

Análisis de intervenciones farmacéuticas utilizando ciencia de datos

Por **Damian Gonzalez Valdez** y **Florencia Primerano**



■ Damian Gonzalez Valdez es Farmacéutico (UBA)
Diplomatura en Ciencia de Datos Aplicada al Sector Salud (ISALUD-UNAB)
Farmacéutico coordinador Hospital Aleman



■ Florencia Primerano es Farmacéutica (UBA)
Diplomatura en Ciencia de Datos Aplicada al Sector Salud (ISALUD-UNAB)
Farmacéutica coordinadora Hospital Aleman

Introducción

La transformación digital en el ámbito de la salud ha traído consigo avances significativos en la recopilación y gestión de datos. Las instituciones tienen acceso a volúmenes masivos de datos sobre pacientes, tratamientos y resultados clínicos, sin embargo, se observa que no siempre se utilizan de manera efectiva para mejorar los procesos de atención ni los resultados para los pacientes. La farmacia hospitalaria no queda excluida de este problema.

Los problemas relacionados con los medicamentos (PRM) son una causa importante de morbilidad y mortalidad, que pueden provocar admisiones hospitalarias y prolongar la estadía, aumentando así los costos en el sistema de salud. La farmacia clí-

nica es una disciplina clave en la optimización de la terapia farmacológica de los pacientes internados, abordando PRM y previniendo eventos adversos. La revisión de las prescripciones de medicamentos por parte de un farmacéutico clínico puede ayudar a detectar, resolver y, en el mejor de los casos, prevenir los PRM¹.

El proceso mediante el cual un farmacéutico identifica PRM y hace una recomendación en un intento de prevenirlos o resolverlos se llama intervención farmacéutica (IF)². Se consideran críticas aquellas intervenciones de mayor impacto clínico en el paciente y en las cuales los fármacos no son dispensados con el objetivo de prevenir eventos adversos, mejorar la seguridad del paciente y opti-

mizar los recursos hospitalarios. Por otro lado, las intervenciones no críticas son recomendaciones y/o sugerencias para mejorar la farmacoterapia del paciente, que se consideran de menor impacto clínico, y por lo tanto, en esos casos los fármacos sí son despachados.

La incorporación de tecnologías avanzadas y el análisis de datos en la práctica farmacéutica representa una oportunidad de cambio en la forma en que se toman decisiones y diseñan estrategias de mejora. Este enfoque no solo transforma el trabajo diario de los profesionales de la farmacia, sino que también impacta directamente en la efectividad y seguridad de los tratamientos para los pacientes.

Problema que motiva la investigación

Este trabajo se enmarca en la Diplomatura en Ciencia de Datos Aplicada al Sector Salud y la motivación surge de la necesidad de aprovechar la gran cantidad de datos generados en el ámbito de la farmacia hospitalaria y convertirlos en información útil para la toma de decisiones clínicas. La falta de herramientas eficientes para analizar y visualizar estos datos limita la capacidad de los farmacéuticos para detectar patrones y tendencias en las IF. Por ello, este trabajo busca aportar soluciones basadas en ciencia de datos que permitan mejorar la gestión farmacéutica y contribuir a una práctica más segura y eficiente³.

Objetivos

Objetivo General: Realizar un análisis exploratorio de las IF que permita identificar las características más comunes de los medicamentos y situaciones que las desencadenan.

Objetivos Específicos:

1. Realizar un análisis exploratorio de datos en el registro de IF críticas y no críticas.

2. Desarrollar visualizaciones que presenten los hallazgos de manera clara y comprensible, para facilitar la toma de decisiones.

3. Evaluar la implementación de modelos de clasificación que evalúen qué características de los medicamentos y pacientes aumentan la probabilidad de que un medicamento requiera IF crítica.

Metodología

El presente estudio se llevó a cabo mediante una metodología de tipo descriptiva y temporalidad retrospectiva de corte transversal. Como fuentes de datos para el análisis se utilizaron tres datasets:

- **IF del servicio:** dataset con las intervenciones realizadas por farmacéuticos entre el 1 de julio de 2020 hasta el 31 de septiembre de 2024.
- **Medicamentos:** dataset con características y atributos de los fármacos que se intervienen.
- **Género:** dataset con el género de los pacientes en los que se hicieron intervenciones.

El preprocesamiento de los datasets, el análisis de datos y el modelado estadístico fue llevado a cabo en Google Colab utilizando Python como lenguaje de programación.

Para el análisis exploratorio de datos se calcularon: Medidas de tendencia central de la edad de los pacientes, cantidad de pacientes con IF, cantidad de IF distribuidas por rango de edad y género, porcentaje de IF de pacientes graves y no graves, cantidad de IF críticas y no críticas por trimestre, cantidad de IF separadas por grupo terapéutico, cantidad de IF separadas por medicamento, cantidad de IF críticas y no críticas separadas por motivo de intervención. Para la visualización de los resultados se creó un tablero interactivo utilizando Looker Studio.

En el modelado estadístico se emplearon modelos de clasificación con el objetivo de obtener un algoritmo que permita clasificar a una IF como crítica.

ticas o no críticas. Para su entrenamiento se probaron dos algoritmos de clasificación: *KNN (k-nearest neighbors)* y árboles de decisión. En el entrenamiento por árboles de decisión se estimó y se seleccionaron las características más importantes con la finalidad de reducir el número de variables a utilizar.

Para la evaluación se compararon las predicciones de los modelos sobre el conjunto de datos de prueba con los valores reales y se calcularon como métricas de rendimiento: Exactitud (Accuracy), Precisión, Sensibilidad (Recall) y F1-Score.

Resultados

Análisis exploratorio de datos

En el análisis exploratorio de los últimos 4 años se encontraron 27579 IF, en su mayoría críticas (16550). Estuvieron involucrados 8074 pacientes, de los cuales la mayor proporción resultaron no graves de más de 75 años de edad. La distribución de las intervenciones críticas a lo largo del tiempo fue uniforme, mientras que las no críticas fueron aumentando (Figura 1).

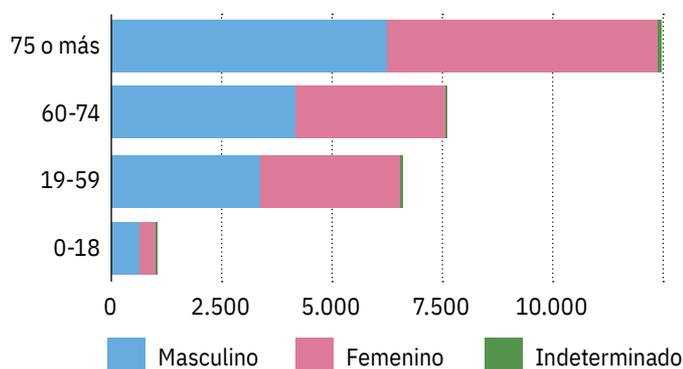
Los grupos farmacológicos más involucrados fueron principalmente los antibacterianos, antivirales

Figura 1. Tablero del análisis exploratorio

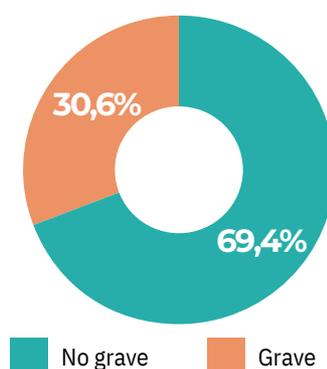
Intervenciones por edad, condición del paciente e intervenciones por trimestre



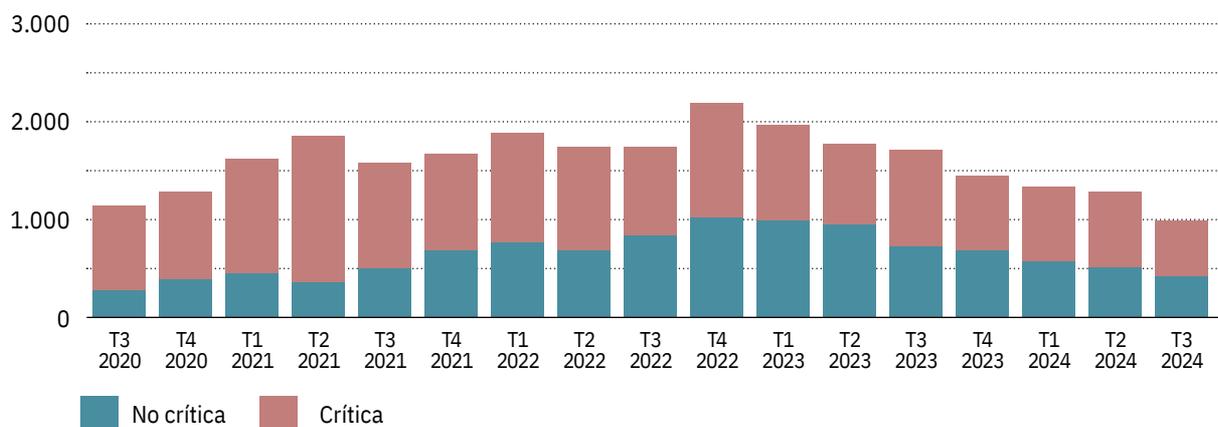
Intervenciones por edad y genero



Condicion del paciente



Cantidad de intervenciones por trimestre



y antifúngicos, antitrombóticos y antiácidos. Estos grupos estuvieron representados principalmente por vancomicina, enoxaparina, omeprazol y cefazolina. Los motivos de intervención más frecuentes fueron

dosis superior o inferior a la recomendada y prescripción incompleta (Figura 2). No se observaron diferencias importantes en las características generales de las intervenciones críticas y las no críticas (Figura 3).

Modelos de clasificación: Los resultados de las métricas utilizadas

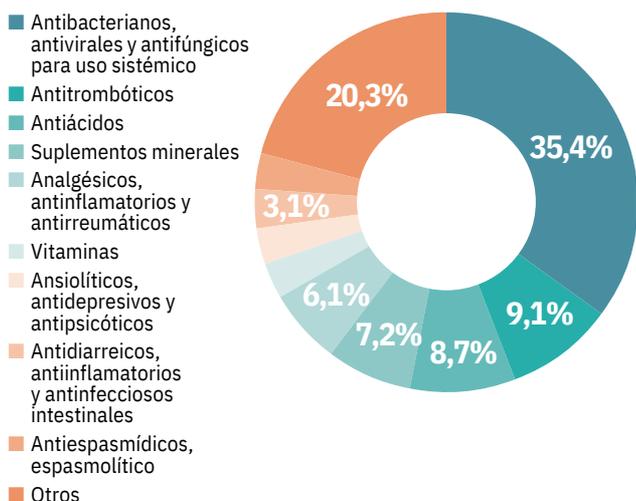
Tipo IF	KNN		Árbol de decisión*	
	No crítica	Crítica	No crítica	Crítica
Precisión	0.66	0.71	0.65	0.70
Sensibilidad	0.49	0.83	0.43	0.85
F1-Score	0.56	0.76	0.52	0.76
Support	2213	3303	2716	4179
Exactitud	0.69		0.68	

* Las características detectadas como las más importantes y utilizadas para entrenar el modelo fueron: grupo terapéutico, motivo de intervención, ajuste renal y vía de administración

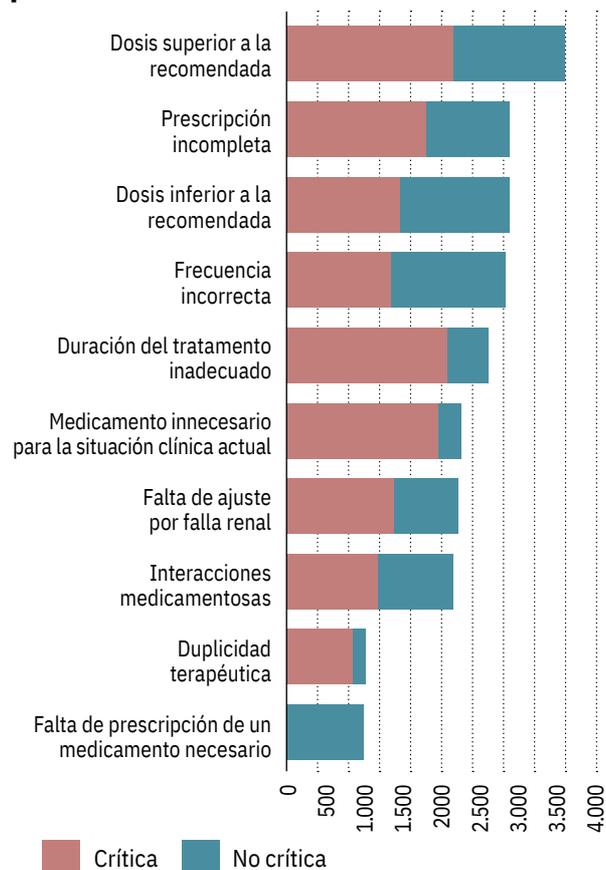
Figura 2. Tablero del análisis exploratorio

Intervenciones por grupo farmacológico, medicamentos más intervenidos e intervenciones por motivo

Porcentaje de intervenciones por grupo



Porcentaje de intervenciones por motivo



Medicamentos más intervenidos



Figura 3. Tablero del análisis exploratorio

Cantidad de intervenciones, distribución por edad, genero y gravedad del paciente, intervenciones por grupo farmacológico y medicamentos mas intervenidos

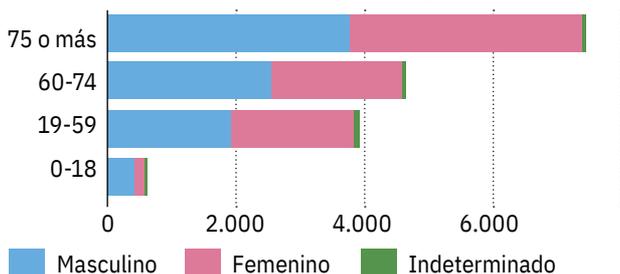
Intervenciones críticas

N° pacientes **6.332**

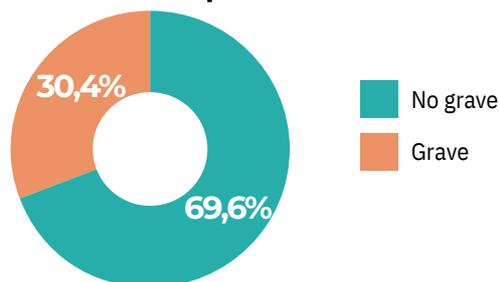
Total de IF **16.550**

IF críticas **16.550**

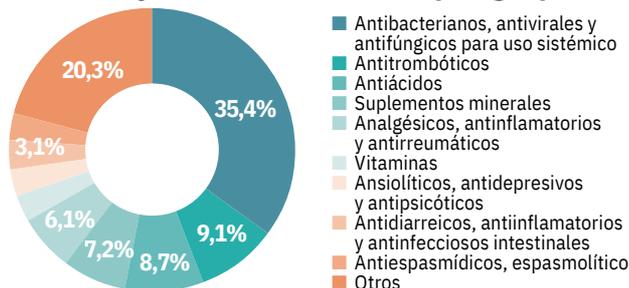
Intervenciones por edad y genero



Condicion del paciente



Porcentaje de intervenciones por grupo



Medicamentos más intervenidos



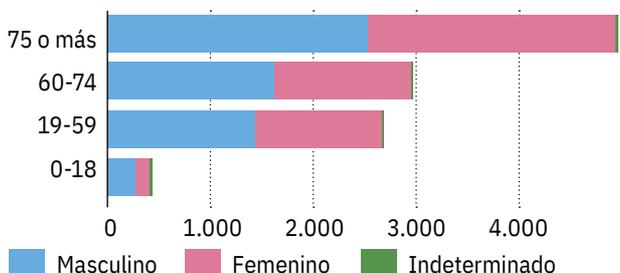
Intervenciones no críticas

N° pacientes **4.689**

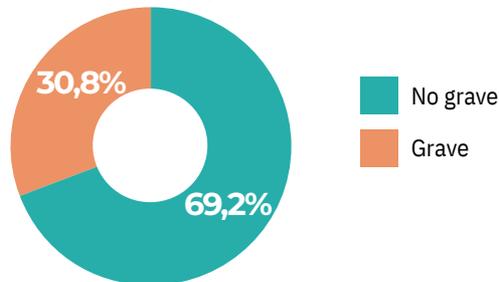
Total de IF **11.029**

IF no críticas **11.029**

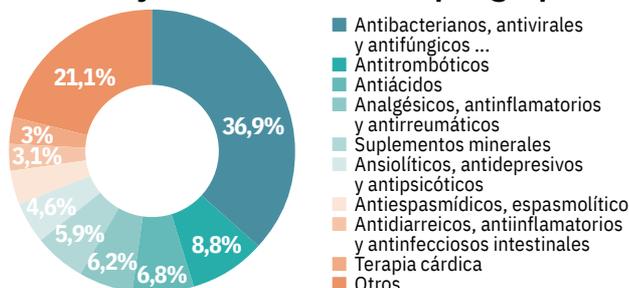
Intervenciones por edad y genero



Condicion del paciente



Porcentaje de intervenciones por grupo



Medicamentos más intervenidos



Discusión

Los resultados obtenidos del análisis exploratorio de los datos evidencian la necesidad de implementar programas de capacitación continua para el personal médico. Este tipo de formaciones resultan ser una estrategia clave para mejorar la calidad del servicio de salud y la seguridad del paciente.

Uno de los aspectos más relevantes que debe ser abordado a través de estos programas es la dosificación correcta de los medicamentos, especialmente aquellos que pertenecen a grupos terapéuticos que presentan mayor riesgo de error en su administración, como los antitrombóticos y antimicrobianos. Una dosificación inadecuada puede dar lugar a efectos adversos graves, lo que pone en riesgo la salud del paciente y genera una carga adicional tanto para los profesionales como para el sistema de salud en general. Las capacitaciones también deben ser diseñadas para ofrecer información actualizada sobre mejores prácticas de prescripción. Es crucial que los médicos comprendan la importancia de emitir una prescripción completa y clara, que incluya no solo la dosis exacta, sino también la frecuencia, la vía de administración y la duración del tratamiento, lo que contribuirá significativamente a evitar errores que podrían surgir durante la administración.

Por otro lado, estos resultados también visibilizan el trabajo de la farmacia clínica y destaca la necesidad de un equipo de farmacéuticos capaces de detectar y prevenir PRM, como parte del equipo interdisciplinario en la atención a los pacientes.

Con respecto a los modelos obtenidos en el análisis mostraron una precisión y exactitud comparables, lo que sugiere que ambos enfoques tienen un rendimiento similar en términos de desempeño general. Sin embargo, al desglosar el desempeño por tipo de intervención, se observa que **KNN** sobresale en la predicción de **intervenciones no críticas**, alcanzando una mayor **sensibilidad** y **F1-SCORE** en esta categoría. Esto significa que KNN fue más

efectivo para identificar correctamente las intervenciones no críticas, minimizando los falsos negativos en este contexto. Por otro lado, el árbol de decisión mostró una ventaja en términos de **sensibilidad** para las intervenciones críticas, lo que implica que este modelo fue más eficaz al identificar intervenciones críticas correctamente, con menos falsos negativos en este caso.

Una observación clave es que el árbol de decisión logró resultados comparables en el reporte de clasificación, analizando solamente **4 variables** (frente a las 13 variables utilizadas en KNN). Esto no solo hace que el modelo sea **menos costoso computacionalmente**, sino que también aporta una ventaja significativa en la práctica, ya que la reducción del número de variables utilizadas para la predicción simplifica el proceso de análisis y la interpretación de los resultados.

Ambos modelos presentaron un **desbalance de clases**, donde las intervenciones críticas dominaron las predicciones. Este sesgo hacia las intervenciones críticas podría afectar la capacidad del modelo para generalizar correctamente sobre intervenciones no críticas, lo que es una consideración importante al evaluar la efectividad de los modelos. Para mitigar este desbalance, podrían explorarse técnicas adicionales como el **ajuste de los umbrales de decisión** con el fin de mejorar el rendimiento del modelo en predicciones de intervenciones no críticas.

Si bien los dos modelos son competitivos en términos de precisión y exactitud, la elección entre ellos podría depender de la **prioridad práctica** que se le dé a la **simplicidad computacional** frente a la **sensibilidad** en la predicción de intervenciones no críticas. Sin embargo, la presencia de un desbalance de clases sugiere que una optimización adicional podría ser necesaria para mejorar la robustez y equidad del modelo frente a todas las clases involucradas.

Una de las limitaciones del análisis fue la falta de datos sobre características de los pacientes

que podrían ser clave para entender las causas subyacentes de las intervenciones farmacéuticas y mejorar su clasificación en críticas o no críticas. También, la falta de datos sobre la frecuencia de prescripción de los medicamentos dificulta evaluar si ciertas monodrogas o grupos terapéuticos tienen más intervenciones debido a su alta tasa de uso. Este sesgo puede generar resultados que reflejen más la distribución de prescripciones que problemas específicos relacionados con los medicamentos. Por otro lado, la gran cantidad de drogas y grupos terapéuticos analizados hace que los datos estén muy dispersos complicando la identificación de patrones claros y reduciendo la potencia estadística de los análisis en categorías específicas de medicamentos. Otro factor clave fue la ausencia de información sobre prescripciones que no generaron intervenciones, esto limitó la capacidad de los modelos para evaluar y validar un modelo predictivo.

Se considera fundamental gestionar la obtención de otros datos adicionales, lo que permitiría incorporar nuevas variables al análisis, y también abordar el desbalance de clases en los modelos de clasificación, con el objetivo de desarrollar un modelo que logre predecir de manera precisa ambos

tipos de intervenciones. Esto facilitaría una comprensión más amplia del funcionamiento del servicio de farmacia y favorecería el desarrollo de modelos orientados a optimizar el flujo de trabajo.

Conclusiones

El análisis de las intervenciones evidencia la necesidad de programas de capacitación continua para el personal, enfocados en mejorar la seguridad del paciente, especialmente en la dosificación de medicamentos como antitrombóticos y antimicrobianos. Además, destaca el papel clave de la farmacia clínica en la detección y prevención de PRM, promoviendo la integración de farmacéuticos en equipos interdisciplinarios. En cuanto a los modelos predictivos, tanto el Árbol de Decisión como KNN mostraron un desempeño similar en la predicción de intervenciones críticas, aunque KNN fue más preciso en intervenciones no críticas y el Árbol de Decisión, con menos variables, resultó más eficiente computacionalmente. Para mejorar la robustez y equidad del modelo, se sugiere la obtención de datos adicionales y el ajuste de umbrales de decisión. Un modelo predictivo optimizado facilitaría la identificación de patrones, optimizando el flujo de trabajo en la farmacia hospitalaria. ■

Referencias bibliográficas

- Farmacia Clínica y Atención Farmacéutica. Bonal J., Alerany C., Bassons T., Gascón P. - Farmacia Hospitalaria 4ta edición digital. Madrid. 2002. Disponible en: <https://www.sefh.es/bibliotecavirtual/fhtomo1/cap21.pdf>
- Guidelines for pharmacists performing clinical interventions. Pharmaceutical Society of Australia Ltd. 2018
- Johns E, Alkanj A, Beck M, Dal Mas L, Gourieux B, Sauleau EA, et al. Using machine learning or deep learning models in a hospital setting to detect inappropriate prescriptions: a systematic review. *Eur J Hosp Pharm Sci Pract.* 2023 Nov 24;ejhpharm-2023-003857.