

# INTELIGENCIA ARTIFICIAL PARA LA DETECCIÓN DE CÁNCER DE INTESTINO

Luisa Valentina de la Hoz Rocha, Jose Ricardo Olarte Pardo

6 de diciembre de 2023

## Resumen

Artificial intelligence is being harnessed as a tool for efficiency and optimization in today's workforce, not only in technological aspects but also in various other fields. In our case, we aspire to make it a valuable asset in the realm of healthcare. The implementation of a technological framework that enables us to monitor, detect, and prevent certain diseases; including the challenging case of colorectal cancer, can yield highly beneficial outcomes for the healthcare sector. Exploring how artificial intelligence can assist us and ensure improved performance is our primary focus in this endeavor.

## 1. Introducción

En el complejo panorama del sector de la salud en Colombia, se enfrentan desafíos significativos en términos de recursos, personal médico insuficiente y problemas de eficiencia, afectando de manera directa a quienes requieren atención especializada, como es el caso de las personas diagnosticadas con cáncer de intestino, el cual emerge como un problema crítico. El diagnóstico y tratamiento temprano del cáncer de intestino es fundamental para mejorar las tasas de supervivencia. Sin embargo, lamentablemente es común que las personas diagnosticadas con esta enfermedad no reciban el nivel de servicio necesario. Esto puede deberse a diagnósticos precipitados o falta de acción adecuada, e incluso a la presencia de posibles falsos positivos.

En Colombia, el cáncer de intestino es el segundo tipo de cáncer con mayor frecuencia después del cáncer de mama, esta situación se agrava por la falta de conocimiento detallado sobre esta enfermedad contribuyendo a una atención deficiente, generando un entorno propenso a diagnósticos erróneos. Ante esta realidad, surge la necesidad de proponer soluciones innovadoras que aborden la insuficiencia de información y la incertidumbre que rodea al cáncer de intestino en el contexto colombiano.

Con el objetivo de enfrentar esta problemática de manera efectiva, se ha realizado un proyecto el cual se enfoca principalmente en la implementación de inteligencia artificial (IA) en la colonoscopia, una de las pruebas de detección del cáncer intestinal, dado que el potencial de la IA al analizar grandes cantidades de datos con mayor rapidez y precisión que los humanos mejora la eficiencia y fiabilidad de la prueba. Empleando técnicas de machine learning, hemos logrado identificar patrones y tendencias por medio de imágenes que mejoran la eficiencia y fiabilidad de las pruebas de detección de esta enfermedad, mejorando la toma de decisiones médicas y contribuyendo a la prevención y tratamiento temprano de esta enfermedad en la población Colombiana, a su vez al ser una herramienta que apoyara a los médicos se puede dar una reducción en los costos de atención médica para esta enfermedad.

Nuestra metodología se apoya en tecnologías que nos ayudan a la recopilación de datos, agregando inteligencia artificial teniendo como principal objetivo evitar falsos negativos, e inclusive falsos positivos, siendo aplicado mediante el servicio Amazon Sage Maker en la nube AWS. Este enfoque garantiza la eficiencia en la recopilación y almacenamiento de datos, así como detección precisa de patrones mediante algoritmos de machine learning. La seguridad, proporcionada por AWS, añade un nivel adicional de confianza en el proceso, siendo uno de los servicios cloud con mejores parámetros de seguridad, crucial en el manejo de información médica sensible.

## 2. Motivación / Problema

El cáncer de intestino se ha convertido en una de las enfermedades más prevalentes y mortales en la sociedad actual. La falta de información precisa y completa sobre esta enfermedad puede llevar a diagnósticos erróneos y dejar a los pacientes en un estado de incertidumbre. En este contexto, consideramos que la aplicación de tecnologías avanzadas, como el machine learning, puede marcar la diferencia. Basándonos en los artículos [6][7] donde describen un estudio y detallan la información sobre cómo va la situación en este año 2023.

### 2.1. Características del cáncer de intestino

El cáncer de intestino, también conocido como cáncer de colon o cáncer colorectal, es uno de los más agresivos de su clase. En Estados Unidos, han estado recolectando datos en el Insituto de Cáncer Nacional de Surveillance, en su forma original, the National Cancer Institute's (NCI's) Surveillance, donde redactan [6], registros desde 1995 a 2019.

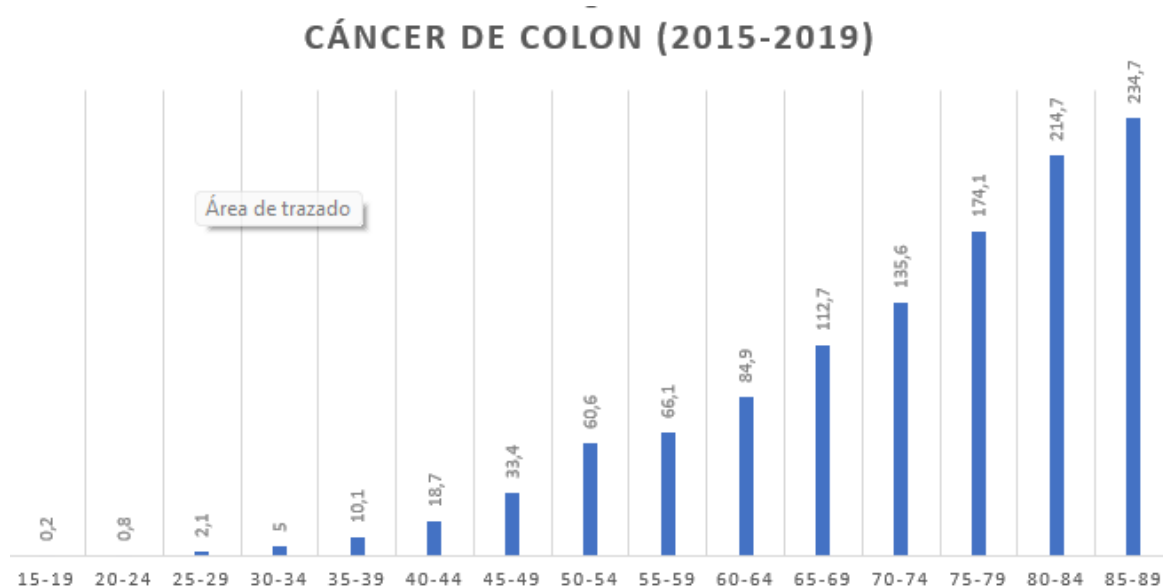


Figura 1: Tasas de incidencia de cáncer colorrectal específicas por edad, 2015-2019, Estados Unidos.[6]

### 2.2. Cáncer de intestino en Colombia

El cáncer de intestino es uno de los tipos de cáncer con mayor frecuencia en Colombia. Según la Cuenta de Alto Costo del Ministerio de Salud y Protección social, entre 2020 y 2021, esté tipo de cáncer ocupó el segundo lugar después del cáncer de mama con un total de 3.267 nuevos casos reportados [1][2]. Además, según los datos reportados por los aseguradores a la cuenta de Alto Costo, la mortalidad y la prevalencia del cáncer de intestino tuvieron un incremento del 27 % y del 14 % respectivamente[3], como podemos observar en la Figura 1.

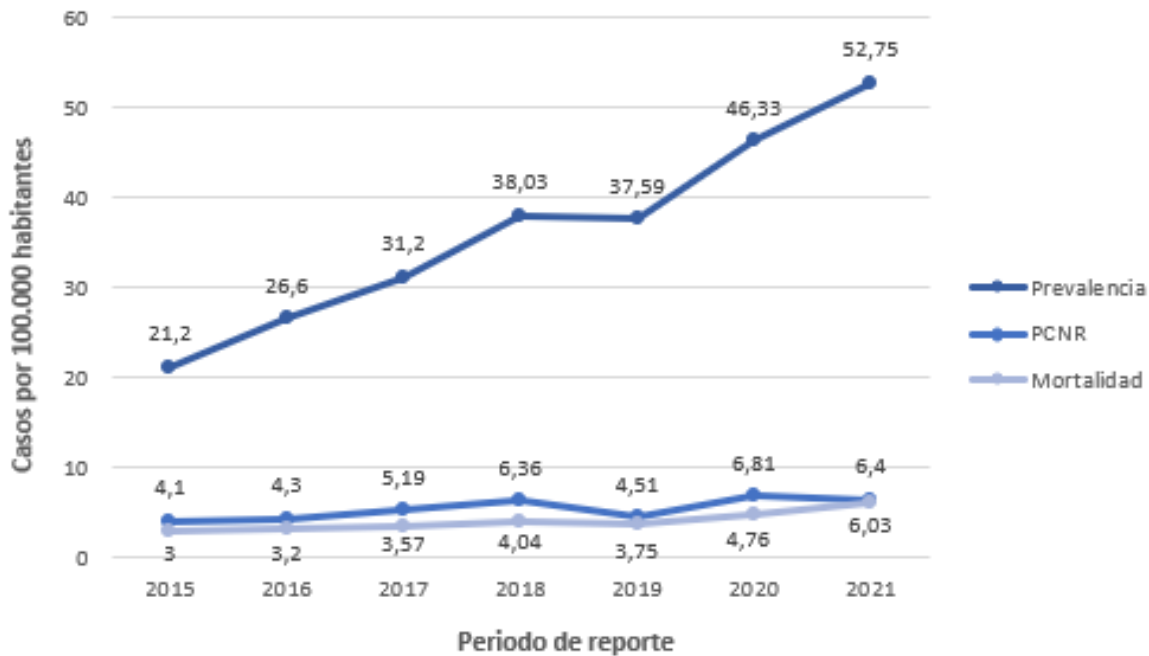


Figura 2: Prevalencia, proporción de casos nuevos reportados (PCNR) y mortalidad de Cáncer de Intestino por cada 100,000 habitantes. Datos basados en Cuenta de Alto Costo. (2022, marzo 30)

### 3. Marco Teórico

#### 3.1. Cáncer de Intestino

El cáncer intestinal es el término general dado para referirse a cualquier tipo de cáncer que se origina en el intestino. Cómo se había explicado anteriormente las características de este tipo de cáncer pueden variar dependiendo de la ubicación del tumor y la etapa que se encuentre. Algunos de sus síntomas son cambios en hábitos intestinales, dolor abdominal, debilidad o fatiga, aparición de masas en el abdomen, sangre en las heces y pérdida de peso sin razón aparente[4]. Además, el diagnóstico temprano y las terapias dirigidas en las características moleculares del cáncer son esenciales para su tratamiento[5].

#### 3.2. Inteligencia Artificial

La inteligencia artificial (IA) es una rama de la informática que desarrolla sistemas informáticos los cuales mediante el uso de matemáticas y lógica logran que simular el razonamiento que siguen las personas para aprender a partir de información nueva y toma de decisiones[9].

Por otro lado, la tecnología de la inteligencia artificial puede ayudar a las organizaciones proporcionando una comprensión más completa de los datos proporcionados. Por ejemplo, los profesionales en desarrollo recurren a la inteligencia artificial para la realización de tareas de una manera más eficaz. Esta tecnología permite la optimización de procesos como la comunicación con los clientes, la identificación de patrones o la resolución de problemas[10].

#### 3.3. Machine learning

El machine learning es un subdominio de la inteligencia artificial que brinda a las computadoras la capacidad de aprender sin una programación explícita, es decir que se centra en el desarrollo de sistemas que mejoran su rendimiento en función a los datos que se le proporcionan [11]. El machine learning comienza con una recolección y preparación de datos, los cuales serán utilizados como datos de entrenamientos, esto quiere decir que, son datos que el modelo utilizará para mejorar su rendimiento[12].

Las aplicaciones de machine learning en la actualidad se utilizan para encontrar soluciones en campos como en los vehiculos inteligentes (temperatura, inclinación del respaldo, entre otras) o la ciberseguridad (mejorar la habilidad de reconocer anomalías) [13].

### 3.4. Deep learning

Deep learning es una rama del machine learning que se basa en redes neuronales artificiales (capas de nodos que trabajan juntos para procesar y aprender datos de entradas) para modelar y resolver problemas complejos, aprendiendo patrones y relaciones complejas dentro de los datos.

Algunas de las aplicaciones de deep learning son para la detección y reconocimiento de objetos dentro de imágenes o videos, la clasificación y segmentación de imágenes. Con ello, se puede afirmar que algunas de sus grandes ventajas son la alta precisión, la escalabilidad y la flexibilidad [14].

### 3.5. Ensemble learning

Ensemble learning es una combinación de varios modelos de aprendizaje automático para la mejora de la precisión de los resultados. Algunas de sus aplicaciones actualmente son la predicción de precios en acciones o la clasificación de imagenes [16].

## 4. Arquitectura de solución

El cáncer de intestino, como se ha mencionado previamente, constituye uno de los tipos de cáncer más comunes en Colombia, ocupando el segundo puesto en incidencia después del cáncer de mama. Dando solución a esta situación se han realizado varias investigaciones e implementaciones donde se utiliza ramas de la inteligencia artificial (IA) como es el deep learning, machine learning y esemble learning para su detección.

En el siguiente artículo [15] se presenta un sistema de deep learning, el cual detecta el cáncer de intestino en muestras de ADN fecal. El sistema fue entrenado en un conjunto de datos de más de 10.000 muestras de ADN fecal, y fue capaz de identificar con precisión los pacientes con cáncer de intestino. Por otro lado, tambien se puede realizar una clasificación del cáncer mediante imagenes, en donde se implementan algoritmos de machine learning y deep learning, los cuales se entrenarán utilizando datos de imágenes para predecir la presencia y gravedad del cáncer de intestinal.

Usando la nube de Amazon Web Services (AWS) como prototipo se utiliza principalmente Amazon SageMaker en donde se realizó una recopilación de datos clínicos, relacionados a la colonoscopia, donde se analiza en tiempo real las imágenes proporcionando resultados instantáneos y apoyo a la detección temprana del cáncer de intestino. Principalmente se desarrollará un algoritmo y se utilizará el bucket S3 de almacenamiento en la nube de AWS almacenar el archivo, esto con el fin de crear un recurso de SageMaker que utilice dicho bucket como fuente de archivos, teniendo en cuenta la seguridad se le dara un rol de IAM en AWS que permite al recurso de SageMaker acceder a los recursos necesarios y finalmente se creara una instancia de SageMaker en AWS que ejecuta el recurso de SageMaker, a su vez se utilizará Amazon API Gateway el cual ofrece ofrece seguridad mediante autenticación (API keys, IAM, Cognito), cifrado SSL/TLS, control de acceso IP a su vez utilizaremos AWS Identity and Access Management (IAM) para gestionar de forma segura el acceso a los recursos de AWS. Cómo se observa en figura 3.

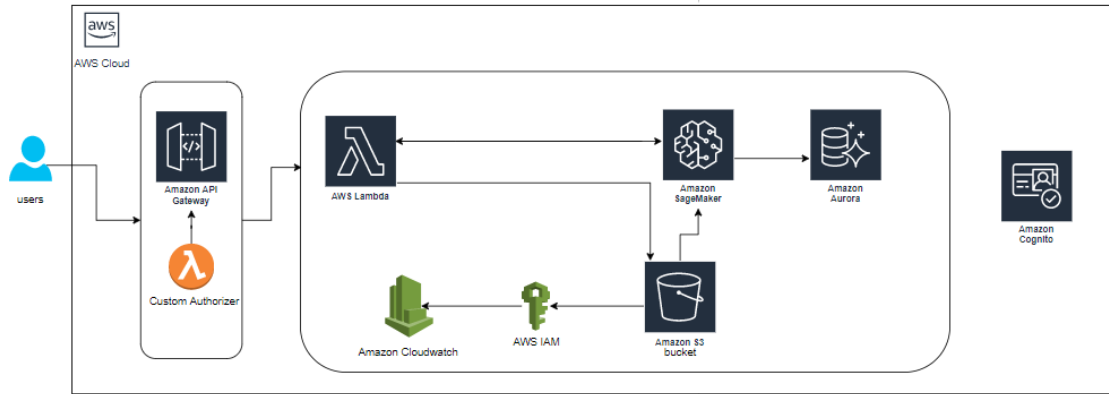


Figura 3: Arquitectura de solución

#### 4.1. Arquitectura Prototipo

La arquitectura del prototipo integra Google Colab, Amazon S3 y Amazon API Gateway con el fin de ofrecer un servicio de aprendizaje automático.

Durante el desarrollo en Google Colab, los resultados, como modelos entrenados, se almacenan en Amazon S3, posteriormente Amazon API Gateway es configurado con medidas de seguridad, como autenticación y autorización, con el fin ofrecer una API web segura, donde los usuarios interactúan con el servicio a través de dicha API, logrando un procesamiento de datos en tiempo real mientras se garantiza la seguridad. Como se observa en la figura 4, en donde se muestra de forma gráfica la arquitectura del prototipo.

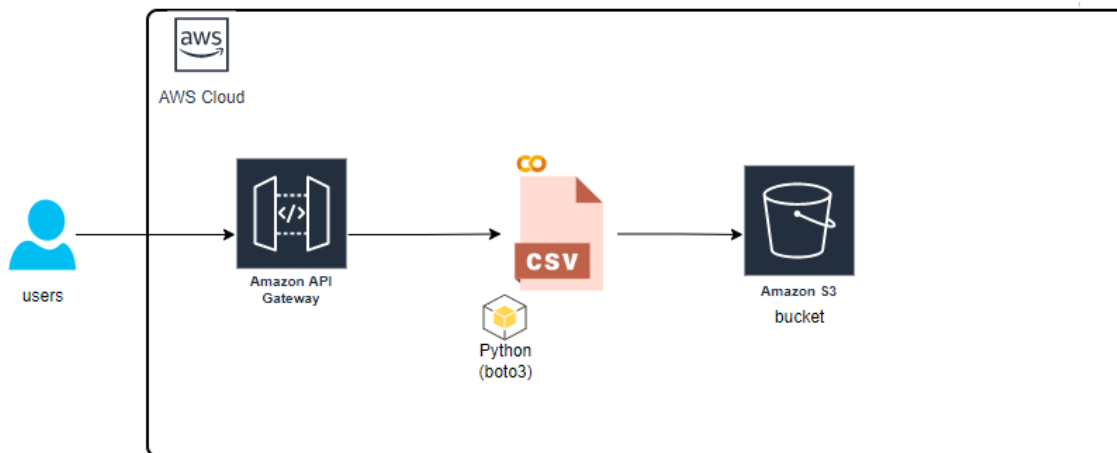


Figura 4: Arquitectura del prototipo

### 5. Experimento

Se espera un impacto positivo de este prototipo, mostrando alguno de sus beneficios, como:

#### 5.1. Diagnóstico Temprano

La detección temprana del cáncer, es una de las mayores eficiencias que puede llegar a tener estos estudios, debido a que, si se integra de forma adecuada la atención médica y utilizando dichos recursos,

se puede esperar considerablemente las tasas de supervivencia de manera positiva.

## 5.2. Recursos

Con la utilización de SageMaker en la arquitectura, junto con AWS S3 permite una administración eficiente de los recursos, reduciendo la carga operativa de los centros médicos. Además, con estas tecnologías, nos permite la escalabilidad del manejo en grandes volúmenes de datos.

## 5.3. Accesibilidad y Descentralización

La implementación en la nube posibilita el acceso a la información médica desde diferentes ubicaciones, promoviendo la colaboración entre profesionales de la salud. Además, esta descentralización podría llevar a una mayor accesibilidad a servicios de detección, incluso en áreas remotas.

En resumen, este prototipo marca un hito en la intersección de la tecnología y la salud. Su impacto inicial puede ser significativo, y su potencial para mejorar con el tiempo mediante la implementación de técnicas más avanzadas de inteligencia artificial ofrece una visión esperanzadora para la detección temprana y efectiva del cáncer de colon.

## 6. Estado del arte

En UCAS, Beijing, China varios investigadores, llegaron hasta un 98 % de precisión, en las muestras de ADN, y un 95 % en el análisis de imágenes, para la detección del cáncer de intestino. Este método al que deseamos implementar, nos basamos de ¿cómo partieron para el análisis de toda la información?. Para ello se utilizaron tres herramientas tecnológicas para la selección de información, y tener una mayor exactitud y precisión en las decisiones: Deep learning, machine learning, ensemble learning.

Este artículo [15] presenta un sistema de deep learning que puede detectar el cáncer de intestino en muestras de ADN fecal. El sistema se entrenó en un conjunto de datos de más de 10.000 muestras de ADN fecal, y fue capaz de identificar con precisión los pacientes con cáncer de intestino. Un gran comienzo, para lo que nos otorga la inteligencia artificial, y por lo que llevamos en el 2023, ha logrado grandes impactos, como se espera en esta propuesta para el desarrollo de la implementación de estas tecnologías. Teniendo en cuenta mejorar tiempos, costos, y mejorar servicios al paciente en Colombia.

Entre más información, podamos llegar a recolectar, así mismo es necesario justificar toda la recolección, pueden haber puntos de inflexión, donde hay colecciones donde no son necesarias, y para ello, es necesario realizar un filtro de la información, y seleccionar la adecuada, para evitar errores; seguimos con la ayuda de este método que realizaron los investigadores en UCAS, Beijing, China.

## 7. Conclusión

Es importante buscar constantemente nuevas tecnologías, en nuestro caso, estamos implementando una arquitectura escalable, sostenible y simple, haciendo que los proveedores de distintas identidades tendrán la facilidad de la colaboración entre ellos, por otro lado, los pacientes, pueden consultar su estado, diagnósticos, y toda la información necesaria, de manera segura y rápida. Comenzamos con la recolección de información, una primera fase que nos ayudará a seleccionar adecuadamente la información, a ir seleccionando adecuadamente la información, por medio de diagnósticos ya sea por imágenes o muestras de ADN, esto nos ayudará a tomar mejores decisiones para detectar este cáncer, esto podría hacerse mediante el uso de algoritmos de aprendizaje automático o inteligencia artificial.

Esto ayuda a la optimización y mejor atención al cliente, resaltando el punto de vista, de ayudar a más personas, y así mismo al sector de la salud en Colombia, ya que, podrán hacer consulta de la información necesaria para los pacientes, independientemente de una identidad pública o privada.

Por lo tanto cada herramienta, una ayuda, para llegar a una evolución donde se puede llegar a mitigar el número de casos reportados, en el cáncer de colon; desde ayudas con Inteligencia artificial, hasta las herramientas digitales, como un portal web, logrando una integración que proporciona una atención más accesible, eficiente y precisa, y, en última instancia, mejora la salud y el bienestar de los pacientes. La telemedicina y la tecnología de detección son aliados poderosos en la lucha contra el cáncer de colon y otras enfermedades.

La visión, lograr que el cáncer sea detectado a tiempo, y sea tratable para la reducción de mortalidad en Colombia, poco a poco, ir llegando a zonas rurales en todo el país, sin que se perjudique la arquitectura trabajada (escalable), y de esta manera, siempre apuntando que este método no sea únicamente para Colombia, también pueda ayudar a Sur América e ir expandiendo este método.

La implementación de inteligencia artificial para la detección del cáncer de colon con una tasa de precisión del 97 % representa un avance significativo en la medicina y la atención sanitaria. Esta tecnología ofrece la capacidad de analizar imágenes médicas de manera rápida y precisa, proporcionando a los profesionales de la salud una herramienta adicional para mejorar la detección temprana de esta enfermedad. La alta precisión del modelo sugiere que puede ser una herramienta valiosa en la identificación de posibles casos de cáncer de colon, lo que podría llevar a diagnósticos más rápidos y tratamientos más efectivos.

La implementación de esta tecnología en la práctica médica cotidiana podría tener un impacto significativo en la eficiencia de los servicios de salud. La detección temprana del cáncer de colon, respaldada por un modelo de inteligencia artificial de alta precisión, podría resultar en tratamientos más oportunos, reducción de costos asociados con tratamientos avanzados y, lo más importante, la mejora de las tasas de supervivencia y calidad de vida de los pacientes.

A medida que esta tecnología se integra más ampliamente en la atención médica, se espera que contribuya a la optimización de recursos y tiempo, permitiendo que los profesionales de la salud se enfoquen en casos más complejos y decisiones de tratamiento estratégicas. Además, la implementación de modelos de inteligencia artificial podría tener un impacto positivo en la reducción de la carga de trabajo de los profesionales médicos, permitiéndoles concentrarse en tareas que requieren juicio clínico y experiencia humana única.

Mirando hacia el futuro, se anticipa que la tecnología de detección de cáncer de colon basada en inteligencia artificial continuará mejorando en términos de precisión y capacidad de procesamiento. A medida que se recopilan más datos y se refinan los algoritmos, es probable que la tasa de precisión del modelo aumente aún más, brindando una herramienta aún más confiable para la detección temprana y el tratamiento personalizado.

Desde una perspectiva estadística, se espera que la implementación de este modelo lleve a una disminución en el número de diagnósticos tardíos y, por lo tanto, una disminución en la tasa de mortalidad por cáncer de colon. Además, se anticipa que los costos asociados con el tratamiento de casos avanzados disminuirán, brindando oportunidades para asignar recursos de manera más eficiente en el sistema de salud.

En resumen, la integración de la inteligencia artificial para la detección de cáncer de colon representa un avance significativo en la mejora de la atención médica. Este enfoque preciso y eficiente puede tener un impacto positivo tanto en la salud de los pacientes como en la eficiencia operativa de los sistemas de atención médica, allanando el camino hacia un futuro donde la tecnología y la experiencia médica colaboran para mejorar los resultados de salud.

## Referencias

- [1] La República. (2023, marzo 31). El cáncer de colon es el segundo más frecuente en Colombia, prevenirlo es posible. Salud Ejecutiva. <https://www.larepublica.co/salud-ejecutiva/prevenir-el-cancer-de-colon-e>
- [2] Cuenta de Alto Costo. (2021, marzo 31). Día internacional del cáncer de colon. <https://cuentadealtocosto.org/cancer/dia-mundial-del-cancer-de-colon/>
- [3] Cuenta de Alto Costo. (2022, marzo 30). Avances y retos en la atención del cáncer de colon y recto en el marco del aseguramiento en Colombia. <https://cuentadealtocosto.org/cancer/avances-y-retos-en-la-atencion-del-cancer-de-colon-y-recto-en-el-marco-del-aseguramiento-en-colombi>
- [4] MedlinePlus en español. (2004, enero 13). Cáncer de intestino: MedlinePlus en español. <https://medlineplus.gov/spanish/intestinalcancer.html>
- [5] Oncol Lett. (2018, Jul 16) The molecular characteristics of colorectal cancer: Implications for diagnosis and therapy. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6006272/>
- [6] CA: A Cancer Journal for Clinicians published by Wiley Periodicals LLC on behalf of American Cancer Society. (2023, Mar 01). <https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21772>
- [7] Markus Schmitt (2022). <https://www.datarevenue.com/en-blog/using-machine-learning-to-detect-bowel->

- [8] MedicalNewsToday (2021, Dic 06). <https://www.medicalnewstoday.com/articles/150496#summary>
- [9] Microsoft Azure. (2023, Junio 13). ¿Qué es la inteligencia artificial?. <https://azure.microsoft.com/es-es/resources/cloud-computing-dictionary/what-is-artificial-intelligence/>
- [10] Oracle. (2023, abril 13) ¿Qué es la IA? Conoce la inteligencia artificial. <https://www.oracle.com/mx/artificial-intelligence/what-is-ai/>
- [11] GeeksforGeeks. (2018, mayo 30). Machine Learning Tutorial. <https://www.geeksforgeeks.org/machine-learning/>
- [12] Sara Brown. (2021, Apr 21). Machine learning, explained. <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>
- [13] Iberdrola. (2020, enero 01). Descubre los principales beneficios del Machine Learning. <https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico>
- [14] GeeksforGeeks (2018, junio 01). Introduction to Deep Learning. <https://www.geeksforgeeks.org/introduction-deep-learning/>
- [15] MDPI (2023, Mar 23) A new method for the extraction of 3D point clouds from 2D images based on a neural network <https://www.mdpi.com/1424-8220/23/3/1225>
- [16] Neptune.ai. (2022, julio 21) A Comprehensive Guide to Ensemble Learning: What Exactly Do You Need to Know. <https://neptune.ai/blog/ensemble-learning-guide>