



REVISTA MÉDICA CLÍNICA LAS CONDES

<https://www.journals.elsevier.com/revista-medica-clinica-las-condes>

ARTÍCULO ESPECIAL

Modelos predictivos en salud basados en aprendizaje de maquina (*machine learning*)

Predictive models in health based on machine learning

Javier Mora Pineda^{a,b,c} ✉

^a Departamento Cirugía Cardíaca, Vascular y Tórax, Clínica Las Condes. Santiago, Chile.

^b Unidad de ECMO, Clínica Las Condes. Santiago, Chile.

^c Unidad de Ciencia de Datos Clínicos, Clínica Las Condes. Santiago, Chile.

INFORMACIÓN DEL ARTÍCULO

Historia del Artículo:

Recibido: 30 09 2022

Aceptado: 08 11 2022

Key words:

Data Sciences; Machine Learning; Artificial Intelligence; Predictive Models; Clinical Decision Support Systems; Cardiac Surgery, Data.

Palabras clave:

Ciencia de Datos; Aprendizaje de Máquina; Inteligencia Artificial; Modelos Predictivos; Sistema de Soporte de Decisión Clínica; Datos.

RESUMEN

Encontrar la causalidad en medicina es el mayor interés de la investigación científica, para luego generar intervenciones que traten o curen la enfermedad. La mayoría de los modelos estadísticos clásicos permiten inferir asociación, y solo pocos diseños logran demostrar causa efecto con una adecuada metodología y sólida evidencia. La medicina basada en la evidencia respalda sus hallazgos en modelos que desde una hipótesis salen a buscar datos para demostrarla o descartarla. Ello también aplica en la elaboración de modelos predictivos que sean confiables y que produzcan algún impacto en la práctica clínica.

La gran cantidad de datos que se están almacenando en los registros clínicos electrónicos y el mayor poder computacional, hacen que las técnicas de aprendizaje de máquina tengan un rol preponderante en el desarrollo de nuevos análisis predictivos y reconocimiento de patrones no conocidos con estos modelos de cómputo, que junto con cambiar la mirada desde los datos a la información, van incorporándose cada vez más en la práctica clínica diaria, con mayor precisión y velocidad para la toma de decisiones.

En el presente artículo, se pretende entregar algunas bases teóricas y evidencia de cómo estas técnicas computacionales modernas de aprendizaje de máquina han permitido llegar a mejores resultados y están siendo cada vez más utilizadas.

ABSTRACT

Finding causality in medicine is of great interest in research, in order to generate interventions that treat or cure the disease. Most classical statistical models allow association to be inferred, and only a few designs are able to demonstrate cause and effect with an adequate methodology and solid evidence. Evidence-based medicine supports its findings in models that go from a hypothesis to search for data to prove or rule it out. This also applies to the development of predictive models to be reliable and to produce impact in clinical practice.

The large amount of data stored in electronic health records and greater computational power mean that machine learning techniques can play a preponderant role in the development of new predictive analyzes and recognition of unknown patterns with these modern computational models. These models, along with

✉ Autor para correspondencia

Correo electrónico: jmorap@clinicalascondes.cl

<https://doi.org/10.1016/j.rmcl.2022.11.002>

e-ISSN: 2531-0186/ ISSN: 0716-8640/© 2021 Revista Médica Clínica Las Condes.

Este es un artículo Open Access bajo la licencia CC BY-NC-ND

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



changing the view from data to information, are increasingly being incorporated into daily clinical practice, providing greater precision and speed for supporting decision making. The intent of this review is to provide theoretical bases and evidence of how these modern computational techniques of machine learning have allowed to achieve better results and they are being widely used. This article will review the most relevant aspects of health data science in Latin America.

INTRODUCCIÓN

Cuando, en 1854, John Snow hacía un análisis crítico del comportamiento de la pandemia del cólera en Londres, nunca pensó que sentaría las bases de la epidemiología moderna. Snow, con mente esclarecida, no sólo asoció muchas variables bio-demográficas de causa-efecto para su investigación, sino que además lo graficó en un mapa (Figura 1). Esto permitió visualizar claramente la distribución de los casos de fallecimiento por cólera en torno a la bomba de agua contaminada más cercana en Broad Street logrando convencer a las autoridades para tomar medidas en torno a ello y comenzar a combatir la pandemia en forma eficiente¹.

A partir de ese modelo, se comenzó a explorar otros datos de salud en diversos escenarios, con la misma lógica de encontrar explicaciones y/o asociaciones entre distintos fenómenos y la salud. Comienza a tomar importancia el registro sistematizado de datos, y su posterior análisis. Modelos matemáticos y estadísticos van siendo gradualmente aplicados, permitiendo formular hipótesis y comprobar su veracidad o rechazarlas en torno a un problema.

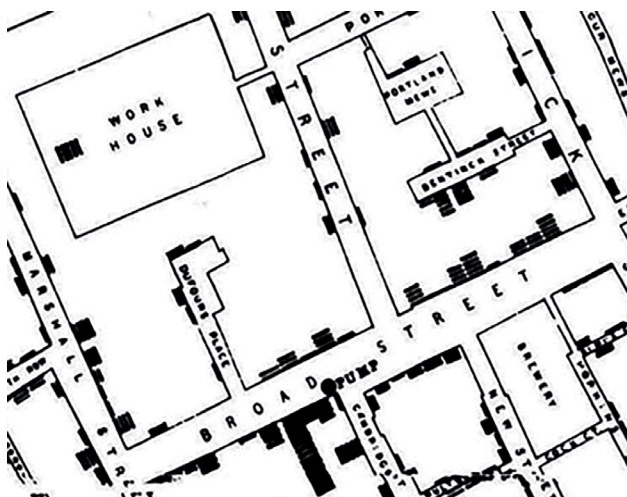
Es conocido, además, cómo en la investigación científica en salud, tanto en ciencias básicas como clínicas, han surgido importantes hallazgos y descubrimientos de forma inesperada o no planifica-

da en el contexto de un modelo de investigación para otro fin. Un ejemplo clásico es el descubrimiento de la penicilina. Esta podría ser una analogía, guardando las proporciones, a lo que en minería de datos se expresa como “dejar que los datos hablen”, es decir, que, mediante el procesamiento analítico de datos, estos entreguen nueva información (agrupación o *clustering*, clasificación, distribución, etc.) no considerada o no planteada previamente.

En la bioestadística clásica (como sustento metodológico de la medicina basada en la evidencia), el modelo esencial consiste en plantear una hipótesis y un método de investigación adecuado que satisfaga de la mejor manera la evidencia que se pueda lograr, y una vez establecido lo anterior salir a buscar los datos que permitan satisfacer el modelo para responder la hipótesis o pregunta de investigación. Dicho sea de paso, existe también un alto porcentaje de mal interpretación de resultados, como la significancia estadística en algunas publicaciones² (Figura 2).

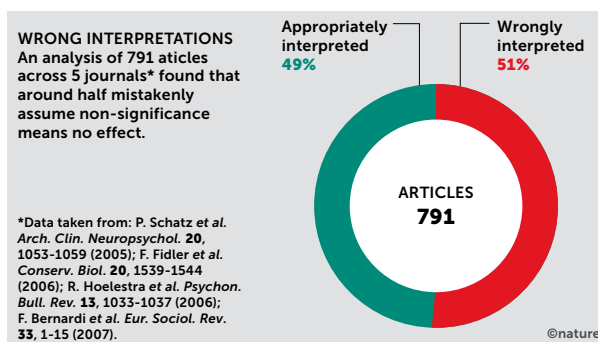
El cambio de paradigma que nos entrega el aprendizaje automático y la ciencia de datos, especialmente el procesamiento de grandes volúmenes de datos, es precisamente una nueva mirada distinta al enunciado para la bioestadística, donde “desde los datos” se aplican modelos computacionales que permiten obtener nueva información relevante al dominio en estudio; cuya definición o determinación de las características del problema que van a influir en la respuesta buscada (predicción o clasificación automática) son entregadas por el propio modelo de aprendizaje de máquina, y no sólo por el investigador, minimizando sesgos y obteniendo significativamente mejores resultados al de los métodos tradicionales. Estas técnicas tienen importantes aplicaciones en medicina y en el campo de la salud en

Figura 1. Mapa confeccionado por John Snow de las muertes por cólera ocurridas en el área de Broad Street



La bomba de agua (*pump*) se ubica en la intersección de Broad y Cambridge Street. Las barras negras corresponden a muertes. Se observan también la cervecería (*Brewery*) y la hospedería (*Work House*). Fuente: Cerda J y Valdivia G¹.

Figura 2. Inadecuada interpretación de la significancia estadística en artículos publicados



Fuente: Amrhein V. et al.².

general; una importante utilidad se ha demostrado en los sistemas de apoyo a la toma de decisiones clínicas³.

Con el aprendizaje automático, los registros clínicos completos son entradas para el aprendizaje de algoritmos. Estos datos no necesariamente deben ser estructurados o tabulares, pudiendo ingresar imágenes, texto libre, audio, videos, series de tiempo, etc. Los modelos resultantes pueden ser subsecuentemente utilizados para ayudar a los profesionales de la salud a precisar diagnósticos a futuros nuevos pacientes (nuevos datos). De esta manera, el estudio y diagnóstico de un paciente puede ser más rápido, preciso y confiable.

En la presente revisión, se pretende entregar algunas bases teóricas y evidencia de cómo estas técnicas computacionales modernas de aprendizaje automático han permitido no sólo llegar a mejores resultados, sino también al cómo y por qué están siendo cada vez más utilizadas en la práctica clínica con el procesamiento en tiempo real, o cuasi real, del inmenso volumen de datos que obtenemos de las distintas fuentes de registros clínicos, biomarcadores, parámetros fisiológicos de pacientes y personas sanas (con sensores específicos o con relojes inteligentes), permitiendo predecir cada vez con más exactitud y celeridad situaciones de salud, facilitando una oportuna intervención; en forma inmediata para las emergencias y en forma preventiva para controles seriados.

(Nota aclaratoria: se intenta utilizar al máximo la traducción o interpretación de términos desde el inglés; sin embargo, dado que la mayoría de las publicaciones de ciencias de la computación y ciencia de datos se publican en inglés, muchos de los principales términos surgen inicialmente en ese idioma, al igual que el uso de librerías y códigos de programación. En el presente artículo, los tér-

minos en inglés se presentan en cursiva y entre paréntesis cuando corresponde, para una correcta interpretación técnica original o de su expresión en español).

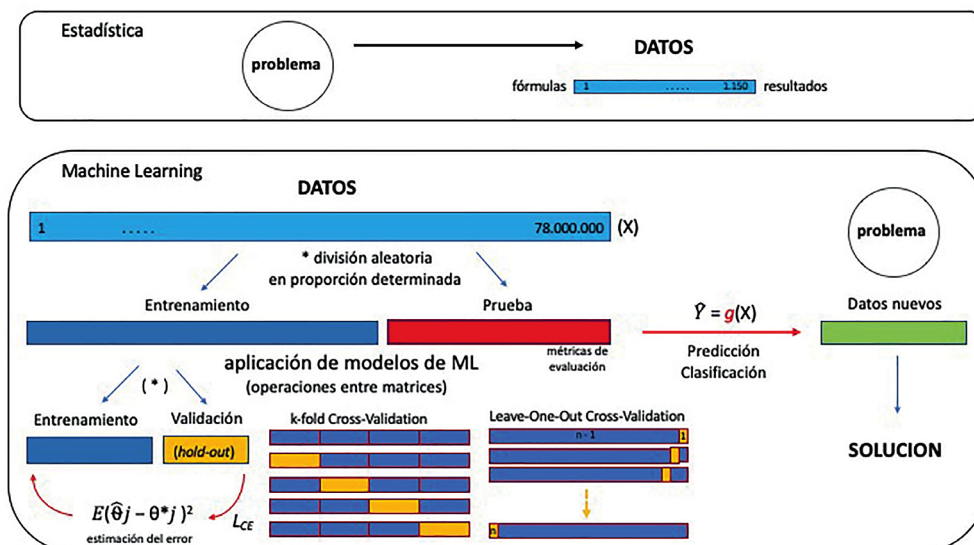
APRENDIZAJE DE MÁQUINA VERSUS ESTADÍSTICA

Basados en conceptos generales de ciencia de datos tratados en un artículo previo de esta revista⁴, y en el libro de investigadores de la Universidad de Chile, "Una Mirada a la Era de los Datos"⁵, esta revisión se centra en el desarrollo de los modelos más utilizados para la tarea de predecir un resultado final en el dominio de la medicina. De aquellos conceptos, es oportuno rescatar el de aprendizaje de máquina de Mitchell: "se dice que un programa computacional aprende de una experiencia E con respecto a algún tipo de tarea T con una medida de rendimiento P , si su rendimiento en la tarea T medido por P mejora con la experiencia E "⁶.

Una forma de expresar la diferencia metodológica entre estadística y aprendizaje de máquina es que la primera implica probar hipótesis, mientras que lo segundo implica la tarea de construir conocimiento desde los datos y almacenarlo en alguna forma de cómputo, como modelos matemáticos, algoritmos o cualquier otra forma computacional que pueda ayudar a determinar patrones o predecir resultados, y que además puedan ser parte de otros modelos que requieran algoritmos ya entrenados para un fin específico (transferencia de aprendizaje o *transfer learning*) (Figura 3).

Es, sin embargo, pertinente dejar en claro que la estadística no es excluyente a la ciencia de datos y aprendizaje automático, por cuanto una amplia variedad de técnicas estadísticas es aplicada en estos modelos computacionales, especialmente durante la evaluación del desempeño de cada uno y al compararlos entre sí, de for-

Figura 3. Esquema del paradigma de los datos en estadística y en las técnicas de aprendizaje de máquina



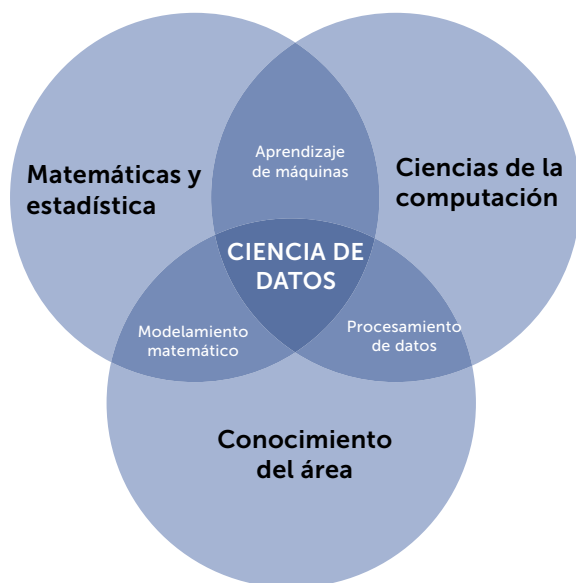
ma dinámica e iterativa. En una tarea de clasificación, por ejemplo, el resultado es una distribución de probabilidades de las distintas posibilidades. Además, el concepto de la ciencia de datos incluye entre sus disciplinas a la estadística, como se esquematiza en la figura 4.

En el pasado, se han desarrollado métodos en forma paralela tanto para estadística como para aprendizaje de máquina; cuatro estadísticos publicaron en 1980 el libro *"Classification and Regression Trees"*, cuyas técnicas estadísticas han sido ampliamente adaptadas por investigadores de ciencias de la computación para mejorar rendimientos de clasificación y para que un procedimiento sea computacionalmente bien organizado⁷.

En este contexto, resulta interesante analizar el por qué estos modelos "aprenden" de los datos, sin tener que necesariamente programar o aplicar instrucciones, fórmulas o reglas explícitas paso a paso y a todo el conjunto de datos (el cambio de paradigma de datos en medicina, enunciado en la introducción).

En un estudio de casos y controles de la bioestadística tradicional, por ejemplo, los datos son registrados en ambos grupos y luego se le aplican fórmulas estadísticas al conjunto total de datos de cada grupo, obteniendo un resultado final, con el cual se analiza el problema en estudio. Similar situación ocurre en las ramas de un estudio prospectivo aleatorizado doble ciego, donde las inferencias estadísticas se realizan igualmente en el conjunto de datos de una sola vez.

Figura 4. Diagrama del concepto de la Ciencia de Datos y las disciplinas que la reúnen



Usado con permiso de los autores. de "Una Mirada a la era de los Datos", Jocelyn Dunstan, Alejandro Maass, Felipe Tobar, Editorial Universitaria 2022. ISBN 9789561128989⁵.

En modelos de aprendizaje automático, por su parte, el conjunto de datos se divide aleatoriamente en 3 grupos: entrenamiento (*train*), validación (*validation*) y prueba (*test*). El primer grupo es la entrada del algoritmo (alguno de los algoritmos utilizados según la tarea a resolver), con lo cual "aprende" a resolver la tarea (básicamente, asigna ciertos valores a ciertas constantes según las funciones utilizadas); con el segundo grupo se evalúa qué tan bien resuelve la tarea con esos valores "aprendidos" (o dicho de otro modo, qué tanto error comete en la tarea), de tal manera de tener la oportunidad de ajustar algunos parámetros del algoritmo (*fine-tuning*) que se puede volver a correr con el grupo de entrenamiento para mejorar la métrica del grupo de validación; y finalmente, con el grupo de prueba o test se ven los resultados de cómo se va a comportar el algoritmo con datos nuevos en el futuro (generalización) (Figura 3).

Uno de los principales objetivos de usar métodos de aprendizaje automático se refiere a la clasificación y predicción. Los términos clasificación y predicción han sido usados indistintamente⁸, pero en general se adopta la predicción para un resultado "binario" (presente o ausente, sí o no, es o no es, etc.) mientras que la clasificación se usa para determinar 1 resultado dentro de un grupo de 3 o más posibles resultados según la tarea a resolver. La clasificación o la predicción pueden ser usados para extraer modelos que describen clases o grupos de datos relevantes frente a otros o para predecir tendencia futura de datos nuevos de un mismo problema⁹. Ambas tareas tienen la capacidad de generalizar sobre un conjunto de datos, lo que significa la capacidad de identificar nuevos resultados con nuevos datos a partir de los datos previos usados en el entrenamiento del algoritmo. A su vez, es importante señalar que, al resolver problemas de clasificación y predicción, éstos son sensibles a los datos faltantes o perdidos de un repositorio¹⁰.

APLICACIONES DE MODELOS PREDICTIVOS CON APRENDIZAJE AUTOMÁTICO

Los principales modelos predictivos de aprendizaje automático aplicados en medicina han sido: redes neuronales (tanto redes *fully connected*, como las convolucionales y recurrentes), support vector machine, árboles de decisión, random-forest, regresiones lineales, modelos bayesianos (*naive bayes*), y vecinos más cercanos, principalmente. En particular, las redes convolucionales, propuestas por Yann LeCun en 1989, son algoritmos de aprendizaje automático basados en el funcionamiento del cortex visual del ojo animal, propuestos por Yann LeCun en 1989, que le permiten "ver" al computador^{11,12}, con una arquitectura de redes neuronales multicapa en que la imagen es dividida en campos receptivos o segmentos de la imagen que van pasando a través de capas convolucionales (operación matemática que hace la integral del producto de dos funciones, o señales, en una tercera función). Esto le permite extraer distintas características de la imagen inicial en cada capa, para luego poder clasificarla. En medicina, esta técnica ha permitido procesar diversos tipos de imágenes para predecir un diagnóstico, desde biopsias hasta radiografías, scanner, resonancia nuclear magnética y fotos de distintas patologías.

Las distintas aproximaciones del aprendizaje automático se basan en la identificación de fuertes asociaciones de datos, pero libres de un fundamento teórico (los modelos denominados “cajas negras”). La confusión hace que sea un salto sustancial en la inferencia causal para identificar factores modificables que realmente puedan tener repercusión en alterar resultados. Como es sabido, asociación no implica causalidad. Muchas predicciones de estos modelos pueden ser altamente precisos, principalmente en casos cuyo resultado probable ya es obvio para el profesional médico. La última milla de la implementación clínica termina siendo la tarea realmente crítica de predecir eventos, lo suficientemente temprano como para que una intervención relevante influya en las decisiones de atención, resultados y pronósticos de salud¹¹. Las técnicas de predicción usando aprendizaje automático ofrecen cada vez más herramientas de apoyo a la toma de decisiones de gestión de riesgos y alertas tempranas¹².

En neurociencias, Liu, del Hospital Universitario de Taiwan, desarrolló un modelo para predecir muerte cerebral usando redes neuronales denominado EANN AAN¹³. Rughani comparó modelos predictivos de redes neuronales usados en neurocirugía con la habilidad de predecir sobrevida con modelos de regresión lineal y con el de los cirujanos, obteniendo resultados significativamente mejores¹⁴. Güler desarrolló un modelo predictivo con redes neuronales para asignar severidad en daño cerebral traumático, con una precisión del 91%¹⁵. Usando árboles de decisión y regresión lineal, Low predijo resultados basados en una escala de Glasgow dicotomizada con una precisión de hasta el 80%¹⁶.

En psiquiatría, un modelo predictivo basado en datos clínicos usando redes neuronales profundas tuvo un resultado de 0,73 y 0,67 área bajo la curva ROC para predecir desorden de ansiedad generalizado y desorden depresivo mayor, respectivamente¹⁷. Algunos de los temas más discutidos en el aprendizaje automático, como el sesgo y la equidad, la limpieza de datos y la interpretación de los modelos, pueden ser poco familiares en el uso de modelos predictivos basados en neuroimágenes en psiquiatría. De manera similar, la investigación diagnóstica y la selección de características basadas en el cerebro para la intervención psiquiátrica son temas modernos que se ha visto que son adecuados abordar con modelos predictivos de aprendizaje automático¹⁸.

En genética y biología molecular, investigadores han desarrollado y evaluado varios nuevos modelos predictivos de información evolutiva en la transcripción del ADN utilizando técnicas de *Support Vector Machine*, para determinar los sitios de unión del material genético (específicamente, “secuencias de residuos de unión”) de ADN y ARN. Sus hallazgos mostraron que estos clasificadores tuvieron 77,3% de sensibilidad y 79,3% de especificidad para la predicción de residuos de unión a ADN, y un 71,6% de sensibilidad y 78,7% de especificidad para las predicciones del sitio de unión al ARN, importante para estimar posibles mutaciones genéticas. Este hallazgo

demostró que el clasificador de Support Vector Machine era mejor y más preciso que los modelos existentes para predecir los sitios de unión molecular en la transcripción genética¹⁹.

En oftalmología, un estudio comparó tres modelos de aprendizaje automático para predecir la necesidad de una intervención quirúrgica en el glaucoma primario de ángulo abierto usando registros clínicos, siendo en este trabajo la regresión logística multivariada el más efectivo para discriminar pacientes con una enfermedad progresiva que requieren cirugía (área bajo la curva ROC 0,67)²⁰.

En gestión clínica, aplicaciones para disminuir la no presentación de un paciente a una cita médica (consulta o examen) prediciendo su comportamiento al agendamiento, ha sido de relevante importancia, ya que este ausentismo asciende entre un 10 a un 20% a nivel mundial, y hasta un 30% en algunos hospitales en China²¹. Con técnicas basadas en redes neuronales profundas (*deep learning* o *representation learning*) se ha podido reducir entre un 6 a un 14% la tasa de paciente no presentado (*no show patient*) y aumentar la asistencia efectiva de pacientes, optimizando el uso de recursos y tiempo de atención de los centros de salud. Estos modelos predictivos se alimentan de técnicas de comunicación automatizada de mensajes de texto, entregando a su vez información adecuada y precisa a los pacientes en tiempo y forma para evitar retrasos y ausentismos, pudiendo también automatizar el re-agendamiento de espacios libres generados por los pacientes o que el sistema predice. En Chile, importantes investigaciones se han desarrollado en esta área por el grupo del Centro de Modelamiento Matemático de la Universidad de Chile, estudiando y aplicando estas tecnologías en tres centros de salud pública²²; así también, a través de soluciones basadas en redes neuronales una start-up chilena ha logrado mejorar el acceso a la salud en hospitales públicos y privados optimizando la gestión del agendamiento²³.

El problema de evitar una re-hospitalización no planificada luego del alta de un paciente ha sido estudiado usando redes neuronales. En 2007 Medicare de EE. UU. reportó un 17% de este tipo de readmisiones hospitalarias dentro de 30 días del alta, estimándose que un 76% era potencialmente evitable, representando, en ese entonces, \$15 mil millones de dólares en gastos para Medicare e incidiendo también como una medida de calidad de atención²⁴. Éstas son algunas razones por las que resulta interesante encontrar métodos que puedan predecir el riesgo de reingreso²⁵. El 2020 Clínica Las Condes publicó un estudio que por primera vez en la literatura reportó resultados de esta tarea con un modelo, en español, de datos clínicos altamente no estructurados; el trabajo arrojó un área bajo la curva ROC de 0,76 para la predicción, utilizando el modelo de redes neuronales denominado *Long Short Term Memory* (LSTM)²⁶. Se usaron como entradas datos demográficos, motivos de consulta, procedimientos, diagnósticos, prescripciones y notas clínicas de registros de 9 años²⁵. Este resultado fue muy similar a lo publicado

con registros clínicos en inglés para la época²⁷, lo que le entrega un gran valor y ha motivado a continuar trabajando actualmente en esta línea, buscando mejorar las métricas del modelo y su prueba en sub-áreas del ámbito hospitalario.

El 2021 una librería del lenguaje de programación Python (el más utilizado en ciencia de datos y aprendizaje de máquina) fue especialmente elaborado para modelos predictivos en salud: PyHealth²⁸. Consiste en módulos especiales para el preprocesamiento de datos, modelamiento predictivo (con 30 modelos disponibles) y de evaluación, de tal manera de diagramar modelos complejos de datos médicos de una forma simple y con pocas líneas de código.

PROYECCIONES FUTURAS

La investigación en ciencias de la computación y ciencias de datos aplicadas a salud continúa mejorando la precisión de las predicciones clínicas, pero es posible que incluso un modelo de predicción perfectamente ajustado (fine-tuning) no se traduzca en una mejor atención clínica. Incluso una predicción muy precisa del resultado de un problema en salud no dice qué hacer o cómo hacerlo si se desea cambiar ese resultado; de hecho, ni siquiera se puede asumir que es posible cambiar los resultados previstos.

Modelos predictivos basados en aprendizaje automático proveen actualmente un gran soporte al conocimiento y experiencia clínica de profesionales. Para reducir la subjetividad, muchos sistemas expertos han sido creados para codificar y combinar el conocimiento médico. Métodos predictivos pueden ser integrados a estos sistemas, contribuyendo a reducir sesgos y subjetividad, a la vez que puede proveer potencialmente nuevo conocimiento médico en distintas áreas de la medicina, o en diversas zonas geográficas donde se aplique (ajustando un modelo de aprendizaje automático con datos “locales” puede ser más eficiente que aplicar una fórmula realizada con población de otros países). Por lo tanto, la predicción cada vez más precisa y acuciosa de los posibles resultados de un paciente durante las distintas etapas del proceso de atención plantea un desafío constante y cambiante en salud. Además, las consideraciones éticas de las distintas aplicaciones de la inteligencia artificial y el aprendizaje de máquina es motivo actualmente de importante discusión y análisis en la comunidad científica especializada, tema que además se extiende a la percepción de la población general de estas “nuevas” tecnologías aplicadas a la salud.

Realizar estudios de predicciones de resultados de problemas clínicos es muy relevante hoy en día porque pueden ayudar al equipo médico a tomar decisiones más precisas con técnicas de aprendizaje de máquina que sean de fácil implementación y de bajo costo. La evolución de estos modelos y la aparición de otros nuevos en el futuro proporcionarán sin duda predicciones cada vez más precisas y dinámicas en la práctica clínica diaria.

La librería PyHealth, con la solidez y la escalabilidad en mente, está en una constante evolución para lograr mejores prácticas de desarrollo, pruebas e integración interactiva, con miras a permitir la aplicación cada vez mayor de algoritmos y otras librerías de Python ampliamente usadas a los datos clínicos²⁸.

CONCLUSIONES

La predicción no es algo nuevo en medicina. Desde puntajes de riesgo para guiar la anticoagulación (CHADS2) y el uso de fármacos para el colesterol alto (ASCVD), a la estratificación de riesgo de pacientes de unidades de cuidados intensivos (APACHE), entre muchas otras, las predicciones clínicas basadas en datos son rutinarias en la práctica clínica. Actualmente, en combinación con técnicas modernas de aprendizaje automático, diversas fuentes de datos clínicos nos permiten generar rápidamente o en tiempo real modelos de predicciones para miles de tareas o problemas clínicos similares. Además, se ha podido incluir en la elaboración de estos modelos distintos tipos de datos no estructurados que antes no era posible procesar como parte de un modelo predictivo con métodos tradicionales (por ejemplo, gran cantidad de texto libre ampliamente usado en los registros clínicos, imágenes directas desde su captura, visión por computadora, etc.). En salud, el tiempo del diagnóstico es un factor crucial en el pronóstico del paciente. El potencial de aplicabilidad de esta aproximación del uso de los datos clínicos es sustancial, tanto para sistemas de alertas críticas tempranas como en el diagnóstico por imágenes de alta precisión, optimizar la gestión de citas clínicas, entre otros.

A pesar de que evidentemente los algoritmos predictivos no pueden eliminar la incertidumbre de la toma de decisiones en medicina, sí logran mejorar u optimizar la asignación de recursos en la atención médica, tanto humanos como físicos. Por ejemplo, el índice de gravedad de embolia pulmonar (PESI, *Pulmonary Embolism Severity Index*) es un modelo predictivo ampliamente validado de riesgo de muerte por esta condición²⁹ y otras priorizan de una manera adecuada (por no decir justa) a pacientes en espera de trasplante de hígado por medio del puntaje MELD (*Model for End-stage Liver Disease*). Sistemas de alerta temprana, que con técnicas tradicionales hubieran demorado años en desarrollarse, ahora pueden aplicarse y optimizarse de forma rápida a partir de datos del mundo real y de forma continua a medida que ocurren. Igualmente, las redes neuronales de aprendizaje profundo son capaces de reconocer de manera rutinaria imágenes patológicas con una alta precisión que antes se creía imposible.

Estimo necesario también enfatizar que estos modelos, como tantos otros en las distintas áreas del aprendizaje automático, no son infalibles y al igual que cualquier humano se equivocan en sus resultados en un porcentaje de veces. De hecho, las métricas con las que se entrenan y evalúa el rendimiento de un algoritmo en desarrollo se basan principalmente en “minimizar el error” (como el pro-

medio de error cuadrático (RMSE, *Root Mean Square Error*), entropía cruzada, entre otros). En la transición contemporánea que estamos viviendo de incorporar cada vez más estas técnicas a la práctica clínica diaria, se tiende a esperar que no cometan errores por el sólo hecho de ser realizadas por máquinas o computadores modernos denominados inteligentes, y que pueden llegar a ser erróneamente desestimados por no tener una precisión del 100%.

Discutir si los modelos predictivos de aprendizaje de máquina van a llegar a ser “más inteligentes” que los mismos profesionales de salud pareciera ser irrelevante. Es indiscutible que, en la salud, como en todas las otras industrias, los expertos del dominio son un eslabón fundamental e irremplazable en el modelo de trabajo de la ciencia de datos, que le dan sentido y modulan tanto la

creación como el uso clínico de estos algoritmos (Figura 4). Sin duda, cada vez más, el resultado final de estas herramientas puede asemejarse o acercarse a un “comportamiento humano”, integrando varios modelos predictivos (eligiendo incluso cuáles usar en un caso u otro) en un mismo momento y en muy corto tiempo para tomar decisiones complejas (de diagnósticos o tratamientos) de forma casi automática o en tiempo real. No obstante, no hay que olvidar que en su lógica más básica no son más que operaciones matemáticas de matrices y tensores de alta dimensionalidad que no piensan, no sienten, ni tienen la intangible interacción humana ni sus virtudes (“ojo clínico”) que ninguna máquina ha reemplazado ni reemplazará, desde los trabajos de Alan Turing a nuestros días y en el futuro.

Declaración de conflicto de interés

El autor certifica que no hay conflicto de interés en relación con lo presentado en este artículo.

REFERENCIAS BIBLIOGRÁFICAS

1. Cerda J, Valdivia G. John Snow, la epidemia de cólera y el nacimiento de la epidemiología moderna. [John Snow, the cholera epidemic and the foundation of modern epidemiology]. *Rev Chil Infectol*. 2007;24(4):331-334. doi: 10.4067/S0716-10182007000400014
2. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019;567(7748):305-307. doi: 10.1038/d41586-019-00857-9
3. Alanazi HO, Abdullah AH, Qureshi KN. A Critical Review for Developing Accurate and Dynamic Predictive Models Using Machine Learning Methods in Medicine and Health Care. *J Med Syst*. 2017;41(4):69. doi: 10.1007/s10916-017-0715-6
4. Mora J. Proyecciones de la ciencia de datos en la cirugía cardíaca. [Projections of data science in cardiac surgery]. *Rev Med Clin Condes*. 2022;33(3):294-306. doi: 10.1016/j.rmcl.2022.05.007
5. Dunstan J, Maass A, Tobar F. Una Mirada a la Era de los Datos. Ed. Universitaria 2022.
6. Mitchell TM. *Machine Learning*. McGraw-Hill. 1997.
7. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and regression trees*. Routledge; 2017.
8. Carreno A, Inza I, Lozano JA. Eventos raros, anomalías y novedades vistas desde el paraguas de la clasificación supervisada. IX Simposio de Teoría y Aplicaciones de la Minería de Datos. 2018:925-930. Disponible en: https://sci2s.ugr.es/caepia18/proceedings/docs/CAEPIA2018...paper_192.pdf
9. de Hond AAH, Leeuwenberg AM, Hooft L, Kant IMJ, Nijman SWJ, van Os HJA, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *NPJ Digit Med*. 2022;5(1):2. doi: 10.1038/s41746-021-00549-7.
10. Alanazi HO, Abdullah AH, Qureshi KN. A Critical Review for Developing Accurate and Dynamic Predictive Models Using Machine Learning Methods in Medicine and Health Care. *J Med Syst*. 2017;41(4):69. doi: 10.1007/s10916-017-0715-6
11. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput*. 1989;1(4):541-551.
12. LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. In: Arbib, Michael A, (ed.). *The handbook of brain theory and neural networks* (Second ed.). The MIT press. 1995:276-278.
13. Liu Q, Cui X, Abbod MF, Huang SJ, Han YY, Shieh JS. Brain death prediction based on ensemble artificial neural networks in neurosurgical intensive care unit. *J Taiwan Inst Chem Eng*. 2011;42(1):97-107. <https://doi.org/10.1016/j.jtice.2010.05.00>
14. Rughani AI, Dumont TM, Lu Z, Bongard J, Horgan MA, Penar PL, Tranmer BJ. Use of an artificial neural network to predict head injury outcome. *J Neurosurg*. 2010 Sep;113(3):585-90. doi: 10.3171/2009.11.JNS09857
15. Güler İ, Gökçil Z, Gülbandır E. Evaluating of traumatic brain injuries using artificial neural networks. *Expert Syst. Appl*. 2009;36(7):10424-10427. doi: 10.1016/j.eswa.2009.01.036
16. Low D, Kuralmani V, Ng SK, Lee KK, Ng I, Ang BT. Prediction of outcome utilizing both physiological and biochemical parameters in severe head injury. *J Neurotrauma*. 2009 Aug;26(8):1177-82. doi: 10.1089/neu.2008.0841
17. Nemesure MD, Heinz MV, Huang R, Jacobson NC. Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence. *Sci Rep*. 2021;11(1):1980. doi: 10.1038/s41598-021-81368-4

18. Tejavibulya L, Rolison M, Gao S, Liang Q, Peterson H, Dadashkarimi J, et al. Predicting the future of neuroimaging predictive models in mental health. *Mol Psychiatry*. 2022. doi: 10.1038/s41380-022-01635-2
19. Ma X, Wu J, Xue X. Identification of DNA-binding proteins using support vector machine with sequence information. *Comput Math Methods Med*. 2013;2013:524502. doi: 10.1155/2013/524502
20. Baxter SL, Marks C, Kuo TT, Ohno-Machado L, Weinreb RN. Machine Learning-Based Predictive Modeling of Surgical Intervention in Glaucoma Using Systemic Data From Electronic Health Records. *Am J Ophthalmol*. 2019;208:30-40. doi: 10.1016/j.ajo.2019.07.005
21. Fan G, Deng Z, Ye Q, Wang B. Machine learning-based prediction models for patients no-show in online outpatient appointments. *Data Sci Manage*. 2021;2:45-52. doi: 10.1016/j.dsm.2021.06.002
22. Ramírez H, Dunstan J, Montenegro H. Soluciones tecnológicas, basadas en técnicas matemáticas avanzadas de aprendizaje de máquinas. FONDEF IDEAL+D ID19110271
23. Modelo de IA a través de chatbot logra mejorar la agenda de citas médicas en Chile. *Salud Digital, Fundación Carlos Slim*. Disponible en: <https://saluddigital.com/es/noticias/modelo-de-ia-a-traves-de-chatbot-logra-mejorar-la-agenda-de-citas-medicas-en-chile/>
24. Covacevich Stipicich T. Exploring Representations of Icd Codes for Patient Readmission Prediction (Doctoral dissertation, Pontificia Universidad Católica de Chile (Chile)). 2021. Disponible en: <https://repositorio.uc.cl/xmlui/handle/11534/58535>
25. Fierro C, Pérez J, Mora J. Predicting unplanned readmissions with highly unstructured data. 2020. Workshop paper at AI4AH, ICLR 2020. doi: 10.48550/arXiv.2003.11622
26. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9(8):1735-1780. doi: 10.1162/neco.1997.9.8.1735.
27. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med*. 2018;1:18. doi: 10.1038/s41746-018-0029-1
28. Zhao Y, Qiao Z, Xiao C, Glass L, Sun J. PyHealth: A Python Library for Health Predictive Models. arXiv:2101.04209 <https://doi.org/10.48550/arXiv.2101.04209>
29. Wells PS. Integrated strategies for the diagnosis of venous thromboembolism. *J Thromb Haemost*. 2007;5 Suppl 1:41-50. doi: 10.1111/j.1538-7836.2007.02493.x