

**DESARROLLO DE UN MODELO PREDICTIVO DE LA RESISTENCIA A LOS ANTIBIÓTICOS  
CARBAPENÉMICOS EN *Klebsiella pneumoniae* MEDIANTE EL USO DE DATOS PROTEÓMICOS E  
INTELIGENCIA ARTIFICIAL EN EL LABORATORIO CLÍNICO SYNLAB-COLOMBIA**

Salazar Valentina<sup>1\*</sup>, Gómez Juan<sup>1</sup>, Mojica Iván<sup>1</sup>, Estrada Ana<sup>1</sup>, Fernández Geysson<sup>2</sup>

<sup>1</sup>. Laboratorio SYNLAB-Colombia, Carrera 14#10173 Bogota.D.C.

<sup>2</sup>. Grupo de investigación Biología y Control de Enfermedades, Universidad de Antioquia.

**RESUMEN**

El uso de antibióticos ha sido fundamental en la medicina para tratar infecciones, pero su uso inadecuado ha conllevado a la selección de microorganismos resistentes. *Klebsiella pneumoniae* es uno de los patógenos más comunes a nivel mundial, ocasionando infecciones y complicaciones asociada a su resistencia, se han implementado en algunos lugares tecnologías innovadoras como la espectrometría de masas (MALDI-TOF) que permite realizar la identificación del microorganismo en menor tiempo. En la búsqueda de la identificación de resistencia a ciertos antibióticos se ha explorado alternativas en la inteligencia artificial.

Este trabajo plantea un modelo predictivo de resistencia a los carbapenémicos en aislados de *Klebsiella pneumoniae* usando espectros proteómicos obtenidos a partir del MALDI-TOF. Se evaluaron 5 modelos de aprendizaje automático supervisado para un total de 75 muestras (21 resistentes y 54 sensibles) los modelos fueron: regresión logística, random forest, máquina de vectores de soporte, k vecinos más cercanos y clasificador bayesiano ingenuo. El modelo que obtuvo las mejores métricas de desempeño con una especificidad del 100% y exactitud del 93% fue máquina de vectores de soporte.

**Palabras clave:** Resistencia, aprendizaje automático, predicción.

## INTRODUCCIÓN

Desde sus inicios, los antibióticos se han convertido en una de las estrategias terapéuticas más ampliamente usadas en la medicina (Huemer et al., 2020). Sin embargo, esta práctica ha conllevado riesgos, como la aparición de la resistencia bacteriana. Los niveles cada vez mayores de resistencia a los antibióticos amenaza la salud lo cual está dirigiendo a una crisis mundial (Laws et al., 2019). De hecho, la Organización Mundial de la Salud (OMS) declaró que la resistencia a los antimicrobianos es una de las 10 principales amenazas para la salud pública a nivel mundial (Widmer, 2022). Microorganismos resistentes son responsables de alrededor de 700,000 muertes por año en todo el mundo y se estima que para el año 2050 serán responsables de más de 10 millones de muertes por año (Huemer et al., 2020).

Entre las bacterias de mayor importancia está *Klebsiella pneumoniae* (Breijyeh et al., 2020), evade los antibióticos mediante diferentes mecanismos de acción, una de las más importantes es la producción de betalactamasas, enzimas que catalizan antibióticos como la penicilina, cefalosporinas y carbapenémicos. En el año 2014, la OMS informó que las cepas de *K. pneumoniae* han incrementado la tasa de resistencia a carbapenémicos sobre el 50%, lo cual ha conllevado un aumento en la mortalidad y morbilidad en pacientes con infecciones por esta bacteria (Vera-Leiva et al., 2017).

Para detectar la resistencia en el laboratorio, se emplean métodos automatizados que permiten determinar la sensibilidad a los antibióticos según el tipo de bacteria, proceso que tarda aproximadamente entre 3 a 5 días. Actualmente se ha implementado la espectrometría de masas de tiempo de vuelo por ionización por desorción láser asistida por matriz (MALDI-TOF MS) para la identificación microbiana, su principal ventaja es la reducción de tiempo de detección, pasando de días a horas a partir de colonias puras. Con las nuevas tecnologías disponibles en salud, se ha generado una gran cantidad de información que resulta complejo analizar. Herramientas como el machine learning permiten lidiar con estas dificultades. Esta herramienta además de ser útil para realizar predicciones clínicas y epidemiológicas, tiene el potencial de aportar al análisis de los procesos clínicos para la toma de decisiones en el ámbito de la salud. (Pedrero et al., 2021).

## OBJETIVO

Plantear un modelo predictivo de resistencia a los carbapenémicos en aislados de *Klebsiella pneumoniae* usando espectros proteómicos obtenidos a partir del MALDITOF.

## MATERIALES Y MÉTODOS

**Tipos de datos:** Se realizó una búsqueda de la identificación bacteriana de los aislamientos de *Klebsiella pneumoniae* entre los años 2022 y junio del 2023 en el equipo de espectrometría de masas MALDI-TOF MS, este equipo permite la identificación de microorganismos mediante la creación de un espectro basado en el perfil de proteínas (MALDI Biotyper de Bruker).

Para evaluar su perfil de sensibilidad, se realizó la búsqueda de estos perfiles en los informes generados a partir del equipo VITEL 2 XL (BioMérieux) y los antibióticos utilizados para el análisis de la resistencia fueron los carbapenémicos.

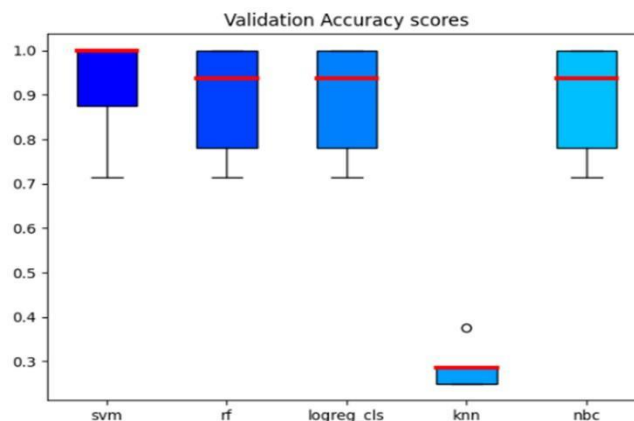
Una vez recolectados los espectros proteómicos y sus perfiles de sensibilidad, se construyó una base de datos con 75 datos en total, 21 de muestras resistentes a los carbapenémicos y 54 de muestras sensibles y se procedió a trabajar en un entorno de ejecución en Google Colab, con lenguaje de programación Python con librerías como numpy, pandas, matplotlib y sklearn.

**Construcción del modelo de predicción:** Se emplearon modelos de aprendizaje supervisado cuyo objetivo es la clasificación; el conjunto de datos se separó aleatoriamente en un set de entrenamiento (80% de los datos totales equivalente a 60 datos) y un set de validación (20% de los datos totales equivalente a 15 datos). Los modelos a evaluar con el set de entrenamiento fueron: regresión logística, random forest (RF), máquina de vectores de soporte (SVM), k-vecinos más cercanos (KNN) y clasificador bayesiano ingenuo (NBC). Los datos de entrada (predictores) serán las variables m/z e intensidad y los resultados esperados será la clasificación de sensible o resistente. Después de realizar el entrenamiento del modelo se procedió a evaluarlo mediante la validación cruzada para observar el rendimiento de cada uno de los modelos y para finalizar se realizó una matriz de confusión evaluando el mejor modelo de clasificación.

## RESULTADOS

Se utilizó algoritmos de aprendizaje supervisado tipo clasificación mencionados anteriormente y se realizó una validación cruzada de todos los modelos para observar el de mejor rendimiento

(clasificación) obteniendo que el mejor modelo de clasificación fue máquinas de vectores de soporte (svm) ya que la mayoría de sus predicciones se encontraron entre el 87 y 100% de exactitud (Figura 1).



**Figura 1.** Validación cruzada de los modelos evaluados; esta técnica se utiliza para evaluar y validar la capacidad predictiva de los modelos, en este caso se evalúa la exactitud de los modelos empleados como máquina de vectores de soporte (svm), random forest (rf), regresión logística (logreg\_cls), k vecinos más cercanos (knn) y clasificador bayesiano ingenuo (nbc).

Se realizó una matriz de confusión, esta matriz es una tabla que se utiliza en la evaluación de modelos de clasificación, describe el rendimiento de un modelo al comparar las predicciones del modelo con los valores reales de los datos. Se evaluó sensibilidad con un porcentaje del 66%, especificidad del 100%, precisión del 100% y exactitud del 93% (tabla 3).

**Tabla 1.** Evaluación de la matriz de confusión para el modelo de clasificación máquina de vectores de soporte (svm)

	Verdadero negativo	Verdadero positivo
Verdadero negativo	12	0
Verdadero positivo	1	2

## DISCUSIÓN

Las enfermedades infecciosas son la principal causa de mortalidad humana en todo el mundo, su identificación y la predicción de brotes de enfermedades son los ejes centrales para combatir esta amenaza. La inteligencia artificial puede ayudar en la identificación y caracterización de microorganismos resistentes (Goodswen et al., 2021).

Hay algunos estudios que han trabajado con inteligencia artificial para predecir resistencia bacteriana. Investigadores han trabajado en la predicción de resistencia en *Klebsiella pneumoniae*, los picos de los espectros obtenidos a partir de la tecnología MALDI-TOF MS en la relación m/z 11109 se correlacionan con una variante de betalactamasa y discriminan algunas cepas de *K. pneumoniae* resistente a carbapenémicos (CRKP). Huang et al. (2020) clasificaron 46 aislados de CRKP y 49 de *K. pneumoniae* sensible a carbapenémicos (CSKP) con una precisión del 97 % utilizando aprendizaje automático en relación a los picos de los espectros analizados.

## CONCLUSIÓN

En general, para 75 datos el modelo tuvo un comportamiento bastante notorio, una exactitud del 93% es bastante alta y sugiere que el modelo tiene un buen rendimiento en la mayoría de las instancias y una especificidad y precisión del 100%. Estos resultados sugieren que el mejor modelo de predicción de resistencia bacteriana es máquina de vectores de soporte (svm). Es fundamental validar estos resultados con un conjunto de datos más robusto e ir observando su comportamiento.

## BIBLIOGRAFÍA

Huemer, M., Shambat, S. M., Brugger, S. D., & Zinkernagel, A. S. (2020). Antibiotic resistance and persistence—Implications for human health and treatment perspectives. *EMBO Reports*, 21(12). <https://doi.org/10.15252/EMBR.202051034>

Laws, M., Shaaban, A., & Rahman, K. M. (2019). Antibiotic resistance breakers: current approaches and future directions. *FEMS Microbiology Reviews*, 43(5), 490. <https://doi.org/10.1093/FEMSRE/FUZ014>

Breijyeh, Z., Jubeh, B., & Karaman, R. (2020). Resistance of Gram-Negative Bacteria to Current Antibacterial Agents and Approaches to Resolve It. *Molecules*, 25(6). <https://doi.org/10.3390/MOLECULES25061340>

Vera-Leiva, A., Barría-Loaiza, C., Carrasco-Anabalón, S., Lima, C., Aguayo-Reyes, A., Domínguez, M., Bello-Toledo, H., González-Rocha, G., Vera-Leiva, A., Barría-Loaiza, C., Carrasco-Anabalón, S., Lima, C., Aguayo-Reyes, A., Domínguez, M., Bello-Toledo, H., & González-Rocha, G. (2017). KPC: Klebsiella pneumoniae carbapenemasa, principal carbapenemasa en enterobacterias. *Revista Chilena de Infectología*, 34(5), 476–484. <https://doi.org/10.4067/S0716-10182017000500476>

Pedrero, V., Reynaldos-Grandón, K., Ureta-Achurra, J., Cortez-Pinto, E., Pedrero, V., Reynaldos-Grandón, K., Ureta-Achurra, J., & Cortez-Pinto, E. (2021b). Generalidades del Machine Learning y su aplicación en la gestión sanitaria en Servicios de Urgencia. *Revista Médica de Chile*, 149(2), 248–254. <https://doi.org/10.4067/S0034-98872021000200248>

Widmer, A. F. (n.d.). *Emerging antibiotic resistance: Why we need new antibiotics!* <https://doi.org/10.57187/smw.2022.40032>

Goodswen, S. J., Barratt, J. L. N., Kennedy, P. J., Kaufer, A., Calarco, L., & Ellis, J. T. (2021). Machine learning and applications in microbiology. *FEMS Microbiology Reviews*, 45(5), 1–19. <https://doi.org/10.1093/FEMSRE/FUAB015>

Huang, T. S., Lee, S. S. J., Lee, C. C., & Chang, F. C. (2020). Detection of carbapenem-resistant Klebsiella pneumoniae on the basis of matrix-assisted laser desorption ionization time-of-flight mass spectrometry by using supervised machine learning approach. *PLoS ONE*, 15(2). <https://doi.org/10.1371/JOURNAL.PONE.0228459>