

# Inteligencia artificial para asistir el diagnóstico clínico en medicina

## RESUMEN

La medicina es uno de los campos del conocimiento que más podrían beneficiarse de una interacción cercana con la computación y las matemáticas, mediante la cual se optimizarían procesos complejos e imperfectos como el diagnóstico diferencial. De esto se ocupa el aprendizaje automático, rama de la inteligencia artificial que construye y estudia sistemas capaces de aprender a partir de un conjunto de datos de adiestramiento y de mejorar procesos de clasificación y predicción. En México, en los últimos años se ha avanzado en la implantación del expediente electrónico y los Institutos Nacionales de Salud cuentan con una riqueza de datos clínicos almacenada. Para que esos datos se conviertan en conocimiento, necesitan ser procesados y analizados a través de métodos estadísticos complejos, como ya se hace en otros países, usando: razonamiento basado en casos, redes neuronales artificiales, clasificadores bayesianos, regresión logística multivariante o máquinas de soporte vectorial, entre otros. Esto facilitará el diagnóstico clínico de padecimientos como: apendicitis aguda, cáncer de mama o hepatopatía crónica. En esta revisión se repasan conceptos, antecedentes, ejemplos y métodos de aprendizaje automático en diagnóstico clínico.

**Palabras clave:** inteligencia artificial, diagnóstico clínico, aprendizaje automático, diagnóstico diferencial, minería de datos, regresión logística, apoyo en decisión clínica.

## Artificial Intelligence to Assist Clinical Diagnosis in Medicine

### ABSTRACT

Medicine is one of the fields of knowledge that would most benefit from a closer interaction with Computer studies and Mathematics by optimizing complex, imperfect processes such as differential diagnosis; this is the domain of Machine Learning, a branch of Artificial Intelligence that builds and studies systems capable of learning from a set of training data, in order to optimize classification and prediction processes. In Mexico during the last few years, progress has been made on the implementation of electronic clinical records, so that the National Institutes of Health already have accumulated a wealth of stored data. For those data to become knowledge, they need to be processed and analyzed through complex statistical methods, as it is already being done in other countries, employing: case-based reasoning, artificial neural networks, Bayesian classifiers, multivariate logistic regression, or support vector machines, among other methodologies; to assist the clinical diagnosis of acute appendicitis, breast cancer and chronic liver

Saúl Oswaldo Lugo-Reyes<sup>1</sup>  
Guadalupe Maldonado-Colín<sup>2</sup>  
Chiharu Murata<sup>3</sup>

<sup>1</sup> Unidad de Investigación en Inmunodeficiencias.

<sup>2</sup> Departamento de Pediatría.

<sup>3</sup> Departamento de Metodología de la Investigación.  
Instituto Nacional de Pediatría, Ciudad de México.

Recibido: enero 2014

Aceptado: marzo 2014

### Correspondencia

Dr. Saúl Oswaldo Lugo Reyes  
Unidad de Investigación en Inmunodeficiencias,  
Torre de Investigación, piso 9  
Instituto Nacional de Pediatría  
Av. del Imán 1  
04530 México, DF  
dr.lugo.reyes@gmail.com

### Este artículo debe citarse como

Lugo-Reyes SO, Maldonado-Colín G, Murata C. Inteligencia artificial para asistir el diagnóstico clínico en medicina. Revista Alergia México 2014;61:110-120.

disease, among a wide array of maladies. In this review we sift through concepts, antecedents, current examples and methodologies of machine learning-assisted clinical diagnosis.

**Key words:** artificial intelligence, clinical diagnosis, machine learning, differential diagnosis, data mining, logistic regression, clinical decision support.

El aprendizaje automático y la minería de datos son ya una realidad en varias áreas del conocimiento. Todos los días, desde hace más de dos décadas, crece el volumen de datos cada vez que alguien captura sus datos de usuario y siempre que se registra actividad o se almacenan *cookies* en la red mundial. Los bancos y las grandes empresas utilizan sus bases de datos para analizar el comportamiento de sus usuarios, empleados y clientes, con el fin de predecir quién va a morir antes, quiénes van a dejar de pagar su hipoteca, quién se va a embarazar pronto, o quién va a dar *click* en una oferta y comprar en línea.<sup>1</sup>

El aprendizaje automático se practica desde hace algunos años, y con buenos resultados, en las ciencias sociales, mercadotecnia, finanzas y ciencias aplicadas. En medicina apenas se ha utilizado,<sup>2</sup> en parte por razones culturales y filosóficas por las que se asume que una computadora nunca será tan capaz como un médico humano; y por el rechazo de algunos médicos a sentirse cuestionados, supervisados o aconsejados por una máquina o por un ingeniero.<sup>3</sup> Así, incluso en las ciencias biológicas y la medicina genómica se usan ya métodos computacionales avanzados; mientras que los clínicos tienen que lidiar con bases de datos cada vez más grandes y complejas recurriendo a métodos estadísticos tradicionales.

Por sus características de complejidad e incertidumbre, la medicina es uno de los campos

del conocimiento que más se puede beneficiar de una interacción con disciplinas como la computación y el aprendizaje automático para fortalecer procesos como el diagnóstico clínico y realizar análisis predictivos acerca de los pacientes y su pronóstico, que resulten en un sistema de salud más eficiente y mejor utilización de recursos. Esta revisión se centrará en la aplicación del aprendizaje automático para asistir en el diagnóstico clínico en medicina.

### **El diagnóstico clínico como proceso cognitivo perfectible**

La tasa de diagnósticos incorrectos en la práctica clínica depende de la fuente y el escenario; así, se ha estimado en 150 de 1,000 pacientes (Health Grades Patient Safety in American Hospitals Study, 2011),<sup>4</sup> o en 10 a 20% (en el servicio de urgencias y en estudios de discrepancias con autopsias).<sup>5</sup> “Error es humano” y en los hospitales causa malestar, daños, gastos y muertes, especialmente en diagnósticos serios como apendicitis, infección y cáncer.

El diagnóstico clínico es un proceso cognitivo complejo que implica capacitación, experiencia, reconocimiento de patrones y cálculo de probabilidad condicional, entre otros componentes menos comprendidos.<sup>6</sup> Los médicos avezados han adquirido con el tiempo una intuición que facilita la tarea, producto de la experiencia de haber visto cientos de casos similares a lo largo

de su carrera.<sup>7</sup> Los clínicos menos experimentados pueden reducir sus errores diagnósticos al hacer consciente y sistemático el proceso (meta-cognición), teniendo en cuenta qué tan confiables son sus apoyos diagnósticos (probabilidad pre-prueba, valor predictivo, razones de verosimilitud) y cuáles son las fuentes más frecuentes de error (fatiga, sesgos, etcétera).<sup>8,9</sup>

¿Cómo piensa un clínico al momento de establecer un diagnóstico? Al enfrentar a un paciente dado, no toma en cuenta de manera aislada y exclusiva el síntoma o signo cardinal (tos crónica, cefalea, dolor abdominal o ictericia) para hacer el diagnóstico diferencial y el abordaje diagnóstico, sino también, y al mismo tiempo, la edad y el sexo, el tiempo de evolución, síntomas acompañantes, fenómenos agravantes y atenuantes, respuesta al tratamiento previo, etcétera.<sup>9</sup> Cada componente de la historia clínica y el examen físico puede considerarse una prueba diagnóstica de mayor o menor utilidad.<sup>10</sup>

Desmitificar y racionalizar el proceso diagnóstico puede conducir a cuestionar el conocimiento empírico y perfeccionar el aprendizaje en las nuevas generaciones. A *grosso modo*, los pasos del proceso incluyen: 1) identificar la información clínica relevante; 2) interpretar su significado; 3) generar hipótesis diagnósticas; 4) probar y refinar dichas hipótesis a partir de coleccionar datos nuevos; y 5) establecer el diagnóstico.<sup>11</sup> En este proceso, la habilidad de integración, la capacidad de introspección para reflexionar acerca del propio desempeño y el desarrollo de la intuición son características que separan a los expertos del resto de sus colegas.

Un médico puede sentirse fatigado o abordar de manera sesgada a un paciente, puede no estar familiarizado con la enfermedad en cuestión o, bien, estar influido por lo que acaba de leer o por la opinión diagnóstica del colega que refirió al paciente (sesgo de disponibilidad)

y, por ende, no considerar un diagnóstico diferencial exhaustivo e incluyente con un abordaje diagnóstico apropiado.<sup>8,11-13</sup> Otro error común es “casarse” con la impresión diagnóstica inicial, y que todos los datos recabados a partir de entonces sólo sirvan para “confirmar” dicha impresión, mientras se descarta información que no es consistente (sesgo de confirmación).<sup>12,14,15</sup>

Contar con una ayuda electrónica o computacional facilita el proceso, porque lo hace sistemático y exhaustivo, no dependiente de la experiencia o del estado actual del clínico; además de que elimina la necesidad de una memoria operacional que sobrepasa la capacidad humana. Bien elaborado, un programa diagnóstico puede ser al menos tan efectivo como un clínico experto y simplificar grandemente el trabajo de los médicos humanos.<sup>16,17</sup>

### El aprendizaje automático en medicina

En 1961, en Utah, Homer Warner desarrolló un sistema de diagnóstico automatizado para diagnosticar correctamente a pacientes con 35 cardiopatías congénitas a partir de 50 atributos clínicos, con un desempeño superior al de los cardiólogos referentes.<sup>18</sup> En la primera mitad de la década de 1970, la Universidad de Stanford desarrolló MYCIN, un sistema experto para ayudar al diagnóstico clínico en infecciones y trastornos de coagulación. Aunque por cuestiones legales nunca se llevó a la práctica, la evaluación de su desempeño mostraba 65% de tratamientos correctamente indicados, lo que era superior al desempeño de cinco humanos expertos en infectología (entre 42.5 y 62.5%).<sup>19</sup> Otros sistemas, como CADUCEUS/INTERNIST-1/Quick Medical Reference e Iliad, se desarrollaron en las décadas de 1970 y 1980 como herramientas educativas electrónicas y para asistir en la toma de decisiones en medicina interna.<sup>20</sup> Su principal dificultad fue una interfaz poco manejable y un tiempo prolongado de consulta, pero en general

han servido como antecedente y han evolucionado hacia nuevas iniciativas y experimentos.

En las últimas décadas se han realizado varios esfuerzos por aplicar el análisis predictivo en hospitales, aseguradoras y sistemas de salud, así como por lanzar sistemas de aprendizaje automático para facilitar el diagnóstico en medicina. Cada vez hay más aplicaciones para el aprendizaje automático, y es un campo en constante expansión. En Estados Unidos, la Fundación Nacional para la Ciencia (NSF) anunció el año pasado una iniciativa presidencial para apoyar la investigación multidisciplinaria de minería de datos y aprendizaje automático, conocida como Big Data, en colaboración con los Institutos Nacionales de Salud, que incluye dinero para investigación, capacitación y esfuerzos colaborativos.<sup>21</sup> El interés en aplicaciones médicas se ha incrementado desde hace tres décadas, y en la Conferencia Internacional de Minería de Datos de 2012 en Siam, las aplicaciones médicas predominaron considerablemente.<sup>22</sup>

### Inteligencia artificial, aprendizaje automático y minería de datos

La estadística tradicional no es suficiente para manejar grandes cantidades de variables, como se encuentran en muchas bases de datos actuales. El aprendizaje automático es conocimiento obtenido al procesar computacionalmente datos de adiestramiento contenidos en esas bases de datos.<sup>23</sup> El reconocimiento de patrones estadísticos es un abordaje para explorar un conjunto de datos y descubrir relaciones previamente insospechadas, sin necesidad de una hipótesis. Los problemas que se plantean y las estrategias para resolverlos se pueden dividir en: agrupamiento (*clustering*), reducción de dimensiones (*dimensionality reduction*) y clasificación.

La *inteligencia artificial* se define como “el campo de la ciencia y la ingeniería que se ocupa

del entendimiento computacional, de lo que comúnmente se conoce como comportamiento inteligente, y de la creación de artefactos que exhiban tal comportamiento”.<sup>24</sup> El *aprendizaje automático* (*machine learning*) es una rama de la inteligencia artificial que trata de la construcción y estudio de sistemas que puedan aprender de los datos. En 1959, Arthur Samuel lo definió como “el campo de estudio que da a las computadoras la habilidad de aprender sin que esto se haya programado explícitamente”. Más formalmente, Tom Mitchell<sup>25</sup> escribió en términos operativos: “Se dice que una computadora aprende de la experiencia (E) con respecto a alguna clase de tareas (T) y medida de desempeño (P), si su desempeño en las tareas T, medida mediante P, mejora con la experiencia E”. La pregunta entonces, según Alan Turing,<sup>26</sup> no es cognitiva (¿pueden las máquinas pensar?), sino operativa: ¿pueden las computadoras hacer lo que nosotros, como entidades pensantes, hacemos?

Una máquina, por ejemplo, puede aprender a clasificar el correo electrónico para distinguir entre *spam* y correo legítimo deseado, y así distribuirlo en las carpetas correspondientes. En el núcleo del aprendizaje automático están la representación y la generalización. Se representan los datos ingresados con base en funciones evaluadas en la instancia de datos. La generalización consiste en que el sistema funcione bien frente a instancias de datos nuevos; es la habilidad de un algoritmo para desempeñarse con exactitud frente a ejemplos nuevos, luego de haberse capacitado en un conjunto de datos de aprendizaje. Su objetivo principal es generalizar a partir de la experiencia: extraer de la distribución de datos de capacitación algo, una función que le permita hacer predicciones útiles en casos nuevos.<sup>27</sup>

La *minería de datos* (*data mining*) consiste en descubrir patrones en bases de datos, en desenterrar propiedades previamente ignoradas de los datos. El aprendizaje automático utiliza minería

de datos como un paso previo para mejorar la exactitud de predicción, o como parte del “aprendizaje no supervisado”. La *minería de textos* (*text mining*) permite extraer meta-datos a partir de textos escritos en lenguaje natural. Por ejemplo, un pdf o cientos de documentos en formato pdf se pueden procesar para extraer meta-datos de interés: ¿en cuántos pacientes de los registrados en la bibliografía el síndrome de Wiskott-Aldrich se manifestó con eccema o con sangrado?

### Los métodos usados y algunos ejemplos

La medicina enfrenta el reto de adquirir, analizar y aplicar una gran cantidad de conocimiento necesario para resolver problemas clínicos complejos.<sup>28</sup> Se han desarrollado programas que ayudan al clínico a formular un diagnóstico, tomar decisiones terapéuticas y predecir el pronóstico o resultado final, y que asisten el aprendizaje médico a partir de la década de 1980.<sup>28,29</sup> Algunos de los abordajes más utilizados que han demostrado mayor efectividad, incluyen: a) redes neuronales artificiales,<sup>30-32</sup> b) métodos de razonamiento basado en casos<sup>33-35</sup> y c) clasificadores bayesianos.<sup>36</sup>

### Análisis discriminante lineal y regresión logística multivariante

Ambos permiten explicar la pertenencia a una categoría a partir de una combinación de sus características o atributos. La regresión logística multivariante es más apropiada cuando las variables dependientes son dicotómicas, mientras que el análisis discriminante lineal se utiliza sobre todo cuando las variables dependientes son de más de dos grupos; sin embargo, las presuposiciones en la regresión son menos rígidas. En el análisis discriminante lineal se asume que las variables independientes son continuas, que tienen una distribución normal y que la variancia es igual en cada grupo.

Mediante regresión logística multivariante, Perkins y su grupo estudiaron en 2012 los registros médicos de 9,023 pacientes a quienes se efectuó amigdalectomía en los siete años previos, en busca de factores de riesgo de hemorragia (casos y controles). Concluyeron que los pacientes operados por amigdalitis crónica tienen el doble de riesgo de sangrar comparados con los operados por apnea obstructiva del sueño y también que los pacientes mayores de seis años sangran más.<sup>37</sup>

### Razonamiento basado en casos

El razonamiento basado en casos se ha utilizado con éxito en el diagnóstico clínico asistido. Se trata de un paradigma de resolución de problemas en el que un problema nuevo se soluciona atendiendo casos similares del pasado.<sup>38</sup> Un abordaje popular es el del vecino más cercano, que consiste en comparar cada caso problema con una serie de casos almacenados; el caso problema nuevo y los casos almacenados consisten en un número de atributos o variables y éstos se comparan para encontrar el caso almacenado que más se parece al caso problema; por ejemplo, en 367 de 388 atributos. De esta manera, se puede seleccionar el material más apropiado para una pieza de aeronáutica, o en el tema que aquí se aborda, la enfermedad conocida que más se parezca a la descripción del paciente. Una vez resuelto, este nuevo caso pasa a formar parte de la biblioteca de casos disponibles para comparar, con lo que el sistema se vuelve cada vez más sólido.

En la Universidad de Tampere, Finlandia, se desarrolla un abordaje de inmunodeficiencias primarias asistido mediante un sistema de conocimiento, una máquina de inferencia y la interfaz del usuario, ya funcional en fase de prototipo, con 708 casos clínicos y 148 atributos o parámetros incluidos. Este grupo cuenta con una amplia experiencia en bioinformática aplicada a inmunodeficiencias primarias ([www.](http://www.)

bioinf.uta.fi); ellos mismos eligieron el agrupamiento no supervisado para reclasificar las inmunodeficiencias en 11 grupos de defectos mediante un consenso de cinco métodos.<sup>39</sup> Su planteamiento es desarrollar un sistema experto asociado con razonamiento basado en casos (*vecino más cercano, leave-one-out strategy*). Los resultados preliminares arrojaron 469 casos correctamente clasificados, una exactitud de 66%, con mediana de exactitud de clasificación por clase (*classwise*) de 50% (Samarghitean y su grupo, 15<sup>th</sup> encuentro bianual de ESID 2012, Florencia, Italia).

El problema con los sistemas expertos es que no son inteligentes porque no aprenden con cada utilización, y no son más capaces que sus creadores de resolver nuevos problemas. Un ejemplo de sistema experto es un árbol de decisiones para apoyar a un usuario cuyo automóvil no arranca (¿encienden los faros? Sí/no; ¿puede hacer sonar el claxon? Sí/no) y así proponer la causa más probable de la avería.

### Redes neuronales artificiales

Las redes neuronales artificiales se han vuelto populares en medicina debido a su flexibilidad y dinamismo. Su estructura consta de varias capas y deben someterse a capacitación antes de ser funcionales.<sup>28</sup> Las neuronas artificiales se comportan como un enjambre, con inteligencia compartida o convergente, que debe llegar a la solución de un problema de la manera más rápida y eficiente. En este sentido, es capaz de evolucionar con el tiempo. El problema con estas redes es que nadie sabe a ciencia cierta cómo funcionan: al principio se promovieron como una imitación de los procesos computacionales cerebrales, luego esto no se pudo verificar y ahora se consideran una especie de caja negra que da un resultado operativamente útil, pero no verificable, por lo que no se pueden usar para enseñar y aprender.<sup>40</sup>

Las redes neuronales artificiales se han utilizado en el diagnóstico clínico, en el análisis de imágenes radiográficas e histopatológicas y en la interpretación de datos usados en terapia intensiva, entre otros campos de la medicina, para optimizar el diagnóstico de, por ejemplo: dolor abdominal agudo, glaucoma, lumbalgia y litiasis vesical.<sup>28</sup> Yoldas y colaboradores probaron la eficacia de estas redes en el diagnóstico de la apendicitis aguda, obtuvieron sensibilidad, especificidad y valor predictivo negativo de 97.2, 88 y 100%, respectivamente,<sup>31</sup> lo que demuestra que esta aplicación es una herramienta diagnóstica útil, especialmente en áreas en las que el ultrasonido y la tomografía axial computada no están disponibles. Prabhudesai y su grupo, en 2008, también utilizaron las redes neuronales artificiales para diagnosticar apendicitis aguda sólo con variables clínicas y de laboratorio,<sup>17</sup> y lograron un desempeño superior al de los clínicos expertos, los estudios de imagen y la escala clínica predictiva de Alvarado.

Seixas y colaboradores, en 2013, usaron redes neuronales artificiales para diagnosticar tuberculosis pleural sólo con parámetros clínicos y el estado de infección por VIH, a partir de una base de adiestramiento de 137 pacientes, considerando patrón de referencia cualquier prueba positiva para tuberculosis.<sup>41</sup> El desempeño fue superior a todas las otras pruebas diagnósticas por separado, incluidas la adenosina deaminasa en líquido pleural, biopsia, cultivo, ELISA y PCR y alcanzó más de 90% de exactitud, sin que se recurriera a procedimientos invasores, con lo que este abordaje podría ser muy útil, sobre todo si hay poca disponibilidad de pruebas y recursos.

### Clasificadores bayesianos

Los clasificadores bayesianos calculan la probabilidad de que un caso con cierto atributo pertenezca a una categoría. Así por ejemplo,

se clasifican de una manera bastante acertada los correos electrónicos como *spam* o correo deseado o legítimo y en medicina se han usado también para predecir recaída de enfermedad o riesgo de cáncer.<sup>28,29</sup> El filtro no sabe esto por adelantado y tiene que capacitarse de la mano del usuario. El programa calcula en lo sucesivo la probabilidad de que cada atributo (la palabra “viagra”, en el ejemplo del correo electrónico) corresponda a una categoría o a otra.

En 2010, Geenen y colaboradores utilizaron clasificadores bayesianos ingenuos (*naïve bayesian classifiers*) para el diagnóstico clínico de la fiebre porcina clásica.<sup>42</sup> Obtuvieron exactitud de 60 a 70% para distinguir pjaras infectadas de las no infectadas, comparable a un constructo conocido como regla diagnóstica.

### **Máquinas de soporte vectorial**

Son modelos de aprendizaje supervisado, con algoritmos asociados, que analizan datos y reconocen patrones para clasificación y análisis de regresión.<sup>43,44</sup> El modelo toma un conjunto de datos ingresados y predice, para cada *input* dado, cuál de dos clases forma el *output*: es un clasificador lineal binario no probabilístico. Con los datos se construye un modelo de representación de los ejemplos como puntos en el espacio, en un hiperplano mapeado con una brecha lo más amplia posible.

En 2011, Tenorio y su grupo utilizaron una base de datos de adiestramiento de 178 pacientes para desarrollar un sistema de apoyo en decisión clínica para el diagnóstico de enfermedad celiaca, mediante cinco técnicas de inteligencia artificial: árboles de decisiones, inferencia bayesiana, vecino más cercano, máquinas de soporte vectorial y redes neuronales artificiales. El método más exacto fue un clasificador bayesiano, con exactitud de 80%, sensibilidad de 0.78, especificidad de 0.80 y área debajo de la

curva (AUC) de 0.84 al medirse con el estándar de referencia, la biopsia intestinal.<sup>36</sup>

### **Sistemas inmunitarios artificiales**

A partir de la década de 1990, se han utilizado algoritmos inspirados en la biología, incluidos las redes neuronales, los algoritmos genéticos y los sistemas inmunitarios artificiales. Éstos imitan procesos complejos, de varias capas, adaptables y tolerantes a ruido o errores, como la selección clonal de los linfocitos, o la hipermutación somática de la generación de anticuerpos.<sup>45,46</sup> Se han utilizado en varias áreas y aplicaciones para procesos como: reconocimiento de patrones, agrupamiento, detección de intrusos y clasificación. Se crean nodos que corresponden a células de memoria y así se facilita el reconocimiento inmediato y consistente en un siguiente encuentro con un problema similar.

En Turquía, en 2009, Kodaz y su grupo desarrollaron un sistema de reconocimiento de patrones por sistemas inmunitarios artificiales basado en la ganancia de información (IG-AIRS) para diagnosticar la función tiroidea a partir de resultados de laboratorio,<sup>47</sup> con exactitud de clasificación de 95.9%.

### **Ensamblados o agregados**

Los métodos compuestos, o ensamblados, que echan mano de dos o más procedimientos, como: Random Forest, Bagging (*bootstrap aggregating*) y Boosting, han alcanzado mejor desempeño que los métodos individuales. Utilizar un método compuesto para trabajar en el mismo problema tiene sentido porque no es más costoso y potencia el resultado al compensar deficiencias de uno u otro método individual.

Un equipo interdisciplinario en Serbia y Polonia<sup>48</sup> puso a prueba seis metodologías: clasificadores bayesianos ingenuos, C4.5, sistemas

inmunitarios artificiales, Random Forest, Bagging y Boosting, para clasificar las cefaleas de acuerdo con las respuestas de un cuestionario clínico. La exactitud de clasificación se mantuvo entre 71 y 81%, y fue superior en los últimos tres métodos, que son ensambles. Estos autores, además, utilizaron algoritmos de selección de atributos para reducir la multidimensionalidad del sistema y trabajar con las variables de mayor peso. Los algoritmos, que funcionan de manera automática y sin prejuicio, tuvieron mejor desempeño que los médicos expertos, quienes seleccionaron “a mano” y de manera intuitiva los atributos clínicos que consideraban más importantes.

En alergología, se han utilizado sistemas inmunitarios artificiales, redes neuronales artificiales y redes bayesianas para agrupar, clasificar y fenotipificar principalmente a pacientes con asma<sup>49</sup> a partir de bases de datos de cohortes desde el nacimiento,<sup>50,51</sup> o en escenarios de primer nivel de atención<sup>52</sup> y urgencias<sup>53</sup> para identificar de manera automatizada a pacientes con probable asma y exacerbaciones.

Existe un repositorio en línea con bases de datos para utilizar en el aprendizaje automático, mantenido por la Universidad de California en Irvine (<http://archive.ics.uci.edu/ml/>), que incluye resultados de biopsias de lesiones mamarias, predictores de cardiopatía, registros de supervivencia posquirúrgica y otras 270 bases de datos de diversas disciplinas para aprender. Varios de los ejemplos disponibles en la bibliografía derivan de equipos que trabajan con estas bases de datos.

## Discusión

Esta revisión es una mirada breve a distintos ejemplos de utilización exitosa de inteligencia artificial en el diagnóstico clínico. La experiencia en biomedicina (proteómica, metabolismo,

microarreglos, imagenología) y en ciencias biológicas es aún más extensa y sólida.

El progreso en el aprendizaje automático para facilitar el diagnóstico clínico ha sido constante y existen varios ejemplos disponibles de aplicaciones, así como resultados de esfuerzos previos desde los cuales construir. Las fronteras actuales incluyen la capacidad para incorporar escenarios del mundo real, el procesamiento de lenguaje natural o reconocimiento de voz, la interconexión de bases de datos y la construcción de algoritmos versátiles. Nunca antes la necesidad de un apoyo computacional ha sido más grande, ya que en estos tiempos se manejan datos en mayor cantidad y complejidad. Nunca, como ahora, había estado disponible la tecnología y la capacidad de almacenamiento para desarrollar sistemas que permitan no sólo distinguir un sujeto enfermo de uno no enfermo, sino también clasificar adecuadamente un diagnóstico, facilitar el diagnóstico diferencial y establecer el pronóstico y riesgo.

Se puede utilizar el conocimiento acumulado en miles de casos de pacientes publicados en los últimos 20 años y en registros detallados de todo el mundo para construir una base de datos con características de cada enfermedad, que sirva como adiestramiento a la máquina para aprender a clasificar correctamente nuevos casos. A partir de la base de datos de adiestramiento se crea un algoritmo de toma de decisiones con diferentes métodos, que extraiga conocimiento de la información disponible, genere ecuaciones para clasificar correctamente los nuevos casos comunicados y que aprenda con cada utilización para incorporar la información de los nuevos casos y se vuelva más sólido con el tiempo.

El ensamble de varios abordajes o modelos garantiza un mejor resultado final; el desempeño del nuevo *über*-modelo se puede evaluar al



ponerlo a prueba con una porción de la base de datos que se reserve para este fin (validación cruzada). El resultado del proceso será una lista de mayor a menor probabilidad con los tres diagnósticos diferenciales más probables. La meta final es una aplicación o *software* con una interfaz cómoda al usuario clínico, en la que pueda ingresar los datos clínicos y de laboratorio del paciente, y obtener un diagnóstico diferencial y un abordaje diagnóstico sugerido.

El aprendizaje automático para asistir en el diagnóstico clínico es un abordaje multidisciplinario que se propone mejorar los procesos cognitivos al depurarlos de sesgos e incorporar un vasto conocimiento disponible de manera sistemática y estructurada. Los antecedentes históricos y los esfuerzos en otras latitudes sugieren que la tarea es realizable y que éste es el mejor momento para comenzar. Existen las mejores condiciones imaginables para garantizar que la empresa resulte interesante, significativa, compleja y fructífera. Los principales retos son manejar el grado de incertidumbre y la calidad de los datos ingresados para construir los algoritmos más indicados con el mejor desempeño matemático y desarrollar así un sistema manejable, actualizable y perfectible.

### Agradecimientos

A José Luis Morales, Edgar Medina, Anahí Anzo y Layla Michán, quienes revisaron y comentaron una versión preliminar del manuscrito.

### Referencias

1. Siegel E. Predictive analytics: The power to predict who will click, buy, lie, or die. New York: John Wiley & Sons, Inc., 2013.
2. Cleophas TJ, Zwinderman AH. Machine Learning in Medicine part 2 [Internet]. Dordrecht Heidelberg New York, London: Springer, 2013. Disponible en: <http://link.springer.com/10.1007/978-94-007-6886-4>
3. Khattree R. Computational methods in biomedical research. Baton Rouge, LA: Chapman & Hall, 2007.
4. Reed K, May R, Nicholas C, Taylor H, Brown A. Health Grades Patient Safety in American Hospitals Study [Internet]. 2011 [cited 2014 Feb 14]. Disponible en: [Healthgrades.com](http://Healthgrades.com)
5. Graber ML. The incidence of diagnostic error in medicine. BMJ Qual Saf [Internet]. 2013 Oct [cited 2014 Feb 9];22 Suppl 2:ii21-ii27. Disponible en: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3786666&tool=pmcentrez&rendertype=abstract>
6. Segal M. How doctors think, and how software can help avoid cognitive errors in diagnosis. Acta Paediatr [Internet]. 2007/09/14 ed. 2007;96:1720-1722. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/17850393>
7. Groves M, O'Rourke P, Alexander H. The clinical reasoning characteristics of diagnostic experts. Med Teach [Internet]. 2003/07/26 ed. 2003;25:308-313. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/12881056>
8. Sherbino J, Dore KL, Siu E, Norman GR. The effectiveness of cognitive forcing strategies to decrease diagnostic error: an exploratory study. Teach Learn Med [Internet]. 2011/01/18 ed. 2011;23:78-84. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/21240788>
9. Gill CJ, Sabin L, Schmid CH. Why clinicians are natural bayesians. BMJ [Internet]. 2005/05/10 ed. 2005;330:1080-1083. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/15879401>
10. Ebell MH, Smith MA, Barry HC, Ives K, Carey M. The rational clinical examination. Does this patient have strep throat? JAMA [Internet]. 2001/01/09 ed. 2000;284(22):2912-2918. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/11147989>
11. Croskerry P. Cognitive forcing strategies in clinical decision making. Ann Emerg Med [Internet]. 2003/01/07 ed. 2003;41:110-120. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/12514691>
12. Pines JM. Profiles in patient safety: confirmation bias in emergency medicine. Acad Emerg Med [Internet]. 2005/12/21 ed. 2006;13:90-94. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/16365325>
13. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. Acad Med [Internet]. 2003/08/14 ed. 2003;78:775-780. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/12915363>
14. Groves M, O'Rourke P, Alexander H. Clinical reasoning: the relative contribution of identification, interpretation and hypothesis errors to misdiagnosis. Med Teach [Internet]. 2004/09/17 ed. 2003;25:621-625. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/15369910>
15. Leblanc VR, Brooks LR, Norman GR. Believing is seeing: the influence of a diagnostic hypothesis on the interpretation of clinical features. Acad Med [Internet]. 2002/10/16 ed. 2002;77:S67-S69. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/12377709>
16. Baxt WG, Shofer FS, Sites FD, Hollander JE. A neural network aid for the early diagnosis of cardiac ischemia in

- patients presenting to the emergency department with chest pain. *Ann Emerg Med* [Internet]. 2002/11/26 ed. 2002;40:575-583. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/12447333>
17. Prabhudesai SG, Gould S, Rekhraj S, et al. Artificial neural networks: useful aid in diagnosing acute appendicitis. *World J Surg* [Internet]. 2007/11/29 ed. 2008;32:301-305. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/18043966>
  18. Warner HR. A mathematical approach to medical diagnosis. *JAMA* [Internet]. American Medical Association; 1961 Jul 22 [cited 2014 Apr 8];177(3):177. Disponible en: <http://jamanetwork.com/article.aspx?articleid=331443>
  19. Yu VL, et al. Antimicrobial selection by a computer: a blinded evaluation by infectious disease experts. *J Am Med Assoc* 1979;242:1279-1282.
  20. Siegel JD, Parrino TA. Computerized diagnosis: implications for clinical education. *Med Educ* [Internet]. 1988 Jan [cited 2014 Feb 6];22:47-54. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/3282154>
  21. (NSF) UNSF. National Science Foundation (NSF) News-NSF Leads Federal Efforts In Big Data [Internet]. *Nsf.gov*. 2012 [cited 2014 Apr 1]. Disponible en: [http://www.nsf.gov/news/news\\_summ.jsp?cntn\\_id=123607](http://www.nsf.gov/news/news_summ.jsp?cntn_id=123607)
  22. (SIAM) S for I and applied M. SIAM International Conference on Data Mining (SDM12) [Internet]. 2012 [cited 2014 Apr 8]. Disponible en: <http://www.siam.org/meetings/sdm12/>
  23. Cleophas TJ, Zwinderman AH. *Machine Learning in Medicine* [Internet]. Vasa. Springer; 2013 [cited 2014 Feb 14]. Disponible en: <http://medcontent.metapress.com/index/A65RM03P4874243N.pdf>
  24. Shapiro SC. Artificial Intelligence. In: Shapiro SC, editor. *Encyclopedia of Artificial Intelligence*. 2<sup>nd</sup> ed. New York: John Wiley & Sons, Inc., 1992.
  25. Mitchell T. *Machine learning*. New York: McGraw Hill, 1997.
  26. BBC News - Alan Turing: The experiment that shaped artificial intelligence [Internet]. Disponible en: <http://www.bbc.co.uk/news/technology-18475646>
  27. Machine learning [Internet]. [en.wikipedia.org: Wikipedia](http://en.wikipedia.org/wiki/Machine_learning). Disponible en: [http://en.wikipedia.org/wiki/Machine\\_learning](http://en.wikipedia.org/wiki/Machine_learning)
  28. Ramesh AN, Kambhampati C, Monson JR, Drew PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl* [Internet]. 2004/08/31 ed. 2004;86:334-338. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/15333167>
  29. Patel VL, Shortliffe EH, Stefanelli M, Szolovits P, Berthold MR, Bellazzi R, et al. The coming of age of artificial intelligence in medicine. *Artif Intell Med* [Internet]. 2008/09/16 ed. 2009;46:5-17. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/18790621>
  30. Atkov OY, Gorokhova SG, Sboev AG, et al. Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters. *J Cardiol* [Internet]. 2012/01/06 ed. 2012;59:190-194. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/22218324>
  31. Yoldas O, Tez M, Karaca T. Artificial neural networks in the diagnosis of acute appendicitis. *Am J Emerg Med* [Internet]. 2011/09/13 ed. 2011. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/21908136>
  32. Dietzel M, Baltzer PA, Dietzel A, et al. Artificial Neural Networks for differential diagnosis of breast lesions in MR-Mammography: A systematic approach addressing the influence of network architecture on diagnostic performance using a large clinical database. *Eur J Radiol* [Internet]. 2011/04/05 ed. 2011. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/21459533>
  33. Chuang CL. Case-based reasoning support for liver disease diagnosis. *Artif Intell Med* [Internet]. 2011/07/16 ed. 2011;53:15-23. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/21757326>
  34. Park YJ, Chun SH, Kim BC. Cost-sensitive case-based reasoning using a genetic algorithm: application to medical diagnosis. *Artif Intell Med* [Internet]. 2011/01/11 ed. 2011;51:133-145. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/21216571>
  35. Huang ML, Hung YH, Lee WM, Li RK, Wang TH. Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. *J Med Syst* [Internet]. 2010/08/13 ed. 2012;36:407-414. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/20703710>
  36. Tenorio JM, Hummel AD, Cohrs FM, Sdepanian VL, Pisa IT, et al. Artificial intelligence techniques applied to the development of a decision-support system for diagnosing celiac disease. *Int J Med Inf* [Internet]. 2011/09/16 ed. 2011;80(11):793-802. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/21917512>
  37. Perkins JN, Liang C, Gao D, Shultz L, Friedman NR. Risk of post-tonsillectomy hemorrhage by clinical diagnosis. *Laryngoscope* 2012;122:2311-2315.
  38. Watson I. Case-based reasoning is a methodology not a technology. 1999;12:303-308.
  39. Samarghitean C, Ortutay C, Vihinen M. Systematic classification of primary immunodeficiencies based on clinical, pathological, and laboratory parameters. *J Immunol* [Internet]. 2009 Dec 1 [cited 2013 Nov 8];183:7569-7575. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/19917694>
  40. Sajda P. Machine learning for detection and diagnosis of disease. *Annu Rev Biomed Eng* 2006;8:537-565.
  41. Seixas JM, Faria J, Souza Filho JBO, et al. Artificial neural network models to support the diagnosis of pleural tuberculosis in adult patients. *Int J Tuberc Lung Dis* [Internet]. International Union Against Tuberculosis and Lung Disease 2013 May [cited 2013 Nov 18];17:682-686. Disponible en: <http://www.ingentaconnect.com/content/iautld/ijtd/2013/00000017/00000005/art00021?token=00581a1cffd6a264d37e41225f40384d576b4628486b253e2c49576b3427656c3c6a333f2566e4ed81d7b599>

42. Geenen PL, Gaag LC, Loeffen WLA, Elbers ARW. Constructing naive Bayesian classifiers for veterinary medicine : A case study in the clinical diagnosis of classical swine fever. *Res Vet Sci* [Internet]. Elsevier Ltd; 2011;91:64-70. Disponible en: <http://dx.doi.org/10.1016/j.rvsc.2010.08.006>
43. Amaral JLM, Lopes AJ, Jansen JM, Faria ACD, Melo PL. Machine learning algorithms and forced oscillation measurements applied to the automatic identification of chronic obstructive pulmonary disease. *Comput Methods Programs Biomed* [Internet]. Elsevier Ireland Ltd; 2012 Mar [cited 2013 Nov 8];105:183-193. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/22018532>
44. Chen H, Yang B, Wang G. Support vector machine based diagnostic system for breast cancer using swarm intelligence. *J Med Syst* 2012;36:2505-2519.
45. Hart E, Timmis J. Application areas of AIS : The past, the present and the future. *J Applied Soft Computer* 2008;8:191-201.
46. Zhao W, Davis CE. A modified artificial immune system based pattern recognition approach-an application to clinical diagnostics. *Artif Intell Med* [Internet]. Elsevier B.V.; 2011 May [cited 2013 Nov 8];52:1-9. Disponible en: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3108456&tool=pmcentrez&rendertype=abstract>
47. Kodaz H, Özşen S, Arslan A, Güneş S. Medical application of information gain based artificial immune recognition system (AIRS): Diagnosis of thyroid disease. *Expert Syst Appl* [Internet]. 2009 Mar [cited 2013 Nov 8];36:3086-3092. Disponible en: <http://linkinghub.elsevier.com/retrieve/pii/S0957417408000171>
48. Krawczyk B, Simić D, Simić S, Woźniak M. Automatic diagnosis of primary headaches by machine learning methods. *Cent Eur J Med* [Internet]. 2012 Nov 16 [cited 2014 Jan 29];8:157-165. Disponible en: <http://www.springerlink.com/index/10.2478/s11536-012-0098-5>
49. Wu W, Bleecker E, Moore W, et al. Unsupervised phenotyping of Severe Asthma Research Program participants using expanded lung data. *J Allergy Clin Immunol* [Internet]. 2014 Feb 28 [cited 2014 Mar 24]. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/24589344>
50. Lazic N, Roberts G, Custovic A, et al. Multiple atopy phenotypes and their associations with asthma: similar findings from two birth cohorts. *Allergy* [Internet]. 2013 Jun [cited 2014 Mar 26];68:764-770. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/23621120>
51. Simpson A, Tan VYF, Winn J, et al. Beyond atopy: multiple patterns of sensitization in relation to asthma in a birth cohort study. *Am J Respir Crit Care Med* [Internet]. 2010 Jun 1 [cited 2014 Mar 21];181:1200-1206. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/20167852>
52. Afzal Z, Engelkes M, Verhamme KMC, JA, et al. Automatic generation of case-detection algorithms to identify children with asthma from large electronic health record databases. *Pharmacoepidemiol Drug Saf* [Internet]. 2013 Aug [cited 2014 Apr 8];22:826-833. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/23592573>
53. Dexheimer JW, Brown LE, Leegon J, Aronsky D. Comparing decision support methodologies for identifying asthma exacerbations. *Stud Health Technol Inform* [Internet]. 2007 Jan [cited 2014 Apr 8];129(Pt 2):880-884. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/17911842>