**CASO DE ESTUDIO DE SALUD**

Piedrahita Allison; Ramírez Anyi; Vergara María

*Departamento de Ingeniería Industrial, Universidad de Antioquia.*

*Medellín, Colombia.*

***RESUMEN*** ~ El Pago Global Prospectivo (PGP) representa un mecanismo de pago anticipado a la prestación de un servicio de salud de una institución prestadora de salud (IPS), dicha modalidad de pago es asignada una población dependiendo de la frecuencia de atenciones o cantidad de población. Con el fin de prestar los servicios de manera efectiva, es necesario asignar los recursos suficientes a los pacientes hospitalizados bajo esta modalidad de pago. Para ello, se propone una solución analítica por medio de modelos de machine learning que permitan predecir la estancia hospitalaria de un grupo de pacientes de PGP para realizar una correcta distribución de recursos y establecer planes de acción que permitan reducir la estancia hospitalaria y a su vez, llevar a cabo las operaciones de la IPS de manera efectiva.

1. **INTRODUCCIÓN**

La prestación del servicio de salud se encuentra a cargo de las instituciones prestadoras de los servicios de salud (IPS), las cuales se encuentran sometidas a dos tipos de mecanismos de pago para la generación de ingresos, entre ellos, el Pago Global Prospectivo (PGP), el cual consiste en realizar el pago antes de la prestación del servicio para una cohorte poblacional determinada de acuerdo con la frecuencia de atenciones o a la cantidad de población. Por tal motivo, dicha modalidad de pago requiere una optimización de recursos para atender a la población asignada de manera efectiva.

En el Hospital Alma Máter de Antioquia, se realiza una atención previa que permite definir la modalidad de tratamiento (ambulatorio o domiciliario) para adecuar la atención a las necesidades, capacidades, habilidades y potencialidades del paciente para optimizar las condiciones de salud. Sin embargo, para mejorar la atención de los pacientes pertenecientes la modalidad de contrato PGP, es de gran utilidad anticiparse a los recursos que se deben destinar en la atención de las personas, por ello, es necesario conocer la estancia hospitalaria (uso de recursos en hospitalización) para hacer una mejor distribución de estos.

Teniendo en cuenta lo anterior, se analizan tres bases de datos que contienen información de los pacientes atendidos en el Hospital Alma Máter de Antioquia, con el fin de establecer una solución analítica que permita establecer la estancia hospitalaria de un grupo relevante de pacientes pertenecientes a la modalidad PGP con la finalidad de identificar y establecer una mejor distribución de recursos y planes de acción y mejora que permitan llevar a cabo las operaciones de la IPS de una manera efectiva.

1. **METODOLOGÍA**

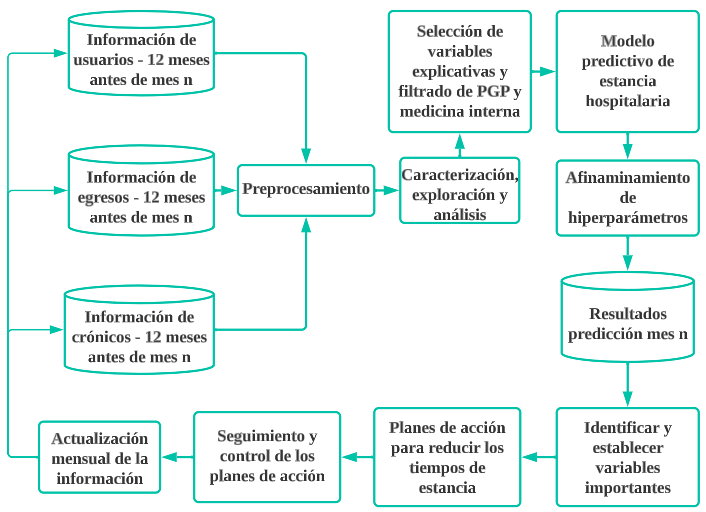
En las siguientes secciones se describe cada uno de los pasos desarrollados.

* 1. **DISEÑO DE LA SOLUCIÓN**

Con las bases de datos proporcionadas por la IPS universitaria, se plantean los siguientes objetivos para desarrollar el caso de estudio:

* Segmentar los pacientes de las bases de datos para seleccionar aquellos que representen un porcentaje significativo de personas con modalidad de contrato PGP, con el fin de enfocar la solución analítica.
* Predecir la estancia hospitalaria del grupo segmentado por medio de un modelo analítico basado en variables significativas de 12 meses históricos.
* Identificar las variables más significativas para establecer planes de acción que permitan mejorar la asignación de recursos y reducir los tiempos de hospitalización.

Con base a lo anterior, se plantea la solución analítica de la Figura 1, en la cual se evidencian los objetivos anteriores y las condiciones de tiempo de la información.



**Figura 1.** Diseño de la solución analítica.

* 1. **LIMPIEZA Y TRANSFORMACIÓN**

Para el desarrollo de la solución, se cuenta con tres bases de datos:

* *“RETO\_df\_egresos”:* contiene la información administrativa y de gestión de los pacientes atendidos en la IPS de los años 2017, 2018 y 2021. Esta base de datos tiene 6.376 registros y 66 variables.
* *“RETO\_df\_cronicos”:* contiene información de variables médicas que describen el estado de salud de los pacientes atendidos en la IPS de los años 2017, 2018, 2021 y 2022. Esta base de datos tiene 38.717 registros y 290 variables.
* *“RETO\_df\_usuarios”:* contiene la información básica de los pacientes atendidos en la IPS de los años 2017, 2018, 2021 y 2022. Esta base de datos tiene 183.911 registros y 16 variables.

Para realizar la limpieza de la información, se realizan los siguientes pasos para cada una de las bases de datos:

* Separación de variables/columnas según el separador de datos (“,”)
* Eliminación de variables con más de 50% de registros nulos y selección de variables significativas que pueden aportar información según el análisis de las autoras.
* Eliminación de registros nulos de la base de datos reducida.
* Tratamiento de registros escritos diferente pero asociados a la misma categoría.
* Tratamiento del tipo de variable con el fin de asignar la clasificación correcta.

Para la base de datos de egresos, se realiza un procedimiento adicional, basado en el cálculo de la estancia hospitalaria para cada uno de los registros haciendo uso de la fecha de ingreso y salida del hospital. Con este procedimiento, se adiciona una variable más a la base de datos mencionada. Asimismo, aquellas variables con información de duración de tiempo se trataron con el fin de unificar la medida a horas.

Con los procedimientos anteriores, se obtienen tres bases reducidas del siguiente tamaño: usuarios con 6 variables y 125.225 registros, egresos con 21 variables y 4.103 registros y crónicos con 22 variables y 37.489 registros.

Después de tener cada una de las bases de datos limpias y uniformes, se realiza la unión de la información para obtener una base de datos general que condensa los datos de los pacientes. La unión se realiza por medio de la función “merge” bajo el método “inner” el cual lleva a cabo una intersección de los datos de las bases para extraer la información contenida en los tres conjuntos. La unión se hizo por medio de la variable “nrodoc”, de manera adicional, se agregaron las claves “mes” y “año” con el fin de obtener al menos un registro por paciente para cada mes de los años informados. Teniendo en cuenta lo anterior, se obtiene una base de datos general con 43 variables y 779 registros.

* 1. **ANÁLISIS EXPLORATORIO**

Después de realizar un análisis para comprender las características de los pacientes y de las variables de la base de datos general, se obtiene que:

Posteriormente, al realizar el análisis exploratorio, se identifican variables explicativas correlacionadas, las cuales son eliminadas teniendo en cuenta el nivel de correlación con la variable objetivo.

Por otra parte, a la base de datos general, se filtra la modalidad de contrato para seleccionar solo aquellos registros pertenecientes a PGP y se identifica que el 68,47% de los registros ha consultado la especialidad de medicina interna. Por tal motivo, se selecciona este grupo de pacientes para desarrollar la solución analítica, debido a que constituyen una mayoría que permite empezar a distribuir los recursos de una manera más efectiva.

* 1. **SELECCIÓN DE ALGORITMOS Y VARIABLES**
  2. **SELECCIÓN DEL MODELO**
  3. **AFINAMIENTO DE HIPERPARÁMETROS**
  4. **ANÁLISIS DEL MODELO**
  5. **DESPLIEGUE DEL MODELO**

1. **RESULTADOS**
2. **CONCLUSIONES**
3. **REFERENCIAS**