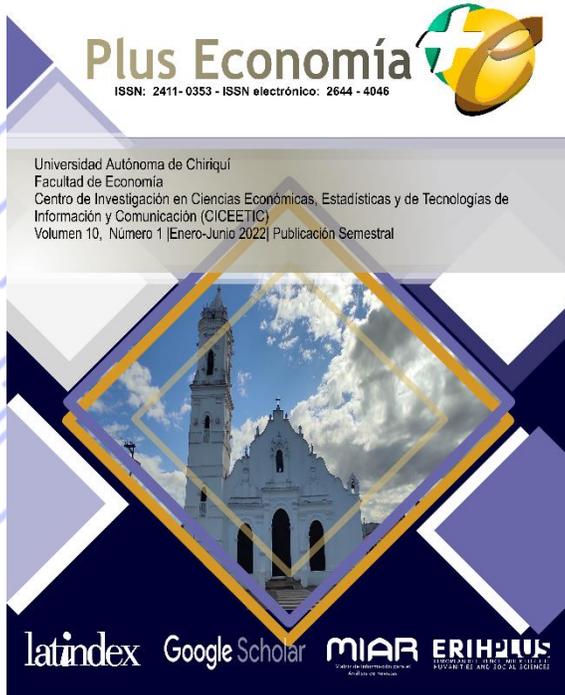




- › Revista Plus Economía
- › ISSN: 2411-0353
- › ISSN electrónico: 2644-4046
- › pluseconomia@unachi.ac.pa
- › Centro de Investigación en Ciencias Económicas, Estadísticas y de Tecnologías de Información y Comunicación, CICEETIC
- › Universidad Autónoma de Chiriquí (UNACHI)
- › República de Panamá



Samuel Saldaña

MODELO DE INTELIGENCIA ARTIFICIAL PARA EL PRONÓSTICO DE DETECCIÓN DE CÁNCER DE MAMA MEDIANTE UNA RED NEURONAL ARTIFICIAL

Vol. 10, Núm. 1, Enero – Junio 2022

pp. 34-42

**Universidad Autónoma de Chiriquí,
Panamá**



MODELO DE INTELIGENCIA ARTIFICIAL PARA EL PRONÓSTICO DE DETECCIÓN DE CÁNCER DE MAMA MEDIANTE UNA RED NEURONAL ARTIFICIAL

Samuel Saldaña

samuel.saldana1@unachi.ac.pa

Universidad Autónoma de Chiriquí

Recibido: Junio de 2021

Aceptado: Noviembre de 2021

Resumen

La Inteligencia Artificial (I.A.) desde su aparición, ha procurado asumir funciones del humano, y por su propia naturaleza, estar al servicio del mismo, la consolidación de la comprensión sobre la naturaleza humana ha permitido generar una ruta de acciones de los cuales la medicina es una ellas.

Las altas tasas de cáncer de mama en el mundo obligan sumar esfuerzos en la detección, análisis y diagnóstico, que estas coadyuven a tareas sensible que exige precisión y tiempo de respuesta alta, por lo que surge la necesidad de desarrollar un modelo algorítmico de Deep Learning y Machine Learning en el campo de la I.A., para detectar con eficiencia el cáncer de mama.

Este proyecto resulta un contraste al comparar las valoraciones de especialistas humanos reportadas en el estado del arte, la propuesta presenta un desempeño muy superior a los procesos convencionales.

Palabras clave: *Inteligencia Artificial, Deep Learning, Machine Learning, Oncológico, Redes Neuronales Artificiales.*

Abstract

Since its appearance, Artificial Intelligence (AI) has tried to assume functions of the human, and by its very nature, to be at its service, the consolidation of the understanding of human nature has allowed to generate a route of actions of which medicine is one of them.

The high rates of breast cancer in the world make it necessary to add efforts in the detection, analysis and diagnosis, the which contribute to sensitive tasks that



require precision and high response time, which is why the need arises to develop an algorithmic model of Deep Learning and Machine Learning in the field of AI, to detect with efficiency the breast cancer.

This project is a contrast when comparing the valuations of human specialists reported in the state of the art, the proposal presents a much higher performance than conventional processes.

Keywords: *Artificial Intelligence, Deep Learning, Machine Learning, Oncology, Artificial Neural Network.*

Introducción

Existe un escenario donde demanda eficiencia y eficacia ante problemática para la detección del cáncer de mama a través los procedimientos y el criterio humano, aunado que los índices de los casos de cáncer a nivel mundial van es progresivo, se ven afectados de acuerdo a las limitaciones humanas.

Este proyecto pretende buscar la solución de la efectividad tanto en la certeza, como en el tiempo de respuesta a los pacientes que buscan de los servicios médicos; para contribuir en este proceso, se emplearán algoritmos de Inteligencia Artificial de altas capacidades con tecnología de procesamiento gráfico con GPU (Graphic Processing Unit), a efectos de lograr un diagnóstico certero por encima de un 90% de

efectividad, y accesible para el servicio telemático.

Enunciado del problema

Cada año el cáncer de mama va en incremento, aunque los criterios de los especialistas realicen el trabajo, hay una imposibilidad en el recurso humano de atender estos volúmenes, así como el grado de certeza con respecto a los errores en los diagnósticos y las consecuencias en la salud mental de aquellos pacientes incurridos.

Mientras que los avances en la Inteligencia Artificial (I.A.) son rápidos, abarcales y cada vez más efectivos, existe una controversia en el sentido de la aceptación para la detección computarizada de cáncer de mama, la confianza del paciente, por un lado, y



la competitividad del algoritmo por parte de la comunidad médica.

Diseñar un modelo en I.A. para lograr ser efectivo en la detección de cáncer de mama, permite no solo acercarse a un procedimiento válido, sino de bajo coste, preciso o confiable, traducido en una mayor capacidad en atención de pacientes.

La implementación para el procesamiento y almacenamiento de la base de datos, son posibles por costos bajos que implica el clouding, muy contrario a la infraestructura tecnológica compleja y costosa.

Asumir una efectividad muy alta para la detección de cáncer de mama (benignos como malignos), es aceptable en la medida técnica, económica y administrativa, otros algoritmos de esta naturaleza, se enfocan en el aspecto académico o limitan sus funciones en lo tecnológico, y no tanto en una respuesta en el campo oncológico.

Diseñar y poner a prueba un algoritmo para la detección del cáncer en el ámbito social, tecnológico y administrativo, basado en Redes Neuronales Artificiales (RNA), son tan

solo algunos aspectos que contribuyen y dinamizan al sector de la salud.

Análisis y discusión del problema

El problema a resolver está dirigido al grado de efectividad en el modelo por pronósticos para la detección del cáncer de mama.

Solución y propuesta

El modelo propone una solución alterna para la labor de los especialistas en el protocolo para la detección del cáncer de mama.

El algoritmo estará basado en Redes Neuronales Artificiales (RNA) en el campo de la Inteligencia Artificial, como una propuesta para promover la resolución ante la problemática de salud pública y los índices muy altos y progresivos del cáncer de mama a nivel mundial.

La primera fase del modelo, parte de la carga y verificación de los datos sobre el estado de pacientes con cáncer de mama, el proceso se hace de forma semi-automática derivando en una base o conjunto de datos de imágenes o *dataset*.



La siguiente fase consiste en la depuración de datos o *Data Cleaning*, a través de técnicas del *Data Mining*, se categorizan cada uno de los componentes para su exploración, manipulación y análisis.

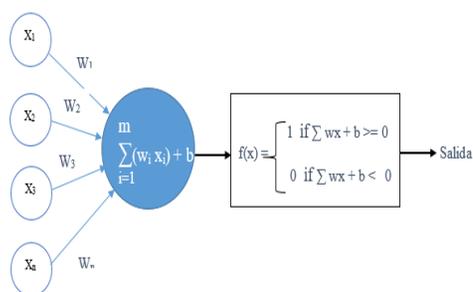
La tercera fase es la Exploración de análisis de los datos o *Exploratory Data Analysis (EDA)*, se evalúan los datos, coherencia y consistencia que inciden sobre los resultados.

La cuarta fase consiste en el entrenamiento, un algoritmo para establecer pautas que acondicionan el escenario para aplicar las RNA, definir entradas, la capa oculta, salida y, el resultado, tal como lo realizaría un perceptrón.

humano, se emplean diferentes capas neuronales compuestas por pesos (*números*) que corresponden a *nodos* diseñados en tres niveles o capas: capa de entrada, capa oculta y capa de salida.

Tal como se muestran en la siguiente figura, la capa de entrada depende de la cantidad de entradas, éstas afectan directamente a los nodos de la capa oculta, cada nodo de entrada está relacionado con cada nodo empleado en la capa oculta, y esta a su vez con los nodos de la capa de salida, es en ésta última capa donde se resumen en señales de salida, típicamente se resume en cantidad hasta el resultante de la salida.

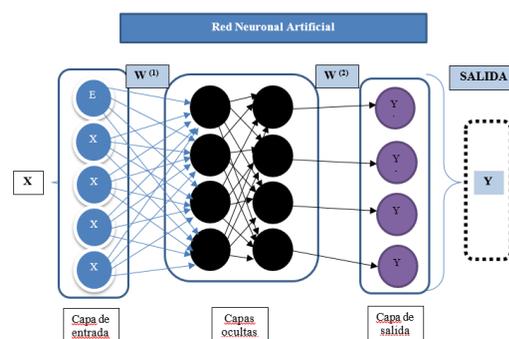
Figura 1
Estructura del perceptrón.



Fuente: Elaboración propia.

Los resultados algorítmicos bajo este enfoque se caracterizan por emular el funcionamiento cerebral

Figura 2
Capas en una Red Neuronal Artificial.



Fuente: Elaboración propia.

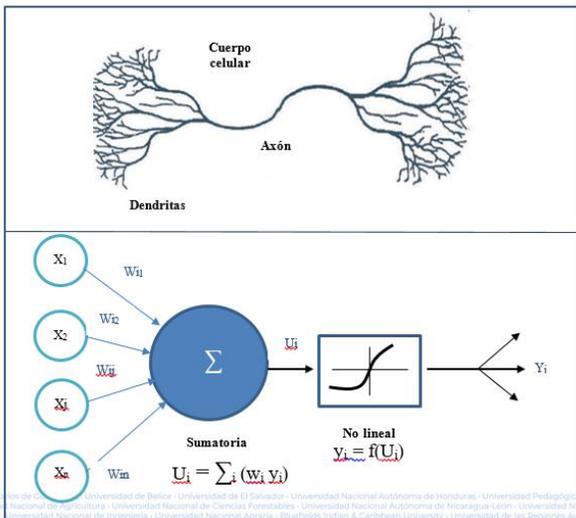


Contraste y similitud

Aunque una RNA tiene a efectos imitar el sistema neuronal del humano, en realidad no funciona idéntico, ni siquiera por el uso de millones de neuronas para realizar las funciones que modela, sin embargo, la interconexión, los pesos y el resultado, en función y logro, es muy similar.

Para comprender un tanto sobre el paralelismo idealizado entre una RNA y una Red Neuronal Humana (RNH), se procede ilustrarlo con la siguiente figura.

Figura 3
Capas en una Red Neuronal Artificial.



Fuente: Elaboración propia.

En cuanto al ciclo de desarrollo de la ingeniería del software, hay una fase de requerimientos que responde un campo vital en el proceso de

pronósticos sobre el cáncer de mama. La etapa inicial implica identificar las necesidades y exigencias técnicas que apuntan a la tercera fase de análisis de sistemas, se elabora un esquema de trabajo o ruta de soluciones validándolo con procesos heurístico y lógico-matemático. Concerniente al diseño del sistema en general, implica la estructura de los datos, refiere la información en atributos del dataset, así:

Figura 4.
Estructura de la base de datos.

- id_numero (int),
- espesor_grupo (10),
- uniformidad_tamano_celular (10),
- uniformidad_forma_celular (10),
- adhesion_marginal (10),
- tamano_una_celula_epitelial (10),
- nucleos_desnudos (10),
- cromatina_suave (10),
- nucleolos_normales (10),
- mitosis (10),
- clase_cancer: (2: Benigno, 4: Maligno).

Fuente: Elaboración propia.

Por otra parte, la fase de diseño implica estructurar donde se almacena, calcula y depura, este proceso considera el almacenamiento puro de la base de datos, así como también el proceso de depuración conocida como Data Mining o Minería de Datos, se



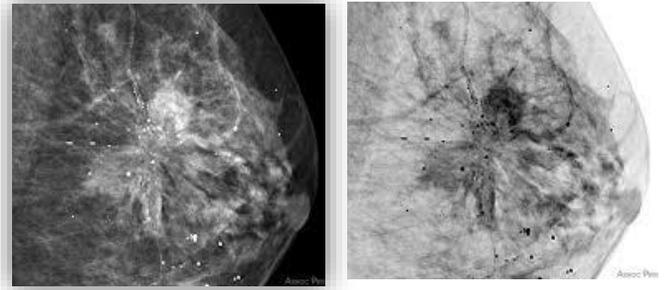
aplica Data Cleaning y Wrangling ya referidos en la segunda fase.

Finalmente, se invocan las librerías Keras y Tensor GPU (Graphic Processing Unit), desde el ambiente Python, dichos paquetes son particulares en el ámbito de las Convolutional Neural Networks (CNN), o Redes Neuronales Convolucionales (RNC).

El resultante de la trama del modelo de la Red Neuronal en respuesta de las capas de Entrada, Oculta y Salida, se asume una *dimensión* de 30 *entradas*, 16 en la primera instancia, luego estas *unidades* se distribuyen en dos capas donde establecen en dos valores de 8, cabe señalar que esta referencia del modelo es el más ajustado de acuerdo a la fuente de datos.

Por otra parte, la activación RELU o Rectified Lineal Unit, que corresponde al ajuste del trazado sigmoidal hacia una unidad más lineal, permite la omisión de regiones de saturación, dicha función en otras palabras permite que la lectura de la fuente sea más lineal, induciendo a localizar fuentes de datos provenientes de regiones con bajos niveles de saturación o ruido.

Figura 5
Tratamiento de la mamografía.

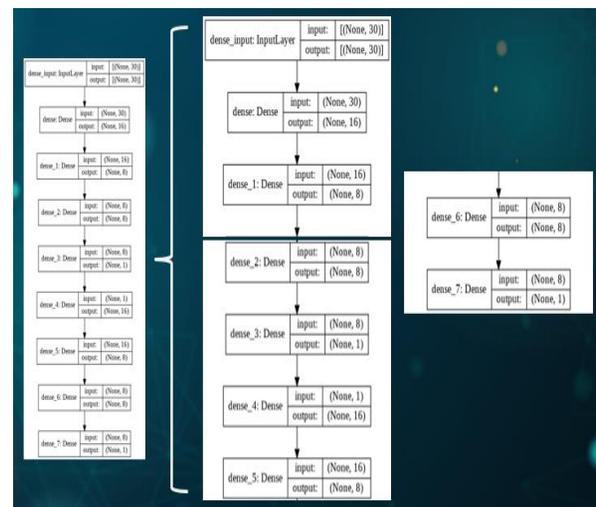


Fuente: Elaboración propia.

Con respecto a la capa de salida, se establece un clasificador de densidad con el valor de uno (1) para la *unidad*, y una *activación* de tipo *sigmoidal*.

La siguiente ilustración representa la salida del modelo bajo las consideraciones previamente referidas.

Figura 6
Trama en capas del modelo de Red Neuronal Artificial.



Fuente: Elaboración propia.



El modelo de esta Red Neuronal Artificial se emplea con la fijación del clasificador, pasando los datos de entrenamiento de X e Y. De forma tal que la ANN podrá detectar los patrones en los datos derivando en la construcción de una Red Neuronal Artificial para realizar predicciones.

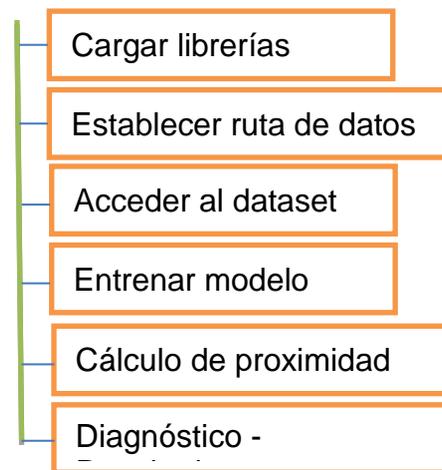
En cuanto al tamaño del batch será de 1, mientras que el número de veces a ejecutar el algoritmo *forwardpropagation* - *backpropagation* será de 50 veces. Cincuenta ciclos (*epoch*) donde todos los datos de entrenamiento pasan por la RNA para que el algoritmo aprenda sobre ellos, si existen 10 ciclos y 1000 datos, cada ciclo los 1000 datos pasarán por la RNA. De hecho, para predecir los nuevos valores, se emplea en la RNA un clasificador predictivo con valores de prueba para X, siendo estos datos desconocidos por la RNA, de forma tal que será el marco referente para devolver el conjunto de predicciones tipificadas con 0 a 1, donde se emplea un criterio de redondeo que, a razón de resultados mayores e iguales a 0,5 se cuentan por 1, si no, se mantienen en 0.

La marcación de 0 y 1, permite comparar en el recorrido del conjunto

de predicciones y valores reales a compararse, etiquetando como *Correcto* o *Incorrecto*; en su recorrido, interviene un *contador* que *totaliza*, de igual manera hay un acumulador para aquellos casos condicionados como *Correcto* o cuando es *Incorrecto*.

Con respecto a la estructura del modelo algorítmico basado en la Red Neuronal Artificial en Inteligencia Artificial, se describe así:

Figura 7
Estructura procedimental.



Fuente: Elaboración propia.

Se procede con un diseño lógico y heurístico como referente en la constitución algorítmica, los aspectos de código y acceso a los servicios procedimentales se desempeña por medio de una infraestructura tecnológica y el hardware que implica los procesamientos de los datos por medio de *clouding*, las cuales



repercuten en la ejecución algorítmica, la valoración del modelo y las vistas de los resultados.

Para el resultado final, se reitera que es resultante de los cálculos de las capas propias de la Red Neuronal Artificial, los resultados se presentarán como diagnóstico que valida las condiciones de: Verdadero o Falso positivo, así como también, Falso positivo o Falso negativo.

Conclusiones

Iniciativas en el desarrollo de modelos de Inteligencia Artificial empleando Redes Neuronales Artificiales para el pronóstico del cáncer de mama, crean una esperanza tanto para los afectados directamente, como para la comunidad científica, abordar temas complejos y con un constante incremento, sirven como coadyuvantes el criterio y valoración del cuerpo médico ante los protocolos existentes.

Las pruebas del modelo actual se realizaron con un dataset real de casos de pacientes con positivos en cáncer de mama, tanto malignos como benignos, un total de 455 pacientes donde el modelo de Inteligencia

Artificial totalizó 114 pacientes de los cuales existen 111 predicciones *Correctas*, y 3 predicciones *Incorrectas*, por lo que deriva en una efectividad en la precisión de 96%.

Figura 8
Grado de efectividad en la predicción.



Fuente: Elaboración propia.

Tal como todo modelo, pueden tanto mejorarse los procesos, como anexarse otras funciones, ampliando la posibilidad de la misión que lleva ya la carga en la salud pública, sin embargo, puntualmente se recomiendan los siguientes aspectos:

- De las pruebas realizadas con el modelo, se empleó un dataset real de casos de pacientes positivos con cáncer de mama, malignos como benignos, un total de 455 pacientes donde el modelo de Inteligencia Artificial totalizó 114 pacientes de los cuales existen 110 predicciones *Correctas*, y 4 predicciones *Incorrectas*, derivando



una grado de efectividad o *precisión/accuracy* de 96%.

- Crear otras alternativas complementarias para abordar pacientes con cáncer como un complemento a la comunidad científica.
- Coadyuvar al criterio y valoración del cuerpo médico ante los protocolos existentes.
- Optimizar el número de atributos de entrada (función extracción) en el conjunto de datos para eliminar la limitación de los conjuntos de datos disponibles relacionados con el cáncer.

Bibliografía

- [1] Taylor, Michael. (2017). Neural Networks A Visual Introduction for Beginners. Canada. Blue Windmill Media.
- [2] Dean, Jared. (2014). Big Data, Data, Mining and Machine Learning. United State. Wiley.
- Kubben, Pieter; Dumontier, Michael y Dekker, Andre. (2019). Fundamentals Of Clinical Data Science. United State. Springer.
- [3] Goodfellow , Ian; Bengio, Yoshua y Courville, Aaron. (2015). Deep Learning. United State. MIT, Press.
- [4] Robert W, Nelson. (2011). New Developments in Artificial Neural Networks Research. United State. Nova Science Publishers.
- [5] Slater, Daniel y Vasilev, Ivan. (2019). Python Deep Learning Exploring techniques, NN architectures and GANs with PyTorch, Keras and TensorFlow. United Kingdom. Packt Publishing.
- [6] Géron, Aurélien. (2019). Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly.
- [7] Chollet, Francois. (2018). Deep Learning with Python. United State. Manning Publications, Co.