



<https://doi.org/10.24245/mim.v40i3.9023>

Validación de un modelo de inteligencia artificial para la predicción de la mortalidad del paciente con sepsis

Validation of an artificial intelligence model for the mortality prediction of the patient with sepsis.

Mayre Alejandra Sierra Juárez,¹ Karen Paola Quintana Barragán,¹ Jesús Alan Hernández Galván,² Luis Bernardo Enríquez Sánchez,⁴ Manuel David Pérez Ruiz,⁵ Carlos Arzate Quintana³

Resumen

OBJETIVO: Validar un modelo de inteligencia artificial que pueda predecir el pronóstico de mortalidad del paciente hospitalizado con sepsis.

MATERIALES Y MÉTODOS: Estudio de cohorte ambispectivo observacional, en el que se analizaron expedientes electrónicos de pacientes adultos del Hospital Central del Estado de Chihuahua, México, de julio de 2018 a marzo de 2020 y de enero de 2021 a enero de 2022. Se analizaron tres modelos: redes neuronales, máquina de soporte de vectores y bosques aleatorios. Para la validación del modelo, el 80% de la muestra fue para adiestramiento y el 20% para la prueba. Del último grupo (20%) se implementó una validación cruzada de 10 pliegues para el cálculo de sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo.

RESULTADOS: Se analizaron 353 expedientes, de los que solo se eligieron 218. El mejor modelo fue el de redes neuronales; sin embargo, su puntaje del área bajo la curva (AUC) apenas alcanzó 0.80; por debajo de este valor estuvieron el algoritmo de bosques aleatorios (AUC 0.667) y el de máquina de soporte de vectores (AUC 0.641). De los 3 modelos, solo se hizo la validación cruzada con el de redes neuronales, del 20% de los datos de prueba se implementaron 10 validaciones. Los puntajes del área bajo la curva obtenidos en cada pliegue fueron de 0.771 a 0.830.

CONCLUSIONES: El modelo es bueno aun trabajando con pocos datos. Se pretende recolectar una mayor muestra para volver a adiestrar y validar el modelo con más datos y mejorar el aprendizaje y rendimiento para finalmente ser válido en pacientes.

PALABRAS CLAVE: Inteligencia artificial; pronóstico; mortalidad; sepsis; redes neuronales; máquina de soporte de vectores; bosques aleatorios.

Abstract

OBJECTIVE: To validate an artificial intelligence model that can predict the mortality prognosis of hospitalized patients with sepsis.

MATERIALS AND METHODS: An ambispective observational cohort study, which included electronic records of adult patients from the Central Hospital of the State of Chihuahua, Mexico, from July 2018 to March 2020 and January 2021 to January 2022. Three models were analyzed: neural networks, support vector machine and random forests. For model validation, the sample was divided into 80% for training and 20% for testing. For the last group (20%), a 10-fold cross-validation was implemented to calculate sensitivity, specificity, positive predictive value, and negative predictive value.

RESULTS: A total of 353 files were analyzed, of which only 218 were chosen. The best model was the neural networks; however, its area under the curve (AUC) score barely reached 0.80, the random forests algorithm (AUC 0.667) and the support vec-

¹ Médico cirujano y partero.

² Ingeniería Biomédica.

³ Docente. Seminario de investigación en la licenciatura de Ingeniería Biomédica.

Facultad de Medicina y Ciencias Biomédicas, Universidad Autónoma de Chihuahua, Chihuahua, Chihuahua, México.

⁴ Médico especialista en Cirugía General. Profesor investigador, jefe del Departamento de Cirugía General, Hospital Central del Estado, Chihuahua, Chihuahua, México.

⁵ Médico adscrito al Departamento de Cirugía General, Hospital Central Universitario, Chihuahua, Chihuahua, México.

Recibido: 2 de agosto 2023

Aceptado: 21 de agosto 2023

Correspondencia

Luis Bernardo Enríquez Sánchez
investigacionhcu@gmail.com

Este artículo debe citarse como: Sierra-Juárez MA, Quintana-Barragán KP, Hernández-Galván JA, Enríquez-Sánchez LB, Pérez-Ruiz MD, Arzate-Quintana C. Validación de un modelo de inteligencia artificial para la predicción de la mortalidad del paciente con sepsis. Med Int Méx 2024; 40 (3): 171-178.

tor machine algorithm (AUC 0.641) were below this value. Of the 3 models, only the cross-validation with the neural networks was done, of 20% of the test data, 10 validations were implemented. The AUC scores obtained in each fold ranged from 0.771 to 0.830.

CONCLUSIONS: The model is good, even working with few data. It is intended to collect a larger sample to retrain and validate the model with more data and improve learning and performance and finally be applicable to patients.

KEYWORDS: Artificial intelligence; Prognosis; Mortality; Sepsis; Neural networks; Support vector machine; Random forests.

ANTECEDENTES

La sepsis es un padecimiento que ha afectado a millones de personas a lo largo del tiempo, por lo que se siguen estudiando y analizando diferentes tratamientos y sistemas de detección con el objetivo de establecer un diagnóstico en etapas más tempranas, implementar un tratamiento más oportuno y mejorar el pronóstico de desenlace. La sepsis es una urgencia médica debido a que puede causar daño a varios órganos al mismo tiempo y llevar a la muerte, además de las secuelas en los pacientes que logran sobrevivir. Se estima que este padecimiento afecta a más de 30 millones de personas en todo el mundo, con más de 6 millones de muertes al año.¹

El diagnóstico de sepsis puede ser todo un reto debido a la gran variedad de signos y síntomas que manifiesta el paciente, así como las causas de base. Para ello se han usado muchos sistemas de puntuación con el objetivo de detectar el padecimiento en etapas tempranas, lo que repercute de manera importante en el pronóstico de los pacientes. Sin embargo, todos estos sistemas tienen variaciones en cuanto a su especificidad y sensibilidad, por lo que no se cuenta con una herramienta única de detección y pronóstico que tenga la suficiente precisión y exactitud para

predecir el padecimiento. Es bien sabido que estos pacientes la mayor parte de las veces necesitan cuidados más especiales, por lo que gran parte del tiempo su estancia es en la unidad de cuidados intensivos, lo que aumenta los costos hospitalarios. Para ser más específicos, un día en cuidados intensivos cuesta alrededor de 34,232 pesos mexicanos y el promedio de estancia de estos pacientes es de 9 días, lo que es una gran carga para el sistema de salud.²

En la actualidad se ha visto la necesidad de usar herramientas más complejas y objetivas que tengan la capacidad de hacer predicciones más exactas y precisas. Una de estas herramientas es la inteligencia artificial. Son incontables los modelos que se han desarrollado para predecir el padecimiento o el pronóstico de los pacientes con sepsis. Sin duda, el uso de la inteligencia artificial ha llegado para revolucionar la forma en que el personal de salud atiende este padecimiento. Los modelos desarrollados por la inteligencia artificial se han comparado con las escalas diagnósticas y de pronóstico usadas en la actualidad y, sorprendentemente, todos estos modelos han superado de forma significativa a estas escalas, con puntajes del área bajo la curva ROC (característica operativa del receptor [AUC-ROC]) superiores a 0.80.^{3,4,5}



Por tanto, este trabajo tiene el objetivo de validar un modelo de inteligencia artificial que fue previamente desarrollado para predecir la mortalidad de estos pacientes, con el propósito de ofrecer una herramienta que permita al personal de salud a cargo identificar a los pacientes con sepsis con mayor probabilidad de tener un desenlace desfavorable durante sus primeras 24 horas de estancia intrahospitalaria.

MATERIALES Y MÉTODOS

Estudio de cohorte ambispectivo observacional, en el que se analizaron expedientes electrónicos de julio de 2018 a marzo de 2020 y de enero de 2021 a enero de 2022. De primera instancia se consideraron los pacientes con procalcitonina mayor de 10 ng/mL, posteriormente se eligieron pacientes de entre 18 y 99 años y que cumplieran con los criterios de Sepsis 3 y estancia intrahospitalaria igual o mayor a 24 horas. Por último, se eligieron 353 expedientes que conformaron la población blanco.

Se depuraron los datos faltantes. Las variables elegidas posterior a la depuración de datos faltantes y que fueron usadas para la realización de las métricas fueron las siguientes:

- Presunto inicio de sepsis (intrahospitalario, extrahospitalario).
- Antecedentes personales patológicos: hábito tabáquico, alcoholismo, consumo de drogas, diabetes mellitus y su tratamiento, hipertensión arterial y su tratamiento, enfermedad respiratoria crónica, insuficiencia cardíaca crónica, enfermedad renal crónica, enfermedad hepática crónica, enfermedad reumática, estado de inmunodepresión (cáncer, SIDA, tratamiento con quimioterapia, administración de esteroides sistémicos, inmunosupresores), embarazo, cirugía en el último mes.
- Medición de los signos vitales al momento del diagnóstico o sospecha de sepsis (presión arterial media, frecuencia cardíaca, índice de oxigenación no invasivo, frecuencia respiratoria, temperatura). Posteriormente se evaluaron los signos vitales a las 24 horas de haber establecido el diagnóstico o sospecha de sepsis. Durante esas primeras 24 horas se tomaron los valores mínimos y máximos de cada signo vital y se sacó la diferencia entre ellos; es decir, diferencia de la presión arterial media mínima y máxima, de la frecuencia cardíaca mínima y máxima, de la frecuencia respiratoria mínima y máxima, del índice de oxigenación no invasivo y diferencia de la temperatura mínima y máxima obtenidas durante su primer día de diagnóstico o sospecha.
- Otras variables que se tomaron en cuenta para los criterios de inclusión fueron: escala de coma de Glasgow, estudios de laboratorio: biometría hemática, leucocitos, neutrófilos, linfocitos, hemoglobina, hematocrito y plaquetas; química sanguínea: ácido úrico, nitrógeno ureico, creatinina sérica y glucosa; electrolitos séricos con la medición de sodio, potasio y cloro. El último análisis de laboratorio que se incluyó fue el hemocultivo.
- También se tomó en cuenta el tratamiento antibiótico elegido de manera inicial (de amplio espectro, con base en la sospecha del sitio de origen de sepsis, o si no se inició), a las cuantas horas del diagnóstico o sospecha se inició el tratamiento antibiótico (ya estaba recibiendo por otra causa, menos de una hora, entre 1 y 3 horas, más de 3 horas, no se inició).

Se tomó un nivel de confianza del 95% y un poder de prueba del 80%.

Para la validación del modelo, la muestra se dividió en 2 grupos, el 80% para adiestramiento y el 20% restante para la prueba. En este trabajo se usó la validación cruzada (*cross validation*) de 10 pliegues, que se hizo sobre el 20% de los datos de prueba.

RESULTADOS

Se eligieron 353 expedientes que conformaron la población blanco. Después de la depuración de datos faltantes se eligieron como muestra 218 expedientes.

Después de la elaboración de la base de datos, limpieza y preprocesamiento, se eligió el 80% de los datos para adiestramiento y el 20% para la puesta a prueba mediante el método “train-test-split”. Se obtuvieron las métricas de 3 algoritmos (**Cuadro 1**) y su puntaje del área bajo la curva (AUC) [**Figura 1**]. Los modelos que se usaron para el adiestramiento y puesta a prueba fueron: máquina de soporte de vectores (SMV), bosques aleatorios (RF) y redes neuronales (NN).

Con base en las métricas anteriores y puntaje del AUC, se observa que el mejor modelo fue el de redes neuronales. Por ello, este modelo fue el que se usó para la validación cruzada. Su puntaje del AUC apenas alcanzó 0.80, por debajo de este valor estuvieron el algoritmo de

bosques aleatorios (AUC 0.667) y el de máquina de soporte de vectores (AUC 0.641), esto podría explicarse porque el tamaño de la muestra no fue lo suficientemente grande. Además, la base de datos no estaba equilibrada, es decir, había 122 pacientes que fueron dados de alta y 96 que fallecieron, esto hace que los modelos no puedan clasificar de la mejor manera.

Al no contar con más datos para hacer una nueva base de datos, se optó por implementar la validación cruzada (*cross validation*). El algoritmo que se usó para esta validación fue el que obtuvo mayor puntaje del AUC en el adiestramiento que fue la red neuronal (0.808).

Para esta validación se usó el 20% de los datos de prueba (20% de los 218 pacientes, n = 43.6 ≈ 44) y se implementaron 10 pliegues, lo que quiere decir que se hizo una validación de 10 veces. Cada validación tiene un puntaje del AUC y métricas diferentes. Los puntajes del AUC obtenidos en cada pliegue variaron entre 0.771 y 0.830; sin embargo, esos valores se promediaron, obteniendo 0.795, es decir, 79.5% de exactitud. **Cuadro 2**

El rendimiento del modelo de redes neuronales fue mejor en la etapa de adiestramiento y prueba que en la validación cruzada; sin embargo, esta diferencia es por milésimas.

Cuadro 1. Métricas de los algoritmos

Exactitud		Máquina de soporte de vectores	Bosques aleatorios	Red neuronal
		0.64	0.67	0.79
Precisión	Defunción	0.63	0.68	0.68
	Alta	0.65	0.65	0.90
Exhaustividad	Defunción	0.63	0.65	0.87
	Alta	0.65	0.68	0.75
F1-Score	Defunción	0.63	0.67	0.76
	Alta	0.65	0.67	0.82

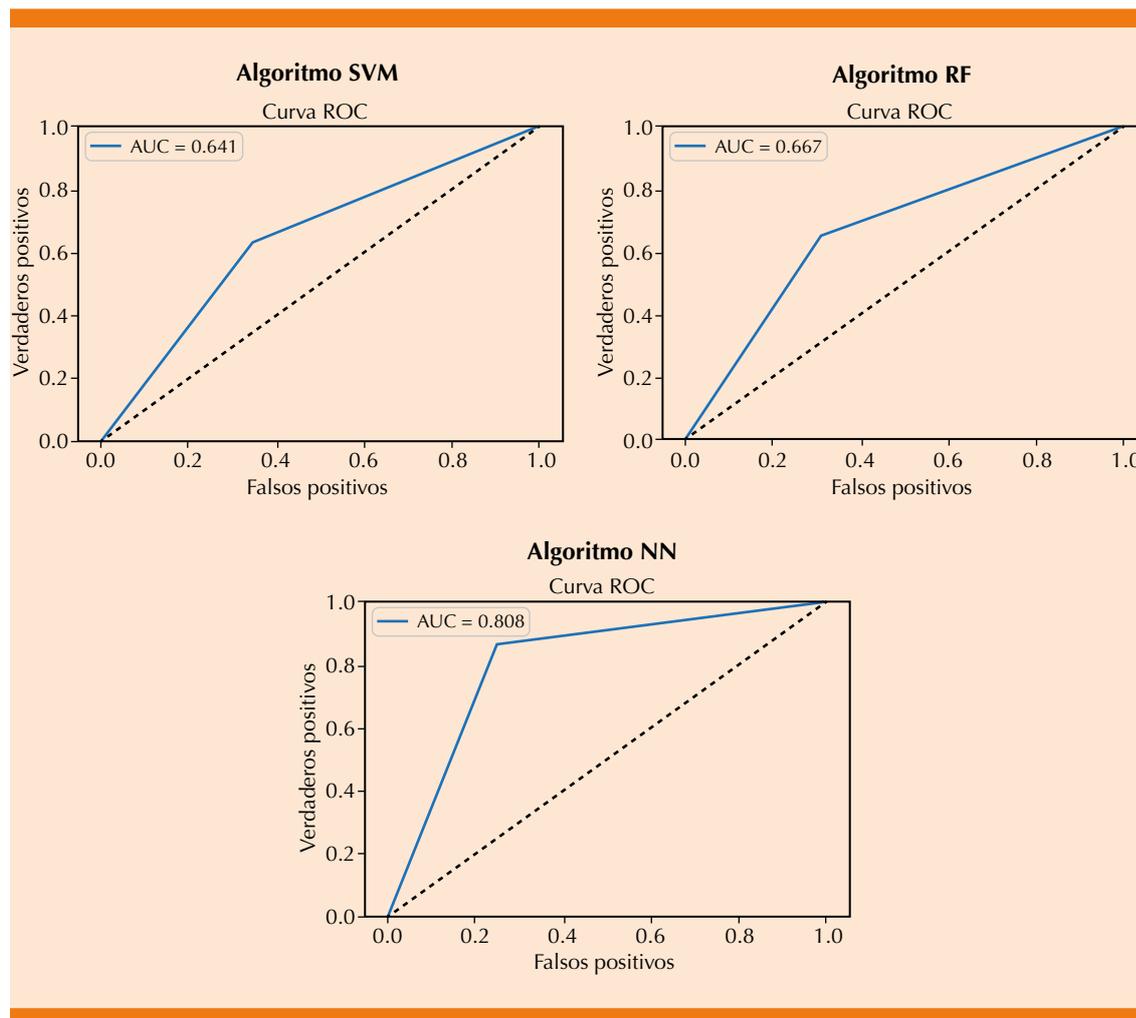


Figura 1. Desempeño de los algoritmos luego del adiestramiento y su puntaje AUC.

AUC: área bajo la curva; ROC: característica operativa del receptor; SVM: máquina de soporte de vectores; RF: bosques aleatorios; NN: redes neuronales.

Nota: las gráficas fueron generadas por medio de la librería Matplotlib.

Además, también se calcularon la sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo usando las variables descritas anteriormente en los criterios de inclusión. Cada pliegue de la validación generaba una matriz de confusión, por lo que al final se tuvieron 10 matrices, de las que esos valores se promediaron para obtener un resultado final. **Cuadro 3**

DISCUSIÓN

La inteligencia artificial aplicada a la sepsis se ha usado con el objetivo de diagnosticar el padecimiento o predecir la mortalidad. Se han usado diferentes herramientas para el pronóstico de los pacientes (MEWS, SOFA, qSOFA, SIRS, APACHE II, etc.) que, al compararse con estudios de

Cuadro 2. Valores promedio del área bajo la curva

Pliegue	Exactitud		Precisión	Exhaustividad	F1-Score
1	0.771	Alta	0.78	0.78	0.78
		Defunción	0.76	0.76	0.76
2	0.771	Alta	0.78	0.78	0.78
		Defunción	0.76	0.76	0.76
3	0.830	Alta	0.88	0.78	0.82
		Defunción	0.79	0.88	0.83
4	0.801	Alta	0.76	0.81	0.79
		Defunción	0.83	0.79	0.81
5	0.770	Alta	0.75	0.75	0.75
		Defunción	0.79	0.79	0.79
6	0.771	Alta	0.76	0.76	0.76
		Defunción	0.78	0.78	0.78
7	0.830	Alta	0.79	0.88	0.83
		Defunción	0.88	0.78	0.82
8	0.801	Alta	0.78	0.82	0.80
		Defunción	0.82	0.78	0.80
9	0.799	Alta	0.81	0.76	0.79
		Defunción	0.79	0.83	0.81
10	0.806	Alta	0.82	0.78	0.80
		Defunción	0.79	0.83	0.81
Promedio	0.795		0.795	0.794	0.793

Cuadro 3. Validación del algoritmo de redes neuronales

Sensibilidad	0.79
Especificidad	0.78
Valor predictivo positivo	0.78
Valor predictivo negativo	0.80

modelos de inteligencia artificial, estos últimos han superado, por mucho, a las escalas usadas en la actualidad.^{3,4,5} De esta forma, el uso que puede darse a la inteligencia artificial en estos pacientes es inimaginable.

El objetivo de este trabajo fue validar un modelo de inteligencia artificial que predijera la mortalidad de los pacientes con sepsis durante sus

primeras 24 horas de estancia intrahospitalaria. De los modelos desarrollados, la red neuronal obtuvo el mejor rendimiento en la etapa de adiestramiento y prueba, por lo que fue el elegido para la validación cruzada.

Sin embargo, el resultado final no fue el esperado y esto puede explicarse porque el tamaño de la muestra fue reducido debido a los datos que tuvieron que eliminarse, ya sea porque no cumplían los criterios de inclusión o porque tenían datos faltantes (pruebas de laboratorio, signos vitales, etc.) y por este motivo quedaban descartados. En consecuencia, los datos con los que se trabajó no fueron del todo suficientes para generar en esta primera prueba (que podría llamarse “prueba piloto”) puntajes del área bajo



la curva superiores a 0.80. Aun así, la validación del modelo fue muy cercana: 0.795.

De lo anterior podemos observar que el modelo es bueno, aun trabajando con pocos datos. Otra de las limitaciones de este estudio es que los casos de defunción y no defunción no estaban parejos (96 y 122, respectivamente), esto hace que la base de datos no esté equilibrada y que el algoritmo no pueda clasificar de forma correcta. Una vez más este problema se soluciona añadiendo más muestra al modelo.

Existen diferentes algoritmos que se han implementado para hacer clasificaciones o predicciones de pacientes con sepsis. Por mencionar algunos está XGBoost, GBM, RN con LSTM, LASSO, InSigth, entre otros.³⁻⁶ Si bien, el área bajo la curva de estos modelos ha sido superior a 0.80, éstos han usado bases de datos con al menos 1000 pacientes. Esto recalca la importancia de tener una base de datos muy grande y completa para que el modelo tenga más datos de donde aprender.

La validación usada en este estudio fue cruzada (*cross validation*) debido a la muestra reducida. Este tipo de validación se ha usado en muchos estudios y es una gran opción porque usa los datos de prueba (20%), con la ventaja de que aprende sobre datos que no fueron presentados al algoritmo durante su adiestramiento. Algunas curvas ROC de estudios que usaron la validación cruzada mostraron altos puntajes: 0.98⁷ y 0.83.⁸

En este trabajo también se tuvo como objetivo obtener, mediante la validación del modelo, la sensibilidad, especificidad, valor predictivo positivo y negativo, que fueron muy similares entre ellos: 0.79, 0.78, 0.78 y 0.80, respectivamente. De éstos, se observa que el modelo tiene mejor capacidad para predecir las “no defunciones”, esto puede explicarse porque en la base de datos había más pacientes que no fallecieron ($n = 122$) en comparación con los que sí fallecieron

($n = 96$), por lo que, al no estar equilibrado, el modelo aprende sobre el alta.

Una gran diferencia de este modelo es que usa una gran cantidad de variables (74 en total) en comparación con los de otros estudios en los que la mayor parte usa menos de 50 variables,^{3,9,10,11} lo que lo hace un modelo muy amplio que toma en cuenta diversos aspectos: comorbilidades, diagnósticos de ingreso, medicamentos administrados, signos vitales, exámenes de laboratorio, soporte ventilatorio, etc.

Entre las ventajas de este modelo es que no existe ningún otro de este tipo que se haya hecho en la población mexicana ni chihuahuense, además, de que no se conoce de forma exacta el pronóstico en nuestro entorno, por lo que su implementación podría llegar a ser de gran ayuda en los hospitales, prediciendo qué pacientes son más propensos a fallecer en las primeras 24 horas de estancia intrahospitalaria, lo que podría ayudar al personal de salud a poner más atención a esos pacientes.

Si bien por el momento no puede implementarse en la práctica debido a las limitaciones mencionadas, se pretende recolectar más muestra para volver a adiestrar y validar el modelo con nuevos datos, que aprenda sobre ellos, mejorar su rendimiento y, finalmente, aplicarse en los pacientes.

Por tanto, se tendrá una nueva y mejor opción de herramienta de pronóstico en pacientes con sepsis, que tome en cuenta múltiples variables, desde signos vitales, estado clínico, exámenes de laboratorio hasta comorbilidades, medicamentos administrados, entre otras. Por último, se tendrá una herramienta novedosa, objetiva, práctica, económica y de gran utilidad para el personal de salud y para el paciente.

Limitantes del estudio

Debido a que es un estudio observacional, los datos que se analizaron fueron recolectados

durante la estancia hospitalaria de los pacientes, por lo que no se tuvo un control tan estrecho de las variables, como la toma de los signos vitales. Hubo que eliminar estudios de laboratorio y pacientes debido a que no todos contaban con todos los estudios. Esto fue causa de un tamaño de muestra reducido para este trabajo, así como una base de datos desequilibrada. Por lo que deberán efectuarse trabajos posteriores para mejorar el rendimiento del modelo.

Consideraciones éticas

Los autores declaran que para esta investigación no se han llevado a cabo experimentos en humanos o animales, se han seguido los protocolos para la publicación de datos de pacientes de acuerdo con la Ley General de Salud, se obtuvo autorización del comité de ética de la institución para el uso, análisis y reporte de la información utilizada.

CONCLUSIONES

Según los resultados obtenidos en esta investigación, no fue posible validar de forma precisa un modelo de inteligencia artificial. En primer lugar, porque no se contaba con una base de datos muy grande y, en segundo, porque estos datos no estaban equilibrados, es decir, había más pacientes que no habían fallecido que los que sí murieron. A pesar de estos obstáculos, el modelo de redes neuronales tuvo un rendimiento bueno, tanto en la primera validación como en la validación cruzada, obteniendo métricas muy parejas o ligeramente superiores a las de las escalas de pronóstico actuales. Por ello, el siguiente objetivo será recolectar más muestra y añadirla al algoritmo, para que aprenda de estos nuevos datos.

A pesar de no obtener los resultados esperados y deseados en la prueba piloto, se espera que al recolectar más datos e introducirlos al algoritmo, éste sea capaz de generar predicciones más pre-

cisas y en un futuro no muy lejano aplicarse en las salas de cuidados intensivos, de piso, etc., y que sea de gran ayuda para mejorar la atención hospitalaria de los pacientes con sepsis.

REFERENCIAS

1. Gyawali B, Ramakrishna K, Dharamoon AS. Sepsis: The evolution in definition, pathophysiology, and management. *SAGE Open Medicine* 2019; 205031211983504. <https://doi.org/10.1177/2050312119835043>.
2. Gorordo-Delsol LA. Sepsis: el enemigo oculto entre líneas. *Rev Med Inst Mex Seg Soc* 2017; 5 5(4): 423.
3. Qingqing M, Jay M, Hoffman JL, Calvert J, Barton C, Shimabukuro D, Shieh L, et al. Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. *BMJ Open* 2018; 1: e017833. doi: 10.1136/bmjopen-2017-017833.
4. Hou N, Li M, He L, Xie B, Wang L, Zhang R, et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. *J Transl Med*. 2020; 18: 462. <https://doi.org/10.1186/s12967-020-02620-5>.
5. Wernly B, Mamandipoor B, Baldia P, Jung C, Osmani V. Machine learning predicts mortality in septic patients using only routinely available ABG variables: a multi-centre evaluation. *Int J Med Inform* 2021; 145: 104312. doi: 10.1016/j.ijmedinf.2020.104312.
6. Kong G, Lin K, Hu Y. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU. *BMC Med Inform Decis Mak* 2020; 20 (1): 251. doi: 10.1186/s12911-020-01271-2.
7. Kok C, Jahmunah V, Oh SL, Zhou X, Gururajan R, Tao X, et al. Automated prediction of sepsis using temporal convolutional network. *Comput Biol Med* 2020; 127 (103957): 103957. <http://dx.doi.org/10.1016/j.compbiomed.2020.103957>.
8. De Alencar Saraiva JL, Becker OM, Silva E, Kadirkamanatan V, et al. Sensitivity analysis-based sepsis prognosis using artificial intelligence. *Res Biomed Eng* 2020; 36: 449-461. <https://doi.org/10.1007/s42600-020-00083-7>.
9. Jaimes F, Farbiarz J, Alvarez D, Martínez C. Comparison between logistic regression and neural networks to predict death in patients with suspected sepsis in the emergency room. *Crit Care* 2005; 9 (2): 150-6. doi: 10.1186/cc3054.
10. Hasegawa D, Yamakawa K, Nishida K, Okada N, Murao S, Nishida O. Comparative analysis of three machine-learning techniques and conventional techniques for predicting sepsis-induced coagulopathy progression. *J Clin Med* 2020; 9 (7): 2113. <https://doi.org/10.3390/jcm9072113>.
11. Rodríguez A, Mendoza D, Ascuntar J, Jaimes F. Supervised classification techniques for prediction of mortality in adult patients with sepsis. *Am J Emerg Med* 2021; 45: 392-7. doi: 10.1016/j.ajem.2020.09.013.