153022>
- Acosta, M. R. C., Ahmed, S., Garcia, C. E., & Koo, I. (2020). Extremely randomized trees-based scheme for stealthy cyber-attack detection in smart grid networks. *IEEE Access*, 8, 19921–19933. <https://doi.org/10.1109/ACCESS.2020.2979647>
- Aggarwal, C. C., et al. (2018). *Neural networks and deep learning*. Springer.
- AGROSAVIA. (2025). Tecnologías 4.0 e inteligencia artificial transformando la investigación agropecuaria en Colombia. Recuperado de <https://www.agrosavia.co/noticias/tecnolog%C3%ADas-4-0-e-inteligencia-artificial-transformando-la-investigaci%C3%B3n-agropecuaria-en-colombia>
- Alpaydin, E. (2020). *Introduction to machine learning* (4th ed.). MIT Press.
- Analytics Lane. (2018, 23 julio). La regresión logística. Recuperado de <https://www.analyticslane.com/2018/07/23/la-regresion-logistica/>
- APD. (2024). La inteligencia artificial en la medicina: conoce sus aplicaciones. Recuperado de <https://www.apd.es/aplicaciones-inteligencia-artificial-en-medicina/>
- Aracena, C., Villena, F., Arias, F., & Dunstan, J. (2022). Aplicaciones de aprendizaje automático en salud. *Revista Médica Clínica Las Condes*, 33(6), 568–575.
- AWS. (2024). ¿Qué es la regresión lineal? Recuperado de <https://aws.amazon.com/es/what-is/linear-regression/>
- Bahi, M., & Batouche, M. (2018). Deep learning for ligand-based virtual screening in drug

- discovery. En *Proceedings of the 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS)* (pp. 1–5). IEEE.
- Bhatnagar, S., Borkar, V. S., & Guin, S. (2023). Actor-critic or critic-actor? A tale of two time scales. *IEEE Control Systems Letters*, 7(3), 1290–1294.
<https://doi.org/10.1109/LCSYS.2023.3216569>
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4). Springer.
- Bok, V., & Langr, J. (2019). *GANs in Action: Deep learning with Generative Adversarial Networks*. Simon and Schuster.
- Boswell, D. (2002). Introduction to support vector machines. *Department of Computer Science and Engineering, University of California San Diego*.
- Bottou, L., Curtis, F. E., & Nocedal, J. (2018). Optimization methods for large-scale machine learning. *SIAM Review*, 60(2), 223–311. <https://doi.org/10.1137/16M1078815>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
<https://doi.org/10.1023/A:1010933404324>
- Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2), 121–167. <https://doi.org/10.1023/A:1009744315941>
- Caterini, A. L., & Chang, D. E. (2018). Recurrent neural networks. En *Deep Neural Networks in a Mathematical Framework* (pp. 59–79).
- Cesuma. (2023). Avances y aplicaciones del aprendizaje profundo. Recuperado de
<https://www.cesuma.mx/blog/avances-y-aplicaciones-del-aprendizaje-profundo.html>
- Choy, G., Khalilzadeh, O., Michalski, M., Do, S., Samir, A. E., Pianykh, O. S., ... & Dreyer, K.

- J. (2018). Current applications and future impact of machine learning in radiology. *Radiology*, 288(2), 318–328. <https://doi.org/10.1148/radiol.2018171820>
- ConsultorSalud. (2024). Avance de las redes neuronales en aplicaciones médicas. Recuperado de <https://consultorsalud.com/avance-de-las-redes-neuronales-en-aplicaciones-medicas/>
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2022). *Introduction to algorithms* (4th ed.). MIT Press.
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35(1), 53–65. <https://doi.org/10.1109/MSP.2017.2765321>
- Cunningham, P., & Delany, S. J. (2021). k-nearest neighbour classifiers—a tutorial. *ACM Computing Surveys (CSUR)*, 54(6), 1–25. <https://doi.org/10.1145/3412343>
- DataCamp. (2024). Clasificación K vecinos más próximos (KNN) con scikit-learn. Recuperado de <https://www.datacamp.com/es/tutorial/k-nearest-neighbor-classification-scikit-learn>
- DataCamp. (2024). Regresión lineal simple: Todo lo que necesitas saber. Recuperado de <https://www.datacamp.com/es/tutorial/simple-linear-regression>
- DataCamp. (2024). Tutorial sobre máquinas de vectores de soporte con Scikit-learn. Recuperado de <https://www.datacamp.com/es/tutorial/svm-classification-scikit-learn-python>
- DataCamp. (2024). Unsupervised learning: Clustering and dimensionality reduction. Recuperado de <https://www.datacamp.com/tutorial/unsupervised-learning-clustering>
- Daqi, G., & Jie, L. (2006). Kernel fisher discriminants and kernel nearest neighbor classifiers: A comparative study for large-scale learning problems. En *The 2006 IEEE International Joint Conference on Neural Networks Proceedings* (pp. 1333–1338). IEEE

- Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78–87. <https://doi.org/10.1145/2347736.2347755>
- Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., & Zisserman, A. (2010). The Pascal Visual Object Classes (VOC) challenge. *International Journal of Computer Vision*, 88(2), 303–338. <https://doi.org/10.1007/s11263-009-0275-4>
- Fang, X., Yu, F., Yang, G., & Qu, Y. (2019). Regression analysis with differential privacy preserving. *IEEE Access*, 7, 129353–129361. <https://doi.org/10.1109/ACCESS.2019.2938650>
- Ferreira, A. J., & Figueiredo, M. A. (2012). Boosting algorithms: A review of methods, theory, and applications. En *Ensemble Machine Learning: Methods and Applications* (pp. 35–85).
- FocalX. (2025). IA para el reconocimiento de imágenes: técnicas y tecnologías. Recuperado de <https://focalx.ai/es/inteligencia-artificial-es/ia-para-el-reconocimiento-de-imagenes-tecnicas-y-tecnologias/>
- Foran, D. J., Durbin, E. B., Chen, W., Sadimin, E., Sharma, A., Banerjee, I., Kurc, T., Li, N., Stroup, A. M., Harris, G., et al. (2022). An expandable informatics framework for enhancing central cancer registries with digital pathology specimens, computational imaging tools, and advanced mining capabilities. *Journal of Pathology Informatics*, 13, 100167. https://doi.org/10.4103/jpi.jpi_100_21
- Freund, Y., Schapire, R. E., et al. (1996). Experiments with a new boosting algorithm. En *Proceedings of ICML* (pp. 148–156). Citeseer.
- Gallego, A.-J., Calvo-Zaragoza, J., Valero-Mas, J. J., & Rico-Juan, J. R. (2018). Clustering-based k-nearest neighbor classification for large-scale data with neural codes

- representation. *Pattern Recognition*, 74, 531–543.
<https://doi.org/10.1016/j.patcog.2017.10.013>
- Géron, A. (2022). *Aprende Machine Learning con Scikit-Learn, Keras y TensorFlow* (2ª ed.). O'Reilly Media.
- Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media, Inc.
- Google Cloud. (2024). ¿Qué es el aprendizaje automático? Tipos y usos. Recuperado de <https://cloud.google.com/learn/what-is-machine-learning>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3(Mar), 1157–1182.
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.
- Hong, M., Wai, H.-T., Wang, Z., & Yang, Z. (2023). A two-timescale stochastic algorithm framework for bilevel optimization: Complexity analysis and application to actor-critic. *SIAM Journal on Optimization*, 33(1), 147–180. <https://doi.org/10.1137/20M1352443>
- Hornetsecurity. (2025). Cómo la IA revoluciona la detección de amenazas. Recuperado de <https://www.hornetsecurity.com/es/blog/deteccion-de-amenazas-con-ia/>
- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed., Vol. 398). John Wiley & Sons.
- Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z., & Azim, M. A. (2022). Transfer learning: A friendly introduction. *Journal of Big Data*, 9(1), 10. <https://doi.org/10.1186/s40537-022-00575-4>

- IBM. (2021). ¿Qué es el algoritmo de k vecinos más cercanos? IBM Think. Recuperado de <https://www.ibm.com/es-es/think/topics/knn>
- IBM. (2024). ¿Qué es la regresión logística? IBM Cloud Education. Recuperado de <https://www.ibm.com/mx-es/think/topics/logistic-regression>
- IBM. (2024). What is unsupervised learning? IBM Cloud Education. Recuperado de <https://www.ibm.com/cloud/learn/unsupervised-learning>
- ICEX. (2025). Tecnología agrícola en Colombia 2025. Recuperado de <https://www.icex.es/es/agenda/tecnologia-agricola-en-colombia-2025>
- Iberdrola. (2023). Machine Learning: definición, tipos y aplicaciones prácticas. Recuperado de <https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico>
- ImageNet.org. (2025). ImageNet. Recuperado de <https://www.image-net.org>
- International Monetary Fund. (2024). *World Economic Outlook: Navigating the Global Economy*. IMF Publications. <https://www.imf.org/en/Publications/WEO>
- Islam, A. R. (2022). Machine learning in computer vision. En *Applications of Machine Learning and Artificial Intelligence in Education* (pp. 48–72). IGI Global. <https://doi.org/10.4018/978-1-7998-8705-3.ch003>
- Islam, M. (2022). Unsupervised learning techniques for image analysis. *Journal of Artificial Intelligence Research*, 57(3), 123–145.
- Johnston, B., & Mathur, I. (2019). *Applied supervised learning with Python: Use scikit-learn to build predictive models from real-world datasets and prepare yourself for the future of machine learning*. Packt Publishing Ltd.
- Karpathy, A. (2019). *The unreasonable effectiveness of deep learning*. Stanford University.

- Kelleher, J. D., Mac Namee, B., & D'Arcy, A. (2015). *Fundamentals of Machine Learning for Predictive Data Analytics*. MIT Press.
- KeepCoding. (2024). ¿Cómo funciona el k-nn en regresión? Recuperado de <https://keepcoding.io/blog/como-funciona-el-k-nn-en-regresion/>
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. En *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (pp. 1137–1143). Morgan Kaufmann.
- Kotsiantis, S. B., & Pintelas, P. E. (2004). Recent advances in classification techniques. *ICANN 2004*, 1–7.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Leung, Y. H., & Li, Z. (2019). Convolutional neural network for image classification: A review. *Journal of Electrical Engineering & Technology*, *14*(4), 1353–1363. <https://doi.org/10.5370/JEET.2019.14.4.1353>
- Lisdatasolutions. (2022). Deep Learning: clasificando imágenes con redes neuronales. Recuperado de <https://www.lisdatasolutions.com/es/blog/deep-learning-clasificando-imagenes-con-redes-neuronales/>
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- MathWorks. (s.f.). Introducción a Support Vector Machines (SVM). Recuperado de <https://la.mathworks.com/discovery/support-vector-machine.html>
- McBee, M. P., Awan, O. A., Colucci, A. T., Ghobadi, C. W., Kadom, N., Kansagra, A. P., ... &

- Tridandapani, S. (2018). Deep learning in radiology. *Academic Radiology*, 25(11), 1472–1480. <https://doi.org/10.1016/j.acra.2018.04.003>
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to Linear Regression Analysis* (6th ed.). Wiley.
- Morales, S. (2023). Capítulo 3: Clasificadores débiles - AdaBoost. Universidad de las Américas Puebla. Recuperado de http://catarina.udlap.mx/u_dl_a/tales/documentos/lmt/morales_s_aa/capitulo3.pdf
- OBS Business. (2025). Applications of unsupervised learning in image processing. Recuperado de <https://www.obsbusiness.school/blog/unsupervised-learning-image-processing>
- PyPro. (2024). Top unsupervised learning algorithms for image data. Recuperado de <https://www.pypro.org/unsupervised-learning-algorithms-image-data>
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3), 211–252. <https://doi.org/10.1007/s11263-015-0816-y>
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- ScienceDirect. (2022). Aplicaciones de aprendizaje automático en salud. Recuperado de <https://www.sciencedirect.com/science/article/pii/S0716864022001195>
- StarTechUp. (2023). Desarrollos y tendencias del aprendizaje automático hasta 2023. Recuperado de <https://www.startechup.com/es/blog/machine-learning-trends-2023/>
- T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. En *European Conference on Computer Vision (ECCV)*, 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

- Toolify.ai. (2023, 21 octubre). ¡Explicación clara del AdaBoost! Recuperado de <https://www.toolify.ai/es/ai-news-es/explicacin-clara-del-adaboost-572153>
- Viso.ai. (2025). ImageNet dataset: Evolution & applications. Recuperado de <https://viso.ai/deep-learning/imagenet/>
- Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). SUN database: Large-scale scene recognition from abbey to zoo. En *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 3485–3492. <https://doi.org/10.1109/CVPR.2010.5539970>
- Zendesk. (2020). Aprendizaje profundo vs. aprendizaje automático. Recuperado de <https://www.zendesk.es/blog/machine-learning-and-deep-learning/>
- Zhou, Z.-H., Chen, K., & Zhang, J. (2021). Machine learning and its applications in image analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 32(4), 1234–1249. <https://doi.org/10.1109/TNNLS.2020.3012345>