



Aprendizaje profundo para escalas pronósticas en la prescripción a pacientes con riesgo de sangrado gastrointestinal

 Carlos Calderón-Vargas¹, José Muñoz Castaño², María Vargas Rincón³,
Víctor Rincón Acosta⁴ y  Miguel Mendieta Hernández⁵

Recepción: 09-06-2021 - Aceptación: 08-11-2021 - En línea: 01-12-2021

MSC: 62P10 - PACS: 42.62.Be

doi:10.17230/ingciencia.17.34.1

Resumen

La evolución de la medicina en los tiempos actuales ha ido de la mano de la tecnología donde cada vez más se implementan soluciones que apoyan ciertos procedimientos médicos con el objetivo de apoyar el ejercicio de los profesionales de la medicina en su oficio. El procesamiento y análisis de datos se ha convertido en un recurso imprescindible en la práctica de cualquier profesión, actualmente, en los hospitales, más puntualmente en el hospital universitario la samaritana, no se posee una herramienta que permita apoyar el diagnóstico para determinar el suministro o no, de los inhibidores de bombas de protones, por lo tanto hemos desarrollado

¹ Hospital Universitario de la Samaritana, caldecal@gmail.com, Bogotá, Colombia.

² Hospital Universitario de la Samaritana, jose.munoz.hospitalario@gmail.com, Bogotá, Colombia.

³ Universidad de la Sabana, marialevr01@gmail.com, Bogotá, Colombia.

⁴ Universidad El Bosque, vrincona@unbosque.edu.co, Bogotá, Colombia.

⁵ Universidad El Bosque, mmendietah@unbosque.edu.co, Bogotá, Colombia.

una aplicación web utilizando un modelo de aprendizaje automático, basado en árboles de decisiones por medio de la aplicación weka, que luego del análisis de los datos recogidos, permita al médico contar con una herramienta para el apoyo de este procedimiento. Esperamos que con la utilización de esta aplicación los médicos puedan realizar un análisis efectivo antes de recetar o no los IBPs.

Palabras clave: Diseño web; machine learning; entrenamiento; árboles de decisión; weka.

Deep Learning for Forecast Scales to Prescribe Patients at Risk of Gastrointestinal Bleeding

Abstract

The evolution of medicine in current times has gone hand in hand with technology where more and more solutions are implemented; those supporting certain medical procedures to serve as base in the field of medical professionals. The process of analyzing data has become an essential resource in the practice of any profession; currently, in hospitals, more specifically in the university hospital La Samaritana. No tool allows the supporting of diagnosis to determine the supply or no, proton pump inhibitors, therefore we have developed an app using a machine learning model, based on decision trees through the weka application, which, after analyzing the data collected, allows the doctor to count with a tool to support this procedure. We hope that with this, doctors can perform an effective analysis before prescribing or not prescribing PPIs.

Keywords: Web design; machine learning; training; decision trees; weka.

1 Introducción

Los inhibidores de bomba de protones (IBP) se han relacionado con el desarrollo de neumonía e infección por *C. difficile*, no obstante, la evidencia es contradictoria en cuanto este riesgo ya que los estudios aleatorizados controlados en comparación a los estudios observacionales no muestran un aumento de riesgo de neumonía en comparación con el otro tipo de estudios [1]. En un metaanálisis de 6 estudios de casos y controles se demostró un aumento en el riesgo de neumonía adquirida en la comunidad con un OR 1.92, 95% IC 1.40 -2.63 p 0.003 para aquellos pacientes que usaron IBPs, incluso en un corto periodo de tiempo [2]. De igual forma

en el metaanálisis realizado por Eom Cs et. al se demostró un aumento del riesgo de neumonía con el uso de IBPs OR 1.27 95 % IC 1.11 - 1.46 [3]. Adicionalmente, el riesgo de *C. difficile* también tiende a ser mucho mayor en aquellos que reciben IBPs fuera de la UCI, así como lo muestra una revisión sistemática publicada por Leonard et. al (OR 2.05, 95 % IC 1.47-2.85) [4]. Estudios como éste obligaron a la FDA (Food and Drug Administration) a realizar una revisión de múltiples estudios (la mayoría observacionales) con el fin de determinar la correlación entre los IBPs y la infección por este Bacilo Gram positivo, encontrando que, aunque la calidad de la evidencia no es fuerte, el riesgo de desarrollar la enfermedad por *C difficile* es 1.4 a 2.75 veces mayor en comparación a aquellos que no reciben IBPs [5]. Basados en esto se han intentado realizar escalas pronósticas que ayuden al médico a identificar aquel paciente que está hospitalizado fuera de la unidad de cuidados intensivos y que tiene mayor riesgo de presentar sangrado gastrointestinal secundario al desarrollo de úlceras por estrés, es por esto que se plantea la utilización del machine learning o el aprendizaje automático, tema que se plantea el estudio de cómo usar computadoras para simular actividades de aprendizaje humano y estudiar métodos de automejora de computadoras para obtener nuevos conocimiento y nuevas habilidades, identificando conocimiento existente, y mejorando continuamente el rendimiento y el logro. En comparación con el aprendizaje humano, el aprendizaje automático aprende más rápido, la acumulación de conocimiento es más fácil. Los resultados del aprendizaje se difunden más fácilmente. Entonces, cualquier progreso humano en el campo del aprendizaje automático, mejorará la capacidad de computadoras, por lo tanto tienen un impacto en la sociedad humana [6]; entendiendo lo anteriormente mencionado se utilizará la técnica de minería de datos o explotación de Información, la cual consiste en el proceso de extraer conocimiento útil, comprensible y novedoso de grandes volúmenes de datos, siendo su principal objetivo encontrar información oculta o implícita, que no es posible obtener mediante métodos estadísticos convencionales. La entrada al proceso de minería está formada generalmente por registros provenientes de bases de datos operacionales o bien bodegas de datos (Datawarehouse) [7].

2 Metodología

Se realizó una búsqueda bibliográfica sistemática en la Web of Science (WoS), bases de datos IEEEExplore y ACM digital biblioteca para identificar todos los estudios potencialmente relevantes desde 2015 hasta el presente. Se incluyeron todos los estudios que evaluaron los modelos de aprendizaje automático (ML) utilizados para el diagnóstico médico.

2.1 Estrategia de búsqueda

Para la calidad de la búsqueda, se llevó a cabo una investigación de los últimos cinco años. Esta investigación se desarrollo en las áreas de la informática y la salud, tomando como referencia principalmente el artículo de Gibbons [8] y el artículo de Vega, Mora y Badilla, [9] se extrajo los datos recolectados para crear un esquema de metodología diagnóstica de la información, lo que nos permite a través de datos y técnicas trabajadas en pasadas investigaciones, crear modelos predictivos aplicados a distintos campos de la medicina, en este caso para el diagnóstico de pacientes. Para obtener un diagnóstico se debe pasar por una predicción y para predecirla se utilizan algoritmos donde se procesa la información, con la información procesada el aprendizaje automático puede actuar para tener un diagnóstico sobre una enfermedad. El objetivo es encontrar la predicción y el algoritmo correctos para el uso de IBP's en diagnostico de sangrado gastrointestinal con aprendizaje automático y mejorar el área de la salud. Para esto, se usó la metodología de investigación Cuantitativa, la cual, frecuentemente cuantifica relaciones entre variables, la variable independiente o predictiva y la variable dependiente o resultado [10].

2.2 Síntesis de datos

El aprendizaje automático identifica patrones de datos procesados por algoritmos. Predice comportamientos con algoritmos y aprende de eso, con los datos de predicción y los datos identificados, se puede hacer un prescripción. Para el análisis de estos patrones, se implementan con una sintaxis diferente y los datos construyen una arquitectura para que

se alimente la predicción [11]. El uso de sistemas de clasificación en la prescripción médica se está volviendo cada vez más popular. La evaluación y recopilación de datos de los pacientes y el análisis de expertos son los factores clave para una prescripción exitosa [12]. Las variables principales para una buena prescripción son: la predicción, el algoritmo y el componente principal. La salida de tres variables de entrada será la prescripción y el proceso de las tres variables mencionadas será el aprendizaje automático.

2.3 Síntesis de desarrollo

Al ser un proyecto de una no tan larga duración por parte del desarrollo y de una continua necesidad de cambios se seleccionó la metodología ágil Scrum, además de que es muy flexible ante los cambios del proyecto, también tiene como ventaja el continuo Feedback acortando el tiempo de entrega, lo cual permite la capacidad de respuesta ante los cambios. El proyecto completo conforma 6 Springs a lo largo de un año, repartiendo la recolección de datos, análisis y validación de los datos, la aplicación de software junto a la generación de resultados, el desarrollo e implementación de pruebas y finalmente, la elaboración de informes y productos finales respectivamente en cada Spring. Para facilitar la organización de las actividades fue usado el tablero kanbanchi con 62 tareas asignadas respectivamente a los investigadores.

3 Marco teórico

3.1 Machine learning: aprendizaje automático

El aprendizaje automático o también llamado aprendizaje de máquina lo podemos definir refiriéndonos a Arthur Samuel,[13] como una rama de la inteligencia artificial que utiliza algoritmos para realizar el procesamiento y análisis de datos, que les permita a los computadores generar resultados basándose en las tendencias que encuentren en los datos de entrada, dando así también la posibilidad de crear modelos independientes sin necesidad de las instrucciones precisas de un programador. Un modelo predictivo o de aprendizaje automático lo podemos segmentar en tres

momentos principales: la alimentación de datos procesados previamente para la ejecución del algoritmo de machine learning que se definió, donde para nuestro caso utilizamos los datos clínicos de los pacientes en 8 variables independientes, y una variable de salida que nos indicaría si es o no requerida la prescripción de los IBPS, luego de esto, el algoritmo se encarga de analizar los datos y encontrar puntos en común entre los diferentes registros ingresados, comparando la entrada de los datos y las salidas esperadas, que para nuestro caso la salida era la necesidad de la prescripción o la no prescripción de los IBPS, luego de que el modelo predictivo ha sido creado por el algoritmo donde en nuestra investigación utilizamos el aplicativo WEKA y el algoritmo (J48), tendremos un modelo final que nos permitirá predecir las nuevas salidas de acuerdo a la entrada de los nuevos datos que se dispongan [14]. Esta rama de la informática esta muy relacionada con aplicaciones mas comunes y actuales como el modelamiento de datos y la simulación de sistemas, debido a que estos últimos mencionados se basan en el análisis superficial de los datos encontrando similitudes, sin entrar mas a fondo a entender cual es la relación que existe entre las variables del modelo. Para entender un poco mas como funciona el aprendizaje automático es importante analizar las diferentes técnicas que encontramos en el mismo, las tres variantes mas conocidas de machine learning son:

3.2 Aprendizaje supervisado:

Cada dato registrado esta relacionado o etiquetado con respecto a una variable objetivo o de salida, para identificar el comportamiento de los datos con respecto a mi salida esperada, por ejemplo, indicar la prescripción de los IBPS[4].

3.3 Aprendizaje no supervisado:

Esta variante de machine learning tiene una diferencia puntual con el aprendizaje supervisado y es que, no definimos una variable de salida para esta técnica, como si lo hacíamos antes en el aprendizaje supervisado, por el contrario, este busca realizar una clasificación de los datos identificando

patrones, sin etiquetas impuestas por el usuario. Este método se utiliza entre otras cosas para la reducción de la cantidad de datos (“Dimension Reduction Technique”)[8].

3.4 Aprendizaje por refuerzo:

Este ultimo modelo nos ofrece un procesamiento de datos mixtos, trabaja tanto con datos etiquetados como el aprendizaje supervisado, y con datos no etiquetados como el no supervisado, interactuando con el entorno para poder estar en una retroalimentación constante de los resultados, ya sea negativos o positivos.

Las categorías de machine learning mencionadas anteriormente son las principales y debemos dividir las en unas variantes del aprendizaje automatizado segun el objetivo que tengamos con el modelo: para el aprendizaje supervisado podemos encontrar: Arboles de decisión, redes neuronales, regresión lineal, entre otros. Para el aprendizaje no supervisado existen: hierarchical clustering, k-means clustering, linear discriminant analysis, con lo anterior se puede tener una vista general de como funciona el aprendizaje automático y su aplicación para el procesamiento de datos, en este caso clínicos, para el cual utilizamos arboles de decisión y aprendizaje supervisado [9].

4 Desarrollo

A través de técnicas de machine learning se puede construir una herramienta que mejora la prescripción y aplicación de las escalas de Herzing et al en el paciente hospitalizado en sala general del Hospital Universitario de la Samaritana. Para esto, antes del propio desarrollo de la herramienta de apoyo se realizaron dos pasos previos, los cuales son el apoyo para la búsqueda del algoritmo de machine learning, en donde permitan conocer cuándo un paciente se beneficia de gastro protección y la implementar pruebas sobre los datos que fueron previamente depurados y configurados para trabajar en los softwares requeridos. Luego de este, se procedio a diseñar y construir una aplicación web, en donde incluya

formularios HTML para el despliegue de la solución y finalmente, se dio el apoyo en la documentación para la construcción de la aplicación, en donde incluya el manual de usuario y técnico para un adecuado manejo para dicha solución.

4.1 Búsqueda del algoritmo

Luego de analizar las principales variantes que tiene machine learning y sus técnicas de análisis se decidió utilizar un aprendizaje supervisado, ya que nuestro modelo dependía de una variable objetivo para la salida de los datos, en este caso la afirmación o negación de la prescripción del medicamento, teniendo en cuenta las demás variables dependientes, por esta razón utilizamos un algoritmo J48 que nos permitía trabajar de manera dicotómica a utilizando arboles de decisión.

4.2 Aplicación del software requerido

Frente a la aplicación del software, se usó WEKA para la implementación de los datos previamente depurados y configurados en el árbol de decisiones seleccionado. Se usó un total de 2174 datos que luego del filtrado se obtuvieron 156 datos que se les logro implementar las escalas diagnosticas como se muestra en la Tabla 1.

En esta tabla se muestra la relación entre los datos y el momento en el tiempo que fueron sacados, los cuales fueron enero, febrero, marzo y abril 2019. En las columnas se encuentra tanto el resultado de la depuración como el cálculo de escala manual de los investigadores, estos últimos con el criterio de si se logro configurar los datos con las escalas o presenta problemas. Finalmente tenemos los datos que responden a cuantos pacientes necesitan realmente Omeprazol y cuales no.

Estos datos finales se colocaron en una hoja de cálculo con el resultado respectivo de los parámetros de las escalas, una por columna. Con la información organizada se procedió a cambiar el formato del documento a uno con valores separados por comas (CSV), esto con el fin de que el aplicativo WEKA logre la lectura de los datos. Finalmente se procedió al uso de machine learning ingresando a la sección de clasificar en el aplicativo,

Tabla 1: Tabla de análisis global de los datos.

Mes		Enero	Febrero	Marzo	Abril	Total
# total de base de datos		510	592	636	436	2174
# Criterios de inclusion		397	433	307	325	1462
mayor a 72 horas pacientes para aplicar la escala		53	29	29	45	156
Base de datos 9pt (Investigadores)	No	45	26	21	37	129
	Si	8	3	8	8	27
Base de datos 8pt (Riesgo bajo)	No	43	25	21	36	125
	Si	10	4	8	9	31
Base de datos 10pt (riesgo intermedio)	No	51	29	26	41	147
	Si	2	0	3	4	9
Aplicacion modelo AI	No	45	26	22	38	131
	Si	8	3	7	7	25

se selección el árbol J48 para este proyecto y una validación cruzada de 10. El resultado es un árbol de decisiones del riesgo baro en las escalas, su modelado se muestra en la Figura 1

4.3 Pagina web

Durante el desarrollo de la investigación se diseñó un sitio web, para llevar un registro actualizado de los avances del proyecto¹. Allí se puede encontrar discriminado por meses los distintos arboles de decisión creados para el modelo, conforme se iba avanzando en el análisis y procesamiento de los datos obtenidos del Hospital. También se encuentra en este espacio las herramientas de análisis de datos utilizadas para el proyecto como lo son el ROC y el PCA, de la misma manera que los arboles seccionados de acuerdo a los meses y los datos recolectados. Luego de la creación del modelo utilizando un algoritmo de aprendizaje automático, se procedió

¹<http://cdtiueb.com.co/HUSB/web/>

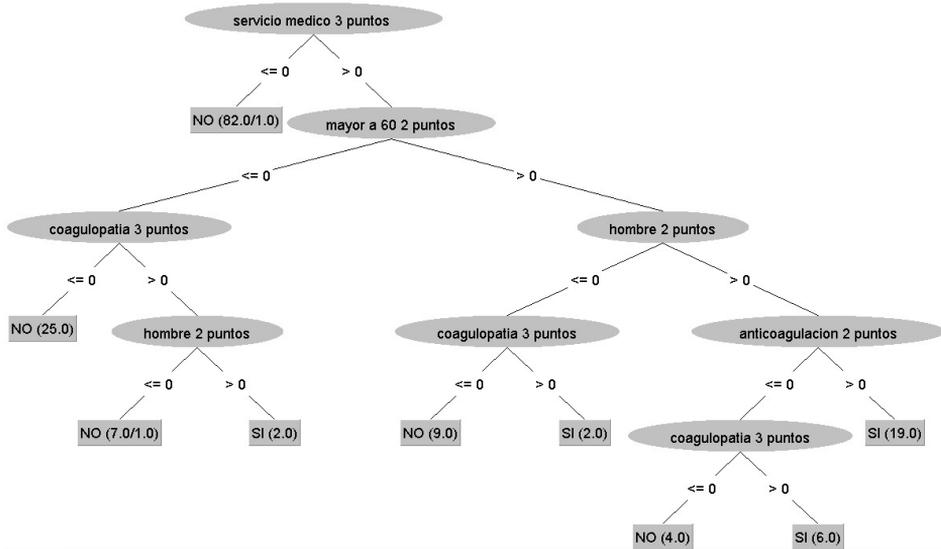


Figura 1: Árbol de decisiones final.

a diseñar nuestra solución web, la herramienta que será utilizada como apoyo por los médicos en el proceso de prescripción de los IBPS, para este aplicativo se utilizó Bootstrap, JavaScript, CSS y HTML, a través de eventos y condicionales se realizó la programación de los botones del formulario HTML para representar el flujo del algoritmo de un árbol de decisiones dicotómico. De acuerdo a la elección de los usuarios de la página se despliega una respuesta que corresponda a las diferentes ramas del árbol resultante del modelo predictivo de la investigación, cada una de las preguntas posee etiquetas relacionadas con el nombre de las variables a tener en cuenta para la prescripción de los medicamentos, de una manera secuencial, el aplicativo guía al usuario a través de las distintas preguntas, para finalmente emitir un mensaje final que indique si es necesario la prescripción de la medicina o si por el contrario, no se requiere como se evidencia en la Figura 2.

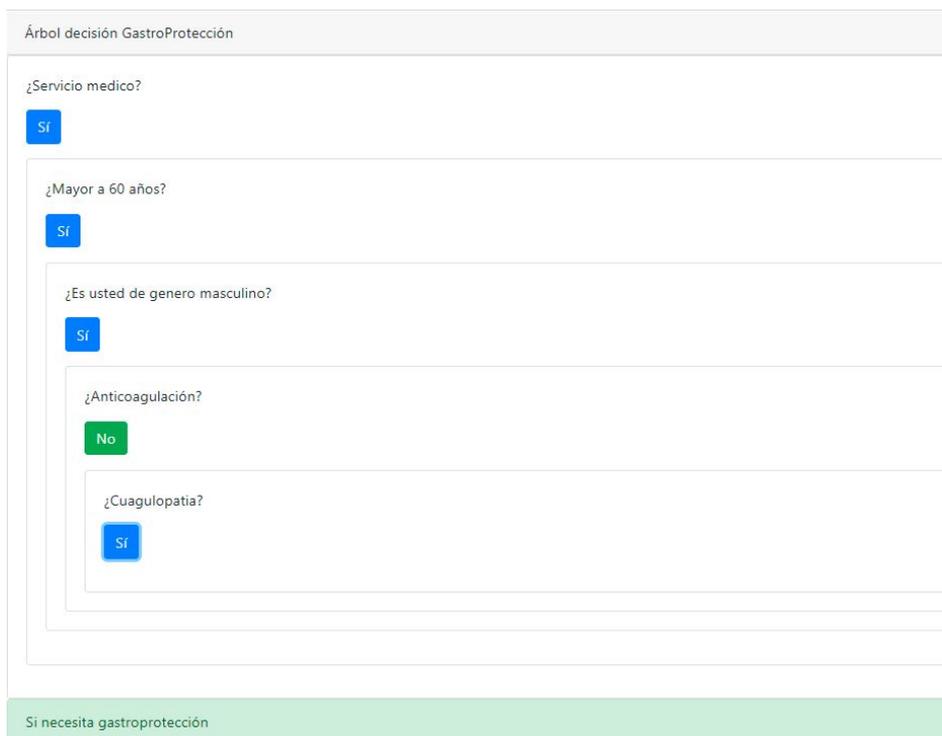


Figura 2: Árbol HTML

4.4 Documentación para la construcción de la aplicación y su uso

En el campo de la ingeniería uno de los apartes fundamentales para garantizar el éxito de un proyecto, es realizar una documentación adecuada, por esto, como parte del proyecto de investigación definimos unos manuales para el uso y entendimiento del aplicativo desarrollado, la documentación se divide en dos partes principales: La documentación orientada al usuario donde se encuentra la información relacionada con el funcionamiento del árbol de decisiones en HTML y el registro de los datos trabajados durante la investigación para la comprensión de los árboles de decisiones y las herramientas utilizadas para el análisis de los datos. En la documentación

técnica orientada a usuarios con conocimientos básicos en informática, describimos todo el proceso de implementación del aplicativo, así como el lenguaje utilizado para el desarrollo y la estructura del código utilizado.

4.5 Implementación de pruebas

Como principal foco de las pruebas fue la variable de sensibilidad del modelo, eso ton el fin de demostrar la veracidad de árbol concluyente, para esto se realizaron 2 tipos de pruebas: validación en la academia y dinámica. Por parte de la academia se dos variables de validación:

- i Nivel umbral:

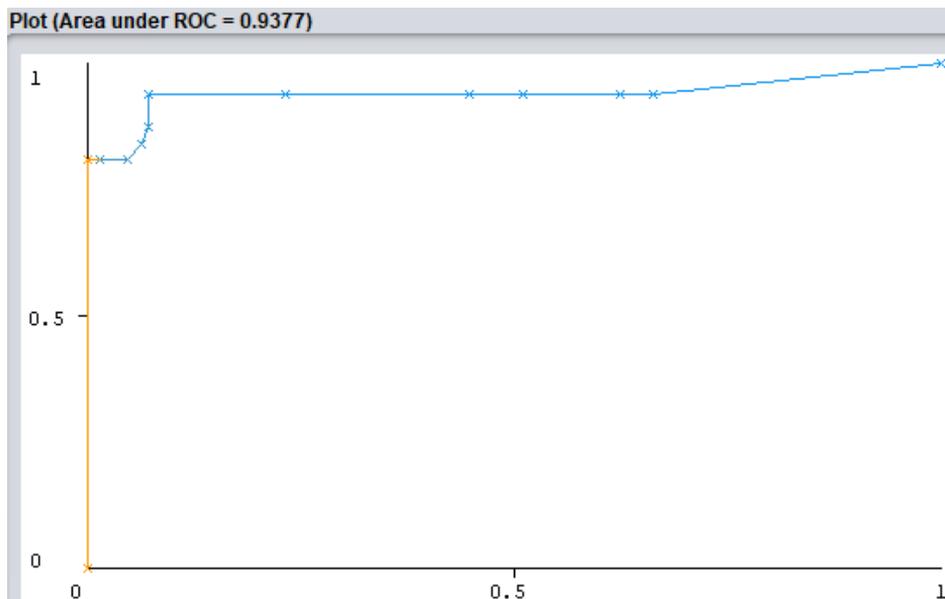


Figura 3: ROC final.

Este apunta directamente a la sensibilidad del modelo a través del punto de corte de la curva ROC, de la cual se estimó un día de duración para la realización de las pruebas. Como resultado, las Características Operativas

del Receptor, el cual es una representación gráfica de la sensibilidad de una prueba diagnóstica en función de los falsos positivos del modelo evaluado. Se alcanzó un 93.77% de área bajo la curva, lo cual representa un gran asertividad del modelo ante los datos suministrados como se muestra en la Figura 3.

ii Precisión:

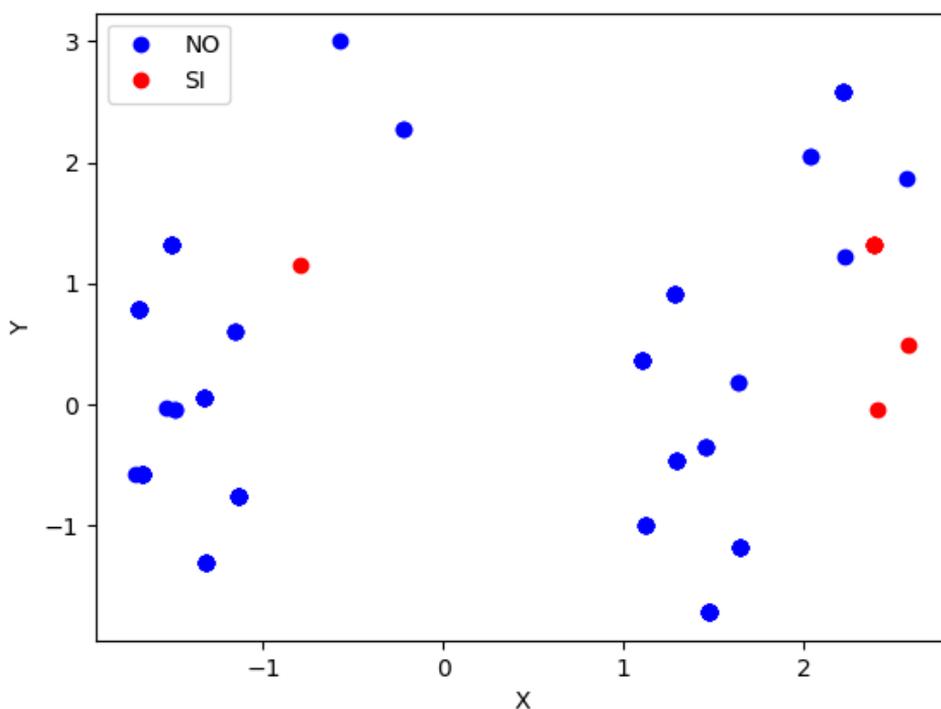


Figura 4: PCA final.

Por otro lado, en esta validación se buscó la especificidad del modelo a través de la probabilidad de que un individuo no necesite IBPs. Se realizó de forma automática esta prueba usando la herramienta WEKA, mediante la creación del modelo el propio software permite analizar diversos componentes como lo es error absoluto relativo, el cual permite conocer la posibilidad de fallo. Este dio como resultado 16.24% en el modelo

final, lo cual permite demostrar una mejora sustancial en el diagnóstico del sangrado gastrointestinal. Esta estadística se ve relacionada en el PCA del modelo, el cual, es análisis de componentes principales es una transformación de los datos en un nuevo sistema de coordenadas, escala los datos y diagrama la dispersión. Como se muestra en la Figura 4, los datos se visualizan en puntos azules como los pacientes que no necesitan Omeprazol y en rojo los que sí, todos estos tras el resultado del diagnóstico manual con las escalas sin haber pasado por el software. La grafica nos muestra un comportamiento no aleatorio que permite reconocer una inclinación de los pacientes que necesitan medicamento, pero con leves datos mezclados que varían la precisión del modelo.

Por parte de la validación dinámica solo se presento una prueba, la cual busca validar el desempeño del modelo a través de los falsos positivos y falsos negativos que arroje el modelo, esto mediante la experimentación con casos de prueba. El resultado final es una matriz de confusión como se muestra en la Tabla 2.

Tabla 2: Matriz de confusión del modelo.

	Positivos	Negativos
Verdaderos	25	6
Falsos	0	125

5 Conclusiones

Se realizó un trabajo en sincronía constante con el director investigador del proyecto, y con el grupo RICAUTA lo que permitió el análisis y mejora constante sobre el producto, lo que permitió que concluyera cual fue el modelo más efectivo para el proyecto luego de la realización de distintos arboles a medida que se iban analizando los datos, finalmente decidiendo que se tomaría el de 8 puntos. Se logró el desarrollar un software para el diagnóstico de riesgo de sangrado gastrointestinal y así apoyar el proceso de diagnóstico realizado por los médicos en el Hospital Universitario La Samaritana, utilizando aprendizaje automático. Se concluyó que el algoritmo más efectivo para el modelo de datos recolectado, es el J48,

ya que nos permitió la creación de un árbol dicotómico con las salidas requeridas para el modelo de datos, a medida que aumentaba la cantidad de registros en los datos procesados, el árbol presentaba variaciones en su extensión y la cantidad de salidas, a través del procesamiento de más datos, logramos llegar al árbol más apropiado para el modelo de datos. La informática ha revolucionado la medicina y los procesos adosados a esta, los diagnósticos aumentan su precisión de acuerdo a la cantidad de información registrada de un suceso determinado, por lo tanto el procesamiento de datos en grandes cantidades, disminuye el error.

Referencias

- [1] F. Alshamsi, E. Belley-Cote, D. Cook, S. A. Almenawer, Z. Alqahtani, D. Perri, L. Thabane, A. Al-Omari, K. Lewis, G. Guyatt *et al.*, “Efficacy and safety of proton pump inhibitors for stress ulcer prophylaxis in critically ill patients: a systematic review and meta-analysis of randomized trials,” *Critical Care*, vol. 20, no. 1, pp. 1–12, 2016. <https://doi.org/10.1186/s13054-016-1305-6> 8
- [2] J. Johnstone, K. Nerenberg, and M. Loeb, “Meta-analysis: proton pump inhibitor use and the risk of community-acquired pneumonia,” *Alimentary pharmacology & therapeutics*, vol. 31, no. 11, pp. 1165–1177, 2010. <https://doi.org/10.1111/j.1365-2036.2010.04284.x> 8
- [3] C.-S. Eom, C. Y. Jeon, J.-W. Lim, E.-G. Cho, S. M. Park, and K.-S. Lee, “Use of acid-suppressive drugs and risk of pneumonia: a systematic review and meta-analysis,” *Cmaj*, vol. 183, no. 3, pp. 310–319, 2011. <https://doi.org/10.1503/cmaj.092129> 9
- [4] J. Leonard, J. K. Marshall, and P. Moayyedi, “Systematic review of the risk of enteric infection in patients taking acid suppression,” *American Journal of Gastroenterology*, vol. 102, no. 9, pp. 2047–2056, 2007. <https://doi.org/10.1111/j.1572-0241.2007.01275.x> 9
- [5] U. Food, D. Administration *et al.*, “Fda drug safety communication: Clostridium difficile-associated diarrhea can be associated with stomach acid drugs known as proton pump inhibitors (ppis),” *Silver Spring, MD: US Food and Drug Administration*, 2013. 9
- [6] H. Wang, C. Ma, and L. Zhou, “A brief review of machine learning and its application,” in *2009 international conference on information engineering*

- and computer science*. IEEE, 2009, pp. 1–4. <https://doi.org/10.1109/ICIECS.2009.5362936> 9
- [7] J. M. Moine, A. S. Haedo, and S. E. Gordillo, “Estudio comparativo de metodologías para minería de datos,” in *XIII Workshop de Investigadores en Ciencias de la Computación*, 2011. http://sedici.unlp.edu.ar/bitstream/handle/10915/20034/Documento_completo.pdf%3Fsequence%3D1 9
- [8] J. A. Sidey-Gibbons and C. J. Sidey-Gibbons, “Machine learning in medicine: a practical introduction,” *BMC medical research methodology*, vol. 19, no. 1, pp. 1–18, 2019. <https://doi.org/10.1186/s12874-019-0681-4> 10, 13
- [9] M. Á. Vega, L. M. Q. Mora, M. V. C. Badilla *et al.*, “Inteligencia artificial y aprendizaje automático en medicina,” *Revista Medica Sinergia*, vol. 5, no. 8, pp. e557–e557, 2020. <https://doi.org/10.31434/rms.v5i8.557> 10, 13
- [10] W. Hong, Z. Xiong, N. Zheng, and Y. Weng, “A medical-history-based potential disease prediction algorithm,” *IEEE Access*, vol. 7, pp. 131 094–131 101, 2019. <https://doi.org/10.1109/ACCESS.2019.2940644> 10
- [11] M. Boehm, D. Burdick, A. Evfimievski, B. Reinwald, P. Sen, S. Tatikonda, and Y. Tian, “Compiling machine learning algorithms with systemml,” in *Proceedings of the 4th annual Symposium on Cloud Computing*, 2013, pp. 1–1. <https://doi.org/10.1145/2523616.2525965> 11
- [12] M. M. Kaklamanis and M. E. Filippakis, “A comparative survey of machine learning classification algorithms for breast cancer detection,” in *Proceedings of the 23rd Pan-Hellenic Conference on Informatics*, 2019, pp. 97–103. <https://doi.org/10.1145/3368640.3368642> 11
- [13] A. N. Reiz, M. A. de la Hoz, and M. S. García, “Big data analysis y machine learning en medicina intensiva,” *Medicina Intensiva*, vol. 43, no. 7, pp. 416–426, 2019. <https://doi.org/10.1016/j.medin.2018.10.007> 11
- [14] S. Chadha, “Clinical oracle: Machine learning in medicine,” *Berkeley Scientific Journal*, vol. 23, no. 2, 2019. <https://doi.org/10.5070/BS3232045344> 12