



ARTICULO DE REVISION

INTELIGENCIA ARTIFICIAL: LA CAJA DE HERRAMIENTAS VIRTUALES AL SERVICIO DE LA BIOINFORMATICA

Roberto Daniel Gordon Graell

Universidad de Panamá, Facultad de Informática, Electrónica y Comunicación del Centro Regional Universitario de Panamá Oeste, Panamá.

roberto.gordon@up.ac.pa 

RESUMEN

Una de las disciplinas científicas que más se ha potenciado con la Inteligencia Artificial (IA) es la Bioinformática. Las herramientas digitales que provee la tecnología permitieron dar un salto cuántico de calidad en el entendimiento del mapa de la vida, su funcionamiento y características que, a su vez, determinó un avance en las ciencias ómicas y en la comprensión y tratamiento de enfermedades como el cáncer. El estudio realizado es una revisión de la literatura en dos documentos relevantes en los procesos de aplicación de la Inteligencia Artificial en la medicina desde las dimensiones de acercamiento de la informática y la medicina, respectivamente. También se hizo una revisión documental en otras áreas de la biotecnología para caracterizar el avance y la interrelación de las dos disciplinas científicas en los procesos productivos. La conclusión más resaltante es el avance en conocimientos y calidad de vida que se ha producido desde la aparición de la bioinformática y el salto cualitativo que le ha significado la Inteligencia Artificial para todos los apartados de la sociedad humana.

PALABRAS CLAVES

Genoma, Gran Data, Algoritmo, Secuenciación, Aprendizaje

ARTIFICIAL INTELLIGENCE: VIRTUAL TOOLBOX AT THE SERVICE OF BIOINFORMATICS

ABSTRACT

One of the scientific disciplines that has been most enhanced by Artificial Intelligence (AI) is Bioinformatics. The digital tools provided by technology allowed a quantum leap of quality in the understanding of the map of life, its functioning and characteristics which, in turn, determined an advance in the omic sciences and in the understanding and treatment of diseases such as cancer. The study conducted is a literature review in two relevant papers on the application processes of Artificial Intelligence in medicine from the approach dimensions of computer science and medicine, respectively. A documentary review was also made in other areas of biotechnology to characterize the progress and interrelation of the two scientific disciplines in the productive processes. The most outstanding conclusion is the advance in knowledge and quality of life that has taken place since the appearance of bioinformatics and the qualitative leap that Artificial Intelligence has meant for all aspects of human society.

KEY WORDS

Genome, Big Data, Algorithm, Sequencing, Learning

INTRODUCCIÓN

La bioinformática es una de las disciplinas que forma parte de la amplia gama de ciencias y técnicas relacionadas con la biología. Es una de las disciplinas que transversalizan ese conjunto de ciencias gracias a su estudio de la información biológica fundamentada en la teoría de la información, las matemáticas y la computación. Nació con la necesidad de interpretar la información contenida en el genoma lo que requirió de algoritmos para catalogar, secuenciar encontrar similitudes y descubrir propiedades, determinar funciones y estructuras en las cadenas genéticas (Balatti y Flórez, 2019).

Como su nombre lo indica es la inserción de características biológicas de los seres vivos en entornos computarizados como variables de investigación o, dicho de otra manera, es la utilización de las facilidades que permiten la informática y sus aplicaciones en el entendimiento del desarrollo de los seres vivos. La gran diferencia es que lo hace a nivel micro, dentro de la célula, Ahí donde se escribe el código de la vida. Las cadenas de Ácidos Ribonucleicos (ARN) y nucleicos (ADN) que no son visibles al ojo humano y cuyo comportamiento, no es totalmente comprensible para los investigadores.

El genoma humano contiene alrededor de 60 gigabits de ‘información’ (sin comprimir). Razón por la que suponemos que se trata de información biológica importante, ya que ha permitido que las especies existan por lo menos 35,000 años y que ha sido cuidadosamente preservada de generación en generación (Biro, 2011).

Desde su aparición la bioinformática ha generado nuevas disciplinas científicas gracias a lo específico de sus aplicaciones. Las ciencias omicas se desprenden de las diferentes especialidades del estudio del genoma en el Código Genético. La disciplina es científica y tecnológica, es el campo de interacción de la Biología Genética y Molecular con la Ciencia de la Computación y la Ingeniería Informática en la gestión del conocimiento genómico y proteómico. Su aplicación trascendió a las ciencias médicas convirtiéndose en apoyo fundamental de la producción agrícola y la industria, entre otros (Pin et al, 2021).

En la disciplina, la computación tiene una función esencial con el manejo y análisis de bases de datos biológicas. El manejo supone registro, resguardo, entornos, combinaciones y otras variables que, otrora, necesitaban de equipos humanos multidisciplinarios para atender todas las dimensiones y determinantes del dato, y la data, y los análisis de estos, corresponde al comportamiento de las secuencias, apreciación de datos biológicos, recuperación, manipulación y enlace con otras fuentes de datos en la importante tarea de entender regularidades en el Código Genético.

La función primordial de las técnicas bioinformáticas es la de ganar tiempo, así parezca pueril. Los procedimientos de la investigación biológica son complejos y muy complejos. El uso de aminoácidos, codones, tasas de sustitución sinónimas y no sinónimas, marcadores de proteínas, etc., encontraron en la disciplina un aliado para que el rango de tiempo necesario para el manejo y análisis de grandes volúmenes de datos. Se verifican informaciones que antes suponían hipótesis sobre poblaciones, de las que se tomaba una muestra representativa, para hacer una especulación científica que, aunque sustentada en el conocimiento y una lógica, no dejaba de ser especulación y que, en el mejor de los casos, tomaba unos pocos años verificar y convertir en conocimiento.

Su impacto en las diferentes ciencias que sirven al ser humano no es poco. Los avances en tecnologías de la salud y la producción se aprecian en muchos campos. La farmacéutica la utiliza para el desarrollo de nuevos fármacos desde la detección de potenciales dianas. Sus recursos son importantes en el desarrollo de otras áreas médicas como la Epidemiología y la Genética.

De igual forma la Investigación clínica y la toma de decisiones en salud, gracias a los biochips o la "biología in silico" que es la construcción del conocimiento a partir de teorías, simulaciones y experimentos realizados con sus tecnologías como la Realidad Virtual (RV), la Realidad Aumentada (RA) o la Realidad Mixta (RM) que permiten suponer que la relación genes y enfermedades es tan cercana que la información genética se hace indispensable en el avance del conocimiento médico (Rodríguez y Vargas, 2019).

Conceptualizar la Inteligencia Artificial (IA) es un trabajo complejo debido a las diferentes áreas y posturas gnoseológicas que abarca, tanto que, en la red, se encuentran millones de documentos que van desde artículos sencillos hasta tratados doctorales. Un acercamiento técnico procedimental es aquel que gira en torno a la idea de que la máquina emule la capacidad de las redes neuronales del ser humano para tomar decisiones para lo que se utilizan métodos de análisis de datos y desarrollo de algoritmos de procedimientos (Arbeláez et al, 2021).

La pequeña gran ventaja de la IA, sobre la capacidad humana que emula, es la ingente cantidad de datos que puede conocer y procesar en diferentes tipos de análisis y en tiempos relativamente cortos. La big data le revela sus secretos, muestra patrones y tendencias de hipótesis de investigación, determina comportamientos sociales y/o naturales en poblaciones de estudio, identifica entornos según los parámetros que requieran los investigadores y muchas otras funciones y aplicaciones más que, sin este avance en las Tecnologías de la Información y la Comunicación (TIC) los investigadores humanos tardaron décadas en establecer en sus diferentes campos. (Gómez, 2021)

La IA es una de esas tecnologías disruptivas que son características de las ciencias de la computación que, aunque dejan en la obsolescencia tecnologías y técnicas anteriores, tiene la ventaja de escalar a un nivel

superior las formas de apoyar el entendimiento y la construcción del conocimiento humano. Son creaciones del entendimiento humano que permiten análisis y perspectivas de análisis de la información desde nuevos enfoques. Es una técnica adaptativa para superar la sociedad de la información y adentrarse en mejores condiciones en la sociedad del conocimiento (Orozco y Arango, 2016).

Como evolución de la informática no puede estar exenta de introducirse en el mundo de la bioinformática. Las capacidades de análisis de la información que suponen la especulación, la hipótesis, las relaciones lógicas o pseudológicas, las proyecciones y demás procesos reservados a la inteligencia humana y científica, se ven abordados con la capacidad superior del manejo de grandes volúmenes de información en tiempos relativamente cortos y con la reducción, al mínimo estadístico, del error. Los científicos cuentan con un asistente de campo súper potenciado.

El propósito de este trabajo es determinar diferentes formas en que la bioinformática se apoya en la IA para servir a las disciplinas del saber humano tanto para producir conocimiento como para traducirlo en bienes de consumo y servicio para la sociedad y su civilización. Para ello se realizó en una revisión documental a través de la red para caracterizar las formas de aplicación de la tecnología en diferentes trabajos científicos. Con los resultados pretende hacer un ejercicio de especulación de cuál será el siguiente paso en que las aplicaciones de la informática disruptiva eleven, un nivel más, la forma de obtener el conocimiento en la bioinformática.

MATERIALES Y MÉTODOS

Los datos utilizados para el análisis fueron extraídos de documentos científicos, de diferentes disciplinas científicas y de aplicación en la producción de bienes, que están presentes en la red. Se revisaron documentos obtenidos en una revisión simple según cadena de búsqueda “Bioinformática, Inteligencia artificial, herramientas, trabajos doctorales”, y derivadas, en el buscador público de Google. Se seleccionaron Investigaciones en diferentes disciplinas relacionada con la informática y la medicina cuyo criterio de inclusión fue la selección y utilización de herramientas digitales que provee la Inteligencia

Artificial y su forma de aplicación en el análisis, generación y presentación de resultados del investigador.

Para la descripción de la selección y uso de herramientas se seleccionaron dos trabajos que, a criterio del investigador, cumplen con los parámetros de búsqueda y permiten dar un ejemplo de las herramientas disponibles para los investigadores y las posibilidades de adaptación o base de diseño de estas a las características particulares de cada investigación.

De igual forma se revisaron documentos de botánica, biología marina, veterinaria e ingeniería de software, para caracterizar las aplicaciones de la IA en que se apoyaron como tecnología disruptiva de la bioinformática y cuál es el beneficio que aportan.

También se creyó necesario investigar la presencia de la bioinformática en la Republica de Panamá por lo que se recurrió a documentos oficiales del ente correspondiente en su portal web como es la Secretaría Nacional de Ciencia, Tecnología e Innovación (SENACYT)

Es una investigación que abarca documentos desde de las dos últimas décadas, 2002-2021 incluidos, de corte documental, cualitativo, descriptivo de las características previamente determinadas.

RESULTADOS

El lenguaje informático de las proteínas

Uno de los campos que potenció la IA en la biotecnología es el entendimiento de cómo funciona las estructuras de las proteínas puesto que, en ocasiones, comparten una misma estructura, pero no la misma función. Significa que sus secuencias pueden ser muy distintas, así estén estructuradas en una arquitectura muy similar, y cumplen funciones diferentes. Entender su comportamiento es un tema complejo por al amplio número de propiedades y relaciones entre sus componentes, los aminoácidos.

El alfabeto del Genoma Humano es tan amplio como permiten tres mil millones de pares de bases o “letras” en una complejidad gramatical que,

para la máquina, puede resultar en redundancia algorítmica. Para hacerlo maleable se necesita una gama de alfabetos reducidos, con diferentes subconjuntos que a su vez deben estar determinados por las propiedades fisicoquímicas de los aminoácidos, como son la hidrofobicidad y el volumen. Para los investigadores representa la identificación correcta de la síntesis, natural o artificial, de una proteína específica (Weidenslaufer et al, 2019).

Cada elemento es un símbolo que tiene su representación en un binario. Cada combinación es una determinante de la arquitectura del compuesto de elementos que, a su vez, determina la función y que en la arquitectura computarizada de la IA es un algoritmo secuencial. Lo importante es entender cómo van las secuencias de algoritmos que representaran una función o un estado que permita reconocer la proteína.

Las secuencias también representan características de los elementos y los parámetros en que se presentan o cumplen una función, en este caso los aminoácidos. Así parámetros físicos de la termodinámica como la pérdida de calor, cambios de energía libre de Gibbs y temperaturas de transición, entre otras, que resultan fundamentales para el entendimiento de la estructura y la estabilidad de las proteínas y, por supuesto, su función.

Los resultados de las combinaciones de cada proteína, en la aplicación IA, se presentan en una matriz compleja, en hojas de cálculo, básicamente en Excel, compuesta a su vez de una matriz superior y una inferior en bases numéricas. Las resultantes permiten la construcción de diferentes herramientas interpretativas según la necesidad de los usuarios como diagramas, esquemas, representaciones estructurales, dendogramas, y mapas de secuencias, genéticas entre otros.

Aplicación en la Protogenómica

El uso de datos ómicos en biomedicina crece muy rápidamente como el estudio global de los sistemas celulares en un nivel concreto. Las principales ciencias ómicas desarrolladas en los últimos años son la Genómica, la Transcriptómica, la Proteómica y la Metabolómica. El concepto proteoma fusiona las palabras proteína y genoma, el primero como dotación génica de una célula u organismo y el segundo entendido

como el total de proteínas expresadas por los genes de una célula, tejido u organismo. “La Proteómica se define como el conjunto de técnicas o tecnologías utilizadas para la obtención de información funcional de las proteínas, y tiene por objetivo el análisis, identificación y caracterización del proteoma celular.” (Franch, 2018)

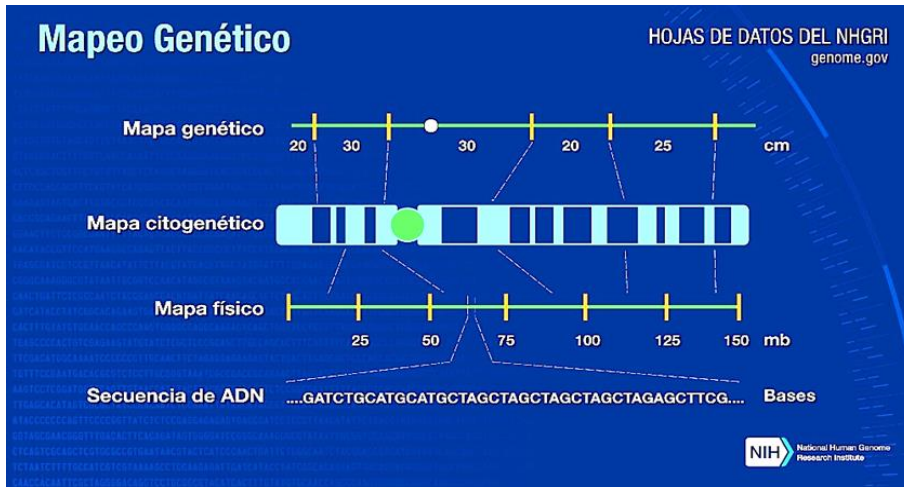


Fig. 1. Relación de hojas de datos del mapeo genético en enfermedades transmisibles del Proyecto del Genoma Humano (HGP) del Instituto Nacional de Investigaciones de Genoma Humano (NHGRI, por sus siglas en inglés) disponibles en bases de datos de los Institutos Nacionales de Salud y el Centro Nacional para la Información Biotecnológica de la Biblioteca Nacional de Medicina (NCBI) (ncbi.nlm.nih.gov), así como en el navegador genómico de la Universidad de California, Santa Cruz. NHGRI (2020)

La disciplina es el resultado de la obtención, proceso, análisis, integración e interpretación de datos de origen transcriptómico y proteómicos, identificados como datos proteogenómicos que provienen de 25000 proteínas, 1000 metabolitos, multitud de elementos proteicos desconocidos y ARN funcionales. Es el interactoma humano, el mapa de las interacciones moleculares entre las proteínas humanas, la red social de las proteínas cuyo estudio es un aliado relevante en la búsqueda de nuevos tratamientos médicos para todas las enfermedades. (Friedrich, 2017).

En la Biología de sistemas, el método de representación y análisis más frecuente es la utilización de redes neuronales o redes biológicas. El autor recurre al metaanálisis del volumen de datos provenientes de linfocitos B y T o de líneas celulares primarias de pacientes con leucemia, en el que cada elemento participante, en un proceso celular particular estudiado, se representará con un nodo y la relación entre ellos se representará con una arista o conexión. Ambos elementos que integrarán partiendo de una red de datos ya creada proveniente de diferentes bases de datos. “Las redes pueden representar (i) procesos biológicos simples; (ii) la combinación de varios procesos o rutas; o (iii) directamente un interactoma, el cuál incluiría todo elemento y procesos biológicos llevados a cabo por un organismo” (Friedrich, 2017).

Para el procesamiento de los resultados utiliza diferentes herramientas bioinformáticas con IA para hacer una conversión de identificadores biológicos a digitales y establecer un único identificador (ID) por cada proteína. El utilizado por el autor en su propuesta de tesis es uno muy utilizado por los investigadores en proteómica que permite hacer tablas de mapeo y muestra si la información de una proteína es, o fue, revisada por expertos. La tabla creada es muy importante porque permite generar bases de datos integrativas y relaciona los ID con los de otros programas.

Los pasos siguientes son la preparación de bases de datos de redes que sean compatibles con otros programas, el procesamiento de set de datos transcriptómicos, la preparación de casos de estudio, visualización de la representación digital, el análisis proteogenómico cualitativo de la línea celular en ramos de linfoma de células B y otra serie de procedimientos inherentes al estudio del cáncer que utiliza una composición de herramientas de IA. La síntesis objetiva de Friederich (2017) y Franch (2019) se pueden ver en la tabla 1.

Tabla 1. Trabajos de Proteogenómica. Síntesis de objetivos

Objetivo	Resultados esperados
Friederich (2017)	Base integrativa de datos biológicos. Traducción de rutas biológicas a redes de
Desarrollo y proteínas.	
aplicación de	Análisis integrativo de calidad de set de datos
algoritmos	proteómicos y set de datos transcriptómicos.
bioinformáticos	Integración de datos proteómicos y
	fosfoproteómicos cuantitativos.
	Flujos de análisis automáticos o semi-automáticos
	de datos biológicos.
Franch (2019)	Identificación el interactoma de las células madre
Desarrollo y hematopoyéticas.	
aplicación de	Herramienta bioinformática para
métodos	reposicionamiento de fármacos.
bioinformáticos	Interfaces de visualización y comparación de
para análisis de	experimentos bioinformáticos.
datos biológicos	Bases de datos Proteogenómicas con variaciones
	y nuevas uniones a las bases de datos de
	referencia.

La Tabla 1 permite observar que los dos trabajos doctorales son de base médica, pero de expresión bioinformática.

Friederich desarrolla algoritmos de aplicación bioinformática y Franch diseña métodos con aplicaciones ya diseñadas. Friederich se orienta a la producción de datos y bases de datos mientras Franch se ocupa del diseño de métodos y herramientas.

Cada autor realizó más de dos procedimientos por objetivo para lo que recurrió una media de diez tareas diferentes utilizando un mínimo de dos herramientas bioinformáticas por lo que, cada uno, utilizó más de veinte herramientas y aplicaciones bioinformáticas al alcance en la Internet. Eso permite a los investigadores profundizar en sistemas complejos a diferentes niveles celulares que antes eran imposibles.

Aplicaciones de la biotecnología

No solo la medicina se aprovecha de las virtudes de la bioinformática con IA. En el universo de la red se encuentra un sinnúmero de trabajos de Informática en las biociencias y viceversa. En la botánica se encuentra el uso de las redes neuronales artificiales para obtener información útil y aplicable en procedimientos como la modelización del enraizamiento y aclimatación de plántulas micropropagadas, la lógica difusa para la detección de factores clave en el enraizamiento y aclimatación, el diseño de medios de cultivo y para la optimización de las condiciones de cultivo para la obtención de plántulas sanas micropropagadas y elucidar los factores críticos de la germinación de especies en peligro de extinción. (Chico, 2017)

En una sociedad que soporta su ingesta proteica en el ganado vacuno la identificación, registro y monitoreo del movimiento de animal es esencial en la industria y eje de la política agrícola para los gobiernos nacionales. El control de enfermedades, la higiene correspondiente, la calidad de los alimentos son esquemas de seguimiento que garantizan la calidad para lo que muchos países construyen bases de datos de la ganadería nacional a través de etiquetas numeradas de identificación.

En Panamá, Villalobos et al. (2020) realizaron una investigación de trazabilidad en bovinos como parte del Proyecto: Conservación y Uso de Bovino Criollo Panameño, con financiamiento del Instituto de Investigación Agropecuaria de Panamá (IDIAP), SENACYT y el Sistema Nacional de Investigación (SNI).), Utilizaron como método la secuenciación de siguiente generación (NGS) y los procedimientos de amplificación en puente y secuenciación por síntesis cuyos datos se analizaron en programas de análisis de secuenciación de genotipos. Todas las herramientas digitales que utilizaron son de uso comercial y de presencia en la red.

Johnson (2017) registró para el SENACYT, en el año 2015, trabajos en tecnología e investigaciones cuyos resultados cuantitativos son:

- Biotecnología animal: 6 en Veterinaria, 2 en Acuicultura y 1 en Mosquitos, control de vectores.
- Biotecnología forestal y agrícola: 1 en Clonación, 2 en Transgénicos, 1 Fitosanitario, 1 genética y 1 Fermentación etílica

- Biotecnología humana: 5 en células madre y 8 en Estudios clínicos.
- Limpieza ambiental, Biorremediación: 3 en hidrocarburos, 1 en energías renovables y 3 en suelos y aguas.
- Biotecnología de alimentos: 3 en alimentos y 2 en energías.

Son 29 trabajos cuya descripción cualitativa no es profunda, pero dejan entrever el uso de bases de datos mundiales y uso de la bioinformática con IA.

El desarrollo de herramientas de IA para la bioinformática es variado y trata de cubrir varias necesidades que van desde lenguajes de programación, actualizaciones de aplicaciones, diseño en software libre, plataformas de diferentes usos y la insuficiente formación en informática de los investigadores biológicos. Eso también implica un problema puesto que “[...] al ser tan diversas y con propósitos tan diferentes se hace indispensable la automatización de la ejecución de estas herramientas, sin perder demasiado tiempo configurándolas” (Orozco y Arango, 2016)

Dentro de la variedad de software libres que existen para diseñar aplicaciones según la necesidad del investigador se encuentran bases de datos biológicas como Perl, Bio* Toolkits, NCBI Toolkit y EMBOSS (European Molecular Biology Open Software Suite. Las bases almacenan ingentes cantidades de información, desde secuencias de genes y genomas, hasta secuencias de proteínas y proteomas además de artículos, literatura relacionada, rutas metabólicas, estructuras proteicas, ontología, o Polimorfismo de un solo nucleótido Single (SNP por sus siglas en inglés), etc., todas según la temática, área o el objetivo al que van enfocados (Alcalde, 2016).

Los investigadores encuentran herramientas fundamentales para el análisis de secuencias, anotación de genomas, Biología evolutiva computacional, medición de la biodiversidad, análisis de la expresión y regulación génica, predicción de la estructura de proteínas, genómica comparativa, modelado de sistemas biológicos y Acoplamiento molecular, todos procesos complejos de múltiples procedimientos y gran volumen de data que no son posibles al cálculo humano tradicional (Alcalde, 2016).

Era un futuro previsto unos años antes con el inicio de la disciplina bioinformática en su etapa estadística. Lo disruptivo de las tecnologías de la ciencia de la computación permitía prever la integración del conocimiento dentro del genoma para conocer, más profundamente, sus componentes, estructuras y funciones. Entre otras herramientas se preveían las representaciones virtuales de organismos, simulaciones ante perturbaciones virtuales, diseños de transmisión vectorial y protocolos más cortos en el tiempo que, en forma experimental, eran imposibles y acceso y análisis de grandes volúmenes de datos (Pin et al, 2021).

Las diferentes aplicaciones bioinformáticas trabajan con la misma materia prima presente en todos los seres vivos como lo es el material genético. León (2015) realiza una investigación en biología marina para establecer la conectividad poblacional en invertebrados y especies demersales del Pacífico mexicano como proceso fundamental de intercambio genético y renovación de población que son características importantes en el mantenimiento de la biodiversidad.

La investigación requirió de análisis genético a nivel larvario para determinar el potencial de dispersión larvario, diferenciación genética entre especies, influencia de mecanismos genéticos y ecológicos a nivel de comunidad sobre la composición genética de las especies del ecosistema, análisis de ADN mitocondrial como en los loci microsatelitales y valores de flujo genético que no hubieran sido posibles sin una herramienta, altamente tecnológica, que permitió los diferentes resultados de los análisis genéticos y el acceso a múltiples bases de datos en tiempo real. (León, 2015)

DISCUSIÓN

Los trabajos analizados muestran un manejo de datos, informaciones y especulaciones con bases matemáticas en tiempos cortos que para el ser humano son imposibles. Las tecnologías de IA tanto Machine Learning (ML) como big data o cadena de bloques coloca el dato donde le ordenan y de inmediato sabe que lo aplicará allí todo el tiempo. No hay que repetir el comando cuando el dato este entre los parámetros del primero. De ahí en adelante aprenderá a reconocer similitudes y diferencias y derivará algoritmos para reconocer patrones y relaciones. Esas funciones son esenciales en la bioinformática.

Todas las aplicaciones se basan en la imitación de las funciones humanas. Guardar el dato en matrices es la función memorística del humano, pero no se trata de solo coleccionar el dato sino asignarlo en el lugar correspondiente y correlacionarlo con otros datos y proyectar los resultados posibles. Las aplicaciones de IA de las investigaciones revisadas se construyeron para su función determinada con base en tecnologías existentes y disponibles en diversas, y diferentes plataformas, presentes en la red, en diferentes lenguajes y sistemas.

Los autores utilizaron aplicaciones disponibles en la red para la construcción de sus propias herramientas a la medida de las necesidades y requerimientos de los estudios. Los resultados no fueron concluyentes por lo que requirió del análisis humano, la redimensión de la herramienta, la experimentación y el ensayo y error con la tecnología y sus usos. No fue la aplicación de una herramienta exacta para un uso específico sino la modulación de una herramienta digital maleable, adaptable y que requirió del reajuste de los parámetros de trabajo.

El ML utiliza algoritmos que “imitan” la estructura y función del cerebro. Las Redes Neuronales Artificiales (RNA) son modelos digitales que procesan la información simulando el funcionamiento de las neuronas biológicas. Están compuestas por nodos o “neuronas” que procesan toda la información y que están conectados como una red por lo que se sirven, y apoyan, del conocimiento de diferentes bases de datos. Ayudan a que los sistemas funcionen como un cerebro humano en el proceso de aprendizaje. Están conformadas por múltiples capas de nodos ocultos que transmiten la información, de entrada y salida entre los nodos, llamadas “capas de aprendizaje”. A mayor cantidad de capas, mayor es la profundidad de la red y mayor es la capacidad de aprendizaje. (Martínez C., 2020)

La IA se adapta perfectamente a los estudios genéticos porque, más que las características particulares de los genes, de las proteínas o cualquiera de los elementos de estudio, el papel que realizan, su comportamiento y sus funciones se determinan en cómo se relacionan o sus secuencias. La IA no usa algoritmos cuantitativos sino secuenciales. Utiliza técnicas de secuenciación que permiten realizar análisis estructurales de comportamientos en forma de encadenamientos o continuidades. A su

vez permite construir modelos, tanto de elementos como de secuencias o cualquiera de las formas que requiera el investigador, con el uso de la RV, La RA y la RM

En la medida en que avanzan las investigaciones las bases de datos se van incrementando. Todos los días alrededor del mundo se van incrementando los sitios con informaciones de investigaciones recientes con sus resultados y perspectivas. De esa forma los investigadores del año 2016 tuvieron a su disposición un volumen de datos superior a los de años anteriores, así como los del año 2022 tendrán a su disposición los resultados del año 2021.

CONCLUSIONES

La bioinformática permite el manejo de grandes volúmenes de datos en los que se verifican informaciones que antes suponían hipótesis sobre poblaciones de las que se tomaba una muestra representativa. También se redujeron los tiempos de entrega. La IA permite mayor fiabilidad en los resultados, mayor calidad en los productos, mayor seguridad en las consecuencias y beneficios de un servicio y por lo tanto mayor calidad de vida.

En la bioinformática la IA es un ejemplo claro de la evolución en términos informáticos. No es un desarrollo que potenció una tecnología sino una aplicación evolutiva que elevó de nivel una disciplina de por sí compleja y en extremo útil para el ser humano. Entender la cadena de la vida ya es un gran avance, pero comprender las secuencias en que se conforman los mapas genéticos y sus resultados es ponerse a un paso, metafórico, de revelar procesos que permitan alargar la o vivirla en ausencia de enfermedades.

El análisis de los virus, como el coronavirus COVID 19, es uno de los mejores ejemplos de la IA en la bioinformática. Gracias a sus recursos didácticos el público conoció la imagen virtual de la forma del virus, su estructura, su comportamiento y sus deferentes metamorfosis en tiempo real. Sin la IA no hubiese sido posible. Las diferentes vacunas no fueron aplicadas, antes, por protocolos éticos de investigación en humanos que exigen tiempos determinados. Los resultados de los experimentos

confirmaron los presupuestos matemáticos, estructurales y de eficiencia de las vacunas que ya se conocían, con niveles insignificantes de error.

Una especulación científica es que, para el próximo virus que enfrente la humanidad, los tiempos de respuesta en la síntesis de un compuesto vacunal serán menores. La prueba mundial del COVID 19 permitirá, en un futuro tal vez no muy lejano, que el compuesto virtual se apruebe aun antes de experimentar con humanos gracias a la IA. Ni siquiera la comunidad científica mundial puede soslayar la fabricación de cinco mil millones de dosis en un proceso de IoT en menos de dos años. También producto de bioinformática y la IA en apoyo al conocimiento humano.

El debate sobre la ética siempre está sobre la mesa, pero poco a poco se acerca el momento en que la IA diseñe un proceso de investigación y lo efectúe a cabalidad sin la intervención humana. Que la misma maquina decida hacia dónde va la experimentación y ponga los límites éticos. Será uno de los grandes avances de las tecnologías disruptivas que, para su desventaja, no posee lo que motiva al ser humano y juega a su favor: La curiosidad.

REFERENCIAS

Alcalde, M. (2016). Software libre enfocado en diversos campos de las ciencias. *Hamut'ay*, 3(1). Enero-junio 2016, 59-70: <http://revistas.uap.edu.pe/ojs/index.php/HAMUT/issue/archive>

Arbeláez, D., Villasmil, J., y Rojas, M. (2021). Inteligencia artificial y condición humana: ¿Entidades contrapuestas o fuerzas complementarias? *Revista de Ciencias Sociales (Ve)*, 28(2), 501-512: <https://www.redalyc.org/journal/280/28066593034/28066593034.pdf>

Balatti, G., y Flórez, N. (2019). Bioinformática para la creación y fortalecimiento de empresas de base tecnológica: conceptos y aplicaciones. *PDG*, 42-62: https://www.researchgate.net/publication/336702599_Bioinformatica_para_la_creacion_y_fortalecimiento_de_empresas_de_base_tecnologica_conceptos_y_aplicaciones

Biro, J. (2011). Biological information: Definitions from a biological perspective. Obtenido de Information : Information Science Journal, 2(1), 117-139: <https://www.mdpi.com/2078-2489/2/1/117>

Chico, J. (2017). Biotecnología e inteligencia artificial. SAGASTEGUIANA, 5(1), enero-julio 2017, editorial: <https://revistas.unitru.edu.pe/index.php/REVSAGAS/issue/view/380>

Franch, M. (2018). Bioinformática funcional y su aplicación en genómica, proteogenómica y reposicionamiento de fármacos. [Tesis Doctoral.] Universidad Complutense de Madrid: <https://eprints.ucm.es/id/eprint/55758/>

Friedrich, C. (2017). Bioinformática para integrar información de la proteína y del gen en un contexto relacional, aplicación a los datos proteómicos y transcriptómicos humanos. [Tesis Doctoral] Universidad de Salamanca: https://gredos.usal.es/bitstream/handle/10366/135785/CIC_DrospeCF_Anexos.pdf?sequence=2&isAllowed=y

Gómez, S. (2021). Bioinformática como recurso educativo para enseñar variabilidad genética mediante la comparación de mapas de restricción. Rev. Asoc. Col. Cienc.(Col.), 2021; 33: 36-45.: doi.org/10.47499/revistaaccb.v1i33.231

Johnson, J. (2017). Levantamiento de Línea Base de Indicadores de Biotecnología de Panamá, 2015. Recuperado el 28 de junio de 2021, de Secretaría Nacional de Ciencia, Tecnología e Innovación (SENACYT): <https://www.senacyt.gob.pa/publicaciones/wp-content/uploads/2019/07/23-Biotecnolog%C3%ADa-L%C3%ADnea.pdf>

León, M. (2015). Análisis comparativo de los patrones de conectividad genética de especies arrecifales del Pacífico mexicano. [Tesis para obtener el grado de Doctor en Ciencias.] Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California: <https://repositorioslatinoamericanos.uchile.cl/handle/2250/2260648>

Martínez, C. (2020). Inteligencia artificial y salud. Centro de Estudios en Tecnología y Sociedad (CETyS): https://www.udes.edu.ar/sites/default/files/inteligencia_artificial_y_salud_1.pdf

NHGRI. (2020). Hoja de datos del mapeo genético. National Human Genome Research Institut. Hoja de datos del mapeo genético: <https://www.genome.gov/es/about-genomics/fact-sheets/Hoja-de-datos-del-mapeo-gen%C3%A9tico>

Orozco, S., y Arango, J. (2016). Aplicación de la inteligencia artificial en la bioinformática, avances, definiciones y herramientas. UGCiencia 22, 159-171.: <https://core.ac.uk/download/pdf/268087858.pdf>

Pin, Á., Holguín, M., Hidalgo, Y., y Villacreses, V. (2021). Uso de aplicaciones bioinformáticas en el área de salud pública en Ecuador y Latinoamérica. Pol. Con. (Ed. 56) 6(3) marzo 2021, 1934-1944: DOI: 10.23857/pc.v6i3.2483

Rodríguez, W., y Vargas, J. (2019). Biochips, aplicaciones convencionales e innovación: Una revisión documental. 7(2), julio - diciembre 2019, 1-11: doi.org/10.17081/invinno.7.2.3086

Villalobos, A., González, R., Murillo, M., y Castillo, H. (2020). Panel reducido de polimorfismos de nucleótido simple para estudios de biodiversidad en bovinos. Ciencia agropecuaria, 31, julio-diciembre 2020, 19-36: <http://www.revistacienciaagropecuaria.ac.pa/index.php/ciencia-agropecuaria/article/view/298/237>

Weidenslaufer, C., Torres, R., y Hernández, R. (2019). Medicina Genómica. Biblioteca del Congreso Nacional de Chile. Asesoría Técnica Parlamentaria: https://www.bcn.cl/obtienearchivo?id=repositorio/10221/27426/1/BCN_Medicina_Genomica_2019.pdf

Recibido el 16 de diciembre de 2022.

Aceptado el 12 de abril de 2022.

Editor Responsable: Dr. Eduardo Camacho Astigarrabia