



**FACULTAD DE INGENIERIA, ARQUITECTURA Y
URBANISMO
ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS
TESIS**

**Análisis comparativo de algoritmos de machine learning
para clasificar enfermedades pulmonares, haciendo uso de
imágenes radiográficas
PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO
DE SISTEMAS**

Autor(es)

Bach. Huancas Chuquipoma Hilder Rody

ORCID <https://orcid.org/0000-0003-4403-2533>

Bach. Renteria Arce Alvaro

ORCID <https://orcid.org/0000-0003-1545-1210>

Asesor(a)

Mg. Forero Vargas Manuel Guillermo

ORCID: <https://orcid.org/0000-0001-9972-8621>

Línea de Investigación

**Ciencias de la información como herramientas multidisciplinares
y estratégicas en el contexto industrial y de organizaciones**

Sublínea de Investigación

**Informática y transformación digital en el contexto industrial y
Organizacional**

Pimentel – Perú

2024

**ANÁLISIS COMPARATIVO DE ALGORITMOS DE MACHINE LEARNING PARA
CLASIFICAR ENFERMEDADES PULMONARES, HACIENDO USO DE IMÁGENES
RADIOGRÁFICAS.**

Aprobación del jurado

Mg. Heber Iván Mejía Cabrera
Presidente del Jurado de Tesis

Mg. Laurita Belén Guevara Alburqueque
Secretario del Jurado de Tesis

Mg. Arcila Díaz Juan Carlos
Vocal del Jurado de Tesis

Huancas Chuquipoma Hilder Rody Renteria Arce Al...

Análisis comparativo de algoritmos de machine learning para clasificar enfermedades pulmonares, haci

My Files
My Files
Universidad Señor de Sipan

Detalles del documento

Identificador de la entrega
trn:oid::26396:409707371

Fecha de entrega
26 nov 2024, 10:12 a.m. GMT-5

Fecha de descarga
26 nov 2024, 10:15 a.m. GMT-5

Nombre de archivo
tesis Final - turnitin.docx

Tamaño de archivo
3.8 MB

57 Páginas

12,465 Palabras

70,477 Caracteres



Página 1 of 68 - Portada

Identificador de la entrega trn:oid::26396:409707371



Página 2 of 68 - Descripción general de integridad

Identificador de la entrega trn:oid::26396:409707371

16% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- ▶ Bibliografía
- ▶ Texto mencionado
- ▶ Coincidencias menores (menos de 8 palabras)

Fuentes principales

- 8% Fuentes de Internet
- 2% Publicaciones
- 12% Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.



Universidad
Señor de Sipán

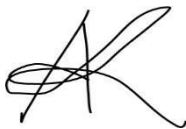

DECLARACIÓN JURADA DE ORIGINALIDAD

Quienes suscribimos la DECLARACIÓN JURADA, somos **Álvaro Rentería Arce, Hilder Rody Huancas Chuquipoma**, del Programa de Estudios de **Ingeniería de Sistemas** de la Universidad Señor de Sipán S.A.C, declaramos bajo juramento que soy (somos) autor(es) del trabajo titulado:

ANÁLISIS COMPARATIVO DE ALGORITMOS DE MACHINE LEARNING PARA CLASIFICAR ENFERMEDADES PULMONARES, HACIENDO USO DE IMÁGENES RADIOGRÁFICAS.

El texto de mi trabajo de investigación responde y respeta lo indicado en el Código de Ética del Comité Institucional de Ética en Investigación de la Universidad Señor de Sipán (CIEI USS) conforme a los principios y lineamientos detallados en dicho documento, en relación a las citas y referencias bibliográficas, respetando al derecho de propiedad intelectual, por lo cual informo que la investigación cumple con ser inédito, original y autentico.

En virtud de lo antes mencionado, firman:

Rentería Arce Alvaro	DNI: 71245286	
Huancas Chuquipoma Hilder Rody	DNI: 72509802	

Pimentel, 12 de septiembre de 2024

Dedicatoria

A mis padres, Pablo Huancas Chinchay y Elsa Chuquipoma Chinchay, cuya dedicación, sacrificio y amor incondicional han sido pilares fundamentales en mi formación. Su apoyo constante y sus enseñanzas me han guiado a lo largo de este camino, permitiéndome superar cada obstáculo y alcanzar este logro. A mi familia y amigos, quienes han estado siempre a mi lado, brindándome aliento y compartiendo cada momento. Este logro es reflejo de la influencia y el cariño de cada uno de ustedes, a quienes dedico con el más sincero reconocimiento y gratitud.

Huancas Chuquipoma Hilder Rody.

Dedico esta investigación, fruto de esfuerzo y dedicación, a Dios, quien me ha guiado en cada paso de mi vida, otorgándome la fuerza y sabiduría necesarias para cumplir mis objetivos. A mi madre, Amparo Arce Barreto, quien con su amor incondicional, apoyo constante y sacrificio diario ha sido mi mayor inspiración. A mis familiares por estar siempre a mi lado, y a mis abuelos que en paz descansen. Gracias a todos ellos por darme fuerzas y aliento en los momentos difíciles. A todos ustedes, mi más profundo agradecimiento y amor eterno

Renteria Arce Alvaro.

Agradecimientos

A la Universidad Señor de Sipán – Chiclayo, por ser una fuente constante de apoyo y proporcionar las herramientas necesarias para nuestro desarrollo académico y profesional. A los profesores y compañeros, por crear un ambiente de aprendizaje y colaboración que enriqueció nuestra formación. Al Mg. Forero Vargas Manuel Guillermo, por su invaluable guía, paciencia y conocimientos compartidos durante todo el proceso de esta investigación. Su dedicación y experiencia fueron esenciales para la culminación de este trabajo y nuestro crecimiento profesional.

A nuestras familias y amigos, por su apoyo incondicional, por estar siempre a nuestro lado con palabras de aliento y por compartir cada paso de este camino. Su cariño, compañía y fortaleza nos dieron la motivación necesaria para seguir adelante incluso en los momentos más desafiantes. Este logro no habría sido posible sin ustedes.

Los Autores.

Índice de Contenidos

Dedicatoria	V
Agradecimientos.....	VI
Índice de tablas	VIII
Índice de figuras.....	IX
Resumen.....	XI
Abstract.....	XII
I. INTRODUCCIÓN.....	1
II. MATERIALES Y MÉTODO	4
2.1. MATERIALES.....	4
2.2. MÉTODO.....	5
III. RESULTADOS Y DISCUSIÓN	34
3.1. Resultados.....	34
3.2. Discusión	53
IV. CONCLUSIONES Y RECOMENDACIONES	56
4.1. Conclusiones	56
4.2. Recomendaciones.....	57
REFERENCIAS.....	59
ANEXOS	67
Anexo 1: Acta de aprobación del asesor	67
Anexo 02: Acta de revisión de similitud de la investigación	68

Índice de tablas

TABLA I: Lista de materiales	4
TABLA II: Conjunto de datos seleccionados.....	8
TABLA III: Algoritmos con rendimientos de revisión sistemática.....	11
TABLA IV: Nomenclatura de términos	32
TABLA V: Promedio de tiempo de entrenamiento	34
TABLA VI: Tiempo de evaluación de algoritmos.....	35
TABLA VII: Indicadores de desempeño.....	46
TABLA VIII: Resultados de evaluación para ResNet50	47
TABLA IX: Resultados de evaluación para CNN-LUNG	49
TABLA X: Resultados de evaluación para Regresión Logística.....	50
TABLA XI: Resultados de evaluación para KNN	51
TABLA XII: Resultados de evaluación para Árbol de Decisión	52

Índice de figuras

Fig 1. Diagrama del método utilizado. Elaboración propia	5
Fig 2: Imágenes radiográficas del dataset creado	10
Fig 3: Subcarpetas de entrenamiento y prueba	10
Fig 4: Estructura de las carpetas de entrenamiento y prueba	10
Fig 5: Arquitectura ResNet50	13
Fig 6.: Imagen convertida a rgb	14
Fig 7: Primera capa del modelo Resnet50.....	15
Fig 8: Representación de Max Pool.....	15
Fig 9: Bloque Residual	16
Fig 10: Arquitectura CNN-LUNG. Elaboración Propia.....	18
Fig 11: Representación gráfica de una convolución	19
Fig 12: Imagen de 224x224 px en escala de grises.....	19
Fig 13: Kernel de 3x3.	20
Fig 14: Ejemplo de aplicación de la función ReLu	20
Fig 15: Ejemplo de funcionamiento de la capa pool en una imagen fotográfica.....	21
Fig 16: Capa flatteng	22
Fig 17: Representación conceptual de una red neuronal de una capa	22
Fig 18: Arquitectura KNN. Elaboración propia	23
Fig 19: Imagen convertida a un vector.....	24
Fig 20: Espacio de características.....	25
Fig 21: Espacio de características.....	25
Fig 22: Imagen convertida a un vector.....	26

Fig 23: Regresion Multinomial	27
Fig 24: Representación de Árbol de decisiones. Elaboración propia	28
Fig 25: Imagen convertida a un vector.....	29
Fig 26: Representación de nodos.....	29
Fig 27: Indice gini	30
Fig 28: Probabilidad que sea Sars-Cov-2	31
Fig 29: Tiempo de entrenamiento de los algoritmos evaluados. Elaboración Propia	35
Fig 30: Tiempo de evaluación de los algoritmos evaluados. Elaboración Propia	36
Fig 31: Matriz de confusión ResNet50. Elaboración Propia	36
Fig 32: Matriz de confusión CNN-LUNG. Elaboración Propia	37
Fig 33: Grafico de entrenamiento y validación.....	38
Fig 34: Matriz de confusión de Regresión Logística. Elaboración Propia.....	39
Fig 35: Matriz de confusión de KNN. Elaboración Propia	40
Fig 36: Imágenes con errores de clasificación: SARS-COV-2, neumonía y normal.	42
Fig 37: Matriz de confusión de Árbol de Decisiones. Elaboración Propia	42
Fig 38: Estructura del árbol de decisión.....	43
Fig 39: Indicadores de desempeño de los algoritmos. Elaboración Propia	46
Fig 40: Resultados por métricas de ResNet50. Elaboración Propia.....	48
Fig 41: Resultados por métricas de CNN-LUNG. Elaboración Propia.....	50
Fig 42: Resultados por métricas de Regresión Logística. Elaboración Propia	51
Fig 43: Resultados por métricas de KNN. Elaboración Propia	52
Fig 44: Resultados por métricas de Árbol de Decisión. Elaboración Propia.....	53

Resumen

Las enfermedades pulmonares como el SARS-COV-2 y la neumonía plantean desafíos críticos para la salud pulmonar, requiriendo una identificación temprana para mejorar el diagnóstico y la atención médica. La presente investigación tiene como objetivo comparar algoritmos de aprendizaje automático en la clasificación de enfermedades pulmonares mediante imágenes radiográficas. Se recopilaron 23,000 imágenes de repositorios públicos distribuidas en conjuntos de entrenamiento y prueba. La evaluación de modelos, incluyendo ResNet50, CNN-LUNG, Regresión Logística, KNN y Árbol de Decisión, reveló una destacada precisión del 97.91%, 97.99%, 91.81%, 91.46% y 86.44% respectivamente. La implementación en Google Colab con Python y Visual Studio Code, utilizando técnicas de aumento de datos, mejoró la generalización. Los resultados, evaluados con métricas como precisión, exactitud, recall y F1-Score, subrayan el rendimiento superior de ResNet50 y CNN-LUNG. Además, se destacó la eficacia de la CNN-LUNG en la identificación de casos de SARS-COV-2 y la necesidad de herramientas interactivas para la interpretación de resultados en entornos clínicos. Estos hallazgos ofrecen perspectivas esenciales para el desarrollo de herramientas diagnósticas y resaltan la eficacia sobresaliente de ResNet50 y CNN-LUNG en este contexto.

Palabras Clave: aprendizaje automático, aprendizaje profundo, enfermedades pulmonares, imágenes radiográficas, clasificadores.

Abstract

Lung diseases such as SARS-COV-2 and pneumonia pose critical challenges to lung health, requiring early identification to improve diagnosis and medical care. The present research aims to compare machine learning algorithms in the classification of lung diseases using radiographic images. A total of 23,000 images were collected from public repositories distributed in training and test sets. Evaluation of models, including ResNet50, CNN-LUNG, Logistic Regression, KNN and Decision Tree, revealed outstanding accuracy of 97.91%, 97.99%, 91.81%, 91.46% and 86.44% respectively. Implementation in Google Colab with Python and Visual Studio Code, using data augmentation techniques, improved generalization. The results, evaluated with metrics such as precision, accuracy, recall and F1-Score, underline the superior performance of ResNet50 and CNN-LUNG. In addition, the effectiveness of CNN-Lung in identifying SARS-COV-2 cases and the need for interactive tools for interpreting results in clinical settings were highlighted. These findings offer essential insights for the development of diagnostic tools and highlight the outstanding efficacy of ResNet50 and CNN-LUNG in this context.

Keywords: machine learning, deep learning, lung diseases, radiographic images, classifiers.

I. INTRODUCCIÓN

En la actualidad, la tecnología continúa avanzando a un ritmo sorprendente en diversos sectores, y estas innovaciones están generando cambios socio-tecnológicos a nivel global, incluyendo el ámbito de la medicina. [1] El aprendizaje de máquinas, conocido en inglés como machine learning, ha tenido un impacto significativo en la práctica médica y se ha convertido en un valioso aliado para la humanidad. Sin embargo, a pesar de estos impresionantes avances tecnológicos, siguen existiendo desafíos considerables en la atención médica, especialmente en lo que respecta al diagnóstico de enfermedades pulmonares. En regiones con recursos limitados, el acceso a los servicios de diagnóstico por imágenes a menudo está restringido y los costos asociados pueden constituir una barrera para numerosos pacientes. Esta situación puede dar lugar a demoras en la identificación y tratamiento, con serias implicaciones para la salud de los pacientes [2]. Además, enfermedades pulmonares como la neumonía y el SARS-COV-2 tienen un profundo impacto en la salud pública, lo que resulta en hospitalizaciones, pérdidas humanas y disminución en la calidad de vida. Hasta abril de 2021, se habían registrado aproximadamente 147 millones de casos de SARS-COV-2 en todo el mundo [3]. Paralelamente, la neumonía es un grave problema global, que afecta principalmente a niños. En el año 2010, se detectaron con 120 millones de casos, de los cuales 14 millones resultaron en fallecimientos [4], lo que subraya la apremiante necesidad de contar con un método de detección más efectiva para estas enfermedades. Además, es posible mejorar la eficacia de los algoritmos de aprendizaje automático en cuanto hace a la clasificación de enfermedades pulmonares [5].

Con base en este contexto, el objetivo principal de esta investigación es analizar comparativamente algoritmos de machine learning para clasificar enfermedades pulmonares, haciendo uso de imágenes radiográficas, a partir de ello surgió la necesidad de plantear la pregunta de investigación ¿Cuál es el mejor algoritmo de machine learning para la clasificación de enfermedades pulmonares haciendo uso de imágenes radiográficas?; esto conllevó a plantear la siguiente hipótesis, que el algoritmo de clasificación CNN-LUNG es el mejor clasificando enfermedades pulmonares haciendo uso de imágenes radiográficas.

Este trabajo de investigación se justificó debido a la necesidad urgente de abordar los desafíos en el diagnóstico de enfermedades pulmonares, especialmente en entornos con recursos limitados y en la importancia de aprovechar el potencial de la inteligencia artificial para mejorar la atención médica y la salud pública en general.

Yari et al. [6] desarrollaron un modelo de aprendizaje de máquina destinado a mejorar la identificación de casos de SARS-CoV-2 a partir de radiografías de tórax. Los modelos propuestos como DenseNet121 y Resnet50 consiguieron un alto nivel de precisión, alcanzando un 97.83%, con muy pocos diagnósticos incorrectos de SARS-CoV-2. Además, se planteó la posibilidad de aplicar este enfoque para la detección de otras enfermedades, con la perspectiva de mejorar aún más su rendimiento mediante el uso de redes neuronales convolucionales. Esta investigación ofreció una solución efectiva para la detección de SARS-CoV-2 en radiografías de tórax y planteó su potencial aplicación en un contexto médico más amplio. También, Qin et al. [7] enfatizaron la importancia de las radiografías de tórax en el diagnóstico de enfermedades, así como el papel del aprendizaje de máquinas, en los sistemas de detección asistida por computadora (CAD). Se encontró que utilizaban métodos considerados más efectivos como la segmentación de imágenes, modelos deformables y aprendizaje profundo, mientras que los enfoques basados en reglas se consideraron menos efectivos y se utilizaron en combinación con otras técnicas para mejorar la precisión en la detección de enfermedades torácicas. Además, Varshni et al. [8] abordaron el problema de la alta mortalidad por neumonía en la India y propusieron un sistema automático de detección diseñado para áreas remotas. En su enfoque, utilizaron extractores de características y clasificadores en radiografías de tórax tanto anormales como normales. Encontraron que la red DenseNet-169 consiguió los mejores descriptores y el SVM fue el clasificador más eficiente entre Regresión logística y Random Forest. En resumen, su investigación se centró en la detección automatizada de neumonía en entornos remotos, con DenseNet-169 desempeñando un papel. Además, Tahir et al. [9] crearon una base de datos de imágenes radiográficas de tórax de SARS-CoV-2, MERS y SARS y desarrollaron un sistema para clasificarlas. Para ello identificaron las regiones pulmonares antes de clasificar las imágenes.

Se encontró que la red InceptionV3 logra sensibilidades del 99,5 %, 93,1 % y 97 % para clasificar imágenes de COVID-19, MERS y SARS, respectivamente, especialmente cuando se segmentaban las imágenes de los pulmones, excluyendo elementos irrelevantes como el corazón o el texto en las imágenes. Asimismo, Joshi et al. [10], desarrollaron un sistema automático para diagnosticar SARS-CoV-2 a partir de radiografías de tórax, utilizando un conjunto diverso de categorías de imágenes, incluyendo SARS-CoV-2 y otras condiciones. Los resultados del modelo DarkNet-53 mostraron una alta precisión en la detección de SARS-CoV-2 frente a otras condiciones, con un 99.61%, y una precisión del 94.79% en la clasificación multiclase de SARS-CoV-2, consiguiendo un modelo eficaz para el diagnóstico rápido de SARS-CoV-2 a partir de radiografías de tórax. De igual manera, Aversano et. al. [11] se centraron en la detección precisa de SARS-CoV-2 a partir de imágenes de tomografía computarizada (TC). Utilizaron redes neuronales profundas, como la VGG19, y aplicaron un algoritmo genético para optimizar los modelos, logrando una precisión de 0.993% y 0.994% en la identificación de casos de SARS-CoV-2 y neumonía en imágenes de TC de tórax utilizando un conjunto de datos construido a partir de bases de datos previamente publicadas. Además, Yao et al. [12] presentaron un algoritmo llamado Pneumonia Yolo (PYolo) para detectar neumonía en imágenes de rayos X de tórax. Este algoritmo permitió localizar áreas afectadas por la neumonía, algo que los enfoques de clasificación binaria no pueden lograr. Utilizaron técnicas como agrupación y muestreo ascendente para ajustar el tamaño de las imágenes y aplicaron mejoras técnicas para aumentar la precisión en la detección de lesiones de neumonía su precisión fue del 0.83%.

II. MATERIALES Y MÉTODO

2.1. MATERIALES

TABLA I: Lista de materiales

Ítem	Nombre	Descripción
1	Conjunto de Datos de Imágenes Radiográficas.	Conjunto de datos que contiene imágenes de enfermedades pulmonares para entrenar y evaluar algoritmos. Construido a partir de las bases de datos [13], [14], [15], [16] y [17]
2	Herramientas de Software	Entorno de desarrollo de Python, bibliotecas de aprendizaje profundo (TensorFlow y Keras), Visual Studio Code para ejecutar el código y Google Colab.
3	Hardware	Una laptop Core i7, con 16gb de RAM, un disco sólido de 512gb con una tarjeta gráfica RTX 760. Una laptop Ryzen 5 5600, con 16gb de RAM, un disco sólido de 256gb con una tarjeta gráfica AMD Radeon.

Nota: Listado de materiales para la investigación. Elaboración propia

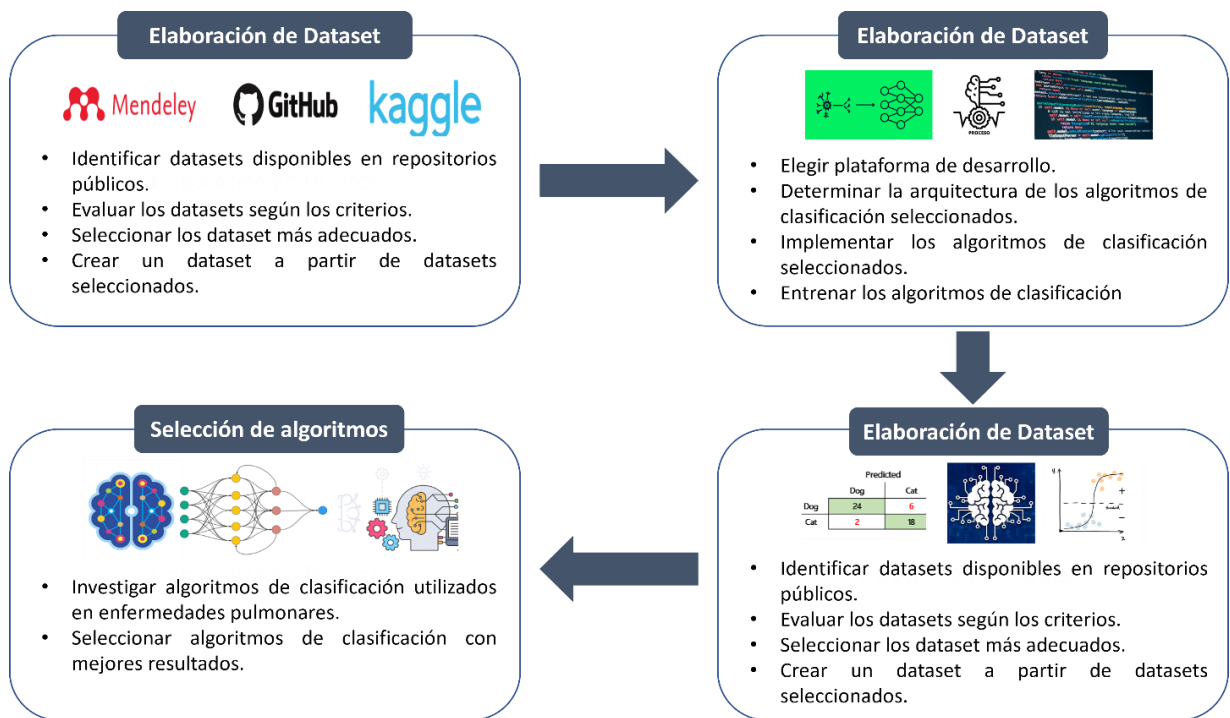


Fig 1. Diagrama del método utilizado. Elaboración propia

2.2. MÉTODO

En la Figura 1, se presenta el método que se utilizó en la investigación, en la primera etapa se seleccionó un dataset de imágenes radiográficas de pacientes con SARS-CoV-2, Neumonía y sujetos sanos, en la segunda etapa se realizó la selección de los algoritmos de clasificación de machine learning, en la tercera etapa se implementó los algoritmos de clasificación, finalmente se evaluaron estos resultados según los indicadores.

Para la selección del Dataset se llevó a cabo una exploración de repositorios públicos como Kaggle, Github y Mendeley con la finalidad de identificar conjuntos de datos que albergaran imágenes radiográficas de pacientes afectados por SARS-COV-2, neumonía y sujetos sanos. Se emplearon términos clave específicos, tales como " SARS-COV-2 thoracic radiograph dataset", " pneumonia radiographic dataset" y " radiographic dataset", con el fin de optimizar la búsqueda. La elección de los repositorios se basó en la identificación de conjuntos de datos que cumplieran con los criterios de interés los cuales fueron el tamaño de las imágenes, la cantidad de imágenes, división del conjunto de datos de entrenamiento,

validación y prueba, prestando especial atención a la descripción del conjunto de datos. Los resultados fueron examinados y filtrados para garantizar su relevancia para la investigación. La evaluación de la calidad del conjunto de datos se centró en aspectos resolución de las imágenes. Además, se consideró el tamaño del conjunto de datos, priorizando conjuntos de datos más extensos para mejorar la capacidad de entrenamiento y evaluación de modelos. La revisión visual de una muestra aleatoria de imágenes se llevó a cabo para asegurar la calidad del conjunto de datos. La fuente del conjunto de datos fue verificada para garantizar su confiabilidad, y se realizó una revisión exhaustiva de la literatura científica para confirmar su utilidad en investigaciones previas.

Con el objetivo de crear el conjunto de datos, se siguieron una serie de criterios rigurosos para garantizar que la información recopilada fuera precisa y relevante. A continuación, se describirán estos criterios de manera detallada

Tabla II: Criterios de selección de conjunto de datos

Criterios de selección de conjunto de datos	
Criterio	Descripción
Calidad de imágenes	Se priorizó la selección de conjuntos de datos que contaran con imágenes radiográficas en formato JPG con una resolución mínima de 600 x 600 píxeles. Este tamaño fue elegido para garantizar que las estructuras pulmonares, como opacidades u otras características relacionadas con SARS-CoV-2, neumonía y casos normales, fueran suficientemente visibles para el análisis y procesamiento por los algoritmos de clasificación. Además, se verificó que las imágenes estuvieran en escala de grises, ya que este formato es esencial para reducir la redundancia de información no relevante (como el color) y resaltar características estructurales clave en los tejidos pulmonares. Para asegurar que estas condiciones fueran cumplidas, se realizó una evaluación

Cantidad de imágenes	<p>inicial en una muestra representativa de cada conjunto de datos, verificando que cumplieran con los criterios definidos. Este enfoque permitió optimizar el proceso sin revisar manualmente todas las imágenes, lo cual sería inviable debido al gran volumen de datos.</p> <p>Se seleccionaron conjuntos de datos con un volumen mínimo de 200 imágenes por clase (SARS-CoV-2, neumonía y normales), asegurando un balance adecuado para el entrenamiento, validación y prueba de los modelos. Una cantidad suficiente de imágenes es crucial para que los algoritmos puedan generalizar correctamente, evitando problemas de sobreajuste. Además, se priorizó la utilización de conjuntos de datos públicos bien documentados, como los disponibles en plataformas como Kaggle y GitHub, que proporcionan información detallada sobre la procedencia y características de las imágenes.</p>
División de conjunto de datos E/V/P	<p>Los conjuntos de datos seleccionados debían estar estructurados o permitir una división en tres subconjuntos: entrenamiento (E), validación (V) y prueba (P), con una proporción de 80%, 10% y 10%, respectivamente. Esta división asegura que el modelo se entrene con la mayor cantidad de datos posibles (80%), se ajuste adecuadamente mediante la validación (10%) y se evalúe su rendimiento en datos no vistos previamente (10%). Este enfoque minimiza el riesgo de sesgos, proporcionando una evaluación confiable del modelo. La proporción 80-10-10 fue seleccionada específicamente porque permite un balance óptimo entre el tamaño de los datos utilizados para el aprendizaje y los requeridos para la evaluación. Además, garantiza que el modelo</p>

tenga suficientes ejemplos para aprender patrones generales mientras se evita el sobreajuste.

En la fase final, se procedió con la selección definitiva de cinco conjuntos de datos de repositorios como GitHub y Kaggle, extrayendo un total de 23,000 imágenes de las carpetas de entrenamiento y prueba. Estos conjuntos de datos se consideran los más adecuados para avanzar en la investigación, proporcionando una base sólida para el análisis de imágenes radiográficas en el contexto de SARS-COV-2, neumonía y condiciones normales.

En la Tabla II se presentan los conjuntos de datos seleccionados y se verifica si cumplen con los criterios para formar parte de la investigación quedando seleccionado las siguientes bases de datos Covid Chestray, Chestray neumonía, Covid19 radiography database, Sars-cov-2ctscan y Covid chest ray x.

TABLA III: Conjunto de datos seleccionados

Nombre	N° de imágenes	Artículo	Criterios			¿Cumple?
			Calidad de imágenes	Cantidad de imágenes	División de conjunto de datos E/N/P	
Covid Chestray	950 imágenes (incluye imágenes de COVID-19, neumonía y otras enfermedades)	[2], [15], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28]	Si	Si	No	Si
Chestray Pneumonia	5863 imágenes (incluye imágenes de neumonía y normales)	[29], [30]	Si	No	Si	Si
Covid19 Radiography Database	21,000 imágenes (incluye imágenes de COVID-19, neumonía y normales)	[31], [32], [33], [34], [35], [36]	Si	Si	Si	Si
Sars-Cov-2-Ctscan	2,000 imágenes de tomografías computarizadas (incluye imágenes de COVID-19)	[37], [38]	Si	Si	No	Si
Covid Chest Ray X	6432 imágenes (incluye imágenes de COVID-19, neumonía y	[39], [40], [41], [42]	Si	Si	No	Si

	normales)						
Covid19 Chest Ray Detection	300 imágenes (incluye imágenes de COVID-19 y otras condiciones)	[43], [44]	Si	No	No	No	No
Actualmed-COVID-Chestxray-Dataset	200 imágenes (incluye imágenes de COVID-19, neumonía y normales)	[5], [17]	Si	No	No	No	No
COVID-Chestxray-Dataset	250 imágenes (incluye imágenes de COVID-19, neumonía y normales)	[45], [46]	Si	No	No	No	No
Covid19-X-Rays	200 imágenes (incluye imágenes de COVID-19, neumonía y normales)	[16], [47], [48], [49], [50], [51]	Si	No	No	No	No
Pneumonia-Sample-Xrays	5000 imágenes (incluye imágenes de neumonía)	[52]	No	Si	No	No	No

Nota: Conjunto de datos seleccionados. Elaboración propia

Luego de seleccionar y unificar los datasets, se recortó el área de interés de las imágenes a un tamaño de 224x224 píxeles. Esta resolución se eligió para optimizar el tiempo de procesamiento y los recursos computacionales, mientras se aseguraba que se mantuvieran los detalles clave necesarios para la clasificación, como las texturas pulmonares y las opacidades asociadas a enfermedades como SARS-CoV-2 y neumonía. Al reducir el tamaño de las imágenes, se conserva la información más relevante para el diagnóstico, como las áreas afectadas por opacidades, lo que permite que el modelo siga siendo eficaz sin comprometer el rendimiento del procesamiento. Con esta reducción, se obtuvieron 7,124 imágenes de SARS-CoV-2, 7,013 imágenes normales y 6,884 imágenes de neumonía para el entrenamiento, mientras que 2,336 imágenes fueron reservadas para la prueba, distribuidas equitativamente entre las tres clases: 778 imágenes de SARS-CoV-2, 779 imágenes normales y 779 imágenes de neumonía.

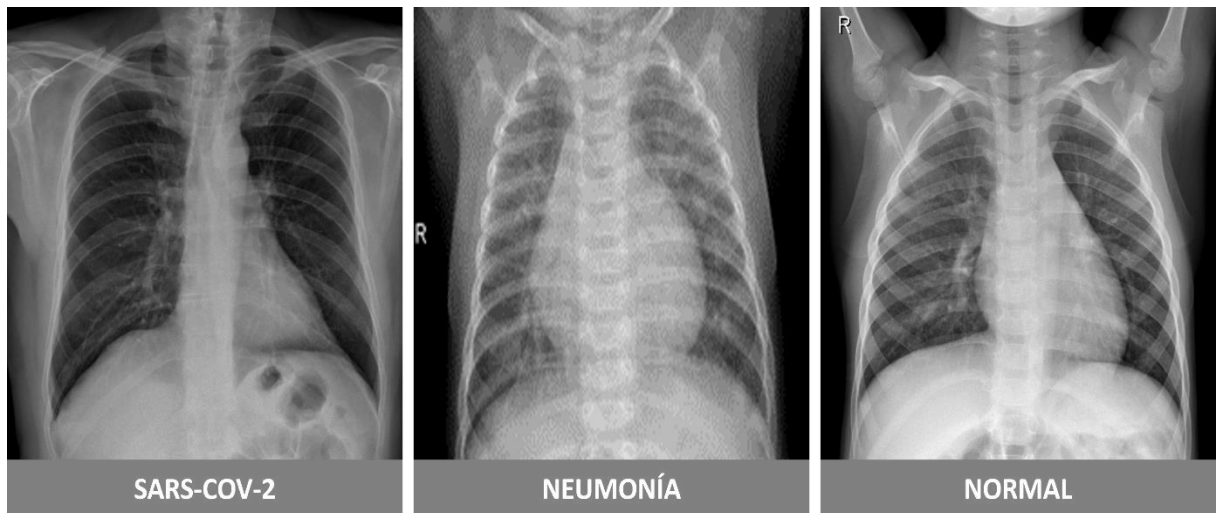


Fig 2: Imágenes radiográficas del dataset creado

Como resultado, se organizó la estructura de carpetas de la siguiente manera: una carpeta principal llamada “Data”, dentro de la cual se crearon dos subcarpetas denominadas “train” y “test”, correspondientes a los conjuntos de datos de entrenamiento y prueba, respectivamente. Esta organización se muestra en la Fig 3.

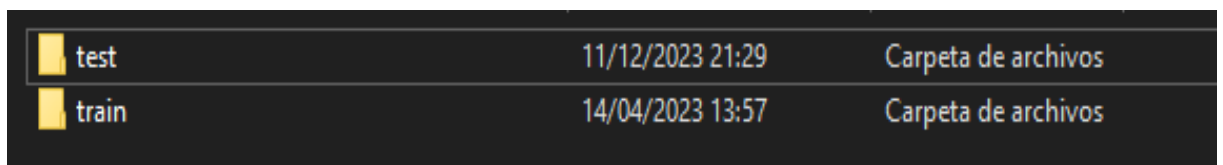


Fig 3: Subcarpetas de entrenamiento y prueba

Dentro de estas carpetas, las imágenes se han organizado en tres subcarpetas correspondientes a las clases Covid-19, Normal y Neumonía, lo que facilita la separación y el acceso a los diferentes conjuntos de datos. Esta estructura se puede observar en la Fig 4.

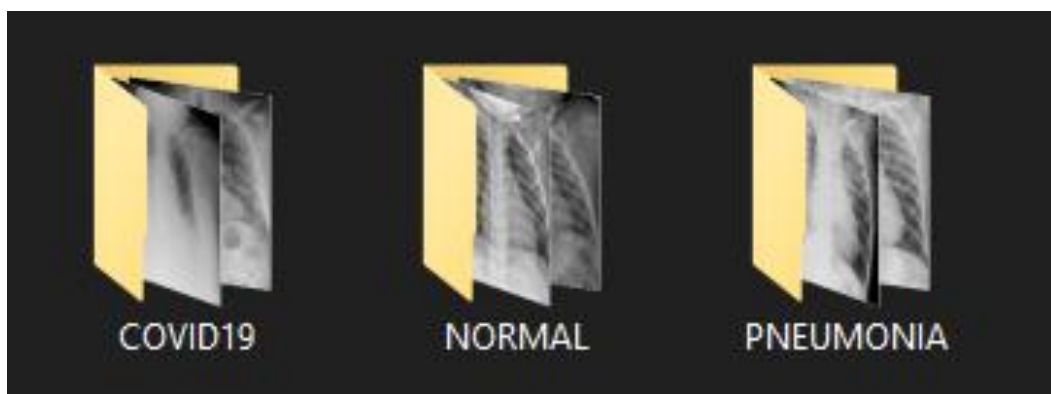


Fig 4: Estructura de las carpetas de entrenamiento y prueba

Asimismo, en la elección de algoritmos para la clasificación de enfermedades pulmonares, se llevaron a cabo exploraciones exhaustivas de opciones empleadas en estudios anteriores, abarcando una diversidad de algoritmos con características distintivas. Inicialmente, se consideraron Redes Neuronales Convolucionales (CNN) para la generación del modelo, ampliamente utilizadas en el procesamiento de imágenes debido a su eficacia en la detección de enfermedades pulmonares mediante capas de convolución. Además, se empleó árbol de decisiones, un método simple de clasificación que toma decisiones en pasos según las características, aunque puede tener problemas de ajuste excesivo en conjuntos de datos pequeños o con ruido. En este contexto, también se evaluó el algoritmo Vecinos más cercanos (KNN), un método basado en instancias comúnmente empleado en clasificación y regresión. Su funcionamiento implica la comparación de una nueva instancia con las instancias en el conjunto de datos de entrenamiento y asigna la clase más frecuente entre las K instancias más cercanas, proporcionando así una comprensión más clara de un punto de vista práctico. Adicionalmente, se consideró la Regresión Logística, un enfoque estadístico aplicable a problemas de clasificación binaria. La elección definitiva de algoritmos se basó en resultados previos, destacando ResNet50, CNN y Árbol de decisión por su eficacia en la clasificación de enfermedades pulmonares, evidenciada en términos de precisión, exactitud y recall.

En la Tabla IV se presenta el rendimiento obtenido por los diez mejores algoritmos encontrados durante la fase de revisión bibliográfica.

TABLA IV: Algoritmos con rendimientos de revisión sistemática

Algoritmo	Artículo	Precisión	Exactitud	Recall	F1-Score	Aplicación en estudios previos
DenseNet12 1	[33], [53]		95.68%	95.28%	95.62%	Clasificación de imágenes de radiografías de tórax para identificar COVID-19, neumonía y casos normales.
ResNet18	[19], [40], [52], [53]		92.86%	96%	95%	Utilizado para distinguir entre COVID-19, neumonía y otros casos

pulmonares en imágenes de rayos X de tórax.

Usado para la clasificación de diversas condiciones pulmonares (COVID-19, neumonía) en imágenes radiográficas.

Clasificación inicial de imágenes de tórax en categorías de COVID-19 y normales, pero con menor precisión en grandes conjuntos de datos.

Usada para distinguir entre imágenes de COVID-19 y normales; también para clasificación binaria en otros estudios pulmonares.

Clasificación de alta precisión para COVID-19, neumonía y normales en imágenes de rayos X de tórax.

Modelo general para la clasificación de COVID-19, neumonía y otras enfermedades pulmonares en radiografías de tórax.

Clasificación binaria entre imágenes de COVID-19 y no COVID-19, a menudo en combinación con extracción de características.

Utilizado para la clasificación de diferentes patologías pulmonares, especialmente en conjuntos de datos limitados.

Empleado en la clasificación de COVID-19 y neumonía en imágenes radiográficas, optimizando precisión en conjuntos de datos grandes.

ResNet50	[16], [19], [33], [34], [40], [45], [54]	98.23%	97.79%	98.75%	98.06%
KNN	[39], [55]	72.66%	79.8%	50.97%	46.92%
Regresión logística	[15], [20], [22], [51], [56]	91.20%	91.10%	95.80%	91.20%
VGG-19	[16], [33], [39], [40], [53], [57]	98.79%	99.14%	99.55%	98.5%
CNN	[19], [33], [35], [38], [39], [50], [52], [58], [59]	94.09%	95.24%	94.62%	94.24%
SVM	[55], [56], [58]		95.38%	93%	98.91%
Árbol de decisión	[5], [22], [33], [56]		93.97%	96.53%	96.50%
XGBoost	[25], [58]		95.46%	94.63%	94.66%

Nota: Algoritmos con rendimientos de revisión sistemática. Elaboración propia

De los algoritmos seleccionados, se ha considerado trabajar específicamente con los siguientes cinco: ResNet50, CNN, Regresión logística, KNN y Árbol de decisión.

a. ResNet50

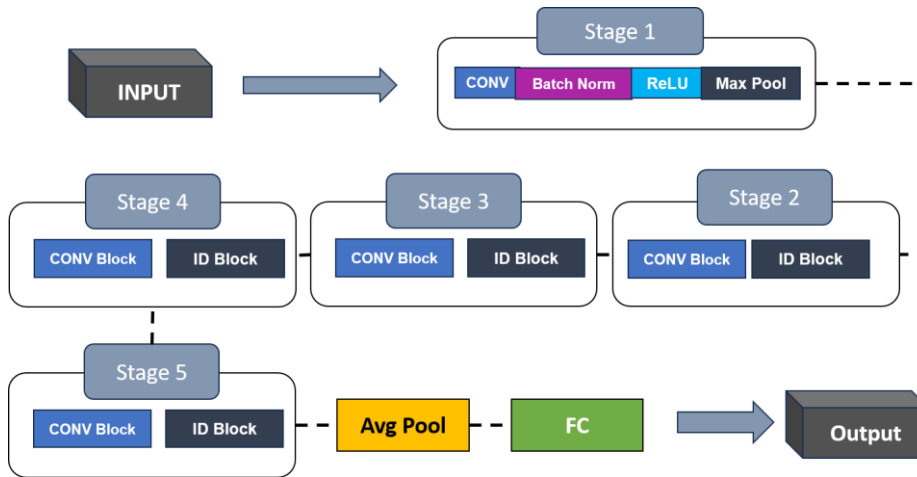


Fig 5: Arquitectura ResNet50

ResNet50 es una arquitectura profunda con 50 capas de redes neuronales convolucionales diseñadas para el procesamiento de imágenes, destacando en visión por computadora por su capacidad de extraer características visuales complejas. Esta red sobresale en tareas de clasificación, detección y segmentación de objetos. Su arquitectura residual facilita el aprendizaje eficiente de patrones visuales, adaptándose a diversos campos como el análisis médico, procesamiento de imágenes satelitales y sistemas de vigilancia [17]. La ResNet-50 espera una entrada de una imagen con 3 canales (RGB). En nuestro caso, hemos redimensionado primero nuestras imágenes radiográficas de su tamaño original a 224x224 píxeles. Luego, los valores de los píxeles se normalizan a un rango entre [0,1] o [-1,1]. Sin esta normalización, los valores de los píxeles pueden ser muy grandes (0-255), lo que provoca inestabilidad en el modelo y hace que el entrenamiento sea más lento e ineficiente. Esto puede dificultar la predicción precisa de la clase a la que pertenece cada imagen radiográfica, ya sea Sars-Cov-2, neumonía o normal.

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}}$$

- I : Valor del píxel en la imagen original.

- I : Valor mínimo posible del píxel (0 para imágenes de 8 bits).
- I : Valor máximo posible del píxel (255 para imágenes de 8 bits).
- I : Valor normalizado del píxel.

Después a nuestra imagen resultante duplicamos su canal de grises en los tres canales RGB, donde se copia el valor de intensidad del canal de escala de grises a los tres canales (R, G, B) donde la imagen resultante visualmente sigue siendo blanco y negro, pero ahora tiene 3 canales. Como ResNet-50 es un modelo preentrenado en ImageNet, esta espera imágenes en formato RGB como en la Fig 6. Si se usa una imagen de 1 canal, el modelo no funcionará correctamente.



Fig 6.: Imagen convertida a rgb

La red comienza con una capa convolucional inicial (Conv1) que tiene 64 kernels de 7x7 cada uno de estos kernels se utiliza para extraer características de nuestra imagen esto ayuda a identificar patrones específicos en las radiografías que son indicativos de diferentes enfermedades pulmonares. En este caso, cada filtro tiene una dimensión de 7 píxeles por 7 píxeles, lo que significa que el filtro analizará una región de 7x7 píxeles de la imagen de entrada a la vez. El tamaño del filtro afecta la resolución de las características que se pueden detectar. Un filtro de 7x7 es relativamente grande, lo que significa que la red podrá capturar patrones más grandes y complejos en la imagen radiográfica.

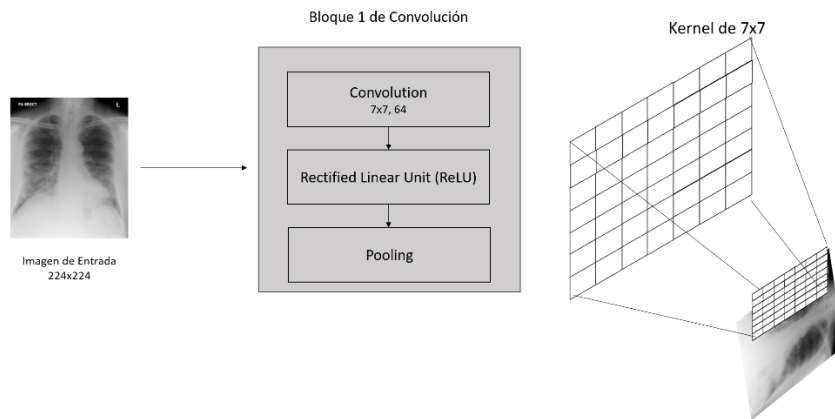


Fig 7: Primera capa del modelo Resnet50

Luego, se aplica una función de activación no lineal, específicamente ReLU (Rectified Linear Unit), que establece en 0 todos los valores negativos. Esto introduce no linealidad en el modelo, permitiendo que la red aprenda representaciones más complejas de las imágenes radiográficas. Esta capacidad de capturar características complejas es esencial para identificar detalles críticos en radiografías, como opacidades intersticiales en casos de neumonía o patrones nodulares en casos de COVID-19. Por último, la capa de max pooling de 2x2 se utiliza para reducir la cantidad de parámetros y conservar las características más relevantes como en la Fig 8. Esta capa selecciona los valores máximos dentro de una ventana específica, lo que ayuda a destacar las características más prominentes de la imagen, como los bordes y las áreas de mayor intensidad. En el contexto de imágenes radiográficas, estas características prominentes pueden incluir opacidades pulmonares, consolidaciones y otros patrones que son cruciales para el diagnóstico.

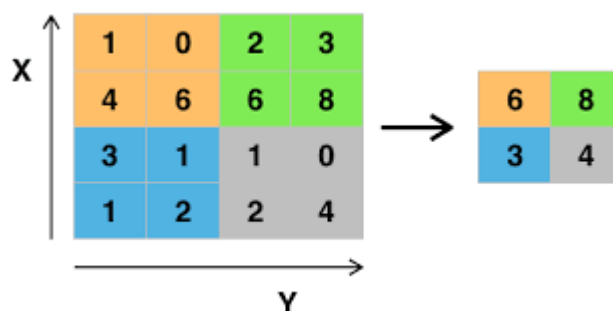


Fig 8: Representación de Max Pool

La ResNet-50 tiene 4 etapas principales, cada una compuesta por varios bloques

residuales con 3 capas convolucionales por bloque. Un bloque residual es un conjunto de capas dentro de la red que incluye un camino directo (skip connection o shortcut connection) que omite una o más capas intermedias y suma su salida directamente al final del bloque. Este camino directo facilita el flujo de la información y las gradientes a través de la red, mitigando problemas como el desvanecimiento del gradiente.

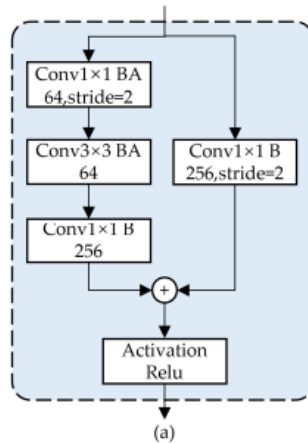


Fig 9: Bloque Residual

La gradiente es esencial porque guía cómo ajustar los pesos de la red para mejorar su precisión de clasificación. Gracias a este proceso, ResNet50 puede aprender a distinguir entre las imágenes de Sars-Cov-2, neumonía y normales de manera más efectiva. Al permitir que las gradientes se propaguen fácilmente, los bloques residuales hacen que el entrenamiento de las redes profundas sea más preciso, donde las skip connection permiten que la información fluya a través de varias capas preservando como bordes y texturas específicos de enfermedades pulmonares (Sars-Cov-2, neumonía y normales) aprendidas en anteriores capas.

Después se utiliza la capa Pooling promedio global (Global Average Pooling) ayuda a capturar las características más relevantes sin introducir el riesgo de sobreajuste que tendrían las capas completamente conectadas. Este enfoque es especialmente útil en la clasificación de enfermedades como Sars-Cov-2 y neumonía, donde es importante identificar patrones específicos en áreas localizadas de los pulmones. La capa Fully Connected (FC) Con 1000 neuronas asegura que cada imagen radiográfica sea asignada a una de las clases posibles

(Sars-Cov-2, Neumonía y Normal) con una probabilidad que refleja la confianza del modelo en su predicción. Después del pooling, el vector resultante de tamaño 2048 pasa a una capa completamente conectada que genera un número de salidas igual al número de clases. Serían 3 clases: SARS-CoV-2, neumonía y normalidad.

La capa softmax en una red neuronal, como ResNet50, es una capa final crucial, especialmente cuando se utiliza para clasificación multiclase. Su principal función es convertir los valores de la capa totalmente conectada (logits) en probabilidades que sumen 1, usando la fórmula softmax.

$$P(i) = \frac{e^{z_i}}{\sum_{j=1}^C e^{z_j}}$$

En esta fórmula, cada logit se transforma exponencialmente y luego se normaliza dividiendo por la suma de todos los valores exponenciados. Esto convierte los logits en un vector de probabilidades, donde cada probabilidad indica la confianza del modelo en que la imagen pertenece a una clase específica. Por ejemplo, si estás clasificando imágenes radiográficas para detectar Sars-Cov-2, neumonía y casos normales, los valores de la capa totalmente conectada se pasan a través de la función de activación softmax, que convierte estos valores en probabilidades que suman 1. Esto es esencial porque permite interpretar los resultados de manera clara y precisa.

La capa softmax facilita la toma de decisiones al seleccionar la clase con la probabilidad más alta como la predicción final. Por ejemplo, una imagen radiográfica de tórax evaluada por la red podría producir una salida como 85% de probabilidad de SARS-CoV-2, 10% de probabilidad de neumonía y 5% de probabilidad de normalidad.

b. CNN-LUNG

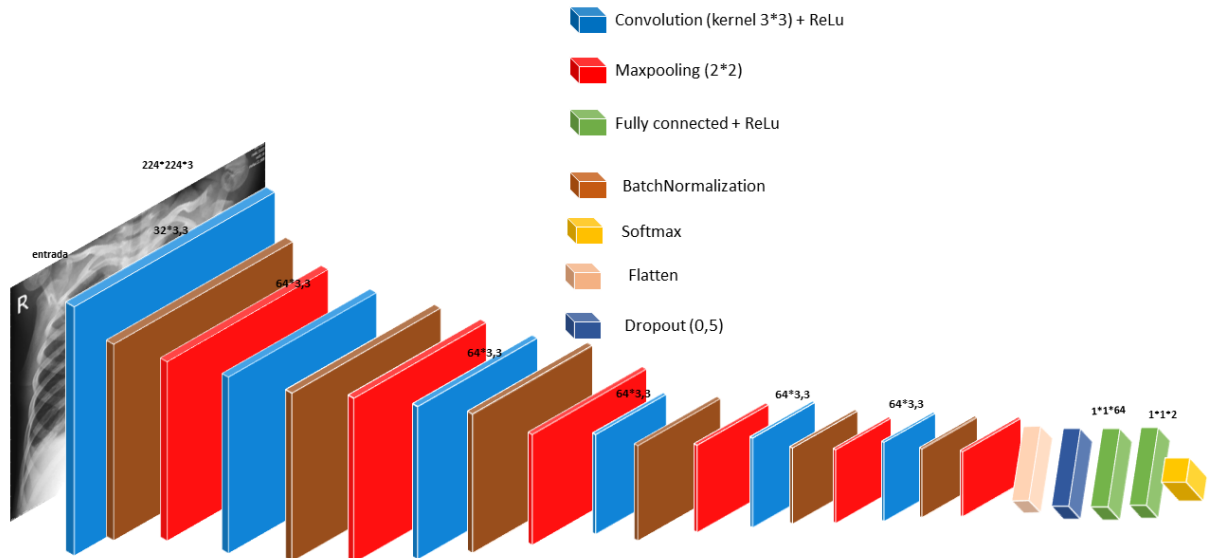


Fig 10: Arquitectura CNN-LUNG. Elaboración Propia

Una red neuronal convolucional (CNN, por sus siglas en inglés) es un tipo de red neuronal diseñada específicamente para procesar datos que tienen una estructura de tipo cuadrícula, como imágenes. Las CNN han demostrado ser extremadamente efectivas para tareas relacionadas con la visión por computadora, como la clasificación de imágenes, detección de objetos y segmentación semántica.

Una convolución es una operación matemática que va desplazando una función por sobre otra y calcula en cada intervalo cuán contenida se encuentra una función en la otra. La definición matemática de convolución se muestra en la siguiente ecuación.

$$y(t) = \int_{-\infty}^{+\infty} f(n) \cdot g(t - n) \cdot d_n$$

Una convolución puede representarse gráficamente como muestra la Fig 11.

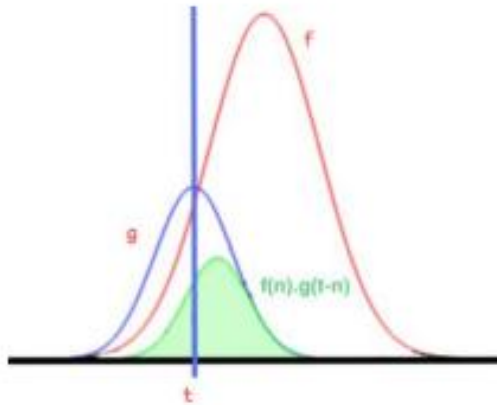


Fig 11: Representación gráfica de una convolución

En este caso, vamos a describir la arquitectura de una red neuronal convolucional (CNN) específica para la clasificación de imágenes radiográficas de enfermedades pulmonares como Sars-Cov-2, neumonía y normal. El modelo espera una imagen de entrada de 224x224 px y de 1 canal que sería en escala de grises.

Imagen en Escala de Grises



Fig 12: Imagen de 224x224 px en escala de grises

La imagen ingresa a la primera capa convolucional de 32 Kernels de 3x3 como muestra la Fig 13, donde identificamos patrones simples como bordes y texturas iniciales de nuestras imágenes radiográficas.

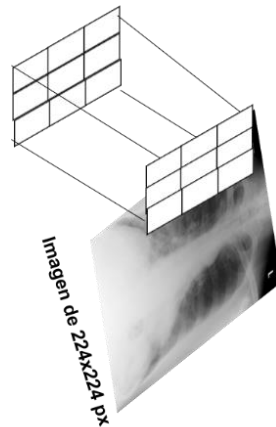


Fig 13: Kernel de 3x3.

Donde a la capa se le agrega una función de activación ReLU consiste en quedarse solamente con los valores positivos de la entrada. Esto significa que dado un número X , la función devuelve el máximo entre X y 0 . Mientras más grandes sean los valores obtenidos más parecidos a la característica buscada. Se define como indica la siguiente ecuación y su resultado se aprecia en la Figura 14 donde se consideran como valores positivos a aquellos de color blanco mientras que aquellos de color negro son considerados negativos.

$$f(x) = \max(x, 0)$$

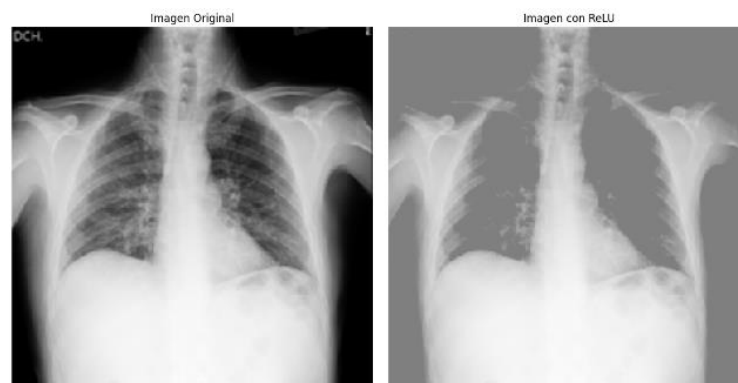


Fig 14: Ejemplo de aplicación de la función ReLU

La red incluye cinco capas convolucionales adicionales, cada una con 64 filtros de tamaño 3×3 , seguidas de procesos de Batch Normalization y Max Pooling con un tamaño de 2×2 . Estas capas están diseñadas específicamente para capturar características detalladas y complejas en radiografías, como patrones que podrían indicar anomalías pulmonares,

consolidaciones, o signos de enfermedades como el SARS-CoV-2 o la neumonía. El Batch Normalization se utiliza para estabilizar y acelerar el entrenamiento al ajustar dinámicamente las activaciones intermedias, reduciendo el riesgo de que la red se desestabilice debido a cambios bruscos en las distribuciones internas. Por otro lado, el Max Pooling ayuda a reducir la dimensionalidad de las características extraídas como en la Fig 16, preservando los detalles más relevantes y permitiendo que el modelo enfoque su atención en las estructuras significativas de las imágenes radiográficas, como los bordes o patrones anómalos en los tejidos.

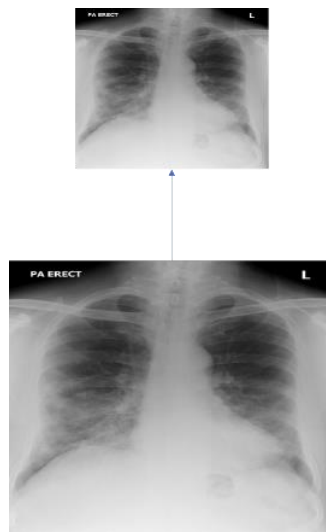


Fig 15: Ejemplo de funcionamiento de la capa pool en una imagen fotográfica

Después de las capas convolucionales, la red aplanada (Flatten) convierte los mapas de características 2D en un vector 1D como en la Fig 16. Este formato es necesario porque las capas densas, o totalmente conectadas, solo pueden procesar datos en forma de vectores, ya que operan mediante multiplicaciones matriciales. Esto permite que la información extraída, como patrones relevantes en las imágenes radiográficas, sea interpretada y combinada para tomar decisiones finales.

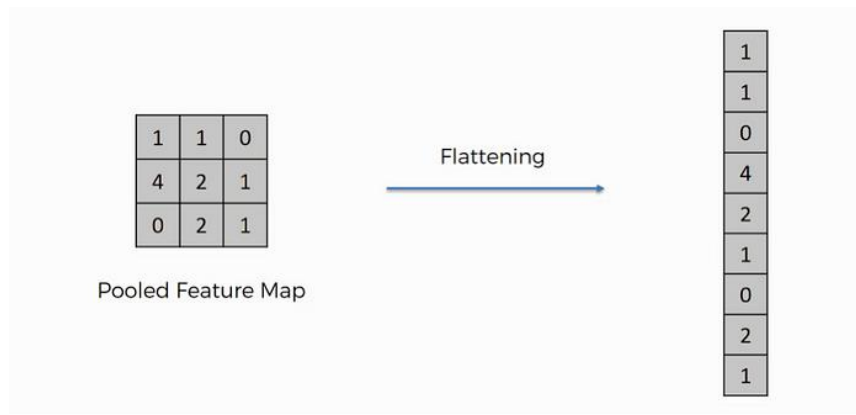


Fig 16: Capa flatteng

Además, se incluye una capa de Dropout con una tasa del 50% para prevenir el sobreajuste, lo cual es esencial en el contexto de imágenes médicas, donde la cantidad de datos de entrenamiento puede ser limitada, y el modelo debe generalizar bien para clasificar correctamente nuevas imágenes radiográficas de Sars-Cov-2, neumonía y pulmones normales.

La primera capa densa, con 64 neuronas y activación ReLU, se encarga de aprender representaciones no lineales complejas. Esta capacidad es vital para capturar las sutilezas y variaciones en las imágenes radiográficas, como diferentes patrones de opacidades pulmonares, consolidaciones y otras irregularidades que pueden indicar enfermedades pulmonares. La activación ReLU introduce no linealidad, permitiendo que el modelo distinga características más complejas y abstractas de las imágenes.

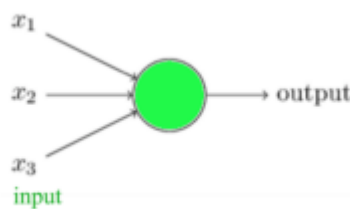


Fig 17: Representación conceptual de una red neuronal de una capa

Finalmente, la capa de salida, está formado por una capa densa donde se le pasa el número de neuronas igual al número de clases que estás clasificando. En este caso, 3 neuronas porque tienes tres categorías: "COVID-19", "Neumonía" y "Normal". La activación

softmax convierte los valores de salida (logits) de las neuronas en probabilidades. Su principal función es convertir los valores de la capa totalmente conectada (logits) en probabilidades que sumen 1, usando la fórmula softmax.

$$P(i) = \frac{e^{z_i}}{\sum_{j=1}^C e^{z_j}}$$

Esta capa es la encargada de transformar las características aprendidas en una predicción final de la categoría, haciendo que el modelo pueda clasificar efectivamente entre "COVID-19", "Neumonía" y "Normal" en imágenes radiográficas.

c. K-Nearest Neighbors (KNN)

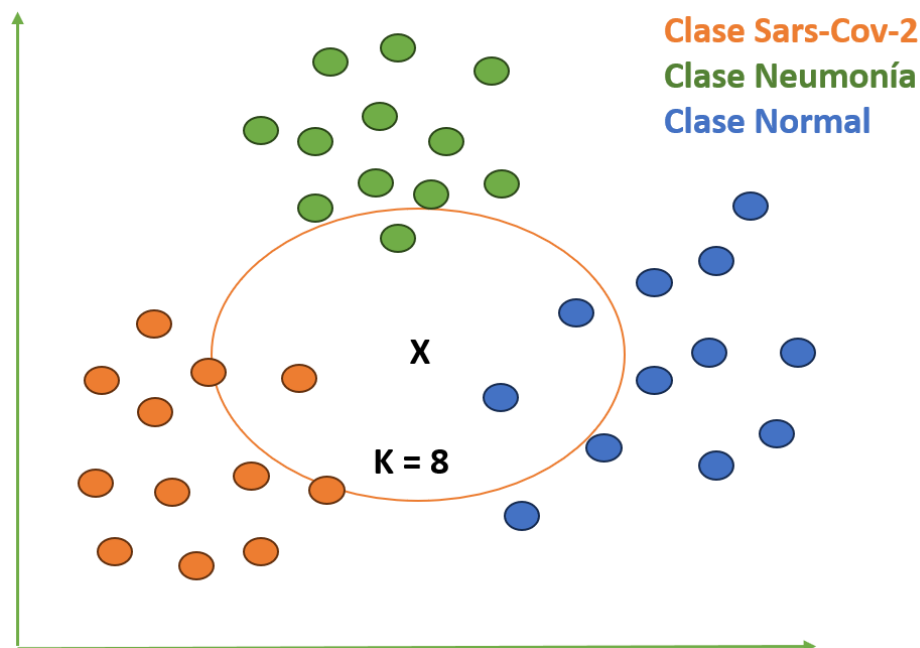


Fig 18: Arquitectura KNN. Elaboración propia

KNN es un algoritmo basado en instancias o memoria. No utiliza un modelo explícito para realizar predicciones, sino que almacena los datos de entrenamiento y los compara directamente con nuevos datos para hacer inferencias. Se clasifica como un método de aprendizaje perezoso (lazy learning) porque no realiza un entrenamiento explícito.

Como entrada pasamos imágenes de 224x224 px, en escala de grises, ya que se reduce la complejidad del procesamiento al tener un solo canal en lugar de tres (RGB).

Normalizamos los valores de los píxeles donde se escalan para que estén en un rango específico, generalmente 0 a 1, para evitar que valores altos dominen la distancia euclidiana. Las imágenes bidimensionales se aplanan en un vector unidimensional. Por ejemplo, una imagen de 224x224 píxeles se convierte en un vector de 50176 (224x224) elementos como en la Fig 19. Este paso es crucial porque KNN requiere que las entradas sean vectores para calcular distancias entre ellas.

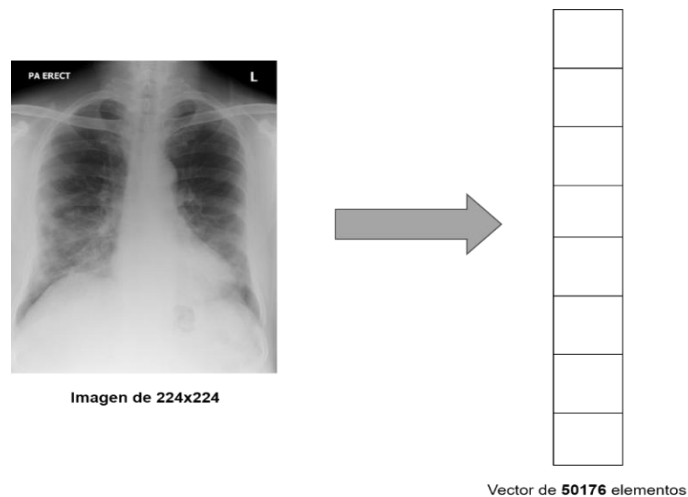


Fig 19: Imagen convertida a un vector

Durante la fase de entrenamiento, el algoritmo KNN no realiza un verdadero proceso de "aprendizaje" como otros algoritmos (por ejemplo, redes neuronales). En lugar de eso, simplemente almacena las imágenes de entrenamiento junto con sus etiquetas correspondientes (Sars-Cov-2, neumonía o pulmones normales).

Lo que hace KNN es memorizar todas las imágenes de entrenamiento y las etiquetas asociadas a ellas. No hay un modelo complejo que se entrene como en los métodos de aprendizaje profundo. En cambio, el entrenamiento consiste en almacenar los datos y las etiquetas, y calcular las distancias entre las imágenes para realizar la clasificación cuando se le da una nueva imagen para predecir.

Una vez que el modelo está entrenado, para clasificar una nueva imagen radiográfica, el algoritmo KNN calcula la distancia entre la imagen de prueba (que también se ha preprocesado y aplanado) y todas las imágenes de entrenamiento. El tipo de distancia más

comúnmente utilizada es la distancia euclidiana.

$$\text{Distancia Euclidiana: } d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

El algoritmo selecciona los K vecinos más cercanos, es decir, las K imágenes de entrenamiento que están más cerca de la imagen de prueba en el espacio de características (el espacio creado por los valores de píxeles) como se muestra en la Fig 20. La "cercanía" se mide a través de la distancia euclidiana entre los vectores de características.

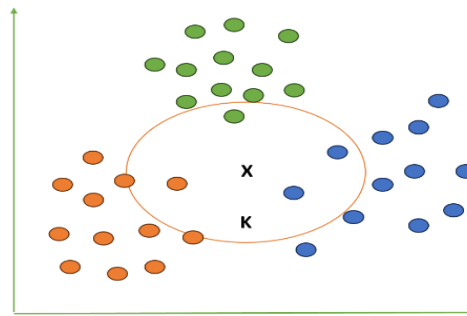


Fig 20: Espacio de características

Una vez que se identifican los K vecinos más cercanos, KNN realiza una votación mayoritaria entre las etiquetas de esas K imágenes vecinas. La clase que aparece con mayor frecuencia entre esos vecinos se asigna como la predicción para la imagen de prueba. Por ejemplo, si se elige K=3 y los tres vecinos más cercanos tienen las siguientes etiquetas: [neumonía, neumonía, normales], el modelo clasificará la imagen de prueba como neumonía.

d. Regresión logística

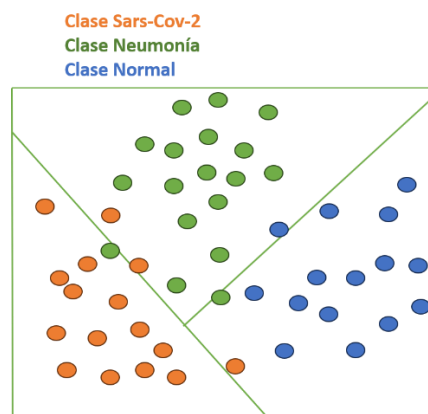


Fig 21: Espacio de características

La regresión logística es un modelo estadístico ampliamente utilizado en machine

learning y estadística para resolver problemas de clasificación binaria o multinomial. Su principal objetivo es modelar la relación entre una variable dependiente categórica y una o más variables independientes (predictoras), sean estas continuas o categóricas. En la regresión logística binaria, la variable objetivo tiene dos categorías (0 y 1). En la regresión logística multinomial, la variable objetivo tiene más de dos categorías.

Al igual que con otros algoritmos de aprendizaje automático, las imágenes radiográficas deben ser preparadas para ser utilizadas como entrada en el modelo. Por lo que a las imágenes radiograficas se ajustan a un tamaño uniforme, por ejemplo, 224x224 píxeles, para facilitar su procesamiento y reducir la complejidad computacional. Para simplificar la representación de los datos, las imágenes se convierten a escala de grises, reduciendo los canales de color de tres (RGB) a uno. Las imágenes bidimensionales (224x224 píxeles) se convierten en vectores unidimensionales de 50176 elementos, creando una representación numérica adecuada para ser procesada por la regresión logística.

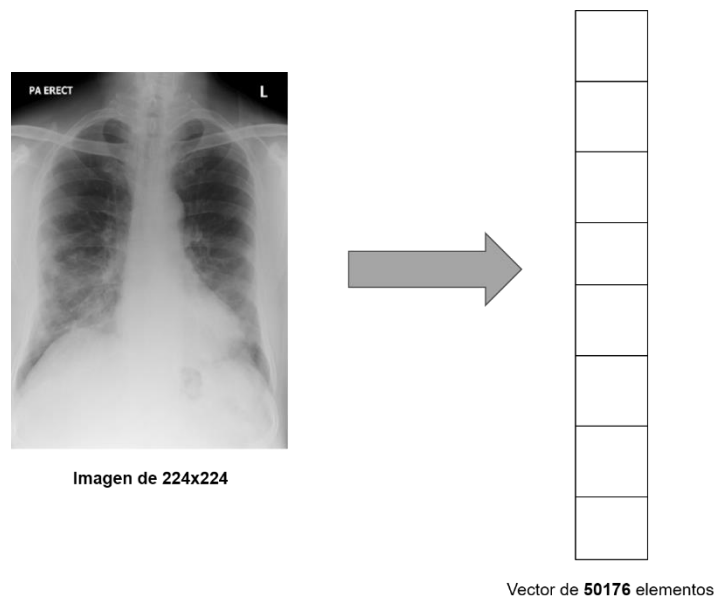


Fig 22: Imagen convertida a un vector

La regresión logística no es un modelo "logístico" en términos de transporte, sino un algoritmo basado en una función sigmoide (para problemas binarios) o una extensión llamada regresión multinomial para problemas multiclase. En este caso, se aplica la regresión logística multiclase como se muestra en la Fig 23 para clasificar entre: Sars-Cov-2, Neumonía,

normales.

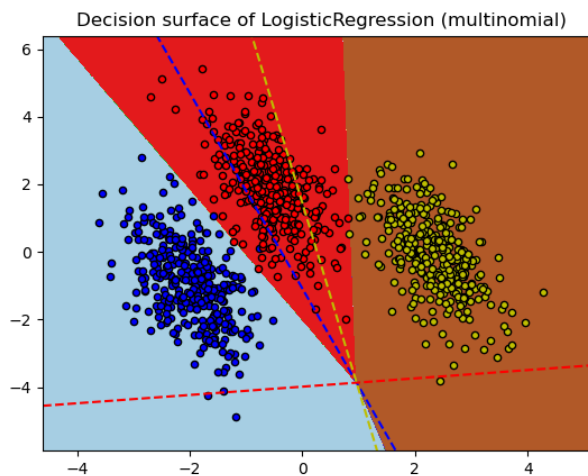


Fig 23: Regresion Multinomial

El modelo toma el vector de características (los valores de los píxeles aplanados) y calcula una puntuación para cada clase. Estas puntuaciones son transformadas en probabilidades mediante la función softmax, de manera que la suma de las probabilidades para todas las clases sea 1.

Formula softmax:

$$P(i) = \frac{e^{z_i}}{\sum_{j=1}^C e^{z_j}}$$

La ecuación básica para una clase es: $P(y = c | X) = \frac{e^{w_c^t X + b_c}}{\sum_{k=1}^C e^{w_k^t X + b_k}}$

Donde:

- X es el vector de entrada (imagen aplanada).
- w_c y b_c son los pesos y el sesgo asociados a la clase c.
- C es el número total de clases.

Durante el entrenamiento, la regresión logística utiliza un conjunto de imágenes de entrenamiento etiquetadas. El modelo ajusta sus parámetros (w y b) minimizando una función de pérdida llamada entropía cruzada. Esta métrica mide cuán bien las probabilidades predichas se alinean con las clases reales.

Después del entrenamiento, el modelo está listo para realizar predicciones sobre imágenes nuevas (de prueba). Cuando se le presenta una nueva radiografía: El modelo procesa la imagen (preprocesada y aplanada). Calcula las probabilidades para cada clase. Asigna la etiqueta de la clase con la mayor probabilidad. Por ejemplo: Si una imagen de prueba tiene una probabilidad del 85% para la clase neumonía, 10% para Sars-Cov-2, y 5% para pulmones normales, el modelo la clasifica como neumonía.

e. Árboles de decisión

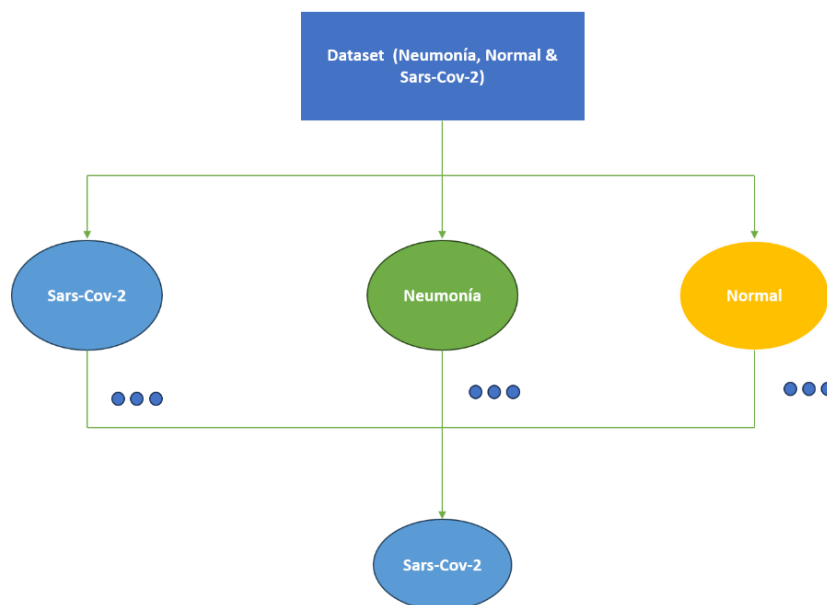


Fig 24: Representación de Árbol de decisiones. Elaboración propia

Un árbol de decisión es una estructura jerárquica que divide los datos en subconjuntos más pequeños utilizando reglas de decisión basadas en las características de los datos. Cada nodo interno del árbol representa una condición (una pregunta sobre los datos), las ramas representan los resultados de esa condición, y las hojas representan la predicción final (la clase asignada).

En el caso de las imágenes radiográficas:

Características: Los valores de los píxeles normalizados o características extraídas de la imagen.

Clases: Las categorías como pulmones normales, neumonía y Sars-Cov-2.

Para aplicar árboles de decisión a la clasificación de imágenes radiográficas, es

necesario convertir las imágenes en un formato adecuado:

Las imágenes se ajustan a un tamaño uniforme, 224x224 píxeles. Se utiliza una sola capa de color para simplificar el modelo y reducir la dimensionalidad. Los valores de los píxeles se escalan a un rango entre 0 y 1, mejorando la estabilidad del modelo.

Las imágenes bidimensionales se convierten en vectores unidimensionales, por ejemplo, una imagen de 224x224 píxeles se convierte en un vector de 50176 elementos.

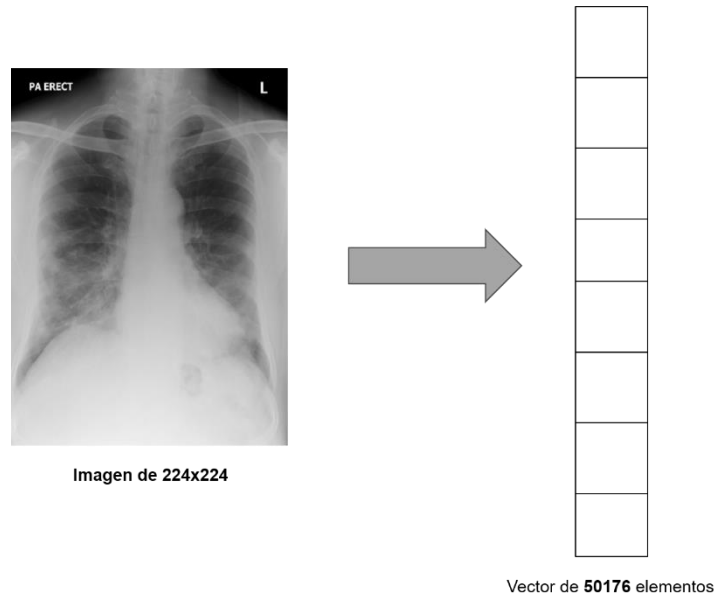


Fig 25: Imagen convertida a un vector

Selección de características: Cada nodo evalúa una característica (como el valor de un píxel o un descriptor) y establece un umbral para dividir los datos en dos subconjuntos. Un "nodo" se refiere a un punto de decisión en el árbol como se muestra en la Fig 26. Por ejemplo, "¿El valor del píxel en la posición (50, 50) es mayor a 0,5?".

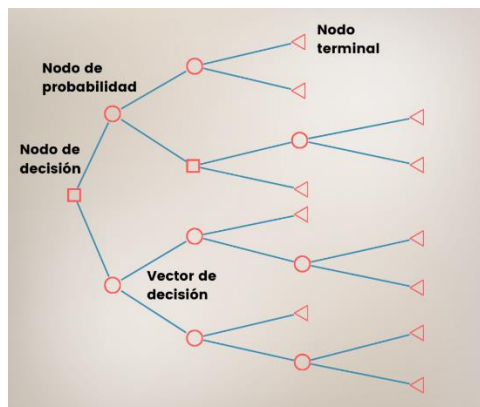


Fig 26: Representación de nodos

Se utiliza un criterio matemático para decidir cómo dividir los datos en cada nodo de un árbol de decisión. Los más comunes incluyen:

El índice de Gini, que mide la impureza de un nodo, indicando qué tan mezcladas están las clases dentro de un grupo de datos. Su fórmula es:

$$Gini = 1 - \sum_{i=1}^k p_i^2$$

El índice de Gini se utiliza para evaluar la impureza o heterogeneidad en un conjunto de datos y es ampliamente aplicado en la construcción de árboles de decisión. Por ejemplo, en la clasificación de imágenes radiográficas para detectar SARS-CoV-2, neumonía y casos normales, el índice de Gini desempeña un papel crucial en la optimización de las divisiones de datos en cada nodo del árbol.

C_1	C_2	Gini
0	6	0
1	5	0.278
2	4	0.444
3	3	0.5

Fig 27: Índice gini

En cada nodo, el árbol busca dividir los datos en dos subconjuntos que minimicen la impureza. Al evaluar diferentes características (como valores específicos de píxeles en las imágenes radiográficas) y diferentes umbrales, se calcula el índice de Gini para cada posible división. La combinación de característica y umbral que produzca la menor impureza promedio en los nodos hijos se selecciona para realizar la división.

Una vez que el árbol de decisiones ha optimizado las divisiones en cada nodo utilizando el índice de Gini, el proceso sigue estos pasos finales para clasificar una nueva imagen radiográfica: Para clasificar una nueva imagen, el árbol de decisiones comienza en el

nodo raíz y sigue las decisiones basadas en las características de la imagen. En cada nodo, se evalúa una característica específica y se decide en qué rama seguir (por ejemplo, opacidad en una región específica del pulmón). La imagen viaja a través de varios nodos internos donde se evalúan diferentes características y se toman decisiones basadas en umbrales previamente optimizados. Por ejemplo, un nodo podría evaluar si la opacidad pulmonar es mayor que un umbral específico, y dependiendo del resultado, la imagen se dirige a la siguiente rama del árbol. La imagen finalmente llega a un nodo hoja, que representa la clasificación final basada en las decisiones anteriores. El nodo hoja indica la clase a la que pertenece la imagen (Sars-Cov-2, neumonía o normal). La clasificación final de la imagen radiográfica permite a los profesionales médicos interpretar el diagnóstico basándose en la clase predicha. Por ejemplo, si una imagen se clasifica como Sars-Cov-2 con alta probabilidad, esto sugiere que las características de la imagen son consistentes con casos conocidos de Sars-Cov-2.

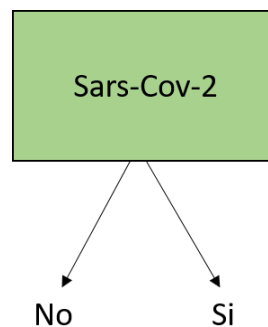


Fig 28: Probabilidad que sea Sars-Cov-2

Para la implementación de los algoritmos de clasificación seleccionados previamente, se optó por utilizar la plataforma Google Colab para la ejecución del código en la nube. Este entorno ofrece colaboración en tiempo real y acceso gratuito a recursos de GPU, lo cual fue crucial para acelerar el entrenamiento de los algoritmos de machine learning. La elección de Google Colab facilitó una implementación eficiente de algoritmos como ResNet50, KNN, Árboles de Decisiones y Regresión Logística, permitiendo una gestión sencilla del código y la integración de bibliotecas como TensorFlow.

Se determinó que Python es el lenguaje de programación más adecuado para el proyecto, debido a su compatibilidad con Google Colab y su capacidad para ejecutarse de

forma gratuita en GPUs, lo cual fue crucial para el procesamiento eficiente de imágenes radiográficas. Como entorno de desarrollo (IDE), se eligió Visual Studio Code, basándose en su versatilidad y en las herramientas específicas disponibles para el aprendizaje automático. Estas incluyen funciones avanzadas de depuración y extensiones que simplifican la manipulación de bibliotecas y la visualización de datos, optimizando así el desarrollo y la implementación de los algoritmos seleccionados.

El código resultante fue estructurado y documentado de manera exhaustiva, lo cual facilitó la comprensión y colaboración entre los miembros del equipo. Se puso un énfasis especial en la gestión de datos, la manipulación de imágenes y la optimización de algoritmos para el procesamiento de imágenes radiográficas. Durante el entrenamiento de los modelos utilizando conjuntos de datos de imágenes radiográficas de SARS-COV-2 y neumonía, se realizaron ajustes cuidadosos de hiperparámetros para mejorar el rendimiento y la precisión. Se empleó validación cruzada y se evaluó el rendimiento de los modelos en términos de sensibilidad, especificidad y precisión. Además, se implementaron técnicas de aumento de datos para mejorar la capacidad de generalización frente a datos no vistos.

Este enfoque, utilizando Python, Google Colab y Visual Studio Code, garantizó un desarrollo eficiente y preciso de algoritmos para la clasificación de enfermedades pulmonares basados en imágenes radiográficas. La combinación meticulosa de herramientas y plataformas proporcionó un sistema confiable, aplicable en entornos clínicos para el diagnóstico temprano y preciso de enfermedades pulmonares, particularmente relevante en situaciones como la pandemia de SARS-COV-2.

De acuerdo al tema de investigación, se empleará el aprendizaje automático para comparar las técnicas de algoritmos de clasificación. En la Tabla 5 se presenta la nomenclatura de términos esenciales para la evaluación, los cuales se utilizan en las fórmulas que miden el rendimiento de los algoritmos.

TABLA V: Nomenclatura de términos

NOMENCLATURA DE TÉRMINOS

Término	Sigla
Verdaderos Positivos	VP
Verdaderos Negativos	VN
Falsos Positivos	FP
Falsos Negativos	FN

a. Precisión

Métrica utilizada en problemas de clasificación para evaluar el rendimiento de un modelo en la clasificación correcta de ejemplos positivos.

$$P = \frac{VP}{FP + VP}$$

b. Exactitud

Métrica utilizada en problemas de clasificación para evaluar el rendimiento general de un modelo al clasificar correctamente ejemplos tanto positivos como negativos.

$$E = \frac{VP + VN}{VP + VN + FP + FN}$$

c. Recall

Métrica estadística utilizada para evaluar el rendimiento de un modelo o prueba diagnóstica en la detección de casos positivos.

$$R = \frac{VP}{VP + FN}$$

d. F1-Score

Métrica de evaluación utilizada en problemas de clasificación para determinar la precisión general del modelo.

$$F = 2 * \left(\frac{precision * recall}{precision + recall} \right)$$

III. RESULTADOS Y DISCUSIÓN

3.1. Resultados

TABLA VI: Promedio de tiempo de entrenamiento

Algoritmo	Tiempo de entrenamiento
ResNet50	20590.8 segundos
CNN-LUNG	36000 segundos
Regresión logística	600 segundos
KNN	300 segundos
Árbol de decisión	600 segundos

Nota: Los datos se obtuvieron considerando las especificaciones técnicas del ordenador mencionado previamente. Fuente: Elaboración propia.

En la Tabla 6 se muestra el tiempo de ejecución de los algoritmos de aprendizaje automático evaluados varía significativamente.

La variación en los tiempos de ejecución de los algoritmos de aprendizaje automático se atribuye a la complejidad de sus arquitecturas. Al comparar modelos más complejos, como ResNet50 y CNN-LUNG, con enfoques más simples como regresión logística, KNN y árbol de decisión, se observan disparidades notables. ResNet50 y CNN-LUNG, con arquitecturas más elaboradas, exhiben tiempos de ejecución significativamente más largos, alrededor de 20,590.8 y 36,000 segundos, respectivamente. En contraste, los algoritmos más simples, como regresión logística, KNN y árbol de decisión, completan su ejecución en aproximadamente 600 segundos cada uno.

La diferencia en los tiempos radica en la complejidad de los algoritmos. Modelos como ResNet50 y CNN-LUNG presentan arquitecturas más profundas con un mayor número de capas y parámetros que requieren ajuste durante el entrenamiento. Esto implica operaciones computacionalmente intensivas, como convoluciones y pooling, especialmente en tareas de visión por computadora. Además, la elección del número de épocas durante el entrenamiento también impacta en el tiempo necesario para completar el proceso. Entrenar la CNN-LUNG con 100 épocas y ResNet con 30 añade una dimensión temporal adicional, ya que más épocas

implican más iteraciones sobre el conjunto de datos de entrenamiento.

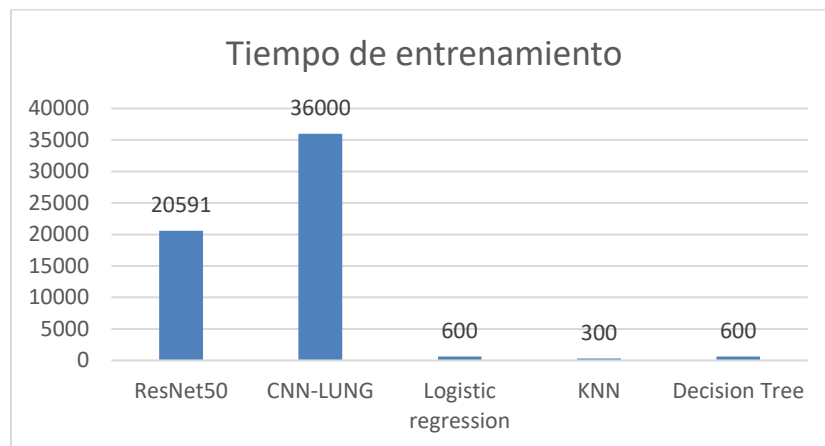


Fig 29: Tiempo de entrenamiento de los algoritmos evaluados. Elaboración Propia

TABLA VII: Tiempo de evaluación de algoritmos

Algoritmo	Tiempo de evaluación
ResNet50	81 segundos
CNN-LUNG	16.8 segundos
Regresión logística	2.3 segundos
KNN	3 segundos
Árbol de decisión	5 segundos

Nota: Los datos se obtuvieron considerando las especificaciones técnicas del ordenador mencionado previamente. Fuente: Elaboración propia.

En la Tabla 7 se presentan los tiempos promedio de evaluación de los algoritmos de aprendizaje automático, evidenciando notables diferencias en los recursos temporales requeridos por cada modelo. ResNet50 y CNN-LUNG, con arquitecturas más complejas, demandan tiempos de evaluación significativamente diferentes. ResNet50, con un tiempo de evaluación promedio de 81 segundos, muestra una eficiencia en la clasificación de enfermedades pulmonares. Por otro lado, la implementación propia de CNN-Lung demuestra un rendimiento aún más rápido, con un tiempo de evaluación promedio de 16.8 segundos, consolidándose como la opción más eficiente en términos temporales. Asimismo, los algoritmos de enfoque más simple, como regresión logística, KNN y árbol de decisión, presentan tiempos de evaluación más cortos, con 2.3, 3 y 5 segundos respectivamente. Estas disparidades reflejan la influencia de la complejidad arquitectónica en los requisitos de tiempo

de evaluación, siendo los modelos más simples inherentemente más rápidos en este aspecto.

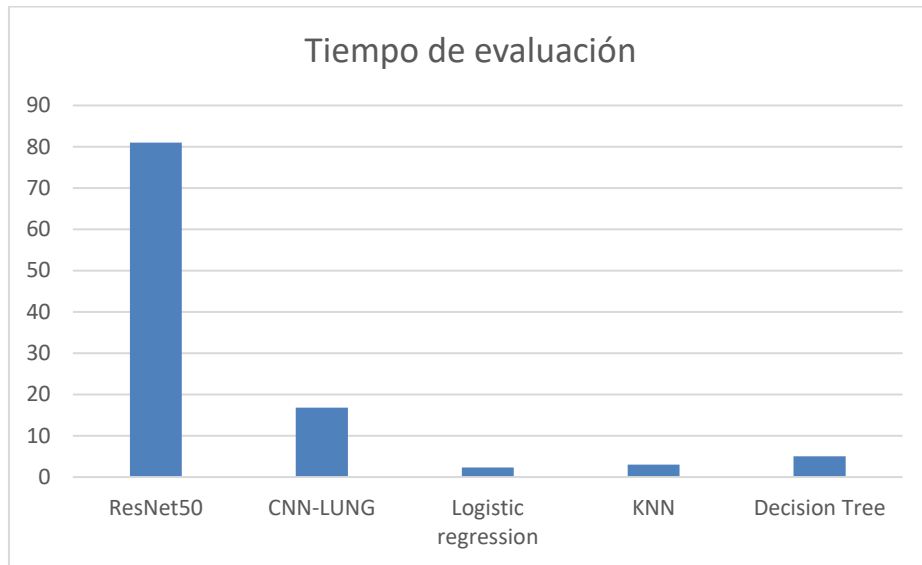


Fig 30: Tiempo de evaluación de los algoritmos evaluados. Elaboración Propia

A continuación, se presentan las matrices de confusión de cada algoritmo evaluado en el presente. Estas matrices proporcionan una visión detallada del rendimiento de cada modelo al clasificar imágenes en clases específicas, permitiendo una evaluación más profunda de su capacidad y efectividad.

3.1.1. RESNET50

En la Figura 14 se muestra la matriz de confusión del algoritmo ResNet50, la cual proporciona una visión detallada de los resultados obtenidos.

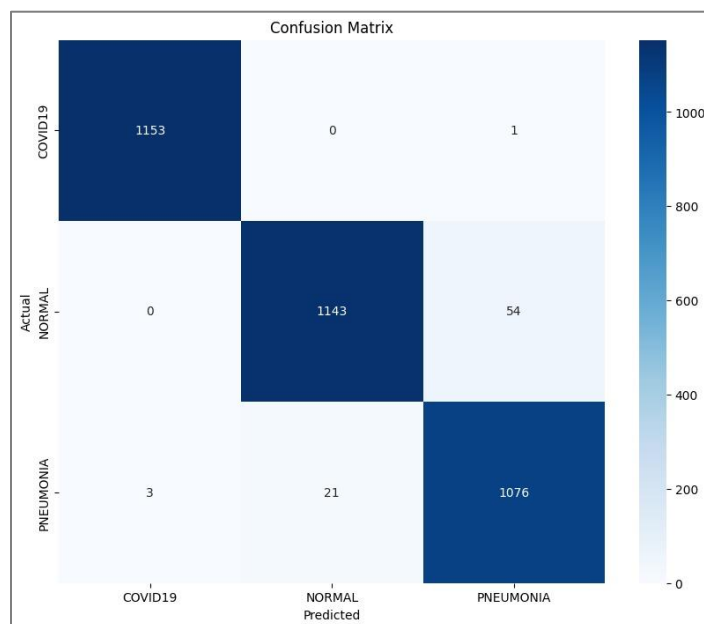


Fig 31: Matriz de confusión ResNet50. Elaboración Propia

Para el modelo ResNet50, se implementó un clasificador con el propósito de asignar imágenes a tres clases: SARS-COV-2, normal y neumonía. La matriz de confusión revela que, aunque el modelo logró identificar correctamente la mayoría de las imágenes, cometió algunos errores. Para la clase SARS-COV-2, el modelo detectó 1,153 imágenes de un total de 1,154, cometiendo un único error al clasificar 1 imagen como neumonía. En cuanto a las imágenes normales, el modelo cometió 54 errores al clasificar imágenes como neumonía, alcanzando una tasa de clasificación correcta del 98% en un conjunto total de 1,197 imágenes normales. Para las imágenes de neumonía, el modelo cometió 24 errores de un total de 1,100 imágenes, clasificando erróneamente 3 imágenes como SARS-COV-2 y 21 imágenes como normales, logrando una tasa de clasificación correcta del 95%.

3.1.2. CNN-LUNG

En la Figura 32 se muestra la matriz de confusión del algoritmo CNN- LUNG, la cual proporciona una visión detallada de los resultados obtenidos.

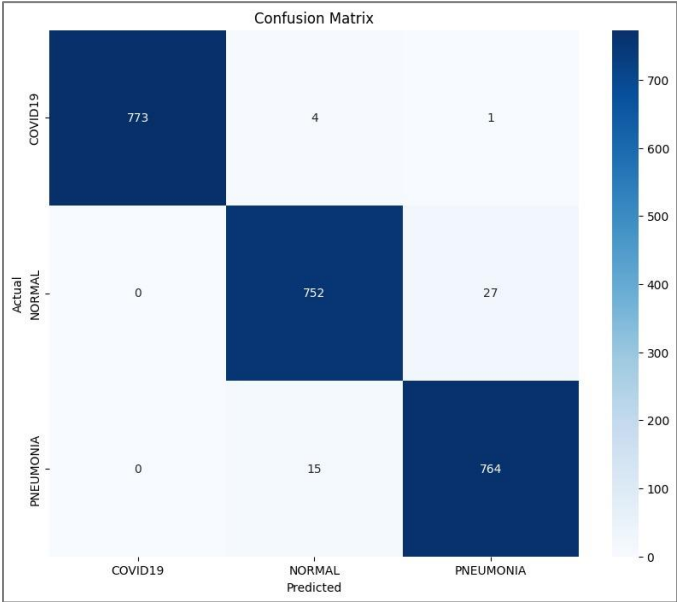


Fig 32: Matriz de confusión CNN-LUNG. Elaboración Propia

Para el algoritmo de CNN-LUNG, se utilizó un clasificador para asignar imágenes a tres clases: SARS-COV-2, normal y neumonía. La matriz de confusión revela que el modelo acertó en la identificación de 773 imágenes como SARS-COV-2, con 4 errores al clasificar

imágenes de SARS-COV-2 como normales y 1 error al clasificar como neumonía. Asimismo, se presentaron 27 errores al clasificar imágenes normales como neumonía, 0 errores al clasificarlas como SARS-COV-2, y logró clasificar correctamente 752 imágenes como normales. En cuanto a neumonía, el modelo incurrió en 15 errores al clasificar imágenes como normal, 0 errores al clasificarlas como SARS-COV-2, y clasificó correctamente 764 imágenes como neumonía.

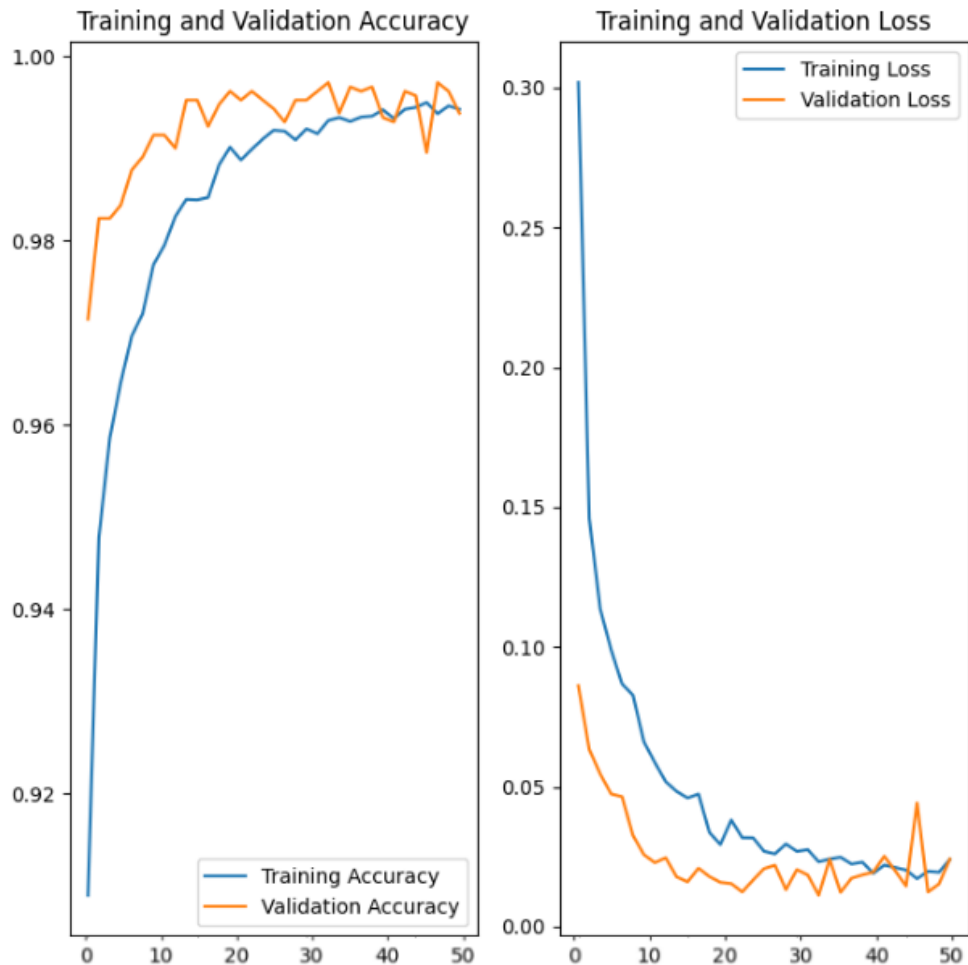


Fig 33: Grafico de entrenamiento y validación

También en el gráfico de resultados del entrenamiento y validación de nuestra red neuronal de 50 épocas, tenemos que, en la parte izquierda, se observa la evolución de la precisión en el conjunto de entrenamiento (línea azul) y en el conjunto de validación (línea naranja) a lo largo de las 50 épocas. Al inicio, la precisión de ambos conjuntos es relativamente baja, pero a medida que avanza el entrenamiento, ambas precisiones aumentan

de forma consistente, alcanzando valores cercanos a 1.0 (o 100%) hacia las últimas épocas. Esto sugiere que el modelo está aprendiendo adecuadamente y mejorando su capacidad de clasificar los datos. En el gráfico de la derecha, se presenta la pérdida del modelo en los conjuntos de entrenamiento y validación. La pérdida en el conjunto de entrenamiento (línea azul) comienza alta y disminuye rápidamente a medida que el modelo ajusta sus parámetros, estabilizándose en un valor cercano a 0.0 al final de las épocas. La pérdida en el conjunto de validación (línea naranja) también disminuye en general, aunque con algunas fluctuaciones. Sin embargo, hacia el final del entrenamiento, ambas pérdidas se mantienen en valores bajos y cercanos entre sí. El comportamiento de estos gráficos sugiere que el modelo ha sido entrenado exitosamente, alcanzando un alto nivel de precisión tanto en el conjunto de entrenamiento como en el de validación sin mostrar signos evidentes de sobreajuste. La baja pérdida y la alta precisión en ambos conjuntos reflejan que el modelo ha aprendido los patrones relevantes de nuestras imágenes de manera efectiva y tiene buen rendimiento en datos no vistos.

3.1.3. REGRESIÓN LOGÍSTICA

En la Figura 34 se muestra la matriz de confusión del algoritmo Regresión Logística, la cual proporciona una visión detallada de los resultados obtenidos.

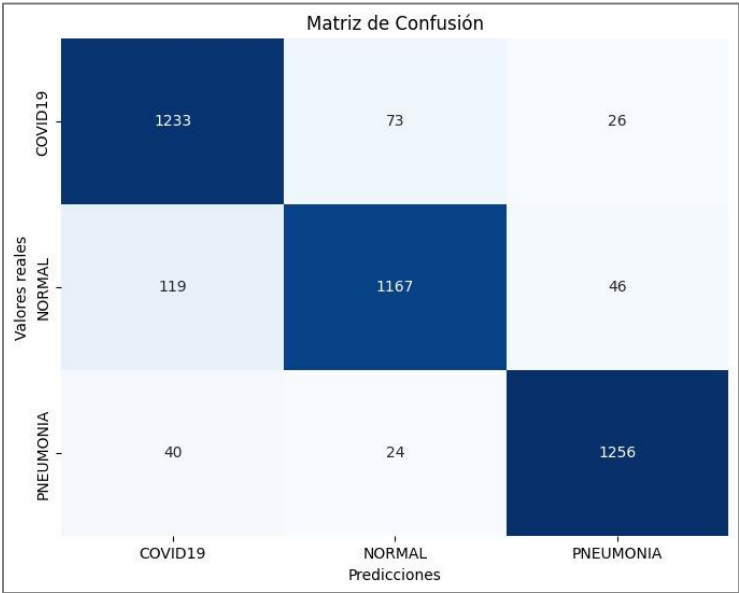


Fig 34: Matriz de confusión de Regresión Logística. Elaboración Propia

Para el algoritmo de Regresión Logística, se empleó un clasificador para asignar imágenes a tres categorías: covid19, normal y neumonía. La matriz de confusión refleja que el modelo acertó en la identificación de 1233 imágenes como covid19, cometiendo 73 errores al clasificar imágenes de covid19 como normales y 26 errores al clasificarlas como neumonía. Además, se presentaron 119 errores al clasificar imágenes de normal como covid19 y 46 errores al clasificarlas como neumonía, aunque logró clasificar correctamente 1167 imágenes como normales. En cuanto a neumonía, el modelo incurrió en 40 errores al clasificar imágenes como covid19 y 24 errores al clasificarlas como normales, pero clasificó correctamente 1256 imágenes como neumonía. Este análisis proporciona una visión detallada de cómo el clasificador de Regresión Logística categorizó las imágenes en cada clase, resaltando los resultados de verdaderos y falsos positivos y negativos. La precisión del modelo fue del 91.81%, y la exactitud alcanzó el 91.76%, con un recall del 92% y un F1-Score del 91.76%.

3.1.4. VECINOS MÁS CERCANOS (KNN)

En la Figura 35 se muestra la matriz de confusión del algoritmo KNN, la cual proporciona una visión detallada de los resultados obtenidos.

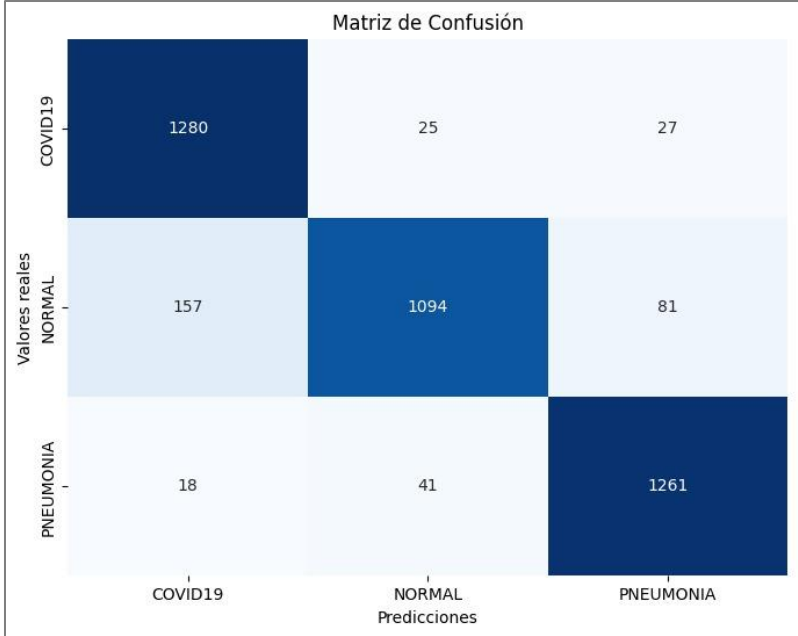


Fig 35: Matriz de confusión de KNN. Elaboración Propia

Para la implementación del algoritmo KNN se estableció k=8 mediante la experimentación con diferentes valores, buscando optimizar el rendimiento del clasificador

para las tres clases de estudio (SARS-COV-2, neumonía y normal). Esta decisión se fundamentó en múltiples criterios técnicos y experimentales: al trabajar con tres clases distintas de imágenes radiográficas, valores reducidos como $k=3$ o $k=5$ hacían el modelo más susceptible al ruido y a las variaciones locales en las características de las imágenes, mientras que valores superiores a diez tendían a difuminar las características distintivas de cada clase, comprometiendo la precisión de la clasificación. Las pruebas realizadas demostraron que la utilización de ocho vecinos proporcionó un equilibrio óptimo, permitiendo al modelo establecer patrones confiables en la clasificación sin perder la capacidad de discriminar detalles importantes. Durante el proceso de clasificación, el algoritmo evaluó las características de estos vecinos cercanos para cada nueva radiografía. Esta configuración demostró ser particularmente efectiva en casos complejos donde las imágenes presentaron características visuales similares entre clases, manteniendo su robustez frente a variaciones en los datos y valores atípicos que podrían comprometer la precisión del diagnóstico

Los resultados del clasificador muestran la distribución de predicciones entre las tres categorías establecidas: Sars-Cov-2, normal y neumonía. Al examinar la matriz de confusión (Fig. 35), se destaca que el modelo acertó en la identificación de 1280 imágenes como covid19, siendo verdaderos positivos. No obstante, se presentaron 25 errores al clasificar imágenes de covid19 como normales y 27 errores al clasificarlas como neumonía. En la categoría normal, el clasificador registró 157 falsos positivos al clasificar erróneamente imágenes como covid19, así como 81 falsos negativos al categorizar imágenes normales como neumonía. A pesar de estos errores, el modelo tuvo éxito en la clasificación de 1094 imágenes como normales. En la clasificación de neumonía, se contabilizaron 18 falsos positivos al clasificar erróneamente imágenes como covid19 y 41 falsos positivos al categorizar imágenes como normales. Sin embargo, el clasificador acertó en la clasificación de 1261 imágenes como neumonía. Estos resultados, incluidos en la Tabla VII, reflejan que la precisión del modelo KNN fue del 91.46%, con una exactitud del 91.23%, un recall del 91%, y un F1-Score del 91.23%, indicando un desempeño satisfactorio en la tarea de clasificación.

En la Fig. 36, se muestran algunas de las imágenes que el modelo clasificó

erróneamente, como aquellas de SARS-COV-2 que fueron asignadas incorrectamente a neumonía o normal, y viceversa. Estas imágenes ilustran los casos en los que el modelo tuvo dificultades para diferenciar entre clases debido a similitudes visuales o características complejas

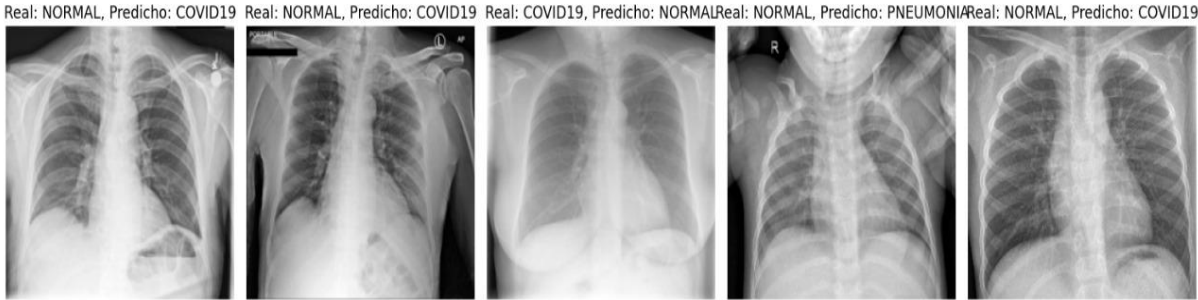


Fig 36: Imágenes con errores de clasificación: SARS-COV-2, neumonía y normal.

3.1.5. ÁRBOL DE DECISIÓN

En la Figura 37 se muestra la matriz de confusión del algoritmo Árbol de decisiones, la cual proporciona una visión detallada de los resultados obtenidos.

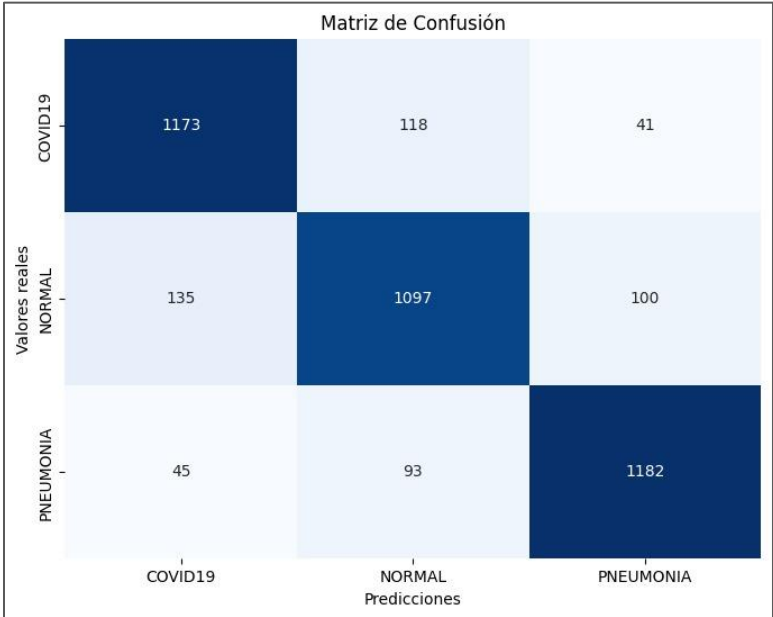


Fig 37: Matriz de confusión de Árbol de Decisiones. Elaboración Propia

Para el algoritmo de Árbol de Decisión, se evidenció que el clasificador identificó correctamente 1173 imágenes como covid19, mostrando verdaderos positivos. Sin embargo, se presentaron errores al clasificar 118 imágenes de covid19 como normales y 41 imágenes

de covid19 como neumonía, constituyendo falsos negativos para la categoría de covid19. En la categoría normal, se cometieron errores al clasificar 135 imágenes como covid19 y 100 imágenes como neumonía, correspondientes a falsos positivos y falsos negativos para la categoría normal. A pesar de estos errores, se logró clasificar correctamente 1097 imágenes como normales, siendo verdaderos negativos. En la categoría neumonía, se identificaron errores al clasificar 45 imágenes como covid19 y 93 imágenes como normales, considerados falsos positivos para neumonía. Sin embargo, se clasificaron correctamente 1182 imágenes como neumonía, siendo verdaderos negativos. Estos resultados ofrecen una visión detallada de la capacidad del clasificador de Árbol de Decisión en la tarea de clasificación de imágenes. La precisión del modelo fue del 86.44%, indicando un desempeño satisfactorio en la tarea general.

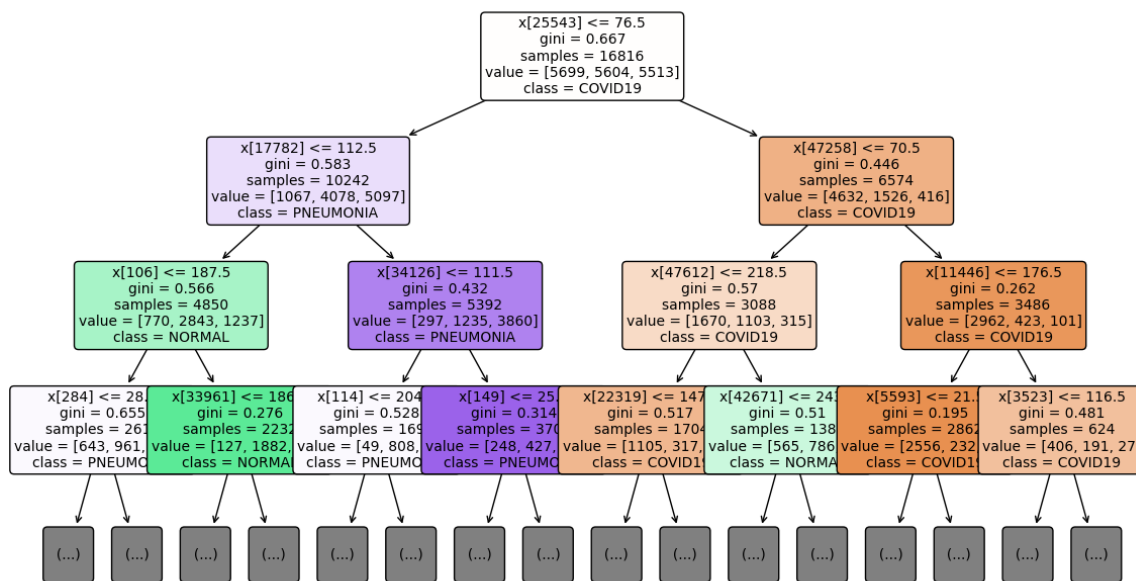


Fig 38: Estructura del árbol de decisión

Tabla VIII: Representación del árbol de Decisión

Representación del Árbol de Decisión						
Nodo	Criterio	Umbral	Gini	Samples	Clase Predominante	
Nodo Raíz	X [25543]	<= 76.5	0.667	16816	COVID19	
Hijo izquierdo	X [17782]	<= 112.5	0.583	10242	PNEUMONIA	

Hijo derecho	X [47258]	≤ 70.5	0.464	6574	COVID19
Hijo izquierdo del nodo 2	X [106]	≤ 187.5	0.565	9650	NORMAL
Hijo derecho del nodo 2	X [34126]	≤ 111.5	0.432	592	PNEUMONIA
Hijo izquierdo del nodo 3	X [47612]	≤ 76.5	0.580	5080	COVID19
Hijo derecho del nodo 3	X [11446]	≤ 176.5	0.296	1494	COVID19

El árbol de decisión presentado se organiza jerárquicamente, donde cada nodo interno representa una condición basada en una característica específica extraída de las imágenes radiográficas, indicada como $x[...]$. Estas características son valores numéricos calculados a partir de las imágenes, y en cada nodo, los datos se dividen según un umbral determinado. El índice de Gini, que se utiliza para medir la impureza de cada nodo, indica cuán equilibradas están las clases en ese punto; un valor cercano a 0 refleja mayor pureza, mientras que valores cercanos a 1 indican mayor impureza. El número de "samples" o muestras indica cuántas imágenes pasan por ese nodo y ayudan a definir la clasificación. Por último, el umbral se refiere al valor específico de la característica que se utiliza para dividir los datos en diferentes ramas. El árbol avanza a través de nodos sucesivos que buscan optimizar la homogeneidad de las clases en cada división, lo que facilita la clasificación de nuevas imágenes basadas en los datos previos.

El nodo raíz utiliza la característica $x[25543]$ con un umbral de ≤ 76.5 . Este nodo inicial contiene las 16,816 imágenes del conjunto completo de datos y tiene un índice de Gini de 0.667, lo que refleja la máxima impureza posible para un problema con tres clases, indicando una distribución balanceada de las imágenes entre las categorías: 5,699 casos de SARS-COV-2, 5,604 casos de Neumonía y 5,513 casos Normales. La clase predominante en este nodo es SARS-COV-2, debido a su leve mayoría en el conjunto de datos.

A partir del nodo raíz, el árbol se divide en dos ramas principales. La rama izquierda evalúa la característica $x[17782]$ con un umbral de ≤ 112.5 , reduciendo el índice de Gini a 0.583 con 10,242 muestras restantes. En esta división, las imágenes comienzan a agruparse hacia la clase Neumonía, que se vuelve predominante en este nodo. Por otro lado, la rama derecha evalúa la característica $x[47258]$ con un umbral de ≤ 70.5 , reduciendo el índice de

Gini a 0.464 y agrupando 6,574 muestras, donde la clase predominante es SARS-COV-2.

En niveles posteriores, cada rama continúa dividiéndose en subconjuntos más específicos. Por ejemplo, el nodo hijo izquierdo de la rama izquierda evalúa la característica $x[106]$ con un umbral de ≤ 187.5 , alcanzando un índice de Gini de 0.565 con 9,650 muestras. Este nodo agrupa principalmente imágenes clasificadas como Normales. Simultáneamente, el nodo hijo derecho de la misma rama evalúa $x[34126]$ con un umbral de ≤ 111.5 , alcanzando un índice de Gini de 0.432 con solo 592 muestras, donde Neumonía sigue siendo la clase predominante.

En la rama derecha del nodo raíz, la división comienza con la evaluación de $x[47612]$ con un umbral de ≤ 76.5 . Este nodo alcanza un índice de Gini de 0.580 con 5,080 muestras, predominando la clase SARS-COV-2. La siguiente división en esta rama evalúa $x[11446]$ con un umbral de ≤ 176.5 , reduciendo el índice de Gini a 0.296 con 1,494 muestras restantes, consolidando aún más la predominancia de la clase SARS-COV-2 en este subconjunto.

Finalmente, los nodos terminales o hojas del árbol representan las decisiones finales del modelo. Cada hoja está etiquetada con la clase más frecuente entre las muestras que alcanzan ese punto, siendo esta la predicción del modelo para cualquier dato nuevo que siga el mismo camino en el árbol. El diseño jerárquico del árbol permite realizar clasificaciones automáticas basadas en características específicas de las imágenes, optimizando la homogeneidad de las clases en cada división y asegurando una evaluación clara y precisa de las categorías médicas presentes en el conjunto de datos.

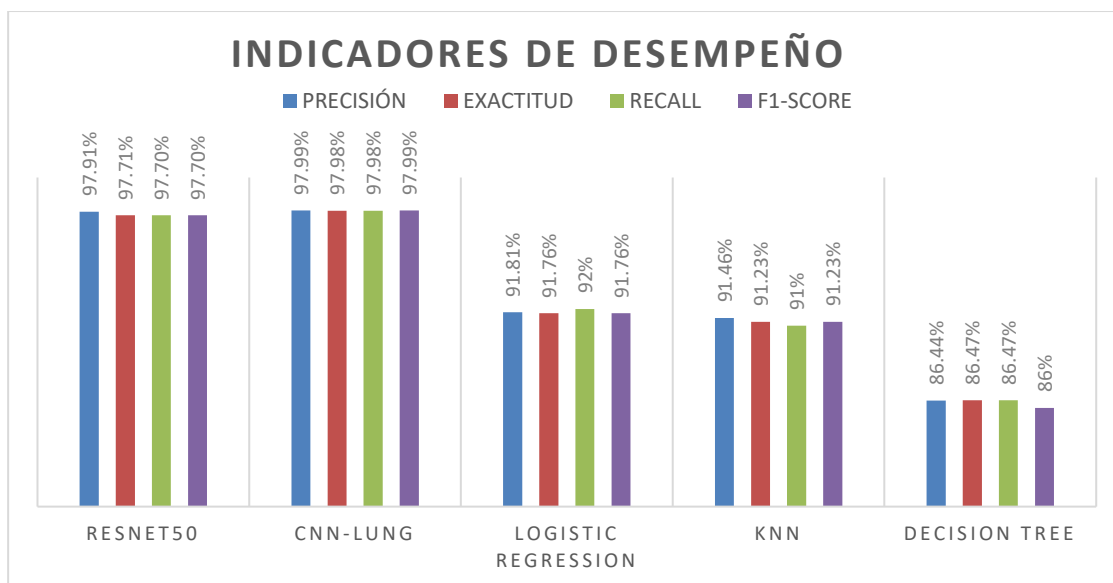


Fig 39: Indicadores de desempeño de los algoritmos. Elaboración Propia

En el Gráfico 39 se presentan la efectividad de los algoritmos plateados para la clasificación de enfermedades pulmonares, en los cuales CNN-Lung obtuvo los mejores resultados, con una precisión del 97,99%. ResNet50 obtuvo resultados muy buenos, con una precisión del 97.91%. Regresión logística, KNN y árbol de decisión obtuvieron resultados más modestos, con una precisión del 91,81%, 91,46% y 86,44%, respectivamente.

Asimismo, se exponen los resultados obtenidos en el marco de nuestra investigación, centrándonos en la evaluación de diversos algoritmos de aprendizaje automático para la clasificación de enfermedades pulmonares a partir de imágenes radiográficas. Este análisis se enfoca en presentar detalladamente las métricas de rendimiento de cada modelo. Específicamente, nos centramos en proporcionar información clave sobre la precisión, recall, F1-Score y exactitud de cada algoritmo, ofreciendo una visión exhaustiva de su desempeño

TABLA IX: Indicadores de desempeño

ALGORITMO	PRECISIÓN	EXACTITUD	RECALL	F1-SCORE
ResNet50	97.91%	97.71%	97.70%	97.70%
CNN- LUNG	97.99%	97.98%	97.98%	97.99%
Regresión Logística	91.81%	91.76%	92%	91.76%
KNN	91.46%	91.23%	91%	91.23%
Árbol de decisión	86.44%	86.47%	86.47%	86%

En la Tabla 9, se muestra los indicadores de desempeño de cada algoritmo seleccionado para llevar a cabo el entrenamiento, se implementaron distintas estrategias de partición de datos. En el caso de ResNet50 y CNN- LUNG, se optó por una división de 60% para entrenamiento, 20% para validación y 20% para pruebas. Esta meticulosa separación permitió obtener resultados altos, como evidencian las cifras de precisión, exactitud, recall y F1-score. En contraste, para los algoritmos regresión logística, KNN y Árbol de decisión, se aplicó la ley de Pareto, asignando el 80% de los datos para el entrenamiento y el restante 20% para las pruebas. Aunque estos modelos presentan cifras ligeramente inferiores en comparación con ResNet50 y CNN-Lung, sus desempeños siguen siendo notables. Particularmente, ResNet50 y CNN-Lung destacan al lograr resultados significativamente superiores en todas las métricas evaluadas. Esto se atribuye, en gran medida, a la capacidad de estas redes neuronales convolucionales de aprender jerarquías complejas de características en conjuntos de datos extensos. Su arquitectura profunda y capacidad para extraer representaciones de alto nivel les confiere una ventaja distintiva, especialmente en tareas complejas como las abordadas en este estudio. Cabe resaltar que, si bien la regresión logística, KNN y árbol de decisión exhiben cifras ligeramente más bajas, su rendimiento sigue siendo bastante satisfactorio. Estos modelos, con enfoques más tradicionales, demuestran su utilidad y eficacia en diferentes contextos

Asimismo, se proporcionan las métricas de rendimiento correspondientes a cada algoritmo, específicamente en relación con cada clase la clasificación de enfermedades pulmonares.

a. RESNET50

En la Tabla 10 se muestra los resultados obtenidos con las métricas de evaluación para el modelo ResNet50, estos datos fueron calculados considerando la matriz de confusión de la Figura 31.

TABLA X: Resultados de evaluación para ResNet50

CLASE	PRECISIÓN	EXACTITUD	RECALL	F1-SCORE
SARS-COV-2	100%	98%	100%	100%
NEUMONÍA	98%	98%	95%	97%
NORMAL	95%	98%	98%	96%

Para la clase SARS-COV-2, el modelo ResNet50 mostró un rendimiento destacado con una precisión del 100%, una exactitud del 98%, un recall del 100%, y un F1-Score del 100%. Estos resultados indican una capacidad destacada para identificar casos de SARS-COV-2 sin errores.

En relación a la neumonía, el modelo demostró una precisión del 98%, una exactitud del 98%, un recall del 95%, y un F1-Score del 97%. Esto implica que el 98% de las imágenes clasificadas como neumonía fueron correctas, y el modelo logró identificar el 95% de las imágenes reales de neumonía. El F1-Score del 97% indica un buen equilibrio entre precisión y recall.

Para las imágenes normales, el modelo mostró una precisión del 95%, indicando que el 95% de las imágenes clasificadas como normales fueron correctas. La exactitud se mantuvo en un 98%, recall fue del 98%, lo que sugiere una recuperación completa de las imágenes reales de normales. El F1-Score del 96% proporciona una medida equilibrada de precisión y recall en esta categoría.

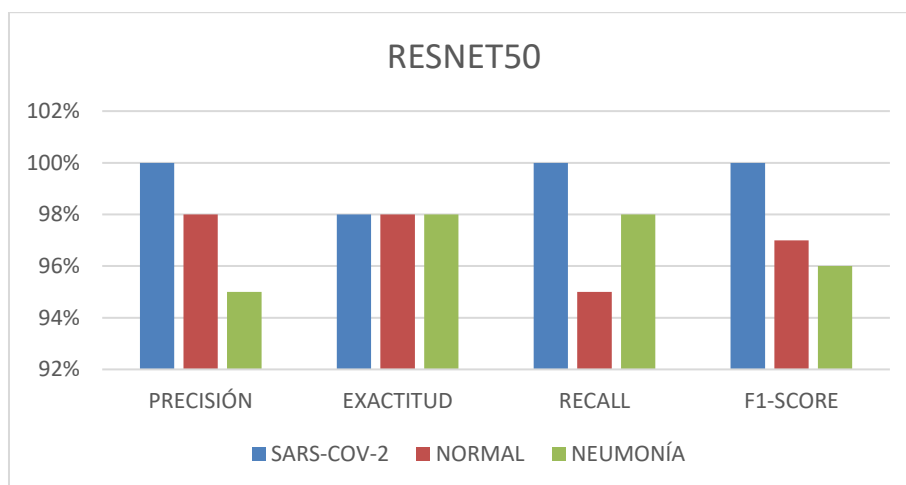


Fig 40: Resultados por métricas de ResNet50. Elaboración Propia

b. CNN-LUNG

En la Tabla 9 se muestra los resultados obtenidos con las métricas de evaluación para el modelo CNN-LUNG, estos datos fueron calculados considerando la matriz de confusión de la Figura 32.

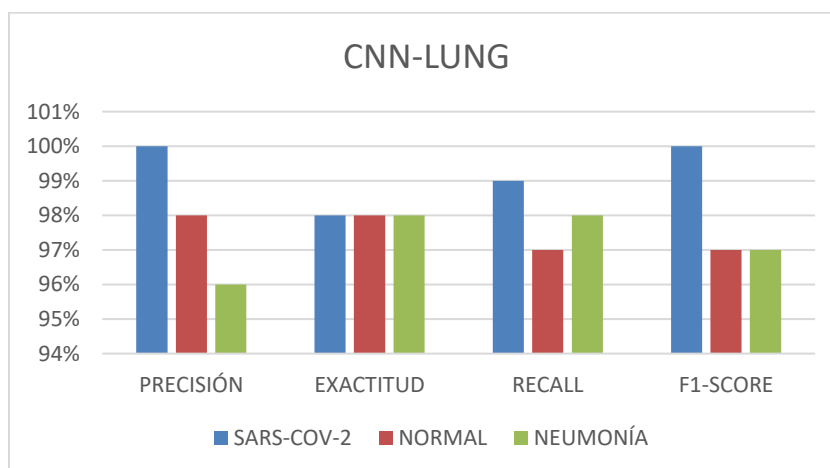
TABLA XI: Resultados de evaluación para CNN-LUNG

CLASE	PRECISIÓN	EXACTITUD	RECALL	F1-SCORE
SARS-COV-2	100%	98%	99%	100%
NEUMONÍA	98%	98%	97%	97%
NORMAL	96%	98%	98%	97%

El modelo CNN- LUNG demostró un rendimiento destacado en la clasificación de tres categorías: SARS-COV-2, neumonía e imágenes normales. En la identificación de casos de SARS-COV-2, logró una precisión del 100%, una exactitud del 98%, un recall del 99%, y un F1-Score del 100%, evidenciando una capacidad alta para esta clase.

En la detección de neumonía, el modelo mostró una precisión alta al clasificar correctamente el 98% de las 779 imágenes asociadas a esta clase. La exactitud y el recall fueron consistentes en un 98% y 97%, mientras que el F1-Score fue del 97%, demostrando la eficacia del clasificador en esta categoría.

Para las imágenes normales, el modelo mostró una precisión del 96% de 779 imágenes que fueron clasificadas como normales fueron correctas. La exactitud se mantuvo constante en un 98%, el recall fue del 98%, el F1-Score fue del 97%.



c. REGRESIÓN LOGÍSTICA

En la Tabla 12 se muestra los resultados obtenidos con las métricas de evaluación para el algoritmo Regresión logística, estos datos fueron calculados considerando la matriz de confusión de la Figura 34.

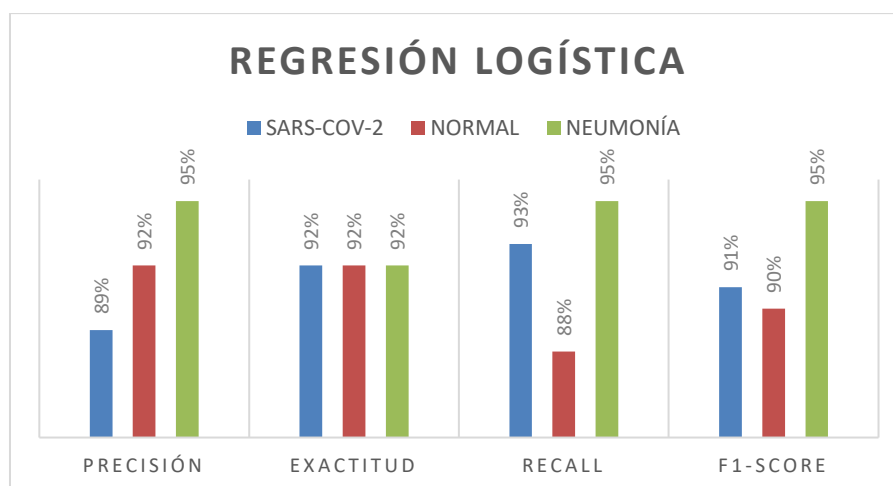
TABLA XII: Resultados de evaluación para Regresión Logística

CLASE	PRECISIÓN	EXACTITUD	RECALL	F1-SCORE
SARS-COV-2	89%	92%	93%	91%
NEUMONÍA	95%	92%	95%	95%
NORMAL	92%	92%	88%	90%

El clasificador de Regresión Logística destacó en la identificación de casos de SARS-COV-2, neumonía e imágenes normales. Para SARS-COV-2, logró una precisión del 89%, con una exactitud del 92% y un recall del 93%. El F1-Score fue del 91%.

En la detección de neumonía, el clasificador alcanzó una precisión del 95%, con una exactitud constante en 92% y un recall del 95%. El F1-Score fue del 95%, indicando eficacia en la clasificación de neumonía.

En la clasificación de imágenes normales, el clasificador mostró una precisión del 92% y una exactitud constante del 92%. Aunque el recall fue del 88%, el F1-Score fue del 90%, proporcionando una medida equilibrada de precisión y recall.



d. VECINOS MÁS CERCANOS (KNN)

En la Tabla 13 se muestra los resultados obtenidos con las métricas de evaluación para el algoritmo KNN, estos datos fueron calculados considerando la matriz de confusión de la Figura 35.

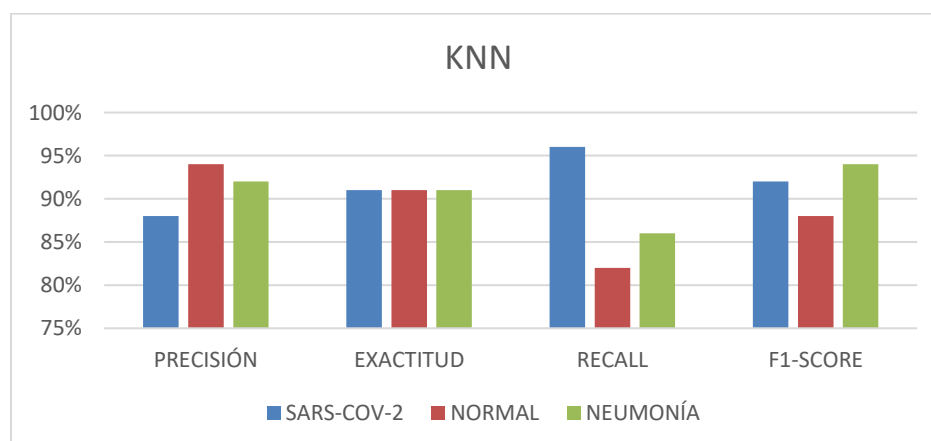
TABLA XIII: Resultados de evaluación para KNN

CLASE	PRECISIÓN	EXACTITUD	RECALL	F1-SCORE
SARS-COV-2	88%	91%	96%	92%
NEUMONÍA	92%	91%	86%	94%
NORMAL	94%	91%	82%	88%

El clasificador KNN, en la clasificación de SARS-COV-2, neumonía e imágenes normales. Para SARS-COV-2, logró una precisión del 88%, clasificando correctamente el 91% de las 1,332 imágenes, con una exactitud del 91%. El recall del 96% y el F1-Score del 92% indicaron una eficaz recuperación de imágenes reales de SARS-COV-2.

En la detección de neumonía, mostró una precisión del 92%, clasificando correctamente el 92% de las 1,320 imágenes con una exactitud del 91%. Aunque el recall fue del 86%, el F1-Score del 94% resaltó la eficacia del clasificador en esta categoría.

Para imágenes normales, presentó una precisión del 94%, identificando correctamente el 94% de las 1,332 imágenes, con una exactitud del 91%. Aunque el recall fue del 82%, el F1-Score del 88% proporcionó una medida equilibrada de precisión y recall.



e. ÁRBOL DE DECISIÓN

En la Tabla 14 se muestra los resultados obtenidos con las métricas de evaluación para el algoritmo Árbol de decisión, estos datos fueron calculados considerando la matriz de confusión de la Figura 37.

TABLA XIV: Resultados de evaluación para Árbol de Decisión

CLASE	PRECISIÓN	EXACTITUD	RECALL	F1-SCORE
SARS-COV-2	87%	86%	88%	88%
NEUMONÍA	88%	86%	89%	89%
NORMAL	84%	86%	82%	83%

En la detección de SARS-COV-2, el modelo de árbol de decisión mostró una precisión del 87%. Clasificó correctamente el 87% de las 1,332 imágenes asociadas a esta clase, con una exactitud del 86%. Además, el recall del 88% indicó una eficiente recuperación de las imágenes reales de SARS-COV-2. El F1-Score, equilibrando precisión y recall, fue del 88%, destacando la efectividad del clasificador en esta categoría.

Para la detección de neumonía, logró una precisión del 88%, clasificando correctamente el 88% de las 1,261 imágenes clasificadas como neumonía. La exactitud, mantenida en un 86%, reflejó un rendimiento confiable, mientras que el recall del 89% indicó una recuperación efectiva de las imágenes reales de neumonía. El F1-Score, evaluando precisión y recall combinados, fue del 89%, resaltando la eficacia del clasificador en esta clase.

En la clasificación de imágenes normales, el algoritmo demostró una precisión del 84%, indicando que el 84% de las 1,320 imágenes clasificadas como normales fueron correctas. Aunque el recall del 82% sugiere que algunas imágenes normales no fueron identificadas, el F1-Score del 83% proporcionó una medida equilibrada de precisión y recall. La exactitud, mantenida en un 86%, confirmó un rendimiento consistente del modelo

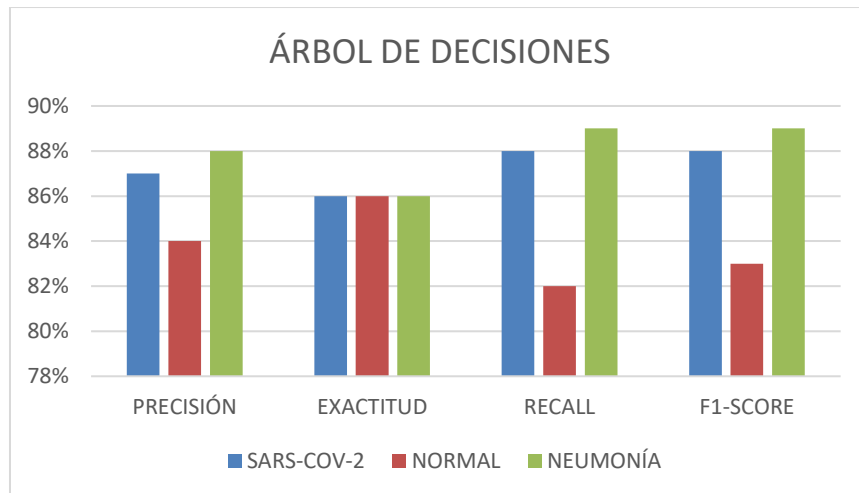


Fig 44: Resultados por métricas de Árbol de Decisión. Elaboración Propia

3.2. Discusión

Los resultados de esta investigación, enfocada en comparar algoritmos de aprendizaje automático para la clasificación de enfermedades pulmonares a partir de imágenes radiográficas, revelan diferencias significativas entre los algoritmos evaluados. Según los hallazgos, CNN- LUNG muestra una precisión del 97.99%, exactitud del 97.98%, recall del 97.98%, y F1-Score del 97.99%, superando a ResNet50, que alcanza un 97.91% de precisión y 97.70% de F1-Score. Ambos algoritmos, a su vez, superan consistentemente a la regresión logística, KNN y el árbol de decisiones en todas las métricas evaluadas. Este rendimiento respalda la hipótesis inicial de que CNN-LUNG es el algoritmo más efectivo para clasificar enfermedades pulmonares, destacando su capacidad para clasificar con mayor precisión entre diferentes clases.

El análisis realizado indica que el modelo de aprendizaje profundo CNN-LUNG se posiciona como una opción viable y eficaz para la clasificación de enfermedades pulmonares. Estos resultados subrayan la capacidad de CNN-LUNG para discernir con precisión casos de enfermedades pulmonares, lo cual podría tener implicaciones significativas en la toma de decisiones clínicas y en el avance científico. La superioridad del modelo puede atribuirse a su habilidad para aprender representaciones complejas y características discriminativas en imágenes radiográficas, aprovechando capas residuales para facilitar el entrenamiento de

redes más profundas y permitir una identificación más precisa de patrones específicos asociados con enfermedades pulmonares.

La complejidad y profundidad de la arquitectura de algoritmos como ResNet50 y CNN-LUNG, que incluyen muchas capas convolucionales, permite captar patrones complejos en las imágenes radiográficas, esenciales para el diagnóstico preciso de enfermedades pulmonares. Sin embargo, esta complejidad también incrementa el tiempo de entrenamiento y evaluación, como se refleja en los resultados. En contraste, algoritmos más simples como la regresión logística, KNN y el árbol de decisión presentan tiempos de procesamiento menores debido a su arquitectura menos compleja, pero su capacidad para capturar patrones detallados es limitada en comparación con modelos profundos.

En cuanto a la precisión, algoritmos de aprendizaje profundo como CNN-LUNG y ResNet50 demostraron ser superiores debido a su capacidad de capturar detalles complejos en las imágenes, crucial para detectar características específicas de enfermedades pulmonares. En comparación, algoritmos tradicionales como KNN y árbol de decisión, aunque útiles en tareas generales de clasificación, no están optimizados para el análisis visual detallado que requieren las radiografías. Estas limitaciones se deben a su incapacidad para captar las características complejas y jerárquicas de las imágenes médicas, lo que explica sus menores índices de precisión en este tipo de problemas. En cambio, modelos como ResNet-50 y CNN LUNG, están diseñados específicamente para el análisis de imágenes, pueden aprovechar estas características visuales detalladas para lograr un desempeño significativamente superior.

En conjunto, los resultados indican que los modelos como ResNet-50 y CNN LUNG está bien ajustado para la tarea de clasificación planteada, logrando un equilibrio adecuado entre la precisión y la capacidad de generalización. Sin embargo, se recomienda realizar una evaluación adicional utilizando un conjunto de prueba independiente para confirmar su rendimiento en datos no vistos y garantizar su robustez en escenarios reales.

En este estudio, algoritmos basados en redes neuronales convolucionales demostraron ser más efectivos en el análisis de imágenes radiográficas, mientras que los

algoritmos tradicionales, aunque menos precisos, presentan una alternativa rápida y eficiente en tareas de menor complejidad visual.

No obstante, es necesario abordar las limitaciones identificadas durante el estudio. La disponibilidad limitada de conjuntos de datos de imágenes pulmonares con anotaciones precisas podría haber afectado la generalización de los modelos. Además, la variabilidad en las condiciones de adquisición de imágenes entre diferentes instituciones médicas representa un desafío significativo que requiere atención en investigaciones y desarrollos futuros.

Al comparar nuestros resultados con investigaciones previas, como los estudios realizados por Aslan et al. [16] donde el modelo ResNet50 mostró una precisión de 98.6%, exactitud de 98.23%, recall de 98.75%, y F1-Score de 97.79%, observamos que nuestros resultados son menores estos valores. Si embargo, en nuestra investigación, utilizamos un conjunto de datos más extenso, compuesto por 23,000 imágenes, lo que contribuyó a obtener una precisión del 97.91%, exactitud del 97.71%, recall del 97.70%, y F1-Score del 97.70%. En su estudio aplicaron técnicas de selección de características INCA y IRF para mejorar la precisión de su modelo Resnet50, la amplitud de nuestro conjunto de datos destaca la superioridad de nuestro enfoque en la clasificación precisa de enfermedades pulmonares. La inclusión de un mayor número de imágenes en nuestra investigación resalta la relevancia y confiabilidad de nuestros resultados mejorados. Asimismo, Rekha et al. [64] presenta que su modelo CNN obtuvo buenos resultados, con una precisión de 97.09%, una exactitud de 95.24%, recall de 94.24% y F1-Score de 94.62%, en los cuales, el modelo evaluado en nuestra investigación, presenta resultados superiores. En el estudio se utilizó 6,432 de radiografía de tórax. Además, el conjunto de datos se dividió en 5 parte con el fin de realizar una validación cruzada, con un 20% de datos de entrenamiento utilizados para la validación, y el 80% restante para el entrenamiento del modelo.

Las implicaciones prácticas de estos hallazgos sugieren que la implementación de algoritmos de aprendizaje automático, especialmente aquellos basados en aprendizaje profundo, podría representar una mejora sustancial en la detección temprana y precisa de enfermedades pulmonares. Sin embargo, a transición de la investigación a la práctica clínica

requerirá superar desafíos adicionales, como la integración con sistemas de información médica y la validación en entornos clínicos del mundo real.

IV. CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

- El proceso de selección del conjunto de datos desempeñó un papel crucial en la calidad y generalización de los modelos. La disponibilidad de imágenes radiográficas anotadas con precisión para casos de COVID-19, neumonía y normales fue esencial para el desarrollo y evaluación de los algoritmos.
- La selección de algoritmos de clasificación abarcó una variedad de enfoques, desde modelos de aprendizaje profundo hasta algoritmos clásicos como regresión logística, k-vecinos más cercanos (KNN) y árboles de decisión. Cada algoritmo fue elegido considerando sus características y aplicabilidad para la tarea de clasificación de enfermedades pulmonares.
- La implementación de los algoritmos permitió comparar su desempeño en la clasificación de casos de SARS-COV-2, neumonía e imágenes normales. La evaluación utilizando métricas clave como precisión, exactitud, recall y F1-Score proporcionó una visión integral del rendimiento de cada modelo, cumpliendo así con el objetivo de implementar y evaluar los algoritmos de clasificación en el contexto específico de la detección de enfermedades pulmonares a partir de imágenes radiográficas.
- Los resultados revelaron que el modelo CNN-LUNG, un enfoque de aprendizaje profundo, superó significativamente a otros algoritmos, alcanzando altos niveles de precisión (97.99%), exactitud (97.98%), recall (97.98%), y F1-Score (97.99%) para cada clase. La eficacia de CNN-LUNG sugiere su potencial para la detección precisa de enfermedades pulmonares en imágenes

radiográficas.

4.2. Recomendaciones

- Dada la continua evolución en el campo del aprendizaje automático, se sugiere la exploración de técnicas adicionales para mejorar el rendimiento de los modelos. Estrategias como el aumento de datos, la optimización de hiperparámetros y la incorporación de técnicas de transferencia de aprendizaje podrían ser consideradas para impulsar aún más la precisión y robustez de los modelos, especialmente en casos donde la disponibilidad de datos es limitada.
- Explorar la combinación de modelos (ensemble methods) y técnicas de aumento de datos podría mejorar la precisión de los sistemas de clasificación en imágenes médicas en el futuro. La integración de datos clínicos con imágenes podría enriquecer los diagnósticos. Es fundamental que el personal médico reciba capacitación sobre el funcionamiento e interpretación de los algoritmos de aprendizaje automático para facilitar decisiones informadas. Antes de una posible implementación clínica, es necesario realizar pruebas exhaustivas en condiciones reales y establecer un monitoreo continuo del desempeño de los modelos.
- Los centros médicos deben ser transparentes sobre el uso de inteligencia artificial en los diagnósticos, informando a los pacientes sobre cómo puede impactar su atención. Es esencial obtener el consentimiento informado de los pacientes para el uso de sus datos, asegurando la protección de su privacidad. Además, debe garantizarse el acceso equitativo a estas tecnologías, considerando factores socioeconómicos. La responsabilidad de las decisiones médicas recaerá en los profesionales de la salud, quienes deben usar los diagnósticos generados por inteligencia artificial como herramientas complementarias. Por último, se deben implementar medidas estrictas para

proteger la privacidad de los datos de los pacientes.

REFERENCIAS

- [1] W. L. Bi *et al.*, “Artificial intelligence in cancer imaging: Clinical challenges and applications,” *CA Cancer J Clin*, Feb. 2019, doi: 10.3322/caac.21552.
- [2] Y. H. Bhosale and K. S. Patnaik, “PuIDi-COVID: Chronic obstructive pulmonary (lung) diseases with COVID-19 classification using ensemble deep convolutional neural network from chest X-ray images to minimize severity and mortality rates,” *Biomed Signal Process Control*, vol. 81, Mar. 2023, doi: 10.1016/j.bspc.2022.104445.
- [3] S. Hamida, O. El Gannour, B. Cherradi, A. Raihani, H. Moujahid, and H. Ouajji, “A novel covid-19 diagnosis support system using the stacking approach and transfer learning technique on chest X-ray images,” *J Healthc Eng*, vol. 2021, 2021, doi: 10.1155/2021/9437538.
- [4] J. Padilla, N. Espíritu, E. Rizo-Patrón, and M. C. Medina, “NEUMONÍAS EN NIÑOS EN EL PERÚ: TENDENCIAS EPIDEMIOLÓGICAS, INTERVENCIONES Y AVANCES,” *Revista Médica Clínica Las Condes*, Jan. 2017, doi: 10.1016/J.RMCLC.2017.01.007.
- [5] B. Sekeroglu and I. Ozsahin, “Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks,” *SLAS Technol*, vol. 25, no. 6, pp. 553–565, Dec. 2020, doi: 10.1177/2472630320958376.
- [6] st Yasin Yari and nd Hieu Nguyen, “Accuracy Improvement in Detection of COVID-19 in Chest Radiography,” 2021.
- [7] C. Qin, D. Yao, Y. Shi, and Z. Song, “Computer-aided detection in chest radiography based on artificial intelligence: A survey,” Aug. 22, 2018, *BioMed Central Ltd*. doi: 10.1186/s12938-018-0544-y.
- [8] D. Varshni, R. Nijhawan, K. Thakral, A. Mittal, and L. Agarwal, “Pneumonia Detection Using CNN based Feature Extraction,” 2019.
- [9] A. M. Tahir *et al.*, “Deep Learning for Reliable Classification of COVID-19, MERS, and SARS from Chest X-ray Images,” *Cognit Comput*, vol. 14, no. 5, pp. 1752–1772, Sep. 2022, doi: 10.1007/s12559-021-09955-1.
- [10] R. C. Joshi *et al.*, “A deep learning-based COVID-19 automatic diagnostic framework

- using chest X-ray images,” *Biocybern Biomed Eng*, vol. 41, no. 1, pp. 239–254, Jan. 2021, doi: 10.1016/j.bbe.2021.01.002.
- [11] L. Aversano, M. L. Bernardi, M. Cimitile, and R. Pecori, “Deep neural networks ensemble to detect COVID-19 from CT scans,” *Pattern Recognit*, vol. 120, Dec. 2021, doi: 10.1016/j.patcog.2021.108135.
- [12] S. Yao, Y. Chen, X. Tian, R. Jiang, and S. Ma, “An improved algorithm for detecting pneumonia based on YOLOv3,” *Applied Sciences (Switzerland)*, vol. 10, no. 5, Mar. 2020, doi: 10.3390/app10051818.
- [13] H. Swapnarekha, H. S. Behera, J. Nayak, and B. Naik, “Covid CT-net: A deep learning framework for COVID-19 prognosis using CT images,” *Journal of Interdisciplinary Mathematics*, vol. 24, no. 2, pp. 327–352, 2021, doi: 10.1080/09720502.2020.1857905.
- [14] J. N. Hasoon *et al.*, “COVID-19 anomaly detection and classification method based on supervised machine learning of chest X-ray images,” Dec. 01, 2021, *Elsevier B.V.* doi: 10.1016/j.rinp.2021.105045.
- [15] N. Benameur, R. Mahmoudi, S. Zaid, Y. Arous, B. Hmida, and M. H. Bedoui, “SARS-CoV-2 diagnosis using medical imaging techniques and artificial intelligence: A review,” Aug. 01, 2021, *Elsevier Inc.* doi: 10.1016/j.clinimag.2021.01.019.
- [16] N. Aslan, G. Ozmen Koca, M. A. Kobat, and S. Dogan, “Multi-classification deep CNN model for diagnosing COVID-19 using iterative neighborhood component analysis and iterative ReliefF feature selection techniques with X-ray images,” *Chemometrics and Intelligent Laboratory Systems*, vol. 224, May 2022, doi: 10.1016/j.chemolab.2022.104539.
- [17] S. Rajpal, N. Lakhyani, A. K. Singh, R. Kohli, and N. Kumar, “Using handpicked features in conjunction with ResNet-50 for improved detection of COVID-19 from chest X-ray images,” *Chaos Solitons Fractals*, vol. 145, Apr. 2021, doi: 10.1016/j.chaos.2021.110749.
- [18] E. J. Mortani Barbosa Jr *et al.*, “Machine learning automatically detects COVID-19 using chest CTs in a large multicenter cohort”, doi: 10.1007/s00330-021-07937-3/Published.

- [19] D. Das, S. K. Biswas, and S. Bandyopadhyay, "Perspective of AI system for COVID-19 detection using chest images: a review," *Multimed Tools Appl*, vol. 81, no. 15, pp. 21471–21501, Jun. 2022, doi: 10.1007/s11042-022-11913-4.
- [20] J. An, Q. Cai, Z. Qu, and Z. Gao, "COVID-19 Screening in Chest X-Ray Images Using Lung Region Priors," *IEEE J Biomed Health Inform*, vol. 25, no. 11, pp. 4119–4127, Nov. 2021, doi: 10.1109/JBHI.2021.3104629.
- [21] S. M. Rezaei, R. Abedi-Firouzjah, M. Ghorvei, and S. Sarnameh, "Screening of COVID-19 based on the extracted radiomics features from chest CT images," *J Xray Sci Technol*, vol. 29, no. 2, pp. 229–243, 2021, doi: 10.3233/XST-200831.
- [22] L. Saba *et al.*, "Six artificial intelligence paradigms for tissue characterisation and classification of non-COVID-19 pneumonia against COVID-19 pneumonia in computed tomography lungs," *Int J Comput Assist Radiol Surg*, vol. 16, no. 3, pp. 423–434, Mar. 2021, doi: 10.1007/s11548-021-02317-0.
- [23] F. Hajje, S. Ayouni, M. Hasan, and T. Abir, "Automatic Detection of Cases of COVID-19 Pneumonia from Chest X-ray Images and Deep Learning Approaches," *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/7451551.
- [24] G. S. Dlamini *et al.*, "Classification of COVID-19 and other pathogenic sequences: A dinucleotide frequency and machine learning approach," *IEEE Access*, vol. 8, pp. 195263–195273, 2020, doi: 10.1109/ACCESS.2020.3031387.
- [25] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. R. Pinheiro, "CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection," *IEEE Access*, vol. 8, pp. 91916–91923, 2020, doi: 10.1109/ACCESS.2020.2994762.
- [26] M. Alruwaili, A. Shehab, and S. Abd El-Ghany, "COVID-19 Diagnosis Using an Enhanced Inception-ResNetV2 Deep Learning Model in CXR Images," *J Healthc Eng*, vol. 2021, 2021, doi: 10.1155/2021/6658058.
- [27] B. Aksoy and O. K. M. Salman, "Detection of COVID-19 Disease in Chest X-Ray Images with capsule networks: application with cloud computing," *Journal of Experimental and*

- Theoretical Artificial Intelligence*, vol. 33, no. 3, pp. 527–541, 2021, doi: 10.1080/0952813X.2021.1908431.
- [28] A. F. Aljouie *et al.*, “Early prediction of COVID-19 ventilation requirement and mortality from routinely collected baseline chest radiographs, laboratory, and clinical data with machine learning,” *J Multidiscip Healthc*, vol. 14, pp. 2017–2033, 2021, doi: 10.2147/JMDH.S322431.
- [29] S. Abdul Gafoor, N. Sampathila, M. Madhushankara, and K. S. Swathi, “Deep learning model for detection of COVID-19 utilizing the chest X-ray images,” *Cogent Eng*, vol. 9, no. 1, 2022, doi: 10.1080/23311916.2022.2079221.
- [30] M. Attaullah *et al.*, “Initial Stage COVID-19 Detection System Based on Patients’ Symptoms and Chest X-Ray Images,” *Applied Artificial Intelligence*, vol. 36, no. 1, 2022, doi: 10.1080/08839514.2022.2055398.
- [31] A. Umar Ibrahim, P. C. Pwavodi, M. Ozsoz, F. Al-Turjman, T. Galaya, and J. J. Agbo, “Crispr biosensing and Ai driven tools for detection and prediction of Covid-19,” *Journal of Experimental and Theoretical Artificial Intelligence*, 2021, doi: 10.1080/0952813X.2021.1952652.
- [32] X. Wang *et al.*, “SSA-Net: Spatial self-attention network for COVID-19 pneumonia infection segmentation with semi-supervised few-shot learning,” *Med Image Anal*, vol. 79, Jul. 2022, doi: 10.1016/j.media.2022.102459.
- [33] D. P. Yadav, A. S. Jalal, A. Goyal, A. Mishra, K. Uprety, and N. Guragai, “COVID-19 radiograph prognosis using a deep CResNeXt network,” *Multimed Tools Appl*, 2023, doi: 10.1007/s11042-023-14960-7.
- [34] Y. Choi and H. Lee, “Interpretation of lung disease classification with light attention connected module,” *Biomed Signal Process Control*, vol. 84, Jul. 2023, doi: 10.1016/j.bspc.2023.104695.
- [35] H. C. Reis and V. Turk, “COVID-DSNet: A novel deep convolutional neural network for detection of coronavirus (SARS-CoV-2) cases from CT and Chest X-Ray images,” *Artif Intell Med*, vol. 134, Dec. 2022, doi: 10.1016/j.artmed.2022.102427.

- [36] N. Sri Kavya, T. shilpa, N. Veeranjanyulu, and D. Divya Priya, "Detecting Covid19 and pneumonia from chest X-ray images using deep convolutional neural networks," *Mater Today Proc*, vol. 64, pp. 737–743, Jan. 2022, doi: 10.1016/j.matpr.2022.05.199.
- [37] F. Xiong *et al.*, "The clinical classification of patients with COVID-19 pneumonia was predicted by Radiomics using chest CT," *Medicine*, vol. 100, no. 12, p. e25307, Mar. 2021, doi: 10.1097/MD.00000000000025307.
- [38] S. Rajaraman and S. Antani, "Weakly labeled data augmentation for deep learning: A study on COVID-19 detection in chest X-rays," *Diagnostics*, vol. 10, no. 6, Jun. 2020, doi: 10.3390/diagnostics10060358.
- [39] Y. Erdaw and E. Tachbele, "Machine learning model applied on chest X-ray images enables automatic detection of COVID-19 cases with high accuracy," *Int J Gen Med*, vol. 14, pp. 4923–4931, 2021, doi: 10.2147/IJGM.S325609.
- [40] P. Aggarwal, N. K. Mishra, B. Fatimah, P. Singh, A. Gupta, and S. D. Joshi, "COVID-19 image classification using deep learning: Advances, challenges and opportunities," May 01, 2022, *Elsevier Ltd*. doi: 10.1016/j.compbiomed.2022.105350.
- [41] R. Vijay *et al.*, "COVIDPRO-NET: a prognostic tool to detect COVID 19 patients from lung X-ray and CT images using transfer learning and Q-deformed entropy," *Journal of Experimental and Theoretical Artificial Intelligence*, 2021, doi: 10.1080/0952813X.2021.1949755.
- [42] M. K. Bohmrah and H. Kaur, "Classification of Covid-19 patients using efficient fine-tuned deep learning DenseNet model," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 476–483, Nov. 2021, doi: 10.1016/j.gltp.2021.08.003.
- [43] D. Lv *et al.*, "A Cascade-SEME network for COVID-19 detection in chest x-ray images," *Med Phys*, vol. 48, no. 5, pp. 2337–2353, May 2021, doi: 10.1002/mp.14711.
- [44] W. Jin, S. Dong, C. Dong, and X. Ye, "Hybrid ensemble model for differential diagnosis between COVID-19 and common viral pneumonia by chest X-ray radiograph," *Comput Biol Med*, vol. 131, Apr. 2021, doi: 10.1016/j.compbiomed.2021.104252.
- [45] J. de Moura, J. Novo, and M. Ortega, "Fully automatic deep convolutional approaches

- for the analysis of COVID-19 using chest X-ray images,” *Appl Soft Comput*, vol. 115, Jan. 2022, doi: 10.1016/j.asoc.2021.108190.
- [46] F. Bougourzi, R. Contino, C. Distante, and A. Taleb-Ahmed, “Recognition of COVID-19 from CT scans using two-stage deep-learning-based approach: CNR-IEMN,” *Sensors*, vol. 21, no. 17, Sep. 2021, doi: 10.3390/s21175878.
- [47] A. Bhattacharyya, D. Bhaik, S. Kumar, P. Thakur, R. Sharma, and R. B. Pachori, “A deep learning based approach for automatic detection of COVID-19 cases using chest X-ray images,” *Biomed Signal Process Control*, vol. 71, Jan. 2022, doi: 10.1016/j.bspc.2021.103182.
- [48] G. Dhiman, V. Chang, K. Kant Singh, and A. Shankar, “ADOPT: automatic deep learning and optimization-based approach for detection of novel coronavirus COVID-19 disease using X-ray images,” *J Biomol Struct Dyn*, vol. 40, no. 13, pp. 5836–5847, 2022, doi: 10.1080/07391102.2021.1875049.
- [49] K. Sengupta and P. R. Srivastava, “Quantum algorithm for quicker clinical prognostic analysis: an application and experimental study using CT scan images of COVID-19 patients,” *BMC Med Inform Decis Mak*, vol. 21, no. 1, Dec. 2021, doi: 10.1186/s12911-021-01588-6.
- [50] Y. Statsenko *et al.*, “Reliability of Machine Learning in Eliminating Data Redundancy of Radiomics and Reflecting Pathophysiology in COVID-19 Pneumonia: Impact of CT Reconstruction Kernels on Accuracy,” *IEEE Access*, vol. 10, pp. 120901–120921, 2022, doi: 10.1109/ACCESS.2022.3211080.
- [51] W. Shi, L. Tong, Y. Zhu, and M. D. Wang, “COVID-19 Automatic Diagnosis with Radiographic Imaging: Explainable Attention Transfer Deep Neural Networks,” *IEEE J Biomed Health Inform*, vol. 25, no. 7, pp. 2376–2387, Jul. 2021, doi: 10.1109/JBHI.2021.3074893.
- [52] N. A. Baghdadi, A. Malki, S. F. Abdelaliem, H. Magdy Balaha, M. Badawy, and M. Elhosseini, “An automated diagnosis and classification of COVID-19 from chest CT images using a transfer learning-based convolutional neural network,” *Comput Biol*

- Med*, vol. 144, May 2022, doi: 10.1016/j.combiomed.2022.105383.
- [53] M. T. Ahemad, M. A. Hameed, and R. Vankdothu, "COVID-19 detection and classification for machine learning methods using human genomic data," *Measurement: Sensors*, vol. 24, Dec. 2022, doi: 10.1016/j.measen.2022.100537.
- [54] S. Rajpal, M. Agarwal, A. Rajpal, N. Lakhyani, A. Saggarr, and N. Kumar, "COV-ELM classifier: An extreme learning machine based identification of COVID-19 using chest X-ray images," *Intelligent Decision Technologies*, vol. 16, no. 1, pp. 193–203, 2022, doi: 10.3233/IDT-210055.
- [55] T. Chauhan, H. Palivela, and S. Tiwari, "Optimization and fine-tuning of DenseNet model for classification of COVID-19 cases in medical imaging," *International Journal of Information Management Data Insights*, vol. 1, no. 2, Nov. 2021, doi: 10.1016/j.jjime.2021.100020.
- [56] S. Gazzah, R. Bayi, S. Kaloun, and O. Bencharef, "A deep learning to distinguish COVID-19 from others pneumonia cases," *Intelligent Automation and Soft Computing*, vol. 31, no. 2, pp. 677–692, 2022, doi: 10.32604/iasc.2022.019360.
- [57] A. Delli Pizzi *et al.*, "Radiomics-based machine learning differentiates 'ground-glass' opacities due to COVID-19 from acute non-COVID-19 lung disease," *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-96755-0.
- [58] D. N. Vinod, B. R. Jeyavadhanam, A. M. Zungeru, and S. R. S. Prabakaran, "Fully automated unified prognosis of Covid-19 chest X-ray/CT scan images using Deep Covix-Net model," *Comput Biol Med*, vol. 136, Sep. 2021, doi: 10.1016/j.combiomed.2021.104729.
- [59] R. M. Pereira, D. Bertolini, L. O. Teixeira, C. N. Silla, and Y. M. G. Costa, "COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios," *Comput Methods Programs Biomed*, vol. 194, Oct. 2020, doi: 10.1016/j.cmpb.2020.105532.
- [60] S. Singh, "Prediction of Depth Maps using Semi-Supervised Learning." [Online]. Available: <https://www.researchgate.net/publication/341671113>

- [61] R. Bhindi, M. Guan, Y. Zhao, K. H. Humphries, and G. B. J. Mancini, "Coronary atheroma regression and adverse cardiac events: A systematic review and meta-regression analysis," May 01, 2019, *Elsevier Ireland Ltd.* doi: 10.1016/j.atherosclerosis.2019.03.005.
- [62] J. Vieira, R. P. Duarte, and H. C. Neto, "Knn-stuff: Knn streaming unit for fpgas," *IEEE Access*, vol. 7, pp. 170864–170877, 2019, doi: 10.1109/ACCESS.2019.2955864.
- [63] Y. Shimokochi and S. J. Shultz, "Mechanisms of noncontact anterior cruciate ligament injury," 2008, *National Athletic Trainers' Association Inc.* doi: 10.4085/1062-6050-43.4.396.
- [64] Rekha Rajagopal, "Comparative Analysis of COVID-19 X-ray Images Classification Using Convolutional Neural Network, Transfer Learning, and Machine Learning Classifiers Using Deep Features," *Pattern Recognition and Image Analysis*, vol. 31, no. 2, pp. 313–322, Apr. 2021, doi: 10.1134/S1054661821020140.

ANEXOS




Anexo 1: Acta de aprobación del asesor



ANEXO 03: ACTA DE APROBACIÓN DEL ASESOR

Yo **FORERO VARGAS MANUEL GUILLERMO**, quien suscribe como asesor designado mediante Resolución de Facultad N° **0426-2023**, del proyecto de investigación titulado **Análisis comparativo de algoritmos de machine learning para clasificar enfermedades pulmonares, haciendo uso de imágenes radiográficas**, desarrollado por el(los) estudiante(s): **Huancas Chuquipoma Hilder Rody, Rentería Arce Álvaro.**, del programa de estudios de **Ingeniería de Sistemas**, acredito haber revisado, realizado observaciones y recomendaciones pertinentes, encontrándose expedito para su revisión por parte del docente del curso.

En virtud de lo antes mencionado, firman:

Manuel Guillermo Forero Vargas	DNI: AV702661	
Huancas Chuquipoma Hilder Rody	DNI: 72509802	
Rentería Arce Álvaro	DNI: 71245286	

Pimentel, 19 de diciembre de 2023

Anexo 02: Acta de revisión de similitud de la investigación



ACTA DE REVISIÓN DE SIMILITUD DE LA INVESTIGACIÓN

Yo **Víctor Alexci Tuesta Monteza** docente del curso de **Investigación II** del Programa de Estudios de **Grado de Sistemas** y revisor de la investigación del (los) estudiante(s), **Huancas Chuquipoma Hilder Rody, Rentería Arce Álvaro**, titulada:

ANÁLISIS COMPARATIVO DE ALGORITMOS DE MACHINE LEARNING PARA CLASIFICAR ENFERMEDADES PULMONARES, HACIENDO USO DE IMÁGENES RADIOGRÁFICAS

Se deja constancia que la investigación antes indicada tiene un índice de similitud del **17%**, verificable en el reporte final del análisis de originalidad mediante el software de similitud **TURNITIN**. Por lo que se concluye que cada una de las coincidencias detectadas no constituyen plagio y cumple con lo establecido en la Directiva sobre índice de similitud de los productos académicos y de investigación en la Universidad Señor de Sipán S.A.C., aprobada mediante Resolución de Directorio N° 145-2022/PD- USS.

En virtud de lo antes mencionado, firman:

Tuesta Monteza Víctor Alexci	DNI: 42722929	
------------------------------	---------------	--

Pimentel, 26 de diciembre de 2023.

Anexo 03: Captura de porcentaje de TURNITIN

The image shows a Turnitin Draft Coach interface. The main document area contains a summary (Resumen) about lung diseases and machine learning models. The right sidebar shows the similarity score and a list of sources.

Document Title: Documento sin título

Menu: Archivo, Editar, Ver, Insertar, Formato, Herramientas, Extensiones, Ayuda

Turnitin Draft Coach: Details, Similarity, Citations, Grammar

Overall Similarity: 17%
Last checked: Dec 24 7:21 PM -05

What should I do with my score?

Source	Similarity
1 Berlin School of Busines... SUBMITTED_WORK	1%
2 Universidad Internacion... SUBMITTED_WORK	1%
3 juandomingofarnos.wor... INTERNET	1%
4 hdl.handle.net INTERNET	1%
5 Grupo IOE on 2023-11-06	10%

Resumen

Las enfermedades pulmonares, como el SARS-COV-2 y la neumonía, plantean desafíos críticos para la salud pulmonar, requiriendo identificación temprana para mejorar el diagnóstico y la atención médica. Esta investigación tiene como objetivo comparar algoritmos de aprendizaje automático en la clasificación de enfermedades pulmonares mediante imágenes radiográficas. Se recopilieron 23.000 imágenes de repositorios públicos, distribuidas en conjuntos de entrenamiento y prueba. La evaluación de modelos, incluyendo ResNet50, CNN-RA, Regresión Logística, KNN y Árbol de Decisión, reveló una destacada precisión del 97.91%, 97.99%, 91.81%, 91.46%, y 86.44%, respectivamente. La implementación en Google Colab con Python y Visual Studio Code, utilizando técnicas de aumento de datos, mejoró la generalización. Los resultados, evaluados con métricas como precisión, exactitud, recall y F1-Score, subrayan el rendimiento superior de ResNet50 y CNN-RA. Además, se destacó la eficacia de la CNN-RA en la identificación de casos de SARS-COV-2, y se resaltó la necesidad de herramientas interactivas para la interpretación de resultados en entornos clínicos. Estos hallazgos ofrecen perspectivas esenciales para el desarrollo de herramientas diagnósticas y resaltan la eficacia sobresaliente de ResNet50 y la CNN-RA en este contexto.

Palabras Clave: aprendizaje automático, aprendizaje profundo, enfermedades pulmonares, imágenes radiográficas, clasificadores.

Run New Similarity Check

View Full Report

How should I use these checks?

Anexo 04: Operacionalización de la variable

Variable de estudio	Definición conceptual	Definición operacional	Indicadores	Instrumento	Valores finales	Tipo de variable	Escala de medición
Algoritmos de clasificación	Métodos utilizados en el campo del aprendizaje automático para categorizar o etiquetar datos en diferentes clases o categorías predefinidas.	Desempeño	Tiempo de entrenamiento $t = \sum_j^n \frac{tf_j - tf_i}{n}$	Técnica: Observación. Instrumento: Ficha de Registro Electrónico	Promedio de los indicadores de desempeño	Independiente	Intervalo: - Bajo: < 1000s - Medio: 1000 - 25000s - Alto: > 25000s
Clasificación de enfermedades pulmonares	Capacidad de poder clasificar enfermedades como Sars-CoV-2 y Neumonía mediante imágenes radiográficas	Eficiencia	Precisión $P = \frac{VP}{VP + FP}$ Exactitud $E = \frac{VP + VN}{VP + FP + FN + VN}$ F1-Score $F = 2 * \frac{P * R}{P + R}$ Recall $R = \frac{VP}{VP + FN}$	Técnica: Observación. Instrumento: Ficha de Registro Electrónico	Promedio de los indicadores de eficiencia	Dependiente	Intervalo: - Bajo: 0 - 0.60 - Medio: 0.60 - 0.90 - Alto: 0.90 - 1.0