

Inteligencia artificial: herramienta diagnóstica para el cáncer de mama

Gianfranco Jesús Curci Robledo, Miguel Mendez Garabeti, Pablo Javier Sandez

Laboratorio de Investigación en Ciencia y Tecnología, Facultad de Ciencias Sociales y Administrativas, Universidad del Aconcagua, Mendoza, Argentina.

gianuniversidad@gmail.com

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 este artículo, se dispone trabajar con datos de pacientes de un padecimiento muy frecuente en la actualidad, que es el cáncer de mama. Para mejorar la precisión diagnóstica precoz de dicha con el uso de redes neuronales.

Keywords: inteligencia artificial · diagnóstico clínico · cáncer de mama · minería de datos.

1 Introducción

La detección temprana de cáncer de mama contribuye a una reducción eficaz de la mortalidad producida por esta enfermedad. Dicha detección se hace en una primera instancia a través de la palpación personal de la mama, por parte del paciente, seguida de estudios imagenológicos y de laboratorio. Entre los estudios de laboratorio, encontramos aquellos dirigidos a la expresión de marcadores tumorales aumentados en sangre y estudios relacionados con la expresión de ciertos genes. Teniendo todos estos estudios, ya el diagnóstico certero depende de la habilidad del oncólogo a cargo del caso. Muchas veces de las cuales, este diagnóstico por falta de análisis del caso de los estudios genéticos anteriormente nombrados, se asume que el paciente, a pesar de tener alterados genes como: BRCA1 y el BRCA2, está sano o bien no es un posible candidato a desarrollar algún tipo de cáncer de mama en el futuro. Para evitar este sesgo diagnóstico, este trabajo plantea de manera teórica la posibilidad de utilizar diferentes herramientas de la inteligencia artificial, de manera de poder plantear algún modelo experimental de IA que permita en base a la expresión de los mencionados los genes alterados y otros estudios positivos. Brindar un porcentaje que refleje la posibilidad de desarrollar o no la enfermedad a futuro, es decir, usar una predicción. Este modelo se utilizaría como una “segunda opinión” que puede utilizar el medico a cargo, para justificar la realización de un tratamiento preventivo o no.

2 Descripción 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 (IA) 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, 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. El problema a resolver está dirigido al grado de efectividad en el modelo por pronósticos para la detección del cáncer de mama.

3 Análisis del problema

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

4 Propuesta de modelo

La iniciativa del 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.

El modelo propone una solución complementaria 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. Se utilizará un perceptrón multicapa con un aprendizaje de tipo supervisado y una función de activación de tipo sigmoïdal. 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 semiautomática derivando en un dataset. La información de susodicho, estaría recopilada de pacientes que tengan aquellos genes específicos alterados y sus derivados, tales como BRCA1 o BRCA2. Esta información se usará entrenamiento del perceptrón y sería provista por la IMBECU.

La segunda 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. Entre las propuestas analizadas, se utilizará Weka como software elegido para esta tarea, debido a su optimización y aval científico por la Universidad de Waikato. Se conformará, además, el grupo de entrenamiento y el conjunto de validación de datos.

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. La quinta 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.

La quinta fase constaría de evaluar la mejor arquitectura posible para el perceptrón, teniendo en cuenta la tasa de aprendizaje, número de capas y neuronas requeridas para el proceso, como bien se sabe este proceso, aunque sea poco ortodoxo, es más de prueba y error; el resultado obtenido se evaluará con el conjunto de validación y se ajustará la arquitectura, tasa de aprendizaje, número de capas y función de activación.

5 Resultado esperado

Se espera que dicho modelo pueda predecir y clasificar de manera correcta los pacientes de manera correcta. En aquellos que si tienen alta probabilidad de tener cáncer de mama de los que no. Ayudando así a la decisión del especialista, de la posibilidad de sugerir un tratamiento preventivo, un seguimiento exhaustivo o bien confirmar con otra herramienta diagnóstica, que efectivamente dicho paciente no desarrollara la enfermedad. Para lograr dicha hazaña, se calcula que al menos la muestra de pacientes debe ser de al menos 100, a mayor cantidad de muestras, diversificación y redundancia de algunos datos, mayor será la credi-

bilidad de los resultados arrojados por la IA. Podemos decir que dicha es eficaz, cuando su porcentaje de precisión supere el 95 por ciento.

6 Conclusiones

La iniciativa del 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.

Este proyecto busca lograr los siguientes objetivos:

- 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.

Referencias

1. Jose Francisco, F., Ávila Tomás Miguel, S., Angel Mayer Pujadas, T., Victor Julio Quesada Varela: La inteligencia artificial y sus aplicaciones en medicina I: introducción antecedentes a la IA y robótica (2020)
2. Jose Francisco, F., Ávila Tomás Miguel, S., Angel Mayer Pujadas, T., Victor Julio Quesada Varela: La inteligencia artificial y sus aplicaciones en medicina II: importancia actual y aplicaciones prácticas (2021)
3. Saúl Oswaldo Lugo-Reyes, F.: Inteligencia artificial para asistir el diagnóstico clínico en medicina. Revista Alergia Mexico (RAM) (2014)
4. Amrita Naika, F Lilavati Samantb, S.: Correlation Review of Classification Algorithm Using Data Mining Tool: WEKA, Rapidminer, Tanagra, Orange and Knime. Conferencia internacional de seguridad y computación (2016)
5. Pedro Isasi, F., Ávila Tomás Miguel, S., Inés M. Galván: Redes neuronales artificiales: un enfoque práctico. 1st ed. Publisher, Pearson-Prentice Hall, Madrid (2004)