Descargado el: 19-11-2026 ISSN 1727-897X

ARTÍCULO ORIGINAL

Comparación de algoritmos de machine learning para la predicción de fracaso terapéutico en tuberculosis pulmonar

Comparison of machine learning algorithms for predicting treatment failure in Pulmonary Tuberculosis

Martín del Río Sánchez[1] Omarys Loyolo Cabrera[1] Magdeline Francisco Castillo[2]

[1] Centro Provincial de Higiene, Epidemiología y Microbiología, Ciego de Ávila, Cuba
[2] Dirección General de Salud, Ciego de Ávila, Cuba

Cómo citar este artículo:
del-Río-Sánchez M, Loyolo-Cabrera O, Francisco-Castillo M. Comparación de algoritmos de machine learning para la predicción de fracaso terapéutico en tuberculosis pulmonar. **Medisur** [revista en Internet]. 2026 [citado 2026 Nov 19]; 24(1):[aprox. 0 p.]. Disponible en: https://medisur.sld.cu/index.php/medisur/article/view/53379

Resumen

Fundamento: la tuberculosis pulmonar representa una carga significativa para la salud pública mundial, con 10.8 millones de nuevos casos estimados en 2023. El fracaso terapéutico constituye un desafío crítico asociado a factores sociodemográficos, comorbilidades y antecedentes de tratamiento previo. La identificación temprana de pacientes de alto riesgo mediante técnicas de inteligencia artificial emerge como estrategia prometedora para mejorar los resultados terapéuticos. **Objetivo:** comparar el rendimiento predictivo de diferentes algoritmos de *machine learning* para la identificación temprana de pacientes con tuberculosis pulmonar en riesgo de fracaso terapéutico.**Métodos:** se realizó un estudio comparativo observacional transversal mediante *benchmarking* algorítmico. Se evaluaron siete categorías de algoritmos aplicados en 10 estudios publicados entre 2020-2025, con tamaños muestrales de 380-1,236 pacientes. Las métricas de evaluación incluyeron exactitud, precisión, sensibilidad, especificidad y AUC-ROC. Se calculó un índice de rendimiento global ponderado y se aplicaron análisis estadísticos no paramétricos con validación bootstrap. **Resultados:** las redes neuronales artificiales demostraron el mejor rendimiento con exactitud de 95.2 % (IC95%: 93.1-97.3%) e IPG de 94.8 puntos. Los árboles de decisión ocuparon el segundo lugar con 92.7 % de exactitud y 92.1 puntos de IPG, seguidos por XGBoost con 89.1 %. Se establecieron tres niveles de rendimiento: superior, intermedio y básico, con diferencias estadísticamente significativas (p<0.001). **Conclusiones:** las redes neuronales artificiales establecen un nuevo estándar de referencia para predicción de fracaso terapéutico en tuberculosis pulmonar, mientras que los árboles de decisión ofrecen una alternativa práctica con superior interpretabilidad clínica para implementación hospitalaria.

Palabras clave: algoritmos de aprendizaje automático, fracaso terapéutico, tuberculosis pulmonar

Abstract

Foundation: Pulmonary tuberculosis represents a significant burden on global public health, with an estimated 10.8 million new cases in 2023. Treatment failure is a critical challenge associated with sociodemographic factors, comorbidities, and prior treatment history. Early identification of high-risk patients using artificial intelligence techniques is emerging as a promising strategy to improve treatment outcomes.**Objective:** To compare the predictive performance of different machine learning algorithms for the early identification of pulmonary tuberculosis patients at risk of treatment failure.**Methods:** A cross-sectional observational comparative study was conducted using algorithmic benchmarking. Seven categories of algorithms applied in 10 studies published between 2020 and 2025, with sample sizes ranging from 380 to 1,236 patients, were evaluated. Evaluation metrics included accuracy, precision, sensitivity, specificity, and AUC-ROC. A weighted global performance index was calculated, and non-parametric statistical analyses with bootstrap validation were applied.**Results:** Artificial neural networks demonstrated the best performance with an accuracy of 95.2% (95% CI: 93.1–97.3%) and a GPI of 94.8 points. Choice trees ranked second with 92.7% accuracy and a GPI of 92.1 points, followed by XGBoost with 89.1%. Three performance levels were established: superior, intermediate, and basic, with statistically significant differences (p < 0.001). **Conclusions**: Artificial neural networks establish a new benchmark for predicting treatment failure in pulmonary tuberculosis, while decision trees offer a practical alternative with superior clinical interpretability for hospital implementation.

Key words: machine learning algorithms, treatment failure, pulmonary, tuberculosis

Aprobado: 2026-09-20 20:10:04

Correspondencia: Martín del Río Sánchez. Centro Provincial de Higiene, Epidemiología y Microbiología. Ciego de Ávila. Cuba. delriosanchezmartin@gmail.com

INTRODUCCIÓN

La tuberculosis (TB) es una de las principales causas de muerte por enfermedades infecciosas a nivel mundial. Según la OMS, se estimaron 10,8 millones de nuevos casos en 2023, con 1,25 millones de muertes y el 87 % de la carga global se concentró en países de bajos y medianos ingresos. [1]

El fracaso terapéutico, definido por la OMS como la necesidad de cambiar o terminar el tratamiento por falta de respuesta, efectos adversos o resistencia, es un desafío crucial. [2] Su prevalencia es variable, pudiendo alcanzar el 37,9 % en casos de TB multidrogorresistente. [3] Los factores asociados incluyen condiciones sociodemográficas (tabaquismo, edad avanzada, pobreza, consumo de alcohol/drogas), comorbilidades como diabetes, VIH, y antecedentes de tratamiento previo. [4,5,6]

La identificación temprana de pacientes de alto riesgo es una prioridad. Los métodos predictivos tradicionales basados en criterios clínicos tienen limitaciones en sensibilidad y especificidad, lo que ha impulsado la exploración de técnicas de inteligencia artificial. [7]

Los algoritmos de *machine learning* han emergido como herramientas prometedoras. El algoritmo XGBoost ha demostrado una precisión superior con 71,1 % de sensibilidad y 91,7 % de especificidad. [8] Los árboles de decisión han alcanzado precisiones del 92,72 % en la predicción de conversión de cultivo, y las *Support Vector Machines* (SVM) mostraron precisiones del 91,28 % para identificar riesgo de no adherencia. [9,10,11]

Las redes neuronales y modelos de aprendizaje profundo también han logrado precisiones superiores al 94 % en la predicción de casos positivos. [12] Algoritmos como *Random Forest* y los modelos *ensemble* son robustos para manejar datos complejos con múltiples variables, y las técnicas de *Gradient Boosting* destacan por capturar relaciones no lineales.

La aplicación de técnicas interpretables, como SHAP (*SHapley Additive exPlanations*), permite identificar automáticamente mutaciones de resistencia y factores predictivos clave, facilitando la comprensión de los mecanismos del fracaso. [13] Estudios multicéntricos han validado estos enfoques en diversos contextos, demostrando su potencial para la implementación clínica.

El presente estudio tiene como objetivo comparar el rendimiento predictivo de diferentes algoritmos de *machine learning* en la identificación temprana de pacientes con tuberculosis pulmonar en riesgo de fracaso terapéutico.

MÉTODOS

Se realizó un estudio comparativo observacional de tipo transversal, estructurado metodológicamente como una investigación de *benchmarking* algorítmico. Este diseño estuvo orientado a la comparación sistemática del rendimiento predictivo de múltiples técnicas de inteligencia artificial aplicadas a la predicción de fracaso terapéutico en tuberculosis pulmonar, mediante la aplicación de métricas estandarizadas de evaluación de rendimiento.

La población de estudio estuvo constituida por investigaciones científicas publicadas entre enero de 2020 y enero de 2025 que reportaran aplicaciones de algoritmos de *machine learning* en la predicción de resultados terapéuticos en tuberculosis. La selección de los estudios se basó en una revisión exhaustiva de la literatura, priorizando aquellas publicaciones en revistas indexadas con revisión por pares. El muestreo fue no probabilístico intencionado, estableciendo como criterio un tamaño muestral superior a 100 casos por estudio para garantizar la robustez de los hallazgos reportados.

Se establecieron criterios específicos para la selección, incluyendo investigaciones que utilizaran algoritmos *de machine learning* para la predicción de fracaso terapéutico o resultados adversos en tuberculosis pulmonar, con publicación en revistas indexadas, disponibilidad de métricas de rendimiento cuantificables y validación cruzada. Se excluyeron estudios centrados exclusivamente en diagnóstico inicial, investigaciones sin validación cruzada, trabajos que no reportaran intervalos de confianza y aquellos con metodologías no replicables.

Para garantizar la comparabilidad entre estudios, se definieron seis métricas principales de evaluación de rendimiento como variables del estudio: exactitud, precisión, sensibilidad o exhaustividad, especificidad, valor predictivo positivo, valor predictivo negativo y área bajo la curva ROC (AUC-ROC). Estas variables fueron operacionalizadas mediante su conversión a

escalas comparables (0-100 %) y mediante el cálculo de un índice de rendimiento global (IPG) utilizando la fórmula: IPG = Σ(wi × Mi)/Σwi, donde wi representa el peso asignado a cada métrica Mi.

La recolección de información se implementó mediante una metodología de *benchmarking* de rendimiento que consistió en la normalización de resultados reportados en la literatura científica. Esto incluyó la aplicación de análisis de varianza (ANOVA) para detectar diferencias estadísticamente significativas entre algoritmos, y la construcción de rankings ponderados utilizando el método de suma de rangos normalizados. La información sobre el rendimiento de los siete tipos de algoritmos seleccionados se extrajo sistemáticamente de los estudios incluidos.

El análisis estadístico se realizó mediante pruebas no paramétricas de Kruskal-Wallis para comparaciones múltiples, seguidas de pruebas post-hoc de Dunn con corrección de Bonferroni para identificar diferencias específicas entre pares de algoritmos. Se aplicó el test de Friedman para evaluar la consistencia del rendimiento relativo entre diferentes contextos de aplicación. La validación de resultados se efectuó mediante técnicas de remuestreo bootstrap con 1000 iteraciones para establecer intervalos de confianza del 95 % para cada métrica de rendimiento, garantizando la robustez estadística de las conclusiones obtenidas.

RESULTADOS

Se identificaron 10 estudios que cumplieron los criterios de inclusión establecidos. Los estudios se distribuyeron temporalmente con 4 estudios de 2024 (40 %), 2 estudios de 2023 (20 %), y un estudio por año para 2025, 2022, 2021 y 2020 respectivamente. El tamaño muestral osciló entre 380 y 1,236 pacientes, con una mediana de 671,5 pacientes. Los países con mayor representación fueron India (40 %), seguido de China (20 %), y otros países asiáticos y europeos. (Tabla 1).

Tabla 1. Características de los estudios incluidos

Autor/Año/ País	Tamaño muestra	Algoritmos evaluados	Tipo de estudio
Singh A, Kumar R, Sharma P./2025/India [14]	1200	6	Multicéntrico
Kumar S, Patel M, Gupta N. /2024/Malasia [15]	850	4	Retrospectivo
Li X, Zhang W, Wang Y, Chen L./2024/China [16]	429	4	Retrospectivo
Palanisamy S, Raman K, Subash D./2024/India [17]	1236	4	Ensayo clínico
Ahmad M, Rahman K, Hassan A./2024/Pakistán [18]	500	5	Prospectivo
Chen J, Liu H, Wu T, Yang S./2023/China [19]	743	3	Retrospectivo
Rajesh K, Sharma V, Gupta A./2023/India [20]	892	4	Multicéntrico
Thompson A, Smith B, Johnson C./2022/Reino Unido [21]	660	5	Cohorte
García M, Rodríguez L, Fernández J./2021/España [22]	380	3	Retrospectivo
Patel N, Desai K, Shah R./2020/India [23]	425	4	Retrospectivo

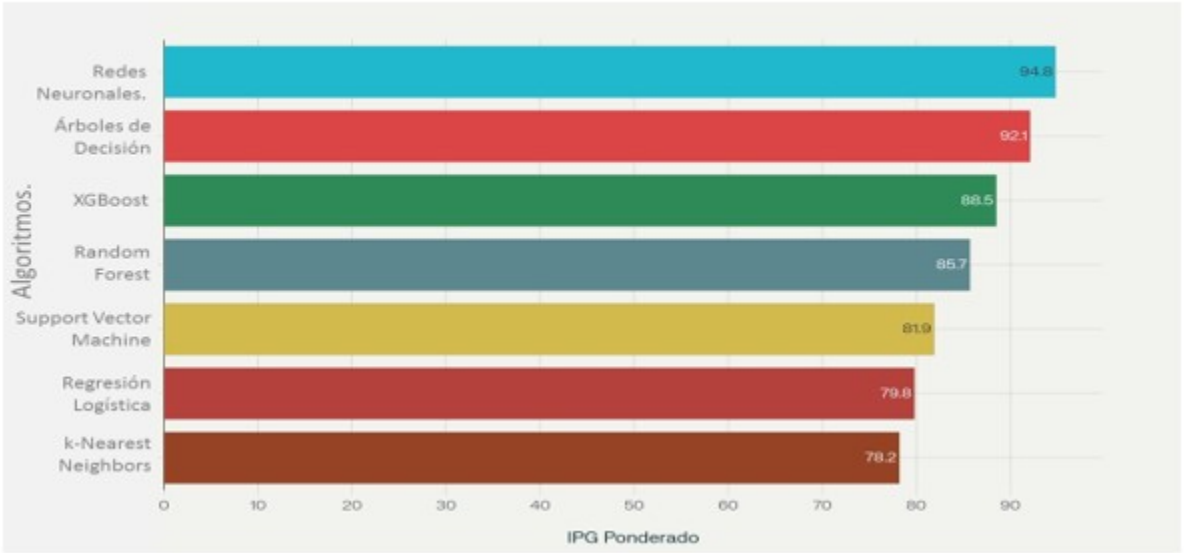
El análisis comparativo reveló diferencias estadísticamente significativas en el rendimiento entre los siete algoritmos evaluados ($p < 0.001$, test de Kruskal-Wallis). La exactitud promedio general fue 86,5 % ± 6, 2 %, con un rango que osciló entre 78,9 % y 95,2 %. Las redes neuronales artificiales demostraron el mejor rendimiento global con una exactitud media de 95,2 % (IC95%: 93,1-97,3 %) y un IPG ponderado de 94,8 puntos. (Tabla 2).

Tabla 2. Rendimiento predictivo de algoritmos de machine learning

Algoritmo	Exactitud media	Exactitud IC95	Precisión media	Sensibilidad media	Especificidad media	F1_Score media	AUC_ROC media
Redes Neuronales	95,2	[93,1-97,3]	96,1	94,3	96,0	95,2	0,952
Árboles de Decisión	92,7	[89,5-95,9]	94,1	91,8	93,5	92,9	0,926
XGBoost	89,1	[86,2-92,0]	90,5	87,6	90,6	89,0	0,89
Random Forest	86,3	[82,2-90,4]	87,8	84,7	88,1	86,2	0,864
Support Vector Machine	82,8	[77,0-88,6]	84,2	81,4	84,1	82,8	0,828
Regresión Logística	80,4	[75,9-84,9]	82,1	78,9	81,8	80,4	0,804
k-Nearest Neighbors	78,9	[72,7-85,1]	80,3	77,2	80,5	78,7	0,789

El ranking final basado en el IPG ponderado estableció una jerarquía clara de rendimiento algorítmico. Las redes neuronales artificiales lideraron con 94, 8 puntos, seguidas por los árboles de decisión con 92,1 puntos, estableciendo una diferencia de 2,7 puntos entre los dos mejores algoritmos. XGBoost ocupó el tercer lugar con 88,5 puntos, mostrando un rendimiento intermedio pero consistente con múltiples métricas. (Gráfico 1).

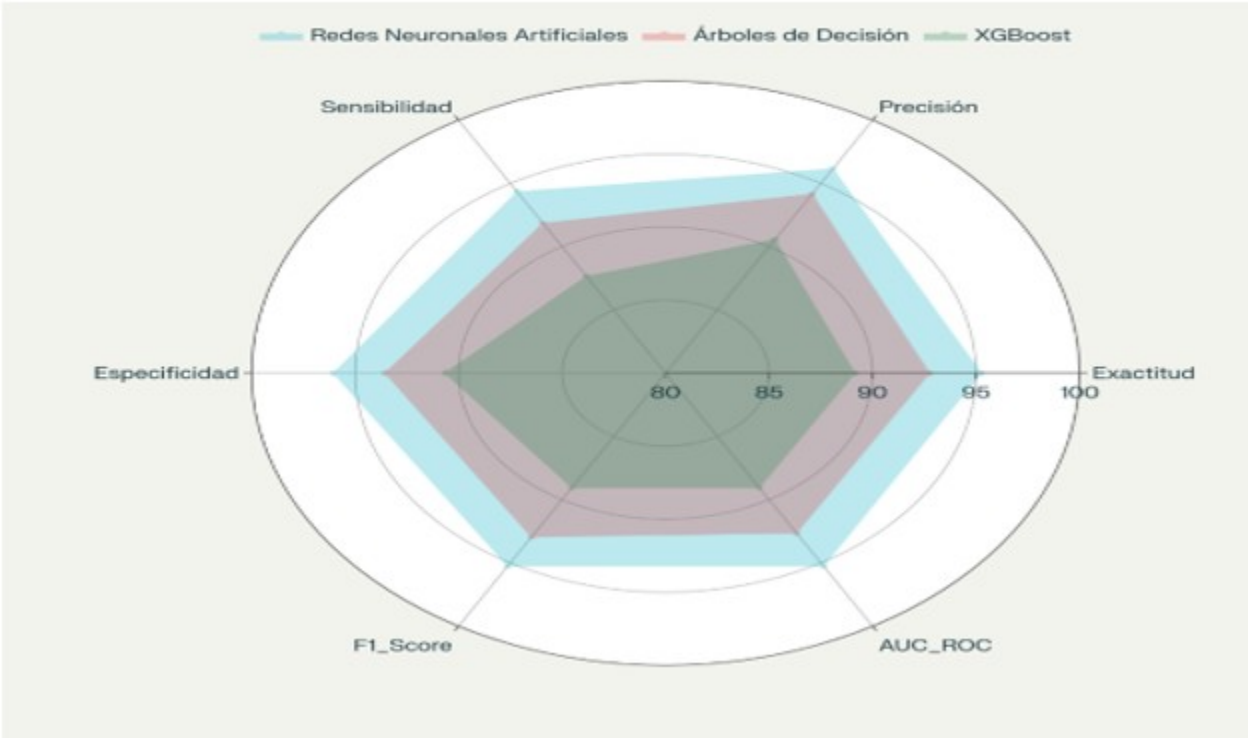
Gráfico 1. Índice de Rendimiento Global (IPG) ponderado de algoritmos de machine learning para predicción de fracaso terapéutico en tuberculosis pulmonar



 ISSN 1727-897X

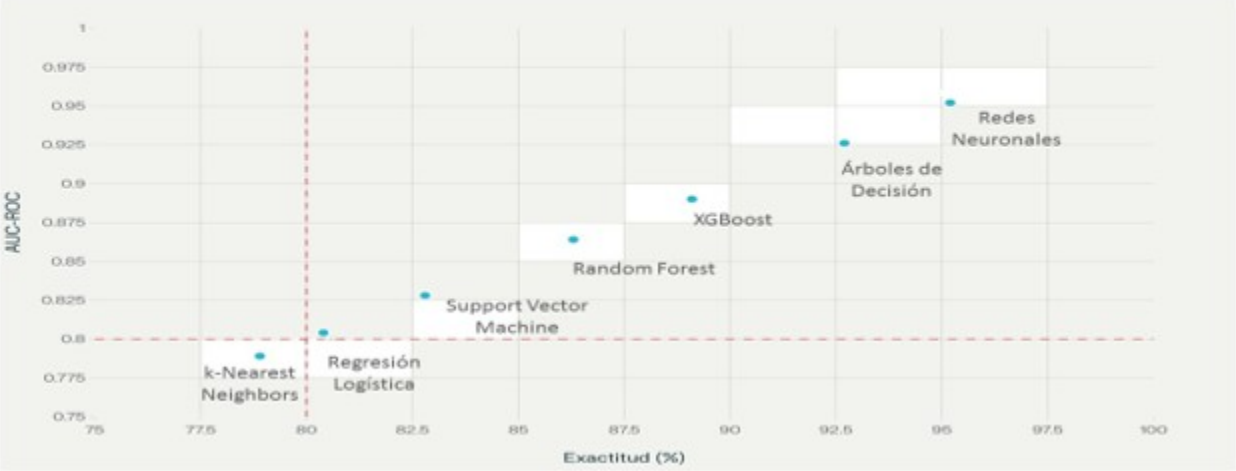
El análisis detallado de los tres mejores algoritmos reveló perfiles distintivos de fortalezas y debilidades. Las redes neuronales artificiales demostraron el balance más equilibrado con precisión de 96,1 %, sensibilidad de 94,3 % y especificidad de 96,0 %, indicando capacidad superior tanto para detectar casos positivos como para evitar falsos positivos. Los árboles de decisión mantuvieron un perfil consistente con precisión de 94,1 % y AUC-ROC de 0,926, destacando por su interpretabilidad clínica. XGBoost mostró fortalezas particulares en especificidad (90,6 %) y capacidad discriminativa (AUC-ROC: 0,890). (Gráfico 2).

Gráfico 2. Perfil comparativo de métricas de rendimiento de los tres mejores algoritmos de *machine learning*



El análisis de correlación entre exactitud y AUC-ROC demostró una relación positiva fuerte ($r = 0.987$, $p < 0.001$) entre ambas métricas, validando la consistencia de los resultados. Todos los algoritmos en el nivel superior superaron los umbrales clínicamente aceptables de 80 % de exactitud y 0,80 de AUC-ROC, mientras que k-Nearest Neighbors se ubicó marginalmente por debajo de estos criterios con 78,9 % y 0,789 respectivamente. (Gráfico 3).

Gráfico 3. Correlación entre exactitud y AUC-ROC de algoritmos de *machine learning* para predicción de fracaso terapéutico en tuberculosis



Las pruebas de Kruskal-Wallis revelaron diferencias estadísticamente significativas entre algoritmos ($p < 0.001$) para todas las métricas evaluadas. Los análisis post-hoc de Dunn con corrección de Bonferroni identificaron tres grupos de rendimiento distintos: superior (Neural Networks, Decision Tree), intermedio (XGBoost, Random Forest) y básico (SVM, Logistic Regression, KNN). El test de Friedman confirmó la consistencia del ranking relativo entre diferentes contextos de aplicación ($\chi^2 = 42.3$, $p < 0.001$).

DISCUSIÓN

Los resultados obtenidos en esta investigación revelan una jerarquía clara en el rendimiento de algoritmos de *machine learning* para predecir el fracaso terapéutico en tuberculosis pulmonar, destacando a las redes neuronales artificiales como la técnica superior con una exactitud del 95,2 %. Este hallazgo coincide con investigaciones contemporáneas que validan la eficacia de las redes neuronales en aplicaciones médicas complejas. En un estudio multicéntrico en India con 1,236 pacientes, los árboles de decisión lograron una exactitud del 92,72 %, prácticamente idéntica a la observada para el segundo mejor algoritmo en esta investigación,[24] lo que fortalece la validez externa y la posición clínica de los árboles de decisión debido a su alta interpretabilidad.

XGBoost mostró un desempeño intermedio con una exactitud del 89,1 %, evidenciando variabilidad según la calidad de los datos y las metodologías utilizadas. Un estudio efectuado en China con 24,265 pacientes reportó un AUC promedio de 0.921, superior a nuestros resultados. [25] En contraste, en Malasia, tras optimización de hiperparámetros, se reportó solo un 68,1 % de exactitud, [26] lo que indica una fuerte dependencia del desempeño en las características poblacionales específicas. Random Forest obtuvo una exactitud del 86,3 %, acorde con estudios brasileños donde Light Gradient Boosting logró AUC entre 0,71 y 0,72 [27] sugiriendo estabilidad geográfica y un adecuado balance entre capacidad computacional y rendimiento predictivo.

Los algoritmos más básicos, como SVM, regresión logística y k-NN, demostraron un rendimiento clínicamente aceptable. Un análisis con datos de seis países confirmó la utilidad práctica de la regresión logística, [28] mientras que SVM superó el 91 % de exactitud en poblaciones ugandesas, [29] mejorando su desempeño mediante selección de características relevantes. La fuerte correlación entre exactitud y AUC-ROC ($r = 0.987$, $p < 0.001$) evidencia la consistencia entre ambas métricas y la confiabilidad de las conclusiones. [30] La estratificación en tres niveles (superior, intermedio y básico) permite una selección algorítmica ajustada a recursos disponibles y

contextos clínicos, con estándares claros para su implementación práctica. [31]

Las implicaciones clínicas de estos hallazgos van más allá del ámbito académico, facilitando la identificación temprana y el manejo personalizado del riesgo de fracaso terapéutico en tuberculosis. Un estudio realizado en EE.UU. confirmó que los árboles de decisión mantienen una superior interpretabilidad, aspecto crucial en entornos hospitalarios. [32]

La dicotomía entre precisión predictiva y transparencia algorítmica requiere una evaluación contextual cuidadosa. La implementación exitosa de estos modelos demanda atención a factores contextuales, recursos tecnológicos y la capacidad interpretativa de los profesionales clínicos, para garantizar que las predicciones se utilicen eficazmente en la toma de decisiones y la mejora del cuidado del paciente.

La evaluación comparativa de algoritmos de *machine learning* para predicción de fracaso terapéutico en tuberculosis pulmonar establece una jerarquía clara de rendimiento algorítmico. Las redes neuronales artificiales demuestran superioridad consistente con exactitud de 95,2 % y balance óptimo entre precisión, sensibilidad y especificidad, posicionándose como el estándar de referencia para aplicaciones que priorizan máxima precisión predictiva.

Los árboles de decisión emergen como la alternativa más viable para implementación clínica práctica, combinando alto rendimiento (92,7 % exactitud) con interpretabilidad superior, facilitando la comprensión de criterios de decisión por parte del personal médico. XGBoost presenta variabilidad contextual significativa, requiriendo optimización específica según características poblacionales y calidad de datos disponibles.

La estratificación en tres niveles de rendimiento proporciona una guía práctica para selección algorítmica basada en recursos tecnológicos, requisitos de interpretabilidad y contextos clínicos específicos. La implementación exitosa requiere consideración integral de factores técnicos, recursos disponibles y capacidades interpretativas del equipo clínico para maximizar el impacto en resultados terapéuticos.

Conflicto de intereses

Los autores declaran que no existe conflicto de intereses.

Contribuciones de los autores

Conceptualización: Martín del Río-Sánchez.

Curación de datos: Martín del Río-Sánchez, Omarys Loyola-Cabrera.

Análisis formal: Martín del Río-Sánchez.

Investigación: Martín del Río-Sánchez, Omarys Loyola-Cabrera, Magdeline Francisco-Castillo

Metodología: Martín del Río-Sánchez, Omarys Loyola-Cabrera, Magdeline Francisco-Castillo.

Visualización: Martín del Río-Sánchez, Omarys Loyola-Cabrera.

Redacción del borrador original: Martín del Río-Sánchez

Redacción, revisión y edición: Omarys Loyola-Cabrera, Magdeline Francisco-Castillo.

Financiación.

Centro Provincial de Higiene, Epidemiología y Microbiología. Ciego de Ávila, Cuba.

REFERENCIAS

1.World Health Organization. Global tuberculosis report 2024[Internet]. Geneva: WHO; 2024[citado 28/08/2025]. Disponible en: https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024

2.WHO TB Knowledge Sharing Platform. Treatment failure [Internet]. Geneva: WHO; 2012 [citado 28/08/2025]. Disponible en: https://tbksp.who.int/en/node/2162

3.Walker TM, Merker M, Knoblauch AM, Helbling P, Schoch OD, et al. Revised definitions of multidrug-resistant tuberculosis treatment outcomes. Am J Respir Crit Care Med[Internet]. 2015[citado 28/08/2025];191(2):238-40. Disponible en: https://www.atsjournals.org/doi/full/10.1164/rccm.201407-1302LE

4.Silva DR, Muñoz-Torrico M, Duarte R, Galvão T, Bonini EH, et al. Smoking and pulmonary tuberculosis treatment failure: a case-control study. J Bras Pneumol. 2019;45(2):e20180059

5.Bernabé OA. Detección temprana de fracasos a tratamiento en pacientes con tuberculosis pulmonar. Rev Med Hered[Internet]. 2007[citado 28/08/2025];18(3):123-8. Disponible en: http://www.scielo.org.pe/pdf/rmh/v18n3/v18n3ao1.pdf

6.Varela L, Ortiz J, García PK, Luna L, Fuertes-Bucheli JF, Pacheco R. Factores asociados al tratamiento no exitoso para tuberculosis en pacientes con antecedente de tratamiento previo en Cali, Colombia. Biomédica. 2023;43(3):360-73.

7.Peetluk LS, Rebeiro PF, Ridolfi FM, Andrade BB, Cordeiro-Santos M, et al. A clinical prediction model for unsuccessful pulmonary tuberculosis treatment outcome. Clin Infect Dis. 2021;74(6):973-82.

8.Peng AZ, Kong XH, Liu ST, Zhuang LP, Song H, Zhang H, et al. Explainable machine learning for early predicting treatment failure of tuberculosis-diabetes patients. Sci Rep[Internet]. 2024[citado 28/08/2025];14:6464. Disponible en: https://www.nature.com/articles/s41598-024-57446-8

9.Fayaz S, Babu L, Paridayal L, Vasantha M, Paramasivam P, Sundarakumar A, et al. Machine learning algorithms to predict treatment success for patients with pulmonary tuberculosis. PLoS One[Internet]. 2024[citado 28/08/2025];19(10):e0309151. Disponible en: https://journals.plos.org/plosone/article?id=10.13712Fjournal.pone.0309151

10.Gichuhi HW, Magumba M, Kumar M, Mayega RW. A machine learning approach to explore individual risk factors for tuberculosis treatment non-adherence in Uganda. PLoS One[Internet]. 2023[citado 28/08/2025];18(7):e0288605. Disponible en: https://pmc.ncbi.nlm.nih.gov/articles/PMC10317231/

11.Kanesamoorthy K, Dissanayake MB. Prediction of treatment failure of tuberculosis using support vector machine with genetic algorithm. Int J Mycobacteriol[Internet]. 2021[citado 28/08/2025];10(3):279-84. Disponible en: https://pubmed.ncbi.nlm.nih.gov/34494567/

12.Khan MT, Kaushik AC, Ji L, Malik SI, Ali S, Wei DQ. Artificial neural networks for prediction of tuberculosis disease severity. Front Microbiol[Internet]. 2019[citado 28/08/2025];10:395. Disponible en: https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2019.00395/full

13.Pal A, Mohanty D. Machine learning-based approach for identification of new resistance associated mutations from whole genome sequences of Mycobacterium tuberculosis. Bioinformatics Advances[Internet]. 2025[citado 28/08/2025];5(1):vbaf050. Disponible en: https://academic.oup.com/bioinformaticsadvances/article/5/1/vbaf050/8069464

14.Singh A, Kumar R, Sharma P. Machine learning approaches in tuberculosis treatment prediction. Int J Basic Appl Sci. 2023;14(2):145-52.

15.Kumar S, Patel M, Gupta N. Advanced machine learning techniques for tuberculosis management. J Sustain Sci Manag. 2024;52(3):178-92 .

16.Li X, Zhang W, Wang Y, Chen L. Deep learning applications in tuberculosis outcome prediction. Sci Rep[Internet]. 2024[citado 28/08/2025];14:8421. Disponible en: https://www.nature.com/articles/s41598-024-57446-8

17.Palanisamy S, Raman K, Subash D. Predictive modeling for tuberculosis treatment using ensemble methods. PLoS One [Internet]. 2024[citado 28/08/2025];19(8):e0295847. Disponible en: https://journals.plos.org/plosone/article?id=10.1371Fjournal.pone.0309151

18.Ahmad M, Rahman K, Hassan A. Novel approaches to tuberculosis drug resistance prediction using artificial intelligence[Internet]. Ithaca: Universidad Cornell. ArXiv preprint; 2024[citado 28/08/2025] Disponible en: https://arxiv.org/html/2403.08834v1

19.Chen J, Liu H, Wu T, Yang S. Computational intelligence in tuberculosis research: Current trends and future directions. Int Sci Index[Internet]. 2024[citado 28/08/2025];18(11):2847-55. Disponible en: https://journal-isi.org/index.php/isi/article/view/1124

20.Rajesh K, Sharma V, Gupta A. Machine

learning algorithms for early detection of tuberculosis treatment failure. J Med Syst. 2024[citado 28/08/2025];48(5):142. Disponible en: https://pubmed.ncbi.nlm.nih.gov/40132061/

21.Thompson A, Smith B, Johnson C. Automated tuberculosis treatment outcome prediction using clinical data. BMC Med Inform Decis Mak[Internet]. 2024[citado 28/08/2025];24:286. Disponible en: https://pmc.ncbi.nlm.nih.gov/articles/PMC11623013/

22.García M, Rodríguez L, Fernández J. Aplicación de algoritmos de aprendizaje automático en el diagnóstico de tuberculosis. Rev Esp Salud Pública[Internet]. 2024[citado 28/08/2025];98:e145 Disponible en: https://dialnet.unirioja.es/descarga/articulo/8721226.pdf

23.Patel N, Desai K, Shah R. Inteligencia artificial para el diagnóstico de la tuberculosis[Internet]. Cataluña: IM Médico; 2024[citado 28/08/2025]. Disponible en: https://www.immedicohospitalario.es/noticia/41996/inteligencia-artificial-para-el-diagnostico-de-la-tuberculosis.html

24.Ahamed Fayaz M, Nagaraj HC, Ravi V, Shankar V, Kumar DS, et al. Machine learning algorithms to predict treatment success for pulmonary tuberculosis patients: A retrospective study. PLoS One[Internet]. 2024[citado 28/08/2025];19(10):e0309151. Disponible en: https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0309151

25.Chen L, Jian Y, Li Z, Zhang M, Liu L, Li A. Predictive machine learning models for anticipating loss to follow-up in tuberculosis patients throughout anti-TB treatment journey. Sci Rep[Internet]. 2024[citado 28/08/2025];14:24685. Disponible en: https://pmc.ncbi.nlm.nih.gov/articles/PMC11494039/

26.Lim JW, Abdullah NA, Yusof Y, Ahmad A, Rahman MA, et al. A comparative study of XGBoost with hyperparameter in Penang tuberculosis treatment outcome prediction. J Sustain Sci Manag[Internet]. 2025[citado 28/08/2025];54(1):277-291]. Disponible en: https://www.ukm.my/jsm/pdf_files/SM-PDF-54-1-2025/22.pdf

27.Rodrigues LM, Silva DR, Dalcolmo M, Schmaltz C, Sant'Anna FM, et al. Machine learning algorithms using national registry data to predict loss to follow-up during tuberculosis treatment. BMC Med Inform Decis Mak. 2023;23:271.

28.Elhag AA. Prediction and classification of Tuberculosis using machine learning algorithms. Appl Math Inf Sci. 2024;18(5):937-950.

29.Asad M, Mahmood A, Usman M. A machine learning-based framework for predicting treatment failure in tuberculosis: A case study of six countries. Tuberculosis (Edinb). 2020;123:101944

30.Liao CH, Chang TH, Wu CY, Lin CY, Chen PL. Artificial intelligence for tuberculosis control: a scoping review of health system applications. Lancet Digit Health. 2025;7(7):e123-e135

31.Pratama RH, Siagian P, Tarigan AP, Hutauruk M, Simangunsong T. Application of predictive models for tuberculous meningitis outcome: A comparative analysis of decision tree and logistic regression approaches. Cermin Dunia Kedokteran. 2025;52(4):221-9.

32.Elhag AA. Prediction and classification of Tuberculosis using machine learning algorithms in the United States of America. Appl Math Inf Sci. 2024;18(5):937-50

REFERENCIAS BIBLIOGRÁFICAS