Influencia del peso corporal como predictor del producto dosis-longitud (DLP) en tomografía computarizada: análisis descriptivo y aplicación de modelos basados en machine learning

Andrés Felipe Giraldo Román

Asesor

Luís Angel Anillo Arrieta

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Sociales Artes y Humanidades ECSAH

Especialización en Ciencia de Datos y Analítica

2025

Nota de Aceptación

Luís Angel Anillo Arrieta

Nombre Director de Trabajo de Grado

_____ _____

Jurado Jurado

2025

Dedicatoria

A mi esposa y a mi hija por regalarme tanto de su amor y paciencia.

Resumen

La predicción de dosis de radiación en estudios de imágenes diagnósticas es un área de gran relevancia en el ámbito médico, ya que contribuye a garantizar la seguridad del paciente y la eficiencia en los procedimientos. Este trabajo de grado se enfoca en la aplicación de técnicas de *machine learning* para predecir las dosis de radiación administradas durante estudios de tomografía computarizada. Para ello, se utilizó un conjunto de datos extenso y diverso, lo que permitió realizar un análisis exploratorio detallado e identificar patrones significativos y variables relevantes.

Durante el proceso, se llevaron a cabo técnicas de limpieza y normalización de datos para asegurar su calidad y consistencia, así como la identificación y corrección de anomalías. Se evaluó la relación entre las dosis estimadas por los modelos y los niveles de referencia establecidos (percentil 75), buscando asegurar predicciones clínicamente seguras.
Aunque los modelos desarrollados no alcanzaron métricas de predicción satisfactorias, los hallazgos permiten comprender mejor las limitaciones actuales en la modelación de este tipo de datos y sientan una base importante para futuros trabajos. Este estudio aporta aprendizajes significativos que pueden orientar investigaciones posteriores hacia el desarrollo de herramientas más precisas y efectivas para la optimización del uso de la radiación en entornos clínicos.

***Palabras clave*:** Dosis de radiación, machine learning, rayos X, tomografía computarizada (TC), mamografía, seguridad del paciente.

Abstract

The prediction of radiation doses in diagnostic imaging studies is a highly relevant area in the medical field, as it contributes to ensuring patient safety and procedural efficiency. This thesis focuses on the application of *machine learning* techniques to predict radiation doses administered during computed tomography. A large and diverse dataset was used, allowing for a detailed exploratory analysis that identified significant patterns and relevant predictive variables. Throughout the process, data cleaning and normalization techniques were applied to ensure quality and consistency, as well as to detect and correct anomalies. The relationship between predicted doses and established reference levels (75th percentile) was evaluated to assess the clinical safety of the predictions.

Although the models developed did not achieve satisfactory predictive performance, the findings offer valuable insights into the current limitations of modeling this type of data and provide a solid foundation for future research. This study contributes important learnings that can guide further efforts toward developing more accurate and effective tools for optimizing radiation use in clinical settings.

***Keywords*:** Radiation dose, machine learning, X – ray, computed tomography (CT), mammography, patient safety.

Tabla de Contenido

Lista de Tablas

Lista de Figuras

Introducción

La medicina moderna ha experimentado avances significativos gracias a la tecnología, especialmente en el campo de las imágenes diagnósticas. Estudios como la tomografía computarizada se han convertido en herramientas indispensables para la detección temprana y el diagnóstico preciso de diversas enfermedades. Sin embargo, el uso de radiación en estos procedimientos plantea un desafío crucial: garantizar que las dosis administradas sean lo suficientemente bajas para proteger al paciente, pero lo suficientemente altas para obtener imágenes de calidad. Este equilibrio es fundamental para asegurar tanto la seguridad del paciente como la eficacia del diagnóstico.

En este contexto, la predicción de dosis de radiación se ha convertido en un área de investigación prioritaria. Los profesionales de la salud necesitan herramientas que les permitan optimizar el uso de la radiación, minimizando los riesgos asociados sin comprometer la calidad de los resultados. Aquí es donde la ciencia de datos y el machine learning emergen como aliados estratégicos. Estas disciplinas ofrecen la posibilidad de analizar grandes volúmenes de datos, identificar patrones y desarrollar modelos predictivos que puedan guiar decisiones clínicas más informadas y seguras.

Este trabajo de grado se enfoca en explorar el potencial de los algoritmos de *machine learning* para predecir las dosis de radiación en estudios de tomografía computarizada y mamografía. Para ello, se recopiló y analizó un conjunto de datos extenso y diverso, que refleja la variabilidad de las prácticas clínicas y las características de los pacientes. El análisis exploratorio de estos datos permitió identificar variables clave y patrones relevantes que podrían influir en las dosis administradas, sentando así una base sólida para la implementación de modelos predictivos.

Uno de los mayores retos en este proceso fue garantizar la calidad y consistencia de los datos. La limpieza y normalización fueron etapas críticas para asegurar que los resultados fuesen representativos. Asimismo, se compararon las dosis estimadas con los niveles de referencia establecidos, basados en el percentil 75, con el fin de evaluar la pertinencia clínica de las predicciones.

Si bien los modelos desarrollados no alcanzaron métricas de predicción satisfactorias, el trabajo permitió identificar limitaciones importantes tanto en los datos disponibles como en la configuración de los algoritmos, lo cual constituye un aporte valioso para investigaciones futuras. El objetivo de contribuir al desarrollo de herramientas que apoyen la optimización del uso de la radiación en estudios diagnósticos se mantiene vigente, y los aprendizajes derivados de esta experiencia ofrecen una base importante para futuros desarrollos más robustos en este campo.

Descripción del Problema

La exposición a radiación ionizante en procedimientos de diagnóstico por imágenes, como la tomografía computarizada (TC), representa una herramienta invaluable en la medicina moderna. No obstante, la creciente utilización de estas tecnologías ha suscitado preocupación por los efectos biológicos adversos asociados a exposiciones innecesarias o excesivas. Como respuesta, se han desarrollado iniciativas regulatorias y científicas para garantizar que las dosis de radiación se mantengan dentro de límites aceptables y estén en consonancia con los niveles de referencia diagnósticos (DRL, por sus siglas en inglés), promoviendo así una cultura de optimización y protección radiológica en beneficio de los pacientes (Togawa et al., 2022).

En este contexto, el presente trabajo se propuso desarrollar modelos predictivos mediante algoritmos de aprendizaje automático (machine learning) para estimar la dosis acumulada de radiación, específicamente el producto dosis-longitud (DLP), a partir de características fácilmente disponibles como el peso del paciente. Inicialmente se asumió que existía una relación significativa y aprovechable entre el peso corporal y las dosis registradas, lo que permitiría automatizar la estimación de la dosis en estudios de TC, optimizando así la toma de decisiones clínicas en tiempo real.

Sin embargo, a pesar de los avances recientes en la aplicación de modelos de machine learning en el campo de la radiología - incluyendo estimaciones automatizadas de parámetros antropométricos y predicción de dosis con base en configuraciones técnicas del escáner (Ichikawa et al., 2024; Morita et al., 2019) -, los resultados obtenidos en este estudio mostraron que el peso corporal, por sí solo, no constituye un predictor robusto ni suficiente de las dosis DLP en tomografía computarizada. La precisión de los modelos fue limitada y la capacidad de generalización resultó deficiente al ser evaluada sobre conjuntos de validación cruzada.

Los modelos entrenados, incluyendo regresiones lineales, árboles de decisión, random forest y redes neuronales, exhibieron desempeños inferiores a lo esperado. Las métricas de error, como el RMSE (Root Mean Squared Error) y el MAE (Mean Absolute Error), se mantuvieron en niveles altos, reflejando una incapacidad del algoritmo para capturar la complejidad del fenómeno solo con el parámetro de peso. Esta situación evidenció que la dosis de radiación en TC es resultado de una interacción multifactorial que no puede ser explicada de manera confiable por una única variable antropométrica.

Estudios previos han señalado la importancia de considerar otros factores, como el tipo de escáner, los parámetros técnicos del protocolo (kVp, mAs, pitch), la región anatómica examinada y características adicionales del paciente (por ejemplo, índice de masa corporal o la circunferencia abdominal), para mejorar la precisión de los modelos predictivos (Morita et al., 2019; Togawa et al., 2022). Ignorar estas variables puede conducir a estimaciones sesgadas o poco útiles en la práctica clínica.

Adicionalmente, investigaciones como la de Ichikawa et al. (2024) demostraron que incluso para estimar el peso del paciente a partir de parámetros de radiación, es necesario incorporar información detallada del protocolo de escaneo y de los parámetros técnicos, lo cual subraya aún más la naturaleza multidimensional del problema. El hecho de que el peso corporal no predijera con precisión la dosis DLP en este estudio concuerda con esa línea de pensamiento y reafirma la necesidad de modelos más integradores y complejos.

En este sentido, los hallazgos del presente trabajo reafirman que la predicción precisa de la dosis en estudios de TC no debe apoyarse en un único parámetro, sino en una visión holística del procedimiento, donde se analicen múltiples entradas simultáneamente. Si bien el peso puede

aportar información útil, su valor predictivo aislado es limitado, y confiar exclusivamente en él puede inducir a errores clínicos o decisiones subóptimas.

Como conclusión, el trabajo no logró desarrollar un modelo de predicción fiable del DLP basado únicamente en el peso corporal del paciente. Esta experiencia ofrece una valiosa lección sobre los límites de los enfoques simplificados en la inteligencia artificial aplicada a la medicina y destaca la necesidad de continuar integrando diversas fuentes de información para construir herramientas realmente útiles y seguras. Futuras investigaciones deberían enfocarse en modelos multivariantes que consideren tanto factores técnicos como antropométricos y clínicos, tal como lo sugieren los estudios previos (Morita et al., 2019; Togawa et al., 2022; Ichikawa et al., 2024).

Justificación

En el ámbito de la medicina diagnóstica, el uso de radiación es una herramienta indispensable para obtener imágenes que permitan identificar y tratar enfermedades de manera efectiva. No obstante, la exposición a la radiación implica ciertos riesgos, ya que dosis elevadas o innecesarias pueden generar efectos adversos en la salud de los pacientes. Por ello, es fundamental lograr un equilibrio entre la calidad de las imágenes obtenidas y la minimización de la dosis administrada. Alcanzar este equilibrio no solo protege al paciente, sino que también optimiza los recursos médicos y mejora la eficiencia de los procedimientos diagnósticos.

Actualmente, la determinación de las dosis de radiación se basa en protocolos estandarizados y en la experiencia del personal médico. Si bien estos métodos han sido eficaces, a menudo no contemplan la variabilidad individual de los pacientes, como el peso, la complexión o condiciones clínicas particulares, lo que puede derivar en dosis no óptimas. En este contexto, la ciencia de datos y el *machine learning* ofrecen una oportunidad prometedora para personalizar y optimizar el proceso de administración de dosis, ajustándose a las características específicas de cada paciente.

Este trabajo de grado se propuso evaluar el potencial de los modelos predictivos basados en *machine learning* para estimar dosis de radiación de manera más precisa. Aunque los modelos implementados no lograron alcanzar métricas de predicción satisfactorias, el proceso permitió identificar desafíos clave en el tratamiento y la calidad de los datos, así como en la configuración de los algoritmos. Estos hallazgos son relevantes, ya que evidencian los límites actuales de la modelación en este campo y señalan aspectos cruciales a considerar en investigaciones futuras.

Además, se mantuvo como eje central la alineación de las estimaciones con estándares internacionales de seguridad, utilizando como referencia el percentil 75 para garantizar que las

dosis propuestas se mantuvieran dentro de rangos clínicamente aceptables. Este enfoque continúa siendo esencial para proteger la salud de los pacientes y fortalecer la confianza en los procedimientos diagnósticos.

Finalmente, aunque los resultados obtenidos no fueron los esperados en términos de predicción, este trabajo aporta un análisis riguroso del problema y sienta las bases para futuras investigaciones más robustas. Representa un esfuerzo valioso por integrar el conocimiento médico con las capacidades analíticas de la ciencia de datos, contribuyendo al desarrollo de una medicina diagnóstica más precisa, segura y centrada en el paciente.

Objetivos

Objetivo General

Aplicar modelos de aprendizaje automático para evaluar la capacidad predictiva del peso corporal en la estimación de la dosis de radiación utilizada en estudios de tomografía computarizada (CT).

Objetivos Específicos

Analizar las relaciones entre las variables disponibles en los datos de dosis de radiación en estudios de tomografía computarizada mediante visualizaciones gráficas (histogramas, boxplots y matrices de correlación), con el fin de identificar anomalías, tendencias y rangos de peso corporal asociados a mayores valores de DLP, estableciendo posibles puntos de corte relevantes para la interpretación clínica o técnica.

Seleccionar e implementar modelos de machine learning (como regresión lineal, random forest, redes neuronales o XGBoost) para estimar la dosis de radiación a partir del peso corporal con el fin de contribuir a la reducción de dosis innecesarias para el paciente.

Evaluar y comparar mediante métricas de evaluación como MSE (Error Cuadrático Medio) y RMSE (Raíz del Error Cuadrático Medio) el desempeño de los diferentes algoritmos de machine learning para seleccionar el que más se acerque a la predicción de la dosis de radiación.

Evaluar si el peso del paciente puede usarse como único predictor para estimar la dosis, sin necesidad de parámetros técnicos complejos.

Marco de Referencia

Estado del Arte

La predicción de dosis de radiación en estudios de imágenes diagnósticas es un área de investigación que ha cobrado relevancia en las últimas décadas debido a la necesidad de equilibrar la calidad diagnóstica con la seguridad del paciente. La radiación ionizante, utilizada en modalidades como la tomografía computarizada (TC) y la mamografía, es esencial para obtener imágenes detalladas que permiten diagnosticar enfermedades como el cáncer, traumatismos y otras condiciones médicas. Sin embargo, la exposición a la radiación conlleva riesgos significativos, como el aumento de la probabilidad de desarrollar cáncer a largo plazo o daños en los tejidos (Brenner & Hall, 2007). Por ello, la comunidad médica y científica ha dedicado esfuerzos considerables a optimizar el uso de la radiación, buscando un equilibrio entre la calidad diagnóstica y la seguridad del paciente.

En las últimas décadas, se han establecido protocolos y estándares internacionales para regular las dosis de radiación en estudios diagnósticos. Organizaciones como la Comisión Internacional de Protección Radiológica (ICRP) y la Administración de Alimentos y Medicamentos de los Estados Unidos (FDA) han emitido directrices que limitan las dosis máximas permitidas (ICRP, 2007; FDA, 2020). Sin embargo, estos estándares suelen ser generales y no siempre consideran las particularidades de cada paciente, como su edad, peso o condición médica. Esta limitación ha llevado a la exploración de enfoques más personalizados, donde la ciencia de datos y el machine learning comienzan a jugar un papel crucial.

El machine learning ha demostrado ser una herramienta poderosa en el ámbito médico, especialmente en tareas de predicción y clasificación. En el contexto de la radiación diagnóstica, los algoritmos de machine learning pueden analizar grandes volúmenes de datos históricos,

identificar patrones y predecir dosis óptimas adaptadas a las características individuales de cada paciente. Estudios recientes han explorado el uso de técnicas como regresión lineal, árboles de decisión y redes neuronales para predecir dosis de radiación en diferentes modalidades de imágenes (Smith et al., 2019; Johnson & Lee, 2021). Por ejemplo, en tomografía computarizada, se han desarrollado modelos que consideran factores como el tipo de escáner, el protocolo utilizado y las características anatómicas del paciente (González et al., 2020).

A pesar de estos avances, aún existen desafíos significativos. Uno de los principales es la calidad y disponibilidad de los datos. Los conjuntos de datos utilizados para entrenar modelos de machine learning deben ser extensos, diversos y representativos de las poblaciones objetivo. Además, es crucial garantizar que los datos estén libres de errores y anomalías, lo que requiere un proceso riguroso de limpieza y preprocesamiento (Wang et al., 2018). Otro desafío es la interpretabilidad de los modelos. En el ámbito médico, es esencial que los profesionales de la salud comprendan cómo se generan las predicciones, lo que ha llevado a un creciente interés en modelos explicables y transparentes (Lundberg & Lee, 2017).

En resumen, el estado del arte en la predicción de dosis de radiación mediante machine learning es prometedor, pero aún en desarrollo. Los avances tecnológicos y la creciente disponibilidad de datos están abriendo nuevas oportunidades para mejorar la seguridad y eficacia de los estudios diagnósticos. Sin embargo, es necesario seguir investigando y refinando los modelos para garantizar que sean precisos, seguros y aplicables en entornos clínicos reales. Este trabajo busca contribuir a este campo en crecimiento, proponiendo un enfoque innovador que combine el conocimiento médico con las capacidades analíticas de la ciencia de datos.

Marco Contextual

En Colombia, el uso de tecnologías diagnósticas como la tomografía computarizada (TC) ha tenido un crecimiento significativo en las últimas dos décadas, en línea con las tendencias globales de modernización en los servicios de salud. No obstante, este aumento también ha generado preocupaciones sobre la exposición a radiación ionizante, especialmente cuando no se cuenta con mecanismos robustos de control de dosis ni con herramientas predictivas que apoyen la toma de decisiones clínicas. La Ley 715 de 2001 y la normatividad del Ministerio de Salud y Protección Social han impulsado la implementación de buenas prácticas en protección radiológica, sin embargo, los retos persisten, particularmente en instituciones que carecen de sistemas automatizados de monitoreo y evaluación dosimétrica.

El marco regulatorio colombiano está alineado con las recomendaciones del Organismo Internacional de Energía Atómica (OIEA) y la Comisión Internacional de Protección Radiológica (ICRP), las cuales promueven la adopción del principio ALARA (As Low As Reasonably Achievable) para la protección del paciente. En este sentido, el Ministerio de Salud expidió en 2020 la Resolución 482 de 2018, que adopta los requisitos esenciales de protección radiológica para instalaciones médicas. No obstante, la implementación efectiva de estas directrices depende en gran medida del acceso a tecnologías de apoyo como los sistemas de seguimiento de dosis (dose-tracking software) y de la formación del talento humano en herramientas analíticas avanzadas como la inteligencia artificial y el aprendizaje automático.

En el país, la mayoría de los centros hospitalarios que cuentan con equipos de tomografía computarizada siguen protocolos estándar, pero la personalización de las dosis según características antropométricas del paciente aún no es una práctica sistemática. La ausencia de sistemas automáticos de estimación y la falta de integración entre los sistemas de información

hospitalaria y los equipos de imagen limitan la posibilidad de realizar ajustes dinámicos de dosis en función de parámetros individuales. Esta situación no solo compromete la calidad del diagnóstico, sino que también puede exponer al paciente a dosis mayores a las necesarias.

La investigación sobre modelos predictivos en este contexto cobra especial relevancia, ya que permitiría a instituciones de salud colombianas —incluso aquellas con recursos limitados— implementar soluciones basadas en inteligencia artificial para mejorar la seguridad radiológica. A pesar de los intentos por desarrollar modelos usando variables simples como el peso del paciente, como se evidenció en el presente trabajo, estas aproximaciones no resultan efectivas por sí solas. La complejidad de los factores que intervienen en la determinación de la dosis requiere un enfoque más holístico y multivariado, respaldado por políticas institucionales y nacionales de innovación tecnológica en salud.

En resumen, el desarrollo de modelos predictivos robustos que puedan ser adoptados por el sistema de salud colombiano debe contemplar tanto las condiciones técnicas locales como la normativa vigente. Asimismo, es necesario fomentar un ecosistema de investigación clínica y computacional que promueva el uso responsable de tecnologías emergentes para apoyar la toma de decisiones médicas, garantizar la protección del paciente y fortalecer las capacidades diagnósticas del país.

Marco Teórico

El marco teórico de este trabajo se sustenta en tres pilares fundamentales: la radiación en imágenes diagnósticas, la ciencia de datos y el machine learning. Cada uno de estos pilares aporta conceptos y herramientas esenciales para comprender y abordar el problema de la predicción de dosis de radiación en estudios médicos.

Radiación en Imágenes Diagnósticas

La radiación ionizante es una forma de energía utilizada en medicina para generar imágenes del cuerpo humano. En estudios como la tomografía computarizada (TC) y la mamografía, la radiación atraviesa los tejidos y es capturada por detectores, creando imágenes detalladas que permiten diagnosticar enfermedades. Sin embargo, la exposición a la radiación conlleva riesgos, como daños en el ADN y un mayor riesgo de cáncer a largo plazo (Brenner & Hall, 2007). Por ello, es fundamental minimizar las dosis de radiación sin comprometer la calidad diagnóstica. Conceptos como la dosis efectiva, el índice de dosis volumétrica (CTDIvol) y el producto dosis-área (DAP) son clave para medir y controlar la exposición a la radiación (ICRP, 2007).

Ciencia de Datos

La ciencia de datos es una disciplina que combina estadística, programación y conocimiento del dominio para extraer insights valiosos a partir de datos. En el contexto médico, la ciencia de datos permite analizar grandes volúmenes de información, como historiales clínicos, resultados de estudios y datos demográficos, para identificar patrones y tendencias (Provost & Fawcett, 2013). Técnicas como el análisis exploratorio de datos (EDA), la limpieza de datos y la normalización son esenciales para garantizar que los datos estén listos para su uso en modelos predictivos. Además, la visualización de datos juega un papel crucial en la comunicación de resultados y la toma de decisiones informadas (Tukey, 1977).

Machine Learning

El machine learning es una rama de la inteligencia artificial que permite a las computadoras aprender de los datos y hacer predicciones sin ser programadas explícitamente (Mitchell, 1997). En este trabajo, se exploran algoritmos como la regresión lineal, los árboles de

decisión y las redes neuronales para predecir dosis de radiación. Cada algoritmo tiene sus ventajas y limitaciones. Por ejemplo, la regresión lineal es simple y fácil de interpretar, pero puede no capturar relaciones complejas en los datos (James et al., 2013). Por otro lado, las redes neuronales son altamente flexibles y pueden modelar relaciones no lineales, pero requieren grandes cantidades de datos y poder computacional (Goodfellow et al., 2016).

El aprendizaje automático (machine learning) es una rama de la inteligencia artificial que permite a los sistemas aprender automáticamente a partir de los datos sin ser programados explícitamente (Mitchell, 1997). Su aplicación en el ámbito médico, especialmente en el control y optimización de dosis de radiación en estudios diagnósticos como la tomografía computarizada y la mamografía, ha demostrado ser una herramienta poderosa para mejorar la seguridad del paciente y la calidad del diagnóstico (Chen et al., 2021). En este marco, se abordan los modelos de aprendizaje supervisado y no supervisado que son clave en esta investigación, proporcionando un fundamento teórico y matemático sólido.

Regresión Lineal. La regresión lineal es uno de los modelos más fundamentales y ampliamente utilizados en estadísticas y aprendizaje supervisado. Su objetivo es modelar la relación lineal entre una variable dependiente y y una o más variables independientes $\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_\mathrm{p}$. El modelo puede expresarse matemáticamente como:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \ \varepsilon \qquad (1)$$

Donde β_0 es el intercepto, β_i son los coeficientes de regresión y ϵ es el término de error (Hastie, Tibshirani & Friedman, 2009).

El proceso de estimación de los coeficientes se realiza comúnmente mediante el método de mínimos cuadrados ordinarios (OLS), el cual busca minimizar la suma de los errores cuadráticos entre los valores observados y los predichos:

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{2}$$

En notación matricial la solución óptima es:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X^T X})^{-1} \boldsymbol{X^T y} \tag{3}$$

donde:

- X es la matriz de diseño.
- y es el vector de observaciones.

La regresión lineal es fundamental como modelo base en muchos procesos predictivos, especialmente por su interpretabilidad (James et al., 2021).

Análisis de Componentes Principales (PCA). El PCA es una técnica de reducción de dimensionalidad que transforma un conjunto de variables posiblemente correlacionadas en un nuevo conjunto de variables ortogonales llamadas componentes principales. Su objetivo es capturar la mayor varianza posible con el menor número de dimensiones.

La transformación se logra encontrando los **autovalores** y **autovectores** de la matriz de covarianza SS:

$$\boldsymbol{S} = \frac{1}{n-1} \sum_{i=1}^{n} (\boldsymbol{x_i} - \overline{\boldsymbol{x}})(\boldsymbol{x_i} - \overline{\boldsymbol{x}})^T \tag{4}$$

Los vectores propios $v_1, \ldots, v_k$ que maximizan:

$$\mathbf{Var}(\boldsymbol{z_i}) = \boldsymbol{v_i^T S v_i} \tag{5}$$

son seleccionados como las direcciones de los componentes principales. La proyección se realiza como:

$$\boldsymbol{Z} = \boldsymbol{X V_k} \tag{6}$$

PCA es útil cuando se busca reducir ruido y multicolinealidad en datos clínicos de alta dimensión (Jolliffe & Cadima, 2016).

Bosques Aleatorios (Random forest). Los bosques aleatorios son una técnica de ensamble basada en la construcción de múltiples árboles de decisión, con la finalidad de mejorar la precisión y reducir el sobreajuste. Cada árbol es construido a partir de una muestra con reemplazo (bootstrap) y, en cada nodo, se selecciona aleatoriamente un subconjunto de variables para realizar la división.

La predicción del modelo se da por votación mayoritaria (clasificación) o por promedio (regresión):

$$\hat{\boldsymbol{y}} = \frac{\mathbf{1}}{\boldsymbol{T}} \sum_{\boldsymbol{t=1}}^{\boldsymbol{T}} \boldsymbol{h_t(x)} \tag{7}$$

donde $h_t(x)$ es la predicción del árbol $t - \acute{e}simo$.

La importancia de las variables puede estimarse por la reducción promedio del índice de impureza, como el índice de Gini:

$$\boldsymbol{Gini} = \mathbf{1} - \sum_{\boldsymbol{k=1}}^{\boldsymbol{K}} \boldsymbol{p_k^2} \tag{8}$$

Random Forest es eficaz para capturar relaciones no lineales y es especialmente robusto a valores atípicos y ruido (Breiman, 2001).

XGBoost (Extreme Gradient Boosting). XGBoost es una implementación optimizada de gradient boosting, que entrena modelos secuenciales minimizando una función de pérdida mediante el método de descenso de gradiente.

Para cada iteración tt, se añade un nuevo árbol que minimiza la función objetivo:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \widehat{y_i^{(t-1)}} + f_t(\boldsymbol{x_i})\right) + \Omega(f_t) \tag{9}$$

donde:

- l es la función de pérdida (ej. MSE),
- $f_t \in f$ es el árbol en la iteración t,
- $\Omega(f_t)$ es un término de regularización:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{10}$$

donde *T* es el número de hojas y w_j los pesos.

Redes Neuronales Artificiales (ANN). Las **redes neuronales artificiales** están inspiradas en el cerebro humano y consisten en capas de nodos (neuronas) conectados mediante pesos sinápticos. Una red con una sola capa oculta se puede expresar como:

$$\hat{y} = f\left(\sum_{j=1}^{m} w_j \cdot g\left(\sum_{i=1}^{n} x_i \cdot v_{ij} + b_j\right) + b\right) \tag{11}$$

donde:

- g es la función de activación (ej. ReLU, sigmoid),
- $w_j, v_{i,j}$ son pesos entre neuronas,
- b_j, b son los sesgos.

El entrenamiento se realiza mediante **backpropagation**, minimizando una función de pérdida L con gradientes:

$$\frac{\partial \mathcal{L}}{\partial w_{ij}} = \frac{\partial \mathcal{L}}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial w_{ij}} \tag{12}$$

Las ANN son potentes para modelar relaciones complejas y no lineales, aunque requieren un ajuste cuidadoso (Goodfellow et al., 2016).

Métodos de Ensamble: Bagging. Bagging (Bootstrap Aggregating) consiste en entrenar múltiples modelos sobre subconjuntos de datos generados mediante muestreo con reemplazo. Luego, sus predicciones son combinadas para reducir la varianza.

Para regresión:

$$\hat{y} = \frac{1}{M}\sum_{m=1}^{M} f_m(x) \tag{13}$$

Para clasificación, se aplica votación mayoritaria:

$$\hat{y} = \text{mode}\big(f_1(x), f_2(x), \ldots, f_M(x)\big) \tag{14}$$

Bagging mejora la estabilidad y precisión de modelos débiles, siendo una técnica base de algoritmos como Random Forest (Dietterich, 2000).

K-means. El algoritmo **k-means** es una técnica de agrupamiento no supervisada que particiona un conjunto de datos en kk grupos, minimizando la suma de las distancias cuadradas dentro de cada grupo:

$$\min_{C_1,\ldots,C_k} \sum_{j=1}^{k} \sum_{x_i \in C_j} |\, x_i - \mu_j|^2 \tag{15}$$

donde:

- μ_j es el centroide del grupo C_j,
- $\|\cdot\|$ denota la norma euclidiana.

El algoritmo alterna entre asignar cada punto al centroide más cercano y actualizar los centroides.

K-means es útil para análisis exploratorio y detección de patrones en datos médicos sin etiquetas (MacQueen, 1967).

Tabla 1

Resumen Modelos de Machine Learning Aplicados

Modelo	Tipo de Modelo	Ventajas	Desventajas	Aplicación en dosimetría
Regresión Lineal	Regresión supervisada	Simplicidad, interpretabilidad	Supone relaciones lineales, sensible a outliers y multicolinealidad	Estimaciones simples de dosis con pocas variables
Random Forest	Ensamble (árboles de decisión)	Robustez, buena precisión sin necesidad de ajustar muchos parámetros	Puede sobreajustar si no se controla la profundidad	Predicción precisa de DLP o CTDIvol usando múltiples variables
XGBoost	Ensamble (gradient boosting)	Alta precisión, maneja bien datos heterogéneos	Requiere más tiempo de entrenamiento y ajuste de hiperparámetros	Optimización de predicción de dosis efectiva con tuning de parámetros
Redes Neuronales	Red neuronal (deep learning)	Capacidad para modelar relaciones complejas no lineales	Requiere gran cantidad de datos y potencia computacional	Modelado de relaciones no lineales complejas en imágenes y parámetros
Bagging	Ensamble (bootstrap aggregation)	Reduce varianza, mejora generalización	Menor interpretabilidad, dependiente de árboles base	Mejora modelos individuales débiles como árboles simples

Nota. En esta tabla se presenta un resumen comparativo de los modelos de machine learning aplicado. Se destaca el tipo de modelo, sus ventajas, desventajas y aplicaciones en la cuantificación de la dosis de radiación para estudios de tomografía computarizada.

Seguridad del Paciente

La seguridad del paciente es un principio fundamental en la práctica médica. En el contexto de la radiación diagnóstica, esto implica garantizar que las dosis administradas sean lo suficientemente bajas para minimizar los riesgos, pero lo suficientemente altas para obtener imágenes de calidad. Estándares como el percentil 75 (tercer cuartil) se utilizan como referencia para establecer límites seguros de dosis de radiación (ICRP, 2007). Estos estándares se basan en estudios epidemiológicos y recomendaciones de organizaciones internacionales, y su cumplimiento es esencial para proteger la salud de los pacientes.

Aplicaciones Clínicas

Finalmente, es importante considerar cómo los modelos predictivos de dosis de radiación pueden integrarse en la práctica clínica. Esto implica no solo desarrollar modelos precisos, sino también garantizar que sean fáciles de usar e interpretar por parte de los profesionales de la salud (Topol, 2019). Herramientas como dashboards interactivos y sistemas de soporte de decisiones pueden facilitar la adopción de estas tecnologías en entornos clínicos reales. Además, es crucial realizar validaciones exhaustivas para asegurar que los modelos sean seguros y efectivos en diferentes contextos y poblaciones (Steyerberg et al., 2010).

Marco Normativo

El uso de radiaciones ionizantes en medicina, particularmente en procedimientos como la tomografía computarizada (TC), está regulado en Colombia con el propósito de proteger tanto a

los pacientes como al personal de salud. El marco normativo parte del principio fundamental de que toda exposición médica debe estar justificada y optimizada. En ese sentido, la Resolución 482 de 2018 del Ministerio de Salud y Protección Social establece los requisitos esenciales de protección radiológica para instalaciones médicas. Esta norma exige, entre otros aspectos, la implementación de programas de garantía de calidad, el control de dosis y la capacitación continua del personal en prácticas seguras de radiología.

Uno de los pilares de esta normativa es la aplicación del principio ALARA (As Low As Reasonably Achievable), que busca mantener las exposiciones a radiación en niveles tan bajos como sea razonablemente posible, sin comprometer la calidad del diagnóstico. La resolución también contempla la utilización de niveles de referencia diagnósticos (DRL), los cuales deben ser determinados y evaluados periódicamente por las instituciones prestadoras de servicios de salud. Sin embargo, aún persisten vacíos en la implementación práctica de estos lineamientos, especialmente en regiones donde el acceso a tecnologías avanzadas o personal especializado es limitado.

Adicionalmente, Colombia adopta las recomendaciones internacionales emitidas por organismos como la Comisión Internacional de Protección Radiológica (ICRP) y el Organismo Internacional de Energía Atómica (OIEA). Estos entes guían a los países en la adopción de políticas públicas y normativas técnicas para la seguridad radiológica. En particular, la publicación ICRP 135 introduce el concepto actualizado de DRL y enfatiza su uso en el monitoreo de las dosis en radiología diagnóstica, incluyendo tomografía computarizada. Estas recomendaciones constituyen el estándar global sobre el cual Colombia ha construido su normativa interna.

En cuanto a la protección del paciente, el marco legal colombiano también se articula con el Sistema Obligatorio de Garantía de la Calidad en Salud (SOGCS), el cual incluye componentes como la auditoría médica, la gestión de la tecnología biomédica y la evaluación de resultados clínicos. Esto implica que los procedimientos diagnósticos, como las TC, deben ser evaluados no solo en términos de efectividad clínica, sino también bajo estándares técnicos y de seguridad. En este escenario, el uso de herramientas como la inteligencia artificial para predecir dosis siempre que se alineen con la normativa vigente puede fortalecer los sistemas de calidad y mejorar la toma de decisiones clínicas.

En resumen, el marco normativo colombiano respalda firmemente la protección radiológica en procedimientos diagnósticos, pero su implementación efectiva requiere no solo del cumplimiento formal de la norma, sino de una apropiación tecnológica e institucional. La integración de modelos predictivos basados en machine learning, como se exploró en este trabajo, representa una oportunidad para avanzar hacia una gestión más precisa, personalizada y segura de la radiación en imágenes diagnósticas, siempre que se respeten las directrices éticas, legales y regulatorias establecidas.

Metodología

Este estudio se desarrolló en varias etapas secuenciales que permitieron la aplicación y evaluación de técnicas de *machine learning* para la predicción de dosis de radiación en estudios de tomografía computarizada (TC).

Primero, se recolectó un conjunto de datos amplio y diverso que incluía variables relacionadas con el protocolo de adquisición, parámetros técnicos de los equipos, características del paciente y dosis reportadas. Esta base de datos fue sometida a un riguroso proceso de limpieza, que incluyó la eliminación de registros incompletos, duplicados o erróneos, así como la identificación y corrección de valores atípicos.

Posteriormente, se aplicaron técnicas de normalización y transformación de variables con el fin de mejorar la calidad y homogeneidad de los datos, facilitando así el proceso de entrenamiento de los modelos. A continuación, se realizó un análisis exploratorio de datos (EDA) que permitió identificar patrones y relaciones entre variables, así como detectar las variables más relevantes para la predicción.

Con base en los hallazgos del EDA, se seleccionaron y entrenaron diversos algoritmos de *machine learning*, incluyendo modelos de regresión lineal, XGBoost, random forest, bagging y redes neuronales. Estos modelos fueron evaluados utilizando métricas estándar como el error cuadrático medio (RMSE) y el coeficiente de determinación (R^2). Además, se aplicó un modelo de clasificación con base en los niveles de referencia diagnósticos (percentil 75), categorizando los estudios como "seguros" o "excesivos" en términos de dosis administrada. Esta clasificación permitió evaluar la capacidad del modelo para discriminar entre exposiciones dentro de los límites aceptables y aquellas que podrían representar un riesgo clínico, utilizando métricas como la precisión, sensibilidad y especificidad para valorar su desempeño.

Finalmente, se analizaron los resultados obtenidos para identificar las limitaciones de los modelos desarrollados, así como las posibles causas del bajo rendimiento predictivo. Este análisis permitió establecer recomendaciones para investigaciones futuras, enfatizando la necesidad de bases de datos más robustas, técnicas de modelado más avanzadas y una mayor integración con criterios clínicos.

Figura 1

Representación de la Metodología Mediante Flujograma

Nota. Representación de la metodología a partir de un flujograma en el cual se mencionan de manera secuencial cada una de las etapas desarrolladas desde la recolección de los datos hasta el análisis de resultados y recomendaciones.

Resultados

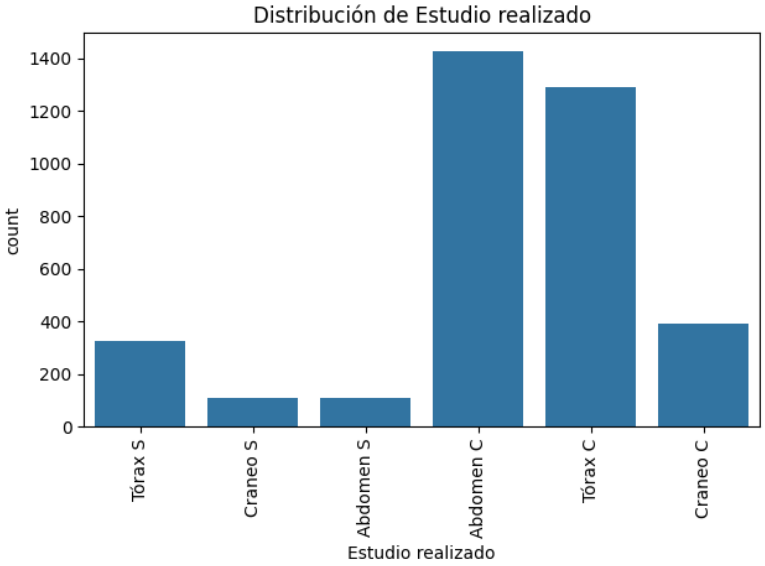
Descripción y Análisis de Relaciones Entre Variables

Inicialmente se realizó una revisión general de los datos disponibles registrados entre mayo de 2023 y febrero de 2025 con el fin de organizarlos, filtrarlos y mejorar sus características con el propósito de eliminar la mayor parte de inconsistencias que puedan comprometer el desempeño de los modelos predictivos a evaluar.

La base de datos con la que se dispone está compuesta por las siguientes variables: Fecha, Nombre del paciente, Documento de identificación, Edad (años), Peso (kg), DLP (mGy.cm), Estudio realizado (Abdomen C, Abdomen S, Tórax C, Tórax S, Cráneo C y Cráneo S) y responsable de realizar el estudio. Los datos personales del paciente junto con el responsable fueron descartados ya que no aportan información relevante al estudio. Después de realizar la respectiva limpieza de datos se obtuvieron las siguientes gráficas y frecuencias:

Figura 2

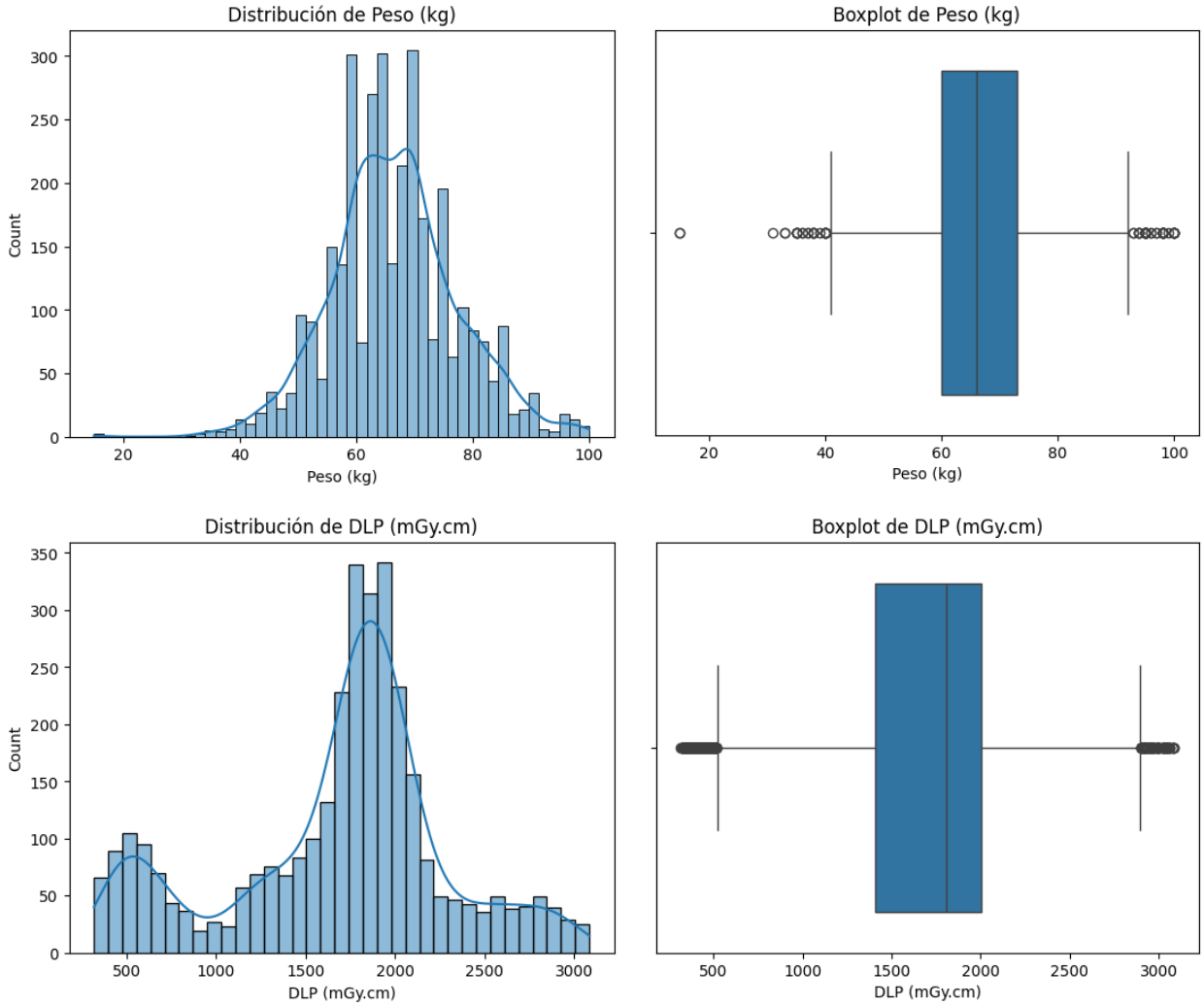
Distribución de Estudio Realizado



Nota. Distribución de la variable estudio realizado después de realizar la respectiva limpieza de los datos. Se puede apreciar la importancia que tiene la cantidad de estudios realizados con medio de contraste intravenoso en comparación con los estudios simples para los mismos grupos anatómicos.

Figura 3

Distribución del Peso (kg) y el DLP (mGy.cm)



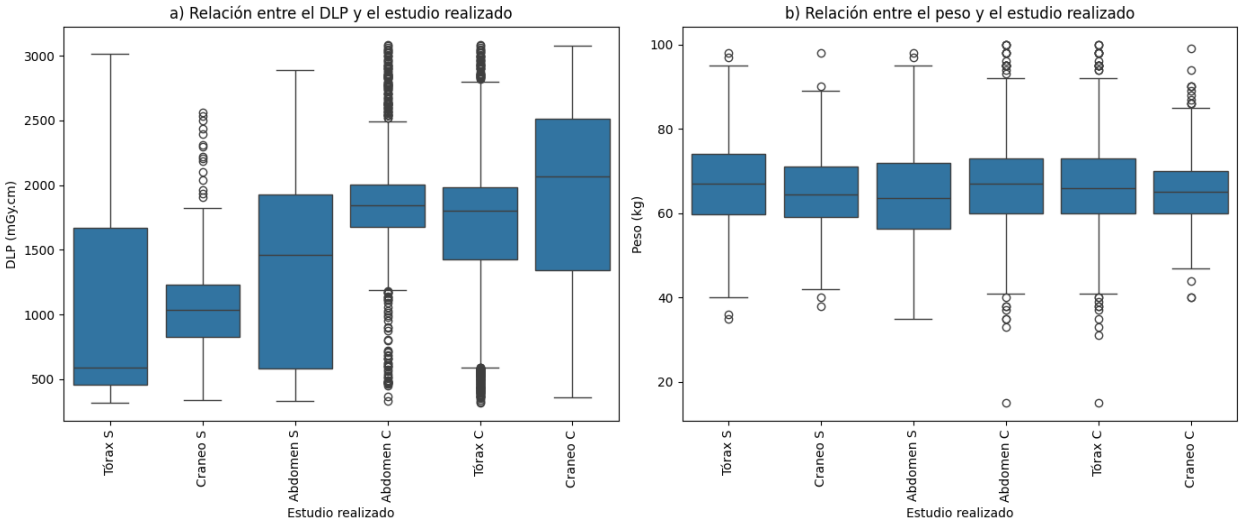
Nota. Histogramas de distribución del peso (kg) y DLP (mGy.cm) junto con sus representaciones mediante gráficos tipo boxplot donde se observa un valor para la mediana de la variable peso alrededor de los 65 kg, indicando la prevalencia de pacientes sometidos a este tipo de estudios en edad adulta.

En la **Figura 3** se observa la prevalencia en el número de estudios de abdomen y tórax que se practicaron durante el periodo de tiempo registrado que obedece a un efecto relacionado con la incidencia de patologías malignas en estas regiones anatómicas.

Para el DLP se aprecia una distribución sesgada hacia la derecha con cola larga hacia valores altos y una concentración de datos en el rango inferior por debajo de 1000 mGy.cm.

Figura 4

Gráficos Tipo Boxplot que Relacionan: a) el DLP con el Estudio Realizado y b) el Peso con el Estudio Realizado



Nota. Relaciones tanto del DLP como del peso con el estudio realizado, para el caso del DLP se observan distribuciones concentradas por encima de la mediana que podrían indicar niveles de exposición o dosis mayores a los niveles de referencia, sin embargo, las dosis de los estudios con contraste tienden hacia valores más altos de DLP que los estudios simples, así como a distribuciones más simétricas. El peso nuevamente indica la prevalencia de pacientes en edad adulta.

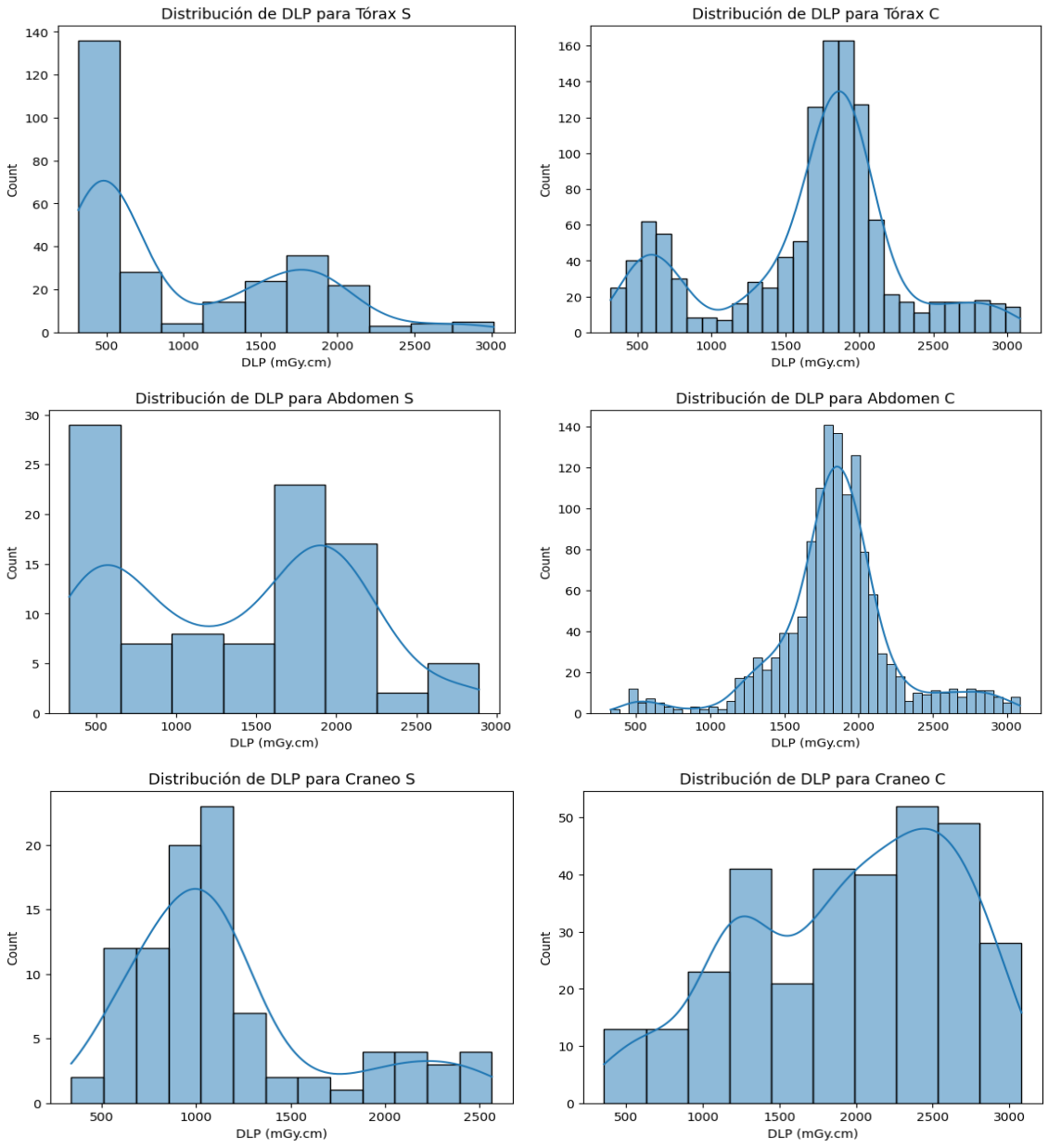
Mientras el DLP presenta alta variabilidad en los datos cuando se analizan estadísticos como la moda y los percentiles, el peso en función del estudio realizado expone un comportamiento más uniforme entre los diferentes grupos, lo que evidencia una distribución aceptable de los datos ya que no se presentan outliers ni valores atípicos importantes dentro de la distribución.

Gráficas Tipo Histograma por Tipo de Estudio

El conjunto de gráficas de la **Figura 4** muestra el comportamiento variable que tiene el parámetro DLP dependiendo del tipo de región anatómica en donde se hizo la exploración y adicionalmente, los estudios con contraste tienen distribuciones más normales con valores más altos de DLP. En el caso de los estudios sin contraste se presentan distribuciones más asimétricas y con valores más bajos de DLP.

Figura 5

Histogramas DLP para cada Categoría Dentro de la Variable Estudio Realizado

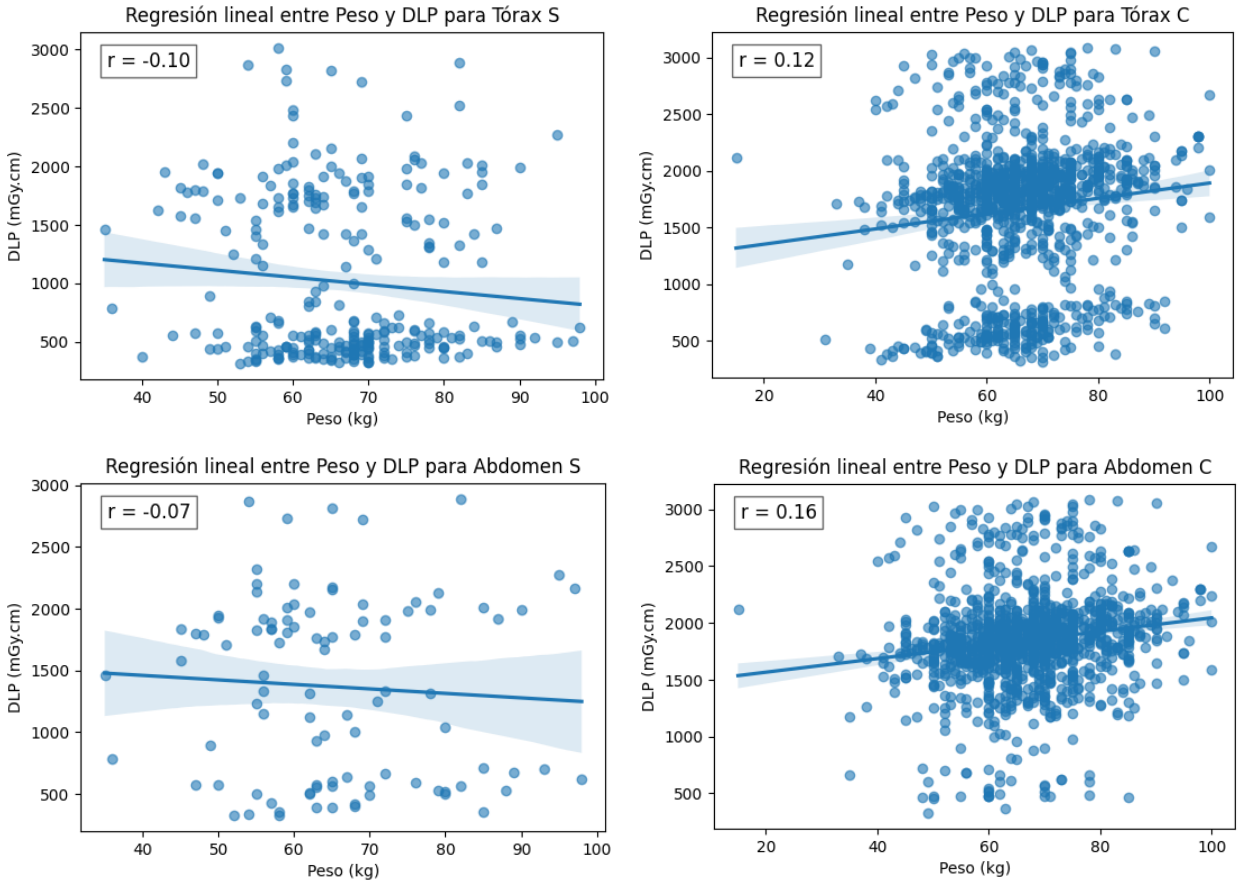


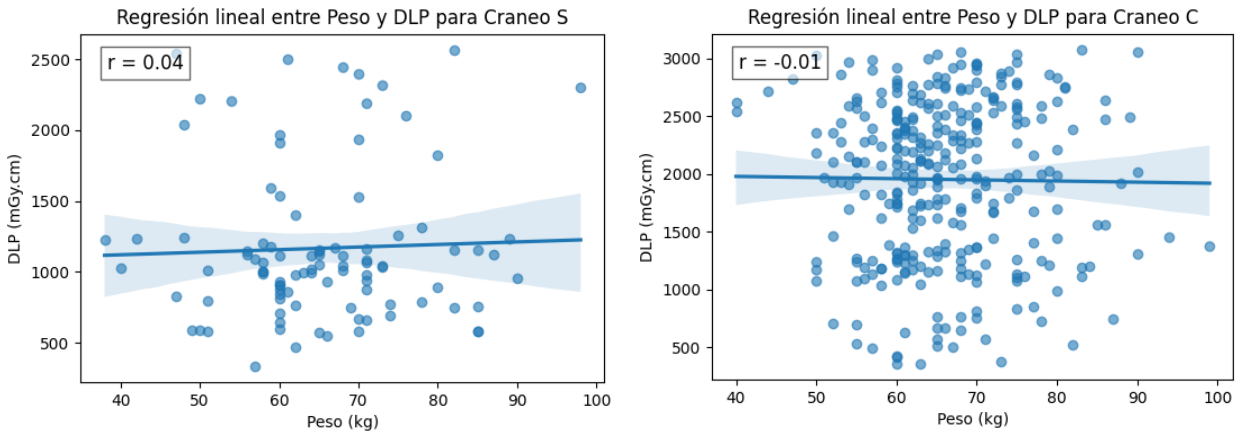
Nota. En las distribuciones de DLP por estudio realizado se observan relaciones o comportamientos similares de las distribuciones entre los estudios de tórax y abdomen para cada caso, simple y con contraste. Para cráneo no es posible evidenciar similitudes entre los datos.

Con el fin de evaluar si existe algún tipo de relación lineal entre los datos, se segmentó cada categoría perteneciente a la variable estudio realizado y se visualizó mediante gráficos de dispersión tomando como variable independiente el Peso (kg). El parámetro estadístico usado en este análisis fue el coeficiente de Pearson ya que permite cuantificar la relación lineal entre las variables en caso de existir.

Figura 6

Representación Tipo Dispersión del DLP en Función del Peso y Cálculo del Coeficiente de Correlación de Pearson





Nota. Representa la relación entre el DLP y el peso por estudio realizado con el valor del coeficiente de correlación para cada caso, sin embargo, no es posible observar una distribución que pueda ser representada por medio de una aproximación lineal.

De acuerdo con las distribuciones presentadas en la **Figura 6** se puede evidenciar la baja correlación ($r < 0.3$) que existe entre el DLP y el peso del paciente, por lo que no se puede concluir que los datos se puedan explicar con base en una representación de tipo lineal.

Identificar Rangos de Peso Asociados con Mayores Valores de DLP

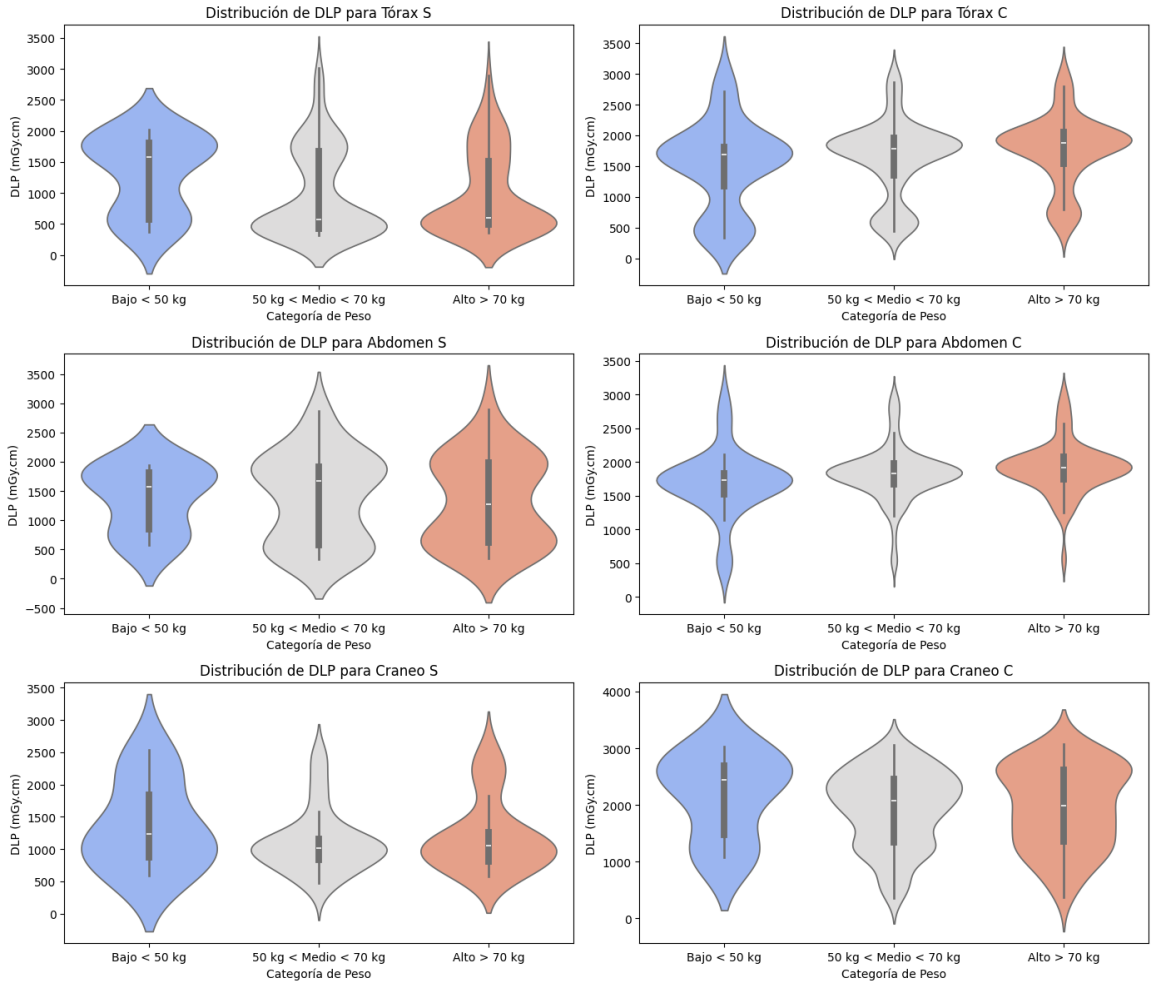
Mediante los gráficos de tipo violín presentados en la **Figura 7** se visualizan las variables DLP, tipo de estudio y peso agrupado en categorías: Bajo < 50 kg, 50 kg < Medio < 70 kg y Alto > 70 kg con el fin de evaluar el comportamiento de la densidad de frecuencia en cada uno de los grupos mencionados. Inicialmente se observa una diferencia considerable entre las dosis reportadas en los estudios simples y aquellos contrastados por región anatómica, siendo los contrastados los que coinciden con niveles de dosis más altos que se explican por la atenuación que produce el medio de contraste al momento de interactuar con el haz de rayos X.

Peso < 50 kg: La densidad de los datos tiene comportamiento similar tanto en tórax como en abdomen con una parte más angosta en el extremo superior que sugiere que son pocos

los pacientes de bajo peso con dosis elevadas de radiación, se aprecian diferencias en cráneo ya que los datos están mayormente concentrados hacia valores más bajos de DLP. Cuando se aplica medio de contraste hay mayor variabilidad y concentración de los datos hacia valores más altos de DLP.

Figura 7

Gráfico Tipo Violín que Representa el DLP en Función del Peso Distribuido en 3 Rangos para Facilitar su Análisis



Nota. Gráficos de tipo violín donde se observan las distribuciones y las densidades de los datos para cada rango de peso, estudio y los valores de mayor y menor dosis correspondientes.

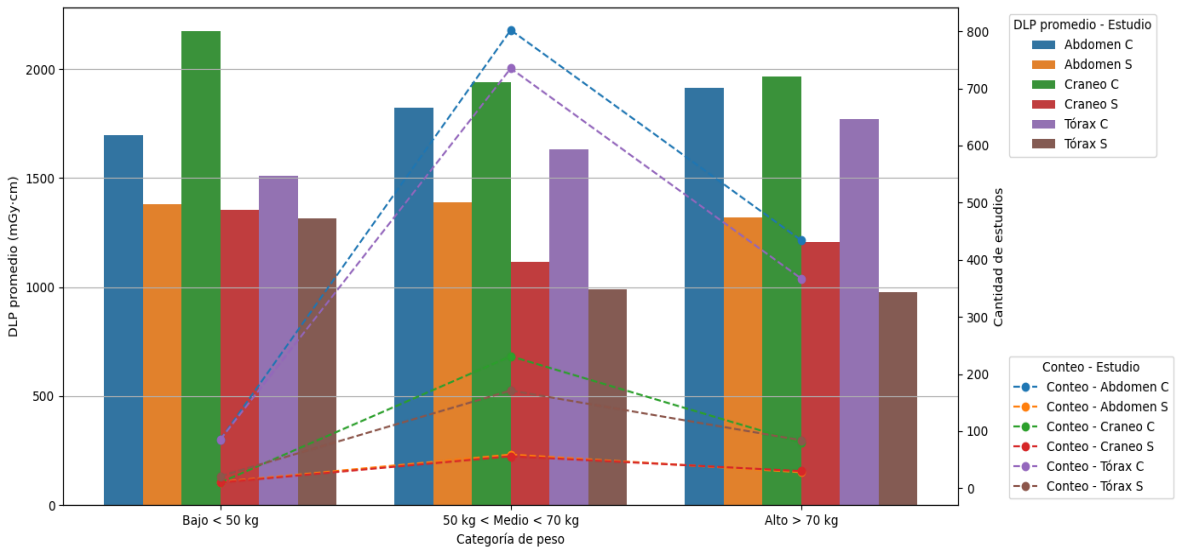
50 kg < Peso < 70 kg: Predominan valores más altos cuando se emplea medio de contraste mostrando una distribución asimétrica de los datos, la densidad en valores bajos de DLP es mayor en los estudios sin medio de contraste. Hay mayor concentración de datos en los valores bajos de DLP en este grupo que en el grupo con menor peso.

Peso > 70 kg: En el caso de cráneo con contraste hay una alta dispersión de los datos ya que se concentran en un rango mucho más amplio de DLP, sin embargo, predomina dicha concentración hacia los valores bajos cuando el estudio es simple. En los grupos tórax y abdomen el comportamiento es similar al grupo clasificado dentro de la categoría peso medio.

Se observa una mayor concentración de valores en la parte superior del rango DLP, esto indica que los pacientes de mayor peso suelen recibir una dosis más alta de radiación debido a la mayor atenuación de los rayos X y a la necesidad de ajustes técnicos en la exposición.

Figura 8.

DLP Promedio Por Rango de Peso y Tipo de Estudio. Se Presentan las Frecuencias para cada Estudio por Rango de Peso.



Nota. Relación entre el DLP y el peso por estudio realizado donde se destaca el aumento de la dosis a medida que aumenta el peso en los estudios con contraste de tórax y abdomen. Los

estudios simples no presentan tendencia, con excepción de tórax simple que presenta una relación inversamente proporcional con la dosis.

En la gráfica de la **Figura 8** se presentan histogramas de frecuencia donde se aprecian tanto los valores de DLP discriminados por tipo de estudio como las frecuencias de cada uno, particularmente se observan dosis mucho más altas en los estudios de cráneo contrastados con respecto a los demás grupos, entre los 50 y los 70 kg que es el grupo donde se ubican la mayor cantidad de pacientes, la principal diferencia se observa en los estudios de cráneo simple.

Se asume el P75 como el nivel de referencia con base en la normativa aplicada en este ámbito relacionado con las dosis recibidas por los pacientes sometidos a estudios de imágenes diagnósticas, sin embargo, dosis por encima de P90 son indeseables ya que aumentan los riesgos derivados de la radiación ionizante en el paciente.

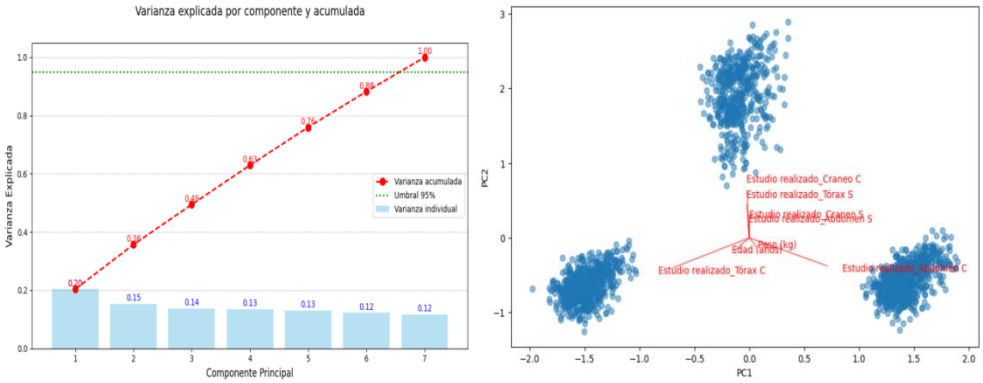
En términos generales, hay una tendencia en el aumento de la dosis (DLP) a medida que aumenta el peso (kg) del paciente.

Aplicación de Modelos de Machine Learning

Habiendo evidenciado anteriormente la dependencia que existe entre la dosis – DLP y el peso del paciente, se hicieron algunas transformaciones de los datos antes de aplicar los modelos: codificación de variables categóricas, análisis de componentes principales (PSA), selección de variables relevantes a partir del análisis de los loadings y validación. Por medio de la **Figura 9** se identifican qué variables originales contribuyen más a cada componente.

Figura 9

Varianza Explicada por Componente y Acumulada



Nota. Varianza explicada y acumulada para los componentes principales que explican más del 95% de la varianza. Se puede apreciar en el gráfico de la derecha la distribución ortogonal de las variables de acuerdo con dichos componentes.

Se aplica el PCA reteniendo el 95% de la Varianza (n_components = 7)

Tabla 2

Loadings de las Variables en los Componmentes Principales

	PC0	PC1	PC2	PC3	PC4	PC5	PC6
Peso (kg)	0.070092	-0.010825	-0.321920	0.527354	0.274655	0.296392	-0.671037
DLP (mGy.cm)	0.010107	0.056203	-0.262146	0.608010	0.157999	0.274600	0.676987
Estudio realizado_Abdomen C	0.913627	-0.397841	-0.030208	0.003092	-0.047742	-0.054205	0.034727
Estudio realizado_Abdomen S	-0.018146	0.206241	0.158911	-0.291885	0.917552	-0.003658	0.071480
Estudio realizado_Craneo C	-0.035955	0.691600	-0.649561	-0.181246	-0.166529	-0.195282	-0.009380
Estudio realizado_Craneo S	-0.018379	0.198402	0.172215	-0.310302	-0.203651	0.890627	-0.001369
Estudio	-0.876874	-0.471279	-0.065829	0.054604	-0.028649	-0.033660	-0.007606

realizado_Tórax C Estudio realizado_Tóraz S	-0.033141	0.522832	0.670391	0.469776	-0.137831	-0.137180	-0.082368

Nota. Loadings de cada una de las variables indicando su participación por componente con base en los valores de los coeficientes.

Tabla 3

Varianza Explicada y Acumulada

	Componente principal	Varianza explicada	Varianza acumulada
0	PC0	0.201394	0.201394
1	PC1	0.152101	0.353495
2	PC2	0.137936	0.491431
3	PC3	0.135678	0.627110
4	PC4	0.129157	0.756266
5	PC5	0.1285542	0.884808
6	PC6	0.115192	1.000000

Nota. Resume la varianza explicada y acumulada de cada componente principal indicando que el 95% de la varianza queda explicado usando 7 componentes.

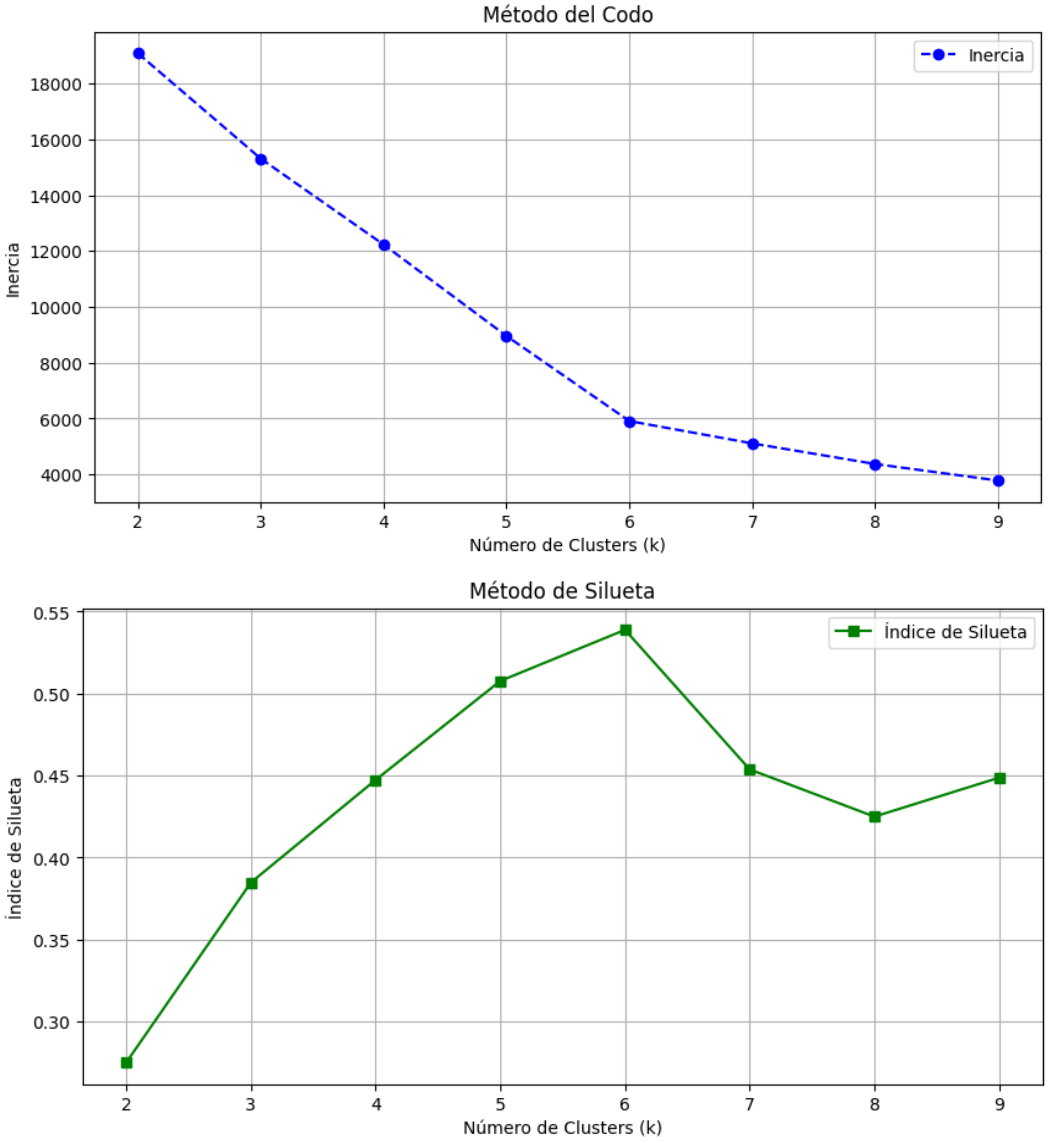
En cada componente principal (PC) se relaciona el peso o importancia que tiene cada una de las variables por lo que se puede determinar con cuantas componentes principales como mínimo se puede representar el conjunto de datos sin que haya pérdida de información relevante, ayudando a los modelos para que usen menor cantidad de información y puedan ser más eficientes sin sacrificar precisión. Debido a que la edad (años) no es un parámetro que guarde

relación con las demás variables se puede suprimir del análisis quedando de esta manera 6 componentes principales.

Método de Clúster Usando k means

Figura 10

Métodos de Codo y Silhuette para Determinar la Cantidad de Clústers que se Deben Seleccionar



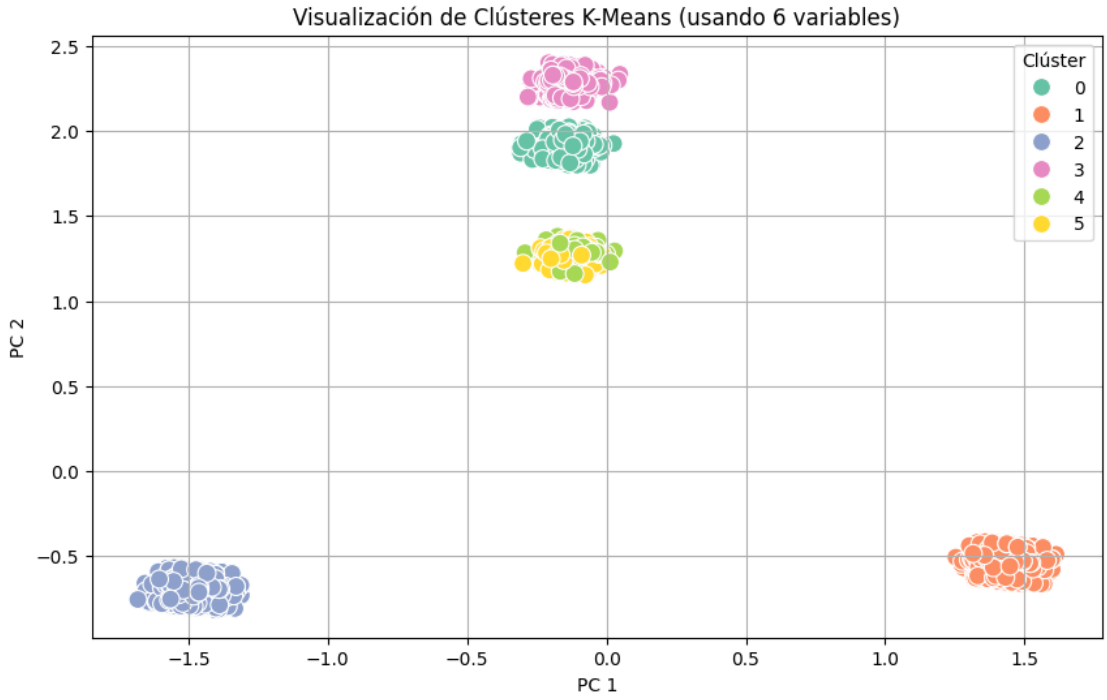
Nota. Métodos de codo y silueta para determinar la cantidad óptima de clústers para el agrupamiento de las variables mediante el método k – means.

Para verificar si es posible algún tipo de agrupación entre las componentes principales se implementó el método k means partiendo de las métricas codo y silhouette para determinar la cantidad de clústers más indicada.

Ambos métodos coinciden en que la cantidad más adecuada de clústers que se deben seleccionar debe ser k = 6.

Figura 11

Representación Gráfica de Clústers



Nota. Con base en los gráficos de codo y silueta se determinaron 6 clústers para la creación de los grupos. En este gráfico se observan los clústers generados por k – means para las primeras dos componentes principales.

Se presentan los siguientes grupos con base en los clúster generados y en las medias de los componentes principales (loadings) por clúster.

Clúster 0: Presencia muy marcada de los Componentes PC1 - PC2 - PC3 por lo que estos componentes contribuyen significativamente a la diferenciación de este grupo: Estudio_realizado Cráneo C y Tórax S. En este grupo cobra especial importancia PC3 ya que tiene buena correlación con DLP y el Peso lo cual puede usarse como predictor. Adicionalmente los pacientes que hacen parte de este grupo tienden a tener altos valores de DLP y Peso.

Clúster 1 y Clúster 2: Tanto el clúster 1 como el clúster 2 están relacionados con la componente PC0, por lo que tanto Estudio realizado_Abdomen C como Estudio realizado_Tórax C pertenecen a este grupo.

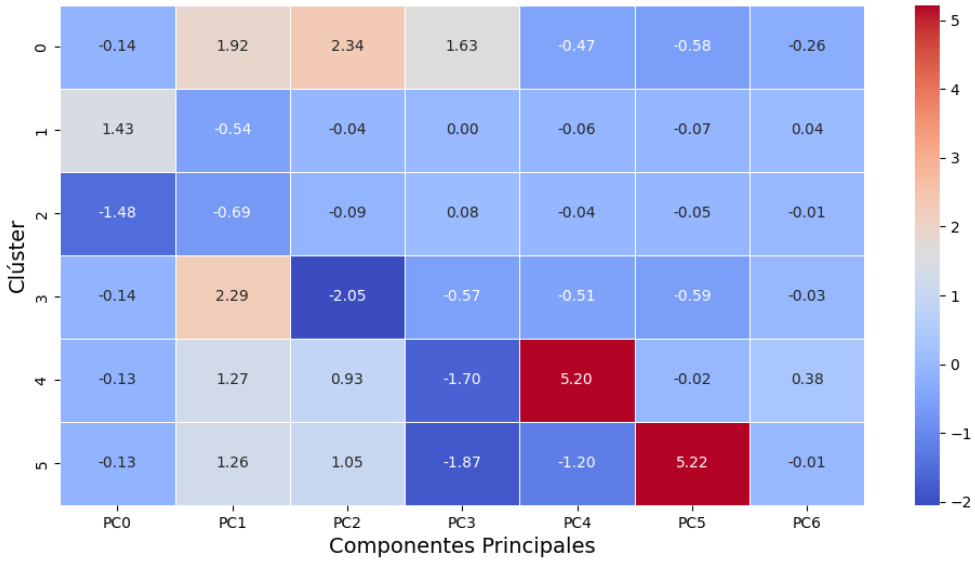
Clúster 3: Formado por PC1 y PC2 y relacionado con las variables Estudio realizado_Cráneo C y Tórax S.

Clúster 4: Formado por PC4 y con alta presencia de Estudio realizado_Abdomen S.

Clúster 5: Formado por PC5 y con alta presencia de Estudio realizado_Cráneo S.

Figura 12

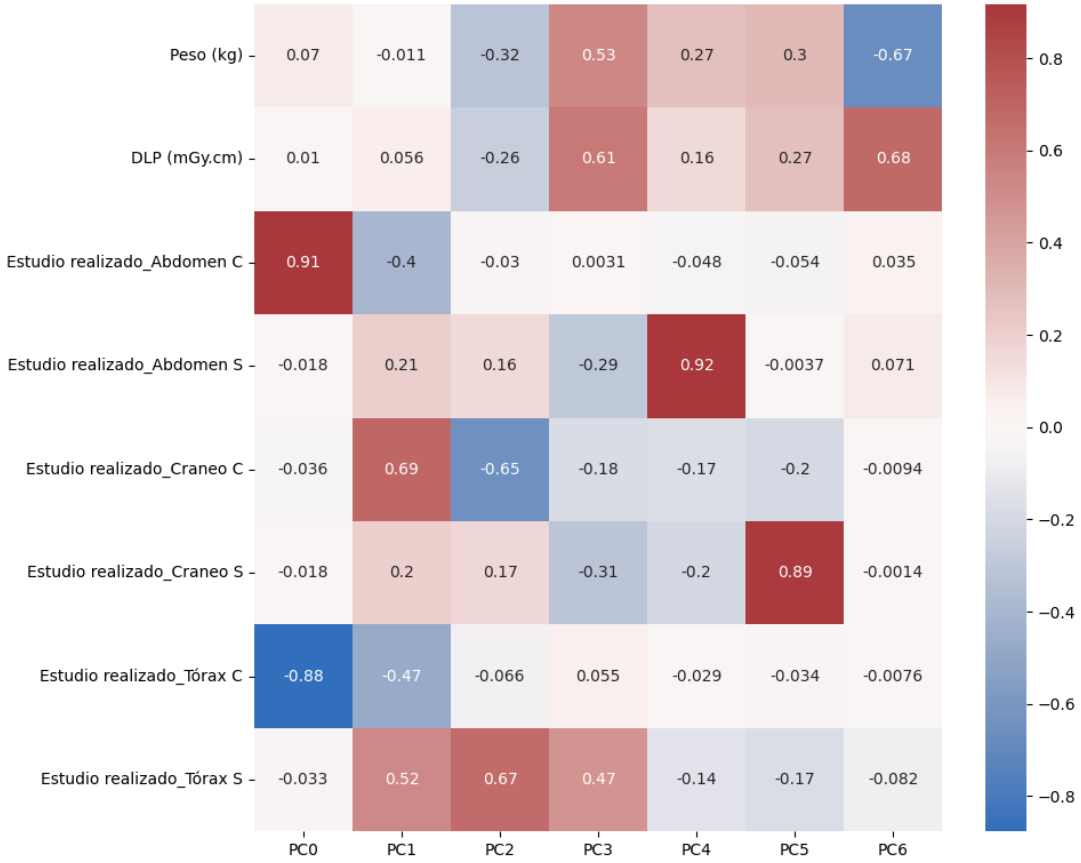
Correlación Entre Variables y Componentes Principales (Loadings)



Nota. Correlación entre los clústers generados por el método k – means y las componentes principales donde se crean relaciones de agrupación en función de los loadings calculados.

Figura 13

Medias de Componentes Principales (PCA) por Clúster



Nota. Representación mediante un mapa de calor de las medias de los componentes principales donde se indica en función del coeficiente el peso que tiene cada variable dentro de cada una de las componentes.

Los modelos seleccionados y aplicados fueron aquellos que poseen mayor robustes al momento de trabajar con datos que no exhiben un comportamiento que pueda ser aproximado de manera lineal, son los casos de bosques aleatorios, XGBoost, redes neuronales con tensor Flow y Bagging. A partir de los resultados obtenidos mediante el método de clustering con k means se seleccionaron solamente las primeras 5 componentes principales ya que poseen valores medios significativos dentro de los clústeres.

Métricas de los Modelos Implementados

Los datos fueron divididos en sets de entrenamiento y prueba con una proporción de 70% - 30%, respectivamente. Las métricas obtenidas para los cinco modelos se presentan en la **Tabla 4**:

Tabla 4

Modelos Aplicados con sus Respectivas Métricas para los Conjuntos de Entrenamiento y Prueba

Nombre del modelo	Métricas de entrenamiento	Métricas de prueba
Regresión lineal	MAE: 474.28	MAE: 478.47
	RMSE: 620.58	RMSE: 621.95
	R2: 0.002	R2: 0
Random forest	MAE: 439.76	MAE: 509.75
	RMSE: 574.76	RMSE: 649.41
	R2: 0.14	R2: - 0.09
XGBoost	MAE: 442.16	MAE: 495.98
	RMSE: 580.33	RMSE: 635.97
	R2: 0.13	R2: - 0.05
Redes neuronales	MAE: 472.24	MAE: 478.81
	RMSE: 618.97	RMSE: 623.17
	R2: 0.007	R2: - 0.004
Bagging	MAE: 439.76	MAE: 509.48
	RMSE: 574.74	RMSE: 649.23
	R2: 0.14	R2: - 0.09

Nota. Resumen comparativo de los modelos aplicados junto con las métricas más relevantes para los conjuntos de entrenamiento y prueba. No existe una diferenciación clara entre los resultados obtenidos con cada modelo, sin embargo, se observa un posible sobre ajuste puesto que los resultados son ligeramente mejores en los datos de entrenamiento que en los de prueba.

Después de revisar cada una de las métricas obtenidas con los modelos entrenados, se evidencia la baja correlación entre los datos de entrenamiento y validación, ya que no se puede explicar la variabilidad de los datos usando como única variable predictora el peso del paciente. Adicionalmente, las métricas obtenidas presentan valores tan bajos que impiden una evaluación robusta en términos comparativos. No obstante, se podría destacar que los modelos Random Forest, XGBoost y Bagging mostraron resultados similares en los datos de entrenamiento, con valores de R^2 superiores a los obtenidos en validación, lo que sugiere un posible sobreajuste. Este comportamiento también se refleja en los valores de MAE y RMSE, que fueron significativamente mejores en entrenamiento que en validación.

Estos hallazgos se alinean con lo reportado en la literatura reciente. Por ejemplo, Ichikawa et al. (2024) demostraron que el peso corporal por sí solo tiene una correlación limitada con parámetros de dosis como el CTDIvol y DLP, especialmente en exámenes abdominales y torácicos, donde la geometría corporal y la atenuación del haz no se representan adecuadamente solo con el peso. De forma similar, Ferrante et al. (2025) evidenciaron que los modelos multivariantes basados en machine learning —alimentados con información del paciente (edad, sexo, altura, peso), parámetros del escáner y dosimetría básica— lograron un desempeño mucho más alto ($R^2 > 0.95$) comparado con modelos que usan únicamente el DLP y factores de corrección (k-factors). Asimismo, García-Sánchez et al. (2019) y Zhang et al. (2019) resaltan la importancia de usar características geométricas derivadas directamente de las imágenes para mejorar la estimación de dosis, mostrando que los modelos univariantes o lineales tienden a subestimar la complejidad de la interacción entre el cuerpo del paciente y la radiación.

Por lo tanto, los resultados obtenidos en este análisis refuerzan lo señalado en los estudios revisados: el peso corporal es una variable informativa, pero insuficiente como predictor único

de la dosis en tomografía computarizada. La inclusión de variables adicionales, tanto anatómicas como técnicas, es esencial para capturar con precisión la variabilidad dosimétrica entre pacientes. Ignorar esta multidimensionalidad no solo compromete la precisión del modelo, sino que también puede poner en riesgo la calidad del diagnóstico y la seguridad radiológica del paciente.

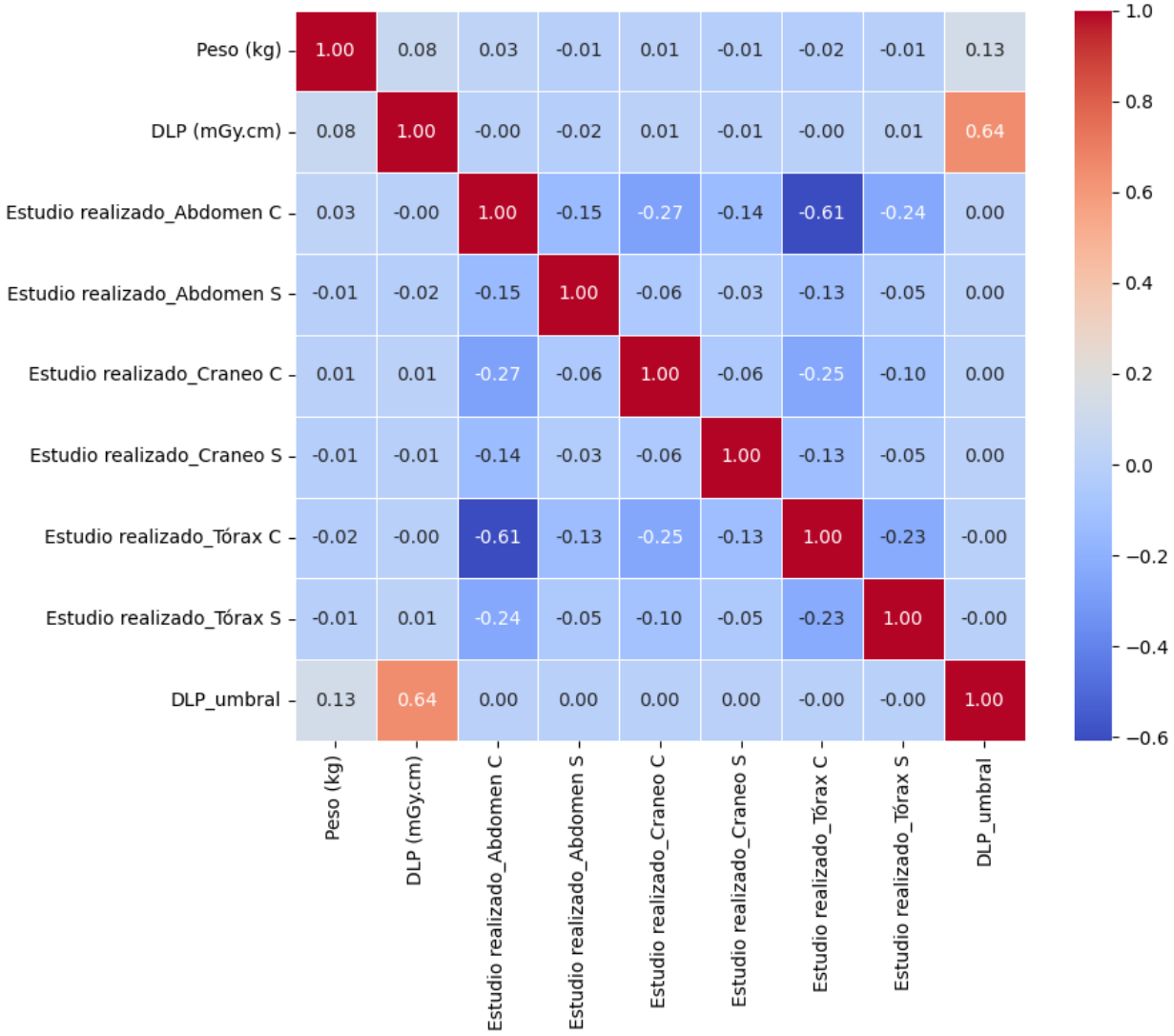
Esto podría deberse a la ausencia de datos adicionales relacionados principalmente con parámetros técnicos del escáner como el kVp (kilo voltaje pico), el mAs (miliamperio segundo) y el pitch de la mesa de exploración. Además, otros factores como el protocolo de adquisición, el tipo de reconstrucción utilizada, la anatomía del paciente, la presencia de artefactos en la imagen y las características específicas del escáner (marca, modelo y algoritmos de optimización de dosis) también pueden influir significativamente en el DLP. Por lo tanto, si se busca mejorar la capacidad de los modelos predictivos para estimar el DLP con mayor precisión, es crucial incorporar una mayor cantidad de variables que representen tanto las características del paciente como los parámetros técnicos del procedimiento.

Abordaje Como un Problema de Clasificación

Debido a la dificultad para ajustar los modelos de machine learning de manera que pudieran predecir los valores de la dosis DLP a partir del peso del paciente, se decidió abordar el problema desde la clasificación creando una variable dicotómica dentro del conjunto de datos de manera tal que tomara el valor de uno en caso de que el DLP superara el tercer cuartil (Q3) y cero en caso de que dicho valor fuera inferior a Q3.

Figura 14

Matriz de Correlación



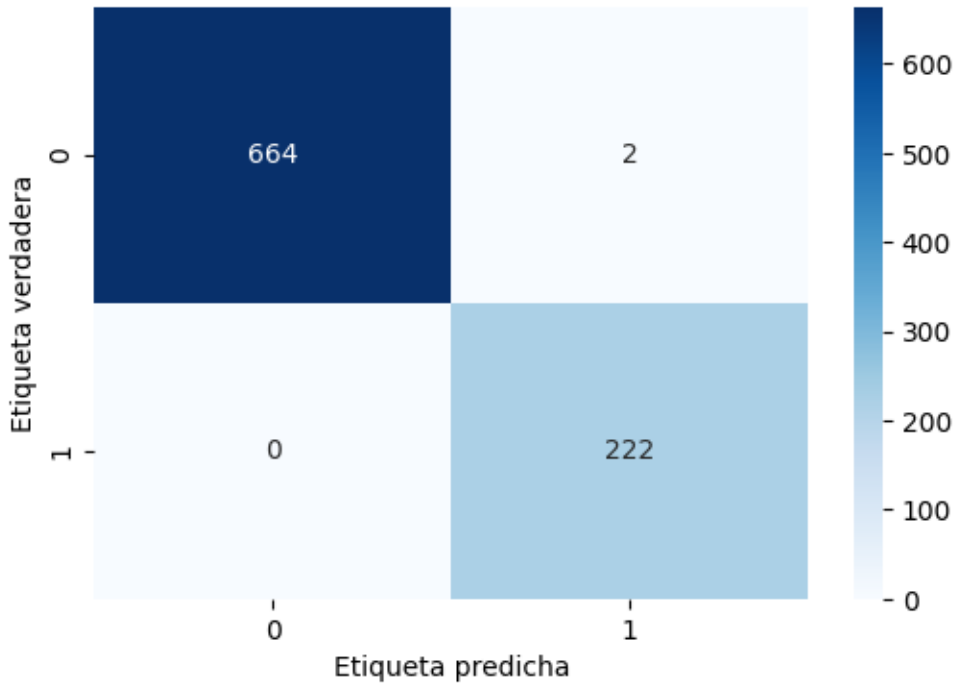
Nota. Habiendo transformado el problema de regresión a clasificación, la matriz de correlación evalúa las posibles variables multi colineales que deben ser descartadas del modelo para evitar posible sobre ajuste.

En la **Figura 14** se presenta la matriz de correlación que muestra precisamente la relación existente entre las diferentes variables que conforman el dataset teniendo en cuenta la nueva variable dicotómica “DLP_umbral”.

Se destaca el valor de 0.64 para el DLP (mGy.cm) y el DLP_umbral puesto que la segunda variable fue creada a partir de un condicional aplicado al DLP (mGy.cm), sin embargo, a pesar de que dicha correlación es medianamente fuerte con un valor del 64%, se decidió usar como parte de las variables predictoras para el entrenamiento del modelo de clasificación logrando las siguientes métricas:

Figura 15

Matriz de Confusión



Nota. Matriz de confusión donde se evalúa la capacidad del modelo para predecir a partir del conjunto de datos de validación la proporción entre los verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos.

De acuerdo con la matriz de confusión se puede observar que el modelo Random Forest captura muy bien la distribución de los valores de las variables presentes en el dataset, lo anterior se puede evidenciar de acuerdo con el reporte de clasificación para los conjuntos de entrenamiento y prueba:

Tabla 5

Métricas para el Conjunto de Entrenamiento del Modelo de Clasificación

Datos: Entrenamiento	Precision	Recall	F1 – Score	Support
0	1.00	1.00	1.00	1553
1	1.00	1.00	1.00	516
Accuracy			1.00	2069
Macro avg	1.00	1.00	1.00	2069
Weighted avg	1.00	1.00	1.00	2069

Nota. Métricas para evaluar la efectividad del modelo para realizar predicciones sobre el conjunto de datos de entrenamiento, el accuracy relaciona el total de predicciones correctas con el número total de predicciones realizadas dando como resultado la visualización del desempeño del algoritmo.

Tabla 6

Métricas para el Conjunto de Prueba del Modelo de Clasificación

Datos: Prueba	Precision	Recall	F1 – Score	Support
0	1.00	1.00	1.00	666
1	0.99	1.00	1.00	222
Accuracy			1.00	888
Macro avg	1.00	1.00	1.00	888
Weighted avg	1.00	1.00	1.00	888

Nota. Métricas para evaluar la efectividad del modelo para realizar predicciones sobre el conjunto de datos de validación o prueba, el accuracy relaciona el total de predicciones correctas con el número total de predicciones realizadas dando como resultado la visualización del desempeño del algoritmo.

Para verificar el desempeño del modelo Random Forest se hizo una validación cruzada con cv = 5 y usando la métrica f1 – score, obteniendo como resultado un F1 promedio = 0.994. Lo anterior evidencia la robustes del modelo para clasificar los valores de acuerdo con el umbral del tercer cuartil definido anteriormente.

Conclusiones

Los resultados del estudio muestran que el peso corporal por sí solo no es un predictor confiable de la dosis absorbida en estudios de tomografía, ya que modelos univariados basados exclusivamente en esta variable presentan baja capacidad explicativa y predictiva, en concordancia con lo reportado por Ichikawa et al. (2024). Asimismo, se evidenció que los estudios que utilizan medio de contraste intravenoso tienden a registrar dosis más elevadas debido al aumento en la atenuación del haz de rayos X (García-Sánchez et al., 2019). En este contexto, se destaca la importancia de emplear enfoques multivariantes que incorporen tanto parámetros del paciente como del escáner, lo cual mejora significativamente la precisión en la estimación dosimétrica (Ferrante et al., 2025).

Además, la inclusión de variables anatómicas derivadas de imágenes junto con la aplicación de algoritmos de *machine learning*, permiten desarrollar modelos más robustos que no solo ofrecen mejores estimaciones continuas de dosis, sino también clasificaciones efectivas entre niveles seguros y excesivos de exposición. En particular, el modelo de clasificación basado en Random Forest demostró un desempeño adecuado al categorizar los valores de DLP según el umbral del percentil 75, resultado alineado con los enfoques propuestos por Ferrante et al. (2025) y García-Sánchez et al. (2019). Estos hallazgos respaldan el uso de la inteligencia artificial como una herramienta clave para la dosimetría personalizada, permitiendo avanzar hacia una gestión más segura y precisa de la radiación en entornos clínicos.

Recomendaciones

Limitaciones

Heterogeneidad en los protocolos de imagen. Las diferencias en parámetros técnicos y protocolos (uso o no de contraste, regiones anatómicas) pueden introducir variabilidad significativa en las dosis, lo que dificulta la generalización del modelo (García-Sánchez et al., 2019).

Calidad y completitud de los datos. La calidad del conjunto de datos empleados, así como la presencia de registros incompletos o inconsistentes, puede haber afectado el desempeño de los modelos entrenados.

Modelos con desempeño moderado. A pesar de emplear técnicas de *machine learning*, los modelos desarrollados no alcanzaron métricas de predicción óptimas, lo cual limita su aplicabilidad clínica inmediata.

Oportunidades Futuras y de Mejora

Incorporación de variables anatómicas derivadas de imagen. La integración de medidas automáticas como el diámetro efectivo o WED, extraídas de scout o imágenes axiales, puede mejorar notablemente la capacidad predictiva (Ichikawa et al., 2024)

Uso de técnicas avanzadas de *machine learning.* Modelos como *gradient boosting* o redes neuronales profundas podrían ofrecer mejoras si se entrenan adecuadamente con conjuntos de datos más amplios y balanceados (Ferrante et al., 2025).

Clasificación por niveles de referencia y riesgo. Continuar desarrollando y refinando modelos clasificatorios que discriminen entre niveles “seguros” y “excesivos” de dosis puede facilitar la vigilancia automática y la auditoría de calidad en servicios de radiología.

Referencias Bibliográficas

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5-32.

Brenner, D. J., & Hall, E. J. (2007). Computed tomography—an increasing source of radiation exposure. *New England Journal of Medicine, 357*(22), 2277-2284.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD*, 785-794.

(ICRP), C. I. (2020). Publication 135: Diagnostic Reference Levels in Medical Imaging. *Annals of the ICRP, 46*(1).

Dietterich, T. G. (2000). Ensemble Methods in Machine Learning. *International Workshop on Multiple Classifier Systems*, 1-15.

FDA. (2020). *Radiation-emitting products: Medical X-ray imaging.* Obtenido de U.S. Food and Drug Administration : https://www.fda.gov

Ferrante, M., De Marco, P., Rampado, O., Gianusso, L., & Origgi, D. (2025). Effective dose estimation in computed tomography by machine learning. *Tomography, 11*(1), 2.

Garcia-Sanchez, A.-J., Garcia Angosto, E., Llor, J. L., Serna Berna, A., & Ramos, D. (2019). Machine Learning Techniques Applied to Dose Prediction in Computed Tomography Tests. *Sensors,, 19*(23), 5116.

ICRP. (2007). The 2007 Recommendations of the International Commission on Radiological Protection. *Radiological Protection*.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning* (2nd ed ed.). Springer.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning.* Springer.

Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2020). *Digital image processing using MATLAB.* Pearson.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT Press.

Ichikawa, K., Doi, K., Matsubara, K., & Abe, H. (2024). *Machine learning-based estimation of patient body weight from radiation dose information and scan parameters in radiographic examinations.* (e14186 ed., Vol. 25). Journal of Applied Clinical Medical Physics.

Ichikawa, K., Sekimoto, M., Tsujii, H., Fujimoto, K., & Kosaka, N. (2024). *Machine learning-based estimation of patient body weight from radiation dose information and scan parameters in radiographic examinations.* (e14199 ed., Vol. 25). Journal of Applied Clinical Medical Physics.

Johnson, A., & Lee, P. (2021). Machine learning in radiology: A review. *Journal of Medical Systems, 45*(3), 1-12.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A, 374*(2065), 20150202.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765-4774.

MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 281-297.

Ministerio de Salud y Protección Social. (2001). Obtenido de Ley 715 de 2001. Por la cual se dictan normas orgánicas en materia de recursos y competencias en salud: https://www.minsalud.gov.co

Ministerio de Salud y Protección Social. (2018). Obtenido de Resolución 482 de 2018. Por la cual se adoptan los requisitos esenciales de protección radiológica en instalaciones médicas: https://www.minsalud.gov.co

Mitchell, T. M. (1997). *Machine learning.* McGraw-Hill.

Morita, K., Yamada, Y., Nagata, M., & Tsurumaki, T. (2019). Deep neural network-based method for dose estimation in computed tomography. *Sensors, 19*(22).

Organismo Internacional de Energía Atómica (OIEA). (2014). Obtenido de Protección radiológica yseguridad de las fuentes de radiación: Normas básicas internacionales de seguridad (Colección de normas de seguridad del OIEA No. GSR Part 3): https://www.iaea.org/publications/8930

Organismo Internacional de Energía Atómica (OIEA). (2018). Obtenido de Radiation Protection and Safety in Medical Uses of Ionizing Radiation: Specific Safety Guide No. SSG-46. : https://www.iaea.org/publications/12283

Provost, F. F. (2013). *Data science for business.* O'Reilly Media.

Smith, J., Brown, T., & Wilson, K. (2019). Predictive modeling of radiation doses in CT seans. *Radiology, 290*(2), 456-463.

Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., & Kattan, M. W. (2010). Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology, 21*(1), 128-138.

Superintendencia Nacional de Salud. (s.f.). (s.f.). Obtenido de Sistema Obligatorio de Garantía de la Calidad del Sistema General de Seguridad Social en Salud (SOGCS): https://www.supersalud.gov.co

Togawa, A., Morita, K., Tsurumaki, T., & al., e. (2022). A study on the optimization of CT scan parameters using dose monitoring systems and AI. *Tomography, 11*(1), 1-12.

Topol, E. J. (2019). *Deep medicine: How artificial intelligence can make healthcare human again.* Basic Books.

Tukey, J. W. (1977). *Exploratory data analysis.* Addison-Wesley.

Wang, Y., Chen, X., & Zhang, Y. (2018). Data preprocessing in machine learning: A comprehensive review. *Journal of Big Data, 5*(1), 1-25.