

# Redes Neuronales Convolucionales para la detección de enfermedades pulmonares mediante radiografías del torax.

Pablo Emiliano Barajas Castellanos<sup>a,1</sup>, Oscar Hernández Torres<sup>b,1</sup>, Emiliano Alí Kick Ramírez<sup>c,1</sup> y Samuel David Pérez Brambila<sup>b,1</sup>

<sup>a</sup>Centro de Enseñanza Técnica Industrial (CETI)

<sup>b</sup>Universidad de Guadalajara

<sup>c</sup>Politécnico Ing. Matute Remus

**Resumen**—Las enfermedades pulmonares representan el 13.8 % de las muertes a nivel mundial, mientras que en México alcanzan el 7.8 %. La aparición de nuevas enfermedades, como la COVID-19, ha modificado significativamente estas cifras. Un diagnóstico oportuno puede marcar la diferencia entre la vida y la muerte.

El uso de modelos de inteligencia artificial, en particular redes neuronales convolucionales, proporciona una herramienta poderosa para mejorar la fiabilidad del diagnóstico. Estos modelos, entrenados con radiografías de tórax, pueden detectar hasta cinco enfermedades pulmonares y estimar la probabilidad de padecerlas. Esta capacidad ofrece un nivel de confianza crucial tanto para los médicos como para los pacientes.

**Keywords**—Pulmones, Enfermedad, Diagnostico, Radiografía, Inteligencia Artificial, Red Neuronal Convolucional, Aprendizaje Profundo, Python.

## 1. Introducción

Las enfermedades pulmonares representan un porcentaje significativo de las muertes tanto en México como en el mundo. Según la Organización Mundial de la Salud (OMS), son responsables del 7.8 % de las muertes en México y del 13.8 % a nivel mundial.

La pandemia de COVID-19 en 2020 transformó por completo el panorama mundial, convirtiendo el diagnóstico en un factor crucial para la supervivencia. Afortunadamente, en la actualidad, la tasa de mortalidad ha disminuido gracias a la inversión en investigación para el desarrollo de vacunas.

Hoy en día, diversas enfermedades pulmonares, incluyendo la COVID-19, la neumonía bacteriana y viral, y la tuberculosis, cuentan con tratamientos eficaces, siempre que el diagnóstico sea preciso y temprano.

Las Redes Neuronales Artificiales (RNN) son una familia de técnicas de procesamiento de información inspirado por la forma de procesar información del sistema nervioso biológico, porque se inspira en el sistema nervioso de un ser vivo, tratando de emular el comportamiento del cerebro, capaz de resolver funciones altamente no lineales en un tiempo corto, porque aprenden de los datos que son difíciles de expresarlas matemáticamente.

### Modelos de IA con un enfoque medico

El desarrollo de modelos de inteligencia artificial (IA) está revolucionando por completo diversas áreas, entre ellas la medicina. La incorporación de estas herramientas tecnológicas en el ámbito de la salud no solo representa un avance, sino que en muchos casos se traduce directamente en vidas salvadas. Un diagnóstico oportuno y confiable puede marcar la diferencia entre la vida y la muerte.

El propósito de este proyecto es desarrollar un modelo de IA con la capacidad de proporcionar diagnósticos médicos precisos, basándose en el conocimiento y la experiencia acumulados por profesionales de la salud. A partir de análisis clínicos y estudios previos que permitieron a médicos emitir diagnósticos, esta IA puede replicar ese proceso, automatizándolo y optimizándolo. De esta manera, se busca mejorar la eficiencia en la labor médica y contribuir a la selección de tratamientos adecuados de forma más rápida y eficaz.

## 1.1. Enfermedades de estudio

Nuestro estudio se centrará en 5 enfermedades pulmonares, las cuales suelen azotar a la mayoría de la población con complicaciones de este tipo, las cuales son:

- La neumonía viral es una inflamación pulmonar causada por virus que infectan los alvéolos, provocando un deterioro respiratorio de gravedad variable. Afecta a personas de todas las edades, pero es especialmente peligrosa para niños pequeños, adultos mayores e individuos con sistemas inmunitarios debilitados. Puede ser originada por distintos virus, como los de la influenza, el virus sincitial respiratorio (VSR), el *citomegalovirus* (CMV) y, más recientemente, el SARS-CoV-2. Sus síntomas incluyen tos persistente, fiebre y dificultad para respirar, pudiendo evolucionar a insuficiencia respiratoria grave que requiera ventilación mecánica. Villasclaras y Bazaga, 2017.
- Neumonía bacteriana: De acuerdo a Cofré et al., 2019 es aquel proceso neumónico en que se documenta la presencia de una bacteria en el pulmón (mediante punción pulmonar), secreción bronquial o en un fluido normalmente estéril (sangre, líquido pleural) o una respuesta inmunológica contra un patógeno respiratorio (anticuerpos específicos tipo IgM o curva ascendente de IgG).
- Enfermedad infecciosa causada por el SARS-CoV-2, detectado en 2019 en Wuhan, China. Se transmite por gotas respiratorias y aerosoles. Sus síntomas van desde fiebre y tos hasta neumonía y casos críticos que pueden ser mortales. Personas asintomáticas contribuyen a su propagación. Los inmunocomprometidos y con comorbilidades tienen mayor riesgo de complicaciones graves. Según Cortés, 2020, 2020, el periodo de incubación es variable, lo que permite que individuos asintomáticos contribuyan a la propagación del virus sin ser detectados de inmediato. Aunque muchas personas presentan síntomas leves o moderados, quienes tienen sistemas inmunológicos comprometidos o padecen enfermedades preexistentes tienen un mayor riesgo de sufrir complicaciones graves.
- La tuberculosis es una infección causada por *Mycobacterium tuberculosis*, del complejo M. Tuberculosis. Puede afectar diversos órganos, aunque el pulmón es el más común. Este microorganismo es resistente al frío y la desecación, pero sensible al calor y la luz ultravioleta. Su crecimiento es lento y ciertas condiciones, como la disponibilidad de oxígeno o el pH, pueden inducir un estado de latencia, dificultando su detección y tratamiento. Bermejo et al., 2007

## 1.2. Radiografía de tórax: Una forma de poder diagnosticar el estado del sistema respiratorio.

Como se dice en Susarte et al., 2018, la radiografía de tórax, con proyecciones posteroanterior y lateral, adquirida en inspiración, es la técnica recomendada para valorar inicialmente la patología torácica en su conjunto. Aunque existen diferentes pruebas de imagen para diagnosticar el neumotórax, la radiología convencional es la técnica de elección como prueba inicial. Sin embargo, existe variabilidad en la práctica clínica habitual para valorar a un paciente con esta sospecha, desde peticiones de radiografía de tórax en inspiración (RTI) o en espiración (RTE) hasta ambas exploraciones, con la consecuente duplicación de costes, radiación y tiempo.

## 2. Marco teórico

### 2.1. Redes Neuronales Artificiales y Convolucionales

Las redes neuronales artificiales (ANN's por sus siglas en inglés) son sistemas de procesamiento computacional inspirados en el funcionamiento de sistemas nerviosos biológicos como el cerebro humano. Se componen principalmente por nodos computacionales interconectados, los cuales trabajan de manera conjunta para aprender colectivamente a partir de una capa de entrada y optimizar su salida.

Las Redes Neuronales Convolucionales (CNNs, por sus siglas en inglés) como se describe en O'Shea y Nash, 2015, son análogas a las ANNs tradicionales en el sentido de que están compuestas por neuronas que se optimizan a sí mismas a través del aprendizaje. Cada neurona sigue recibiendo una entrada y realizando una operación (como un producto escalar seguido de una función no lineal), que es la base de innumerables ANNs.

La única diferencia notable entre las CNNs y las ANNs tradicionales es que las CNNs se utilizan principalmente en el campo del reconocimiento de patrones dentro de imágenes. Esto nos permite codificar características específicas de las imágenes en la arquitectura, haciendo que la red sea más adecuada para tareas enfocadas en imágenes, al tiempo que reduce aún más los parámetros necesarios para configurar el modelo.

### 2.2. PyRadiomics

PyRadiomics es una biblioteca de Python diseñada para la extracción automática de características radiómicas a partir de imágenes médicas. Se basa en SimpleITK y numpy, que permite extraer una amplia variedad de características cuantitativas de imágenes como radiografías, tomografías (CT), resonancias magnéticas (MRI), entre otras.

- Permite extraer múltiples características automáticamente
- Compatible con imágenes en formatos médicos estándar (DICOM, NIFTI, NRRD)
- Se puede usar en flujos de trabajo de aprendizaje automático
- Totalmente personalizable mediante archivos de configuración YAML
- Compatible con frameworks de IA y análisis de datos como scikit-learn y TensorFlow

### 2.3. Selección de características

La selección de características según Kaur et al., 2021, es el proceso de identificar y eliminar las variables irrelevantes o redundantes de un conjunto de datos. Esto mejora el rendimiento del modelo y reduce el sobreajuste (overfitting), además de hacer que el modelo sea más interpretable y eficiente computacionalmente.

Abordar un gran número de imágenes médicas, especialmente radiografías de tórax, puede ser computacionalmente costoso al entrenar y evaluar modelos. Una alternativa eficiente es el uso de bibliotecas como PyRadiomics, que permite procesar imágenes y extraer características relevantes. Esto no solo optimiza el procesamiento, sino que también mejora la estabilidad y la convergencia de los modelos.

## Metodología

### 2.4. Modelo de CNN para diagnóstico de enfermedades pulmonares

Para este proyecto se ha utilizado el conjunto de datos <https://www.kaggle.com/datasets/omkarmanohardalvi/lungs-disease-dataset-4-types/data>, que contiene imágenes de rayos X de tórax. Este conjunto de datos fue preparado a partir de varios datasets combinados de manera adecuada (eliminando las imágenes duplicadas en el dataset

utilizando VisiPics). Tiene 4 tipos de enfermedades pulmonares y una carpeta de pulmones normales. El autor aumentó el conjunto de datos por un factor de 6, por lo que básicamente contiene 10,000 imágenes.

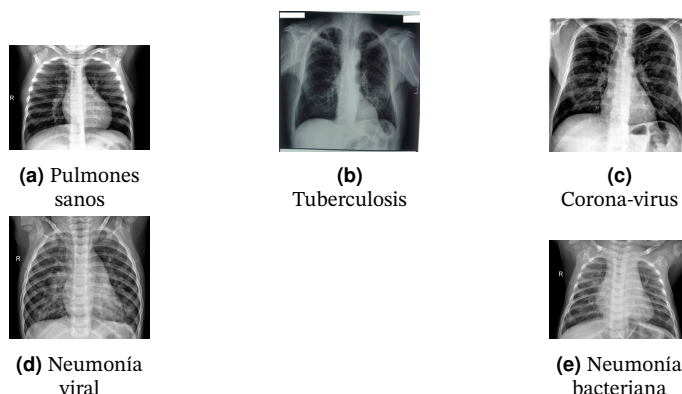


Figura 1. Comparación de radiografías pulmonares con distintas condiciones

Con el fin de crear un modelo de clasificación mediante redes neuronales convolucionales, se optó por implementar el paquete *EfficientNetB1*, una variante de la familia EfficientNet. Esta serie de modelos de redes neuronales convolucionales (CNN) está diseñada para tareas de clasificación de imágenes. EfficientNet, desarrollado por Google, es reconocido por su eficiencia y alto rendimiento en comparación con otras arquitecturas de deep learning, como ResNet o VGG.

### Información

La arquitectura de la CNN funciona así:

- EfficientNetB1 se usa como extractor de características sin su capa de clasificación (*include\_top = False*).
- Se agrega una capa GlobalAveragePooling2D para reducir la dimensionalidad.
- Se usa una capa Dense con 256 neuronas y activación ReLU para aprender patrones específicos.
- Dropout con 0.5 ayuda a prevenir sobreajuste.
- La salida tiene 5 neuronas con activación softmax, adecuada para clasificación multiclase.

A continuación se muestra de manera gráfica como funciona el modelo desarrollado:

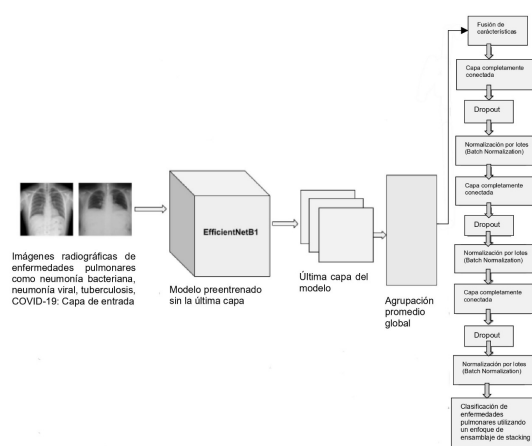


Figura 2. Pseudo-algoritmo del funcionamiento del modelo de clasificación de enfermedades mediante CNN.

El tiempo de entrenamiento varía según los alcances computacio-

nales del equipo de cada integrante de este proyecto, las métricas mejoraron significativamente y en base a ellas, se eligió el número de épocas suficientes para tener un modelo óptimo, que fue de 39 épocas.

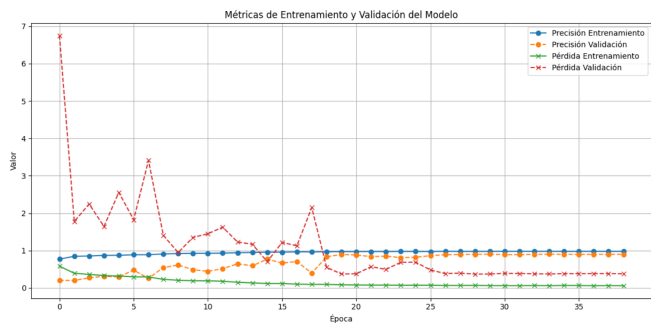


Figura 3. Métricas de validación y entrenamiento del modelo de CNN.

La siguiente tabla contiene métricas como accuracy, precisión, recall, F1-score y loss para el conjunto de entrenamiento y prueba. Donde las clases corresponden a las clasificaciones, siendo: 0 - Neumonía bacteriana, 1 - Coronavirus, 2 - Pulmones sanos, 3 - Tuberculosis, 4 - Neumonía viral.

Cuadro 1. Tabla de Resultados del Modelo

Clase	Precisión	Recall	F1-Score	Support
0	0.8648	0.7618	0.8100	403
1	0.9736	0.9951	0.9842	407
2	0.8407	0.9926	0.9103	404
3	1.0000	0.9804	0.9901	408
4	0.7719	0.7221	0.7462	403
<b>Accuracy</b>	0.8909 (Total: 2025)			
<b>Macro Avg</b>	0.8902	0.8904	0.8882	2025
<b>Weighted Avg</b>	0.8906	0.8909	0.8886	2025

Cuadro 2. Precisión, Recall, F1-Score y Support por clase

El rendimiento se puede visualizar mediante la matriz de confusión:

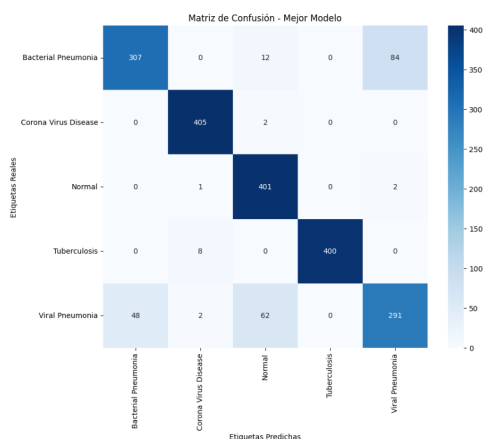


Figura 4. Matriz de confusión

Haciendo un análisis estadístico, se realiza un estudio de la curva ROC, se explica en Marcoulides, 2004 que:

- Curva más cercana a la esquina superior izquierda: Mejor rendimiento. El modelo tiene una alta tasa de verdaderos positivos y una baja tasa de falsos positivos.

- Curva diagonal: Si la curva sigue una línea diagonal de la esquina inferior izquierda a la esquina superior derecha, el modelo está clasificando al azar. El AUC sería de 0.5.
- Curva por debajo de la diagonal: Indica que el modelo está funcionando peor que el azar.

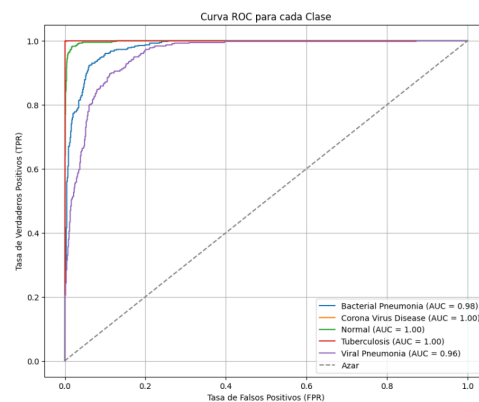


Figura 5. Curva ROC

De esta manera, podemos tener un alto nivel de confianza en nuestro modelo, asegurándonos de que no se presentarán problemas típicos como el sobreajuste (overfitting) o el subajuste (underfitting).

## 2.5. Modelo de regresión para clasificación

Se desarrolló un modelo de clasificación utilizando métodos de *machine learning* con el objetivo de comparar su eficiencia frente al desempeño obtenido con el modelo de red neuronal convolucional (NNC). Para ello, se llevó a cabo un proceso de selección de características (*feature selection*), tras el cual se seleccionaron las 10 mejores características, siguiendo los siguientes pasos:

- Prueba de Kruskal-Wallis: Para cada característica se compararon los valores entre las diferentes clases (en este caso, enfermedades) utilizando el test de Kruskal-Wallis, que genera un p-value. Un p-value bajo indica que la distribución de esa característica difiere significativamente entre las clases.
- Corrección por FDR: Dado que se realizaron múltiples pruebas, se aplicó el método de Benjamini-Hochberg para ajustar los p-values y controlar el error de descubrimiento falso (FDR).
- Selección de las 10 mejores: Se ordenaron las características según sus p-values ajustados (de menor a mayor) y se seleccionaron las 10 características con los p-values más bajos. Estas características se consideran mejores en el sentido de que presentan diferencias estadísticas significativas entre las clases, lo que sugiere que tienen el mejor potencial discriminatorio para diferenciar entre enfermedades.

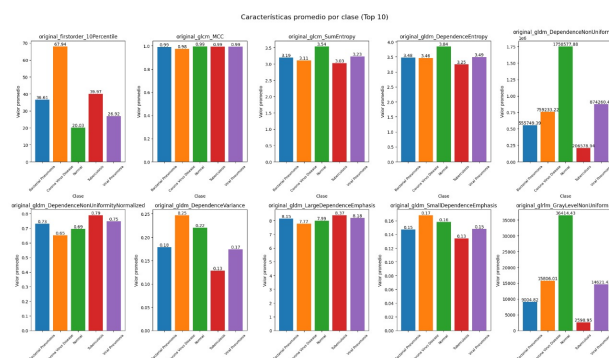


Figura 6. Diez mejores características para la clasificación de enfermedades.

Utilizando estas características, se realiza una reducción a 2 dos componentes principales (PCA), para poder graficar y visualizar el modelo de regresión logística, obteniendo los siguientes resultados:

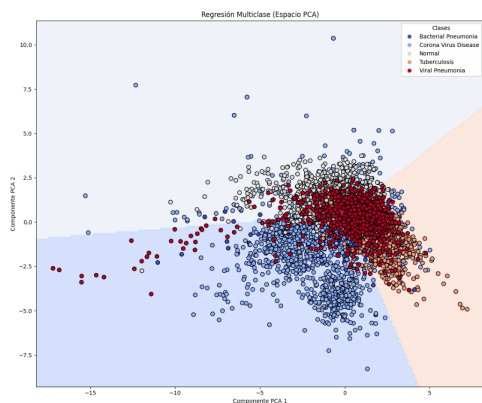


Figura 7. Análisis de componentes principales

## Resultados

### 2.6. Modelo de NNC

El modelo de NNC una vez entrenado es evaluado mediante un conjunto de prueba seleccionado al azar, cada imagen se muestra junto con la etiqueta predicha y los porcentajes de probabilidad para cada clase. Esto facilita la interpretación de las predicciones a nivel individual y permite verificar visualmente el comportamiento del modelo.

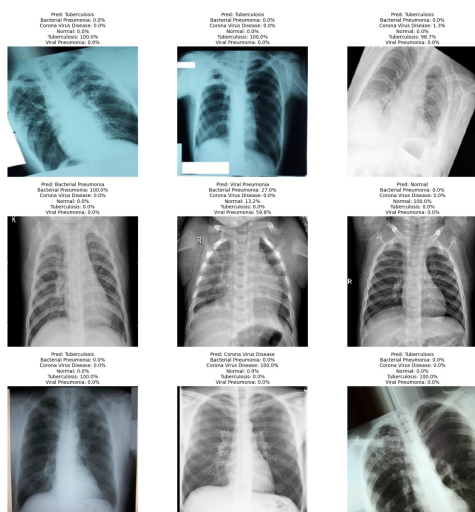


Figura 8. Validación mediante un conjunto de prueba aleatorio.

### 2.7. Modelo NNC vs modelo de regresión logística con selección de características principales.

Aunque el modelo de regresión represento un menor coste computacional y no tuvo un mal desempeño en su principal objetivo que es la clasificación logística, se debe destacar el evidente desempeño del modelo de redes neuronales, que sin duda, gana en todas las métricas de rendimiento con respecto al modelo de machine learning.

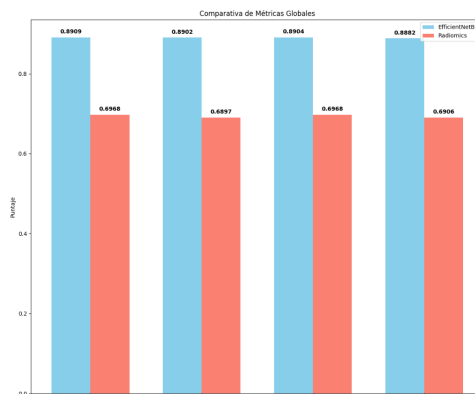


Figura 9. Modelo de redes neuronales convolucionales vs modelo de regresión logística con selección de características.

### 2.8. Aplicabilidad real del modelo de NNC

La aplicabilidad de modelos de inteligencia artificial, como el desarrollado en este proyecto, enfrenta la dificultad de ser implementados directamente debido a la falta de conocimiento en el área de programación. Es decir, un profesional de la salud difícilmente podría entender los parámetros y las estructuras matemáticas que rigen el funcionamiento de una red neuronal convolucional (NNC). Por lo tanto, se desarrolló una interfaz en Python, la cual facilita una interacción intuitiva entre el profesional y la computadora. Esta interfaz permite cargar una radiografía, realiza automáticamente la validación en el modelo desarrollado y, como resultado, proporciona un rango de probabilidad para cada enfermedad que la NNC detecte como más probable.

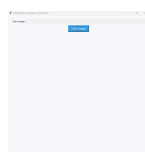


Figura 10.  
(1)Carga de radiografía



Figura 11.  
(2)Evaluación automática



Figura 12.  
(3)Rango de probabilidades de cada enfermedad

Figura 13. Interfaz para interacción medico - computadora.

## 3. Conclusiones

Como se ha podido inferir a partir de las métricas analizadas, es evidente el superior desempeño que ofrecen los modelos de aprendizaje profundo (deep learning), particularmente mediante la implementación de redes neuronales convolucionales para el diagnóstico de enfermedades pulmonares. A pesar de que su uso implica un costo computacional considerablemente mayor en comparación con métodos de machine learning tradicionales, como la regresión logística, los resultados justifican dicha inversión.

La incorporación de estas tecnologías en el ámbito de la salud está transformando la manera en que se busca optimizar el uso de los recursos disponibles. En situaciones donde está en juego la vida de una persona, es preferible contar con diagnósticos altamente confiables, aunque ello implique mayores requerimientos computacionales, antes que optar por soluciones más eficientes pero con un mayor sesgo en sus predicciones.

No obstante, no se plantea reemplazar la labor médica. Si bien los modelos de inteligencia artificial pueden generar diagnósticos con altos niveles de confianza, siempre existirá el riesgo de cometer errores. En particular, un error tipo I (falso positivo) puede tener



consecuencias clínicas importantes, pero un error tipo II (falso negativo) puede ser aún más grave, al no detectar una enfermedad real y retrasar el tratamiento oportuno.

Finalmente, el entrenamiento de estos modelos clínicos debe adaptarse a las características morfológicas y genéticas de la población local. Es decir, es fundamental que los modelos de inteligencia artificial sean entrenados con datos clínicos de personas lo más similares posible a nuestra población, con el fin de mejorar su precisión, relevancia y eficiencia diagnóstica.

Una de las principales limitaciones encontradas en este estudio fue la disponibilidad de recursos computacionales. Aunque los modelos fueron entrenados con un conjunto limitado de parámetros, el proceso requirió tiempos de ejecución considerablemente largos. La implementación de hardware de alto rendimiento permitiría explorar arquitecturas más complejas y validar múltiples configuraciones de manera más eficiente.

Se espera que, con el avance y mayor accesibilidad a tecnologías de cómputo de alto desempeño, esta área emergente —la inteligencia artificial aplicada a la salud en México— pueda consolidarse y evolucionar significativamente en los próximos años.

#### 4. Contactanos

- ✉ samuelbrambila34@gmail.com
- ✉ oscarporazul@live.com
- ✉ kickemiliano@gmail.com
- ✉ pablocast26@gmail.com

#### Referencias

- Marcoulides, G. A. (2004). Book review of *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, by Trevor Hastie, Robert Tibshirani, and Jerome Friedman; and *Generalizability Theory*, by Robert L. Brennan. *Structural Equation Modeling*, 11(1), 150-152.
- Bermejo, M., Clavera, I., Michel de la Rosa, F. J., & Marín, B. (2007). Epidemiología de la tuberculosis. *Anales del sistema sanitario de Navarra*, 30, 07-19.
- O'Shea, K., & Nash, R. (2015). An Introduction to Convolutional Neural Networks. <https://arxiv.org/abs/1511.08458>
- Villasclaras, J., & Bazaga, E. (2017). Neumonía adquirida en la comunidad. *Virus*, 12(11), 5.
- Susarte, I. C., González, A. S., & Martínez, J. P. (2018). ¿ Se debe realizar como primera prueba de imagen una radiografía de tórax en inspiración o en espiración ante la sospecha de neumotórax? *Radiología*, 60(5), 437-440.
- Cofré, J., Pavez, D., Pérez, R., & Rodríguez, J. (2019). Recomendaciones para el diagnóstico y tratamiento antimicrobiano de la neumonía bacteriana adquirida en la comunidad en pediatría. *Revista chilena de infectología*, 36(4), 505-512.
- Cortés, M. E. (2020). Coronavirus como amenaza a la salud pública. *Revista médica de Chile*, 148(1), 124-126.
- Kaur, A., Guleria, K., & Trivedi, N. K. (2021). Feature selection in machine learning: Methods and comparison. *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 789-795.