



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA DE POSTGRADO Y EDUCACIÓN CONTINUA
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**DISEÑO Y EVALUACIÓN DE MODELO DE PREDICCIÓN DE
PREDISPOSICIÓN QUIRÚRGICA PARA GESTIÓN DE PACIENTES EN
VIAJES DE SALUD**

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIA DE DATOS

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

TOMÁS ANDRÉS AGUIRRE MARGERY

PROFESOR GUÍA:
Sebastián Ríos Pérez

MIEMBROS DE LA COMISIÓN:
Juan Neme Gaviola
Diego Martinez Cea
Marcel Goic Figueroa

SANTIAGO DE CHILE
2024

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIA DE DATOS Y MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL
POR: TOMÁS ANDRÉS AGUIRRE MARGERY
FECHA: 2024
PROF. GUÍA: Sebastián Ríos Pérez

DISEÑO Y EVALUACIÓN DE MODELO DE PREDICCIÓN DE PREDISPOSICIÓN QUIRÚRGICA PARA GESTIÓN DE PACIENTES EN VIAJES DE SALUD

El presente trabajo de tesis se realiza en colaboración con un prestador de salud privado, que tiene presencia a lo largo de todo Chile, enfocado en entregar un servicio de calidad a sus pacientes. En el ámbito quirúrgico, se identifica la problemática de que los pacientes enfrentan dificultades para llegar a sus cirugías, debido a que no conocen las distintas acciones que deben realizar, para iniciar y navegar su viaje de salud.

Para abordar esta problemática, se plantea que un correcto acompañamiento orientado a pacientes quirúrgicos podría ayudar a guiarlos a la concreción de sus cirugías. En este trabajo, se propone un paso inicial a esta propuesta, con el desarrollo de una herramienta que permite identificar tempranamente a los pacientes quirúrgicos. Esto es, antes de que los pacientes soliciten un presupuesto a través de la web. Además, la herramienta indica qué cirugía es la más probable que se someterá el paciente, junto con un listado de servicios que, en general, asisten los pacientes que se someten a dicha cirugía. De esta manera, el prestador tiene la oportunidad de apoyar y guiar proactivamente a los pacientes seleccionados por la herramienta propuesta, ayudándoles a navegar un sistema complejo y desconectado, y brindándoles la orientación necesaria para avanzar hacia su cirugía.

La herramienta recién mencionada, se desarrolla utilizando un enfoque de clasificación multiclase, el cual utiliza datos de agenda de los pacientes que ya se realizaron sus cirugías, como variables independientes, y las cirugías particulares como la variable a predecir. Este modelo, a lo largo de las 10 cirugías más frecuentes realizadas por el prestador, logra un *precision macro* de 0.81, un *recall macro* de 0.79, y un *f1 score macro* de 0.8. Además, mediante los *Shap values*, se establece un listado de servicios a los que, en general, asisten los pacientes que se someten a dicha cirugía, para la creación de las respectivas rutas quirúrgicas.

Adicionalmente, la evaluación de la herramienta se realiza implementando un piloto de acompañamiento, con el objetivo de validar su desempeño. Esto se realiza tomando el conjunto de pacientes entregado por el modelo, y contactándolos vía llamada telefónica, buscando validar la predicción y posteriormente ayudando al paciente a reservar su siguiente servicio a asistir. A partir de estos llamados, 313 resultaron ser pacientes quirúrgicos, 776 no quirúrgicos, y el contacto falla en 805 casos.

Con todo lo anterior, se logra dar una primera aproximación a la creación de la herramienta, pues si bien logra un buen desempeño sobre los datos de prueba, en la realidad, donde existen más cirugías que las incluidas en el entrenamiento del modelo, este ve deteriorado su desempeño, por lo que finalmente se plantean distintas metodologías que podrían mejorar la situación anterior en trabajos futuros.



A mi familia y amigos.

Agradecimientos

En primer lugar, quiero agradecer a mi mamá Katina, mi papá Cristian, mis hermanas Sofía y Catalina, mis abuelos Katina y Patricio, y mi tío Cristian. Por ser parte de este proceso, y por todo el apoyo que me han entregado de forma incondicional durante todo este tiempo, en los altos, los bajos, las alegrías, las tristezas, los malos humores y las risas. Todos ustedes han sido parte de esto, y hoy puedo decir que si estoy aquí, es gracias al aporte de cada uno de ustedes.

Agradezco a todos los “Loki-llos”, Chino, Martín, Nacho, Pollo, Santiago y Zeta (por orden en que me aparecen en el grupo de Whatsapp). Mis amigos incondicionales que durante todo este tiempo fueron fuente de energía e inspiración.

Agradezco a mi amiga Javiera N., a quien conocí una fría mañana de semestre beta en laboratorio de Sistemas Newtonianos, donde realizamos nuestro primer trabajo juntos, y que a partir de entonces nos convertimos en una pareja inseparable para todas las tareas y actividades. Además de eso, ese día gané una amiga para la vida. Definitivamente este logro no habría sido posible sin esta amistad.

Agradezco a mi amiga Javiera S., quien desde nuestra primera interacción, se ha encargado de levantarme el ánimo y ser una amiga incondicional a pesar de que nos veamos una vez al año.

Agradezco a todos los miembros de los chokitos digitales, quienes me recibieron con los brazos abiertos, haciéndome sentir como parte de una familia con todo el cariño entregado. En particular, quiero agradecer a Allan, por acompañarme, guiarme y ayudarme durante todo el proyecto. También quiero agradecer a los demanditos, quienes siempre fueron fuente de risas y conversación en la oficina. Más en específico, quiero agradecer a Joachim, quien además siempre estuvo dispuesto a ayudarme con sus tablas y a entregarme sus queries de sql.

Agradezco a mi profesor guía Sebastián, quien apenas le pedí ayuda para encontrar un trabajo de tesis entretenido y desafiante, no dudo en darme su mano, estando siempre disponible a las dudas que pudieran surgir.

Finalmente, quiero agradecer a Antonia, mi pareja, quien a lo largo de estos años ha sido fuente de sabiduría, apoyo emocional, y mucho cariño. Hoy estoy muy agradecido de habernos cruzado en esas clases de Módulo Interdisciplinario.

Gracias a todos ustedes por ser parte de mi vida.

Tabla de Contenido

1. Planteamiento del Problema	1
1.1. Características del prestador	1
1.2. Características del sistema de salud de Chile	3
1.3. Problema identificado	4
1.4. Objetivos	5
1.4.1. Objetivo general	5
1.4.2. Objetivos específicos	5
2. Marco Teórico, Conceptual y Metodología	6
2.1. Aprendizaje Automático	6
2.2. Algoritmos de Clasificación	7
2.3. Optuna	15
2.4. Validación Cruzada	15
2.5. Reducción de Dimensionalidad	16
2.6. Métricas de evaluación	17
2.7. <i>Shap Values</i>	20
2.8. Estado del arte	20
2.9. Metodología	21
2.10. Antecedentes	22
3. Modelamiento y Evaluación de Resultados	24
3.1. Comprensión de los datos	24
3.2. Análisis y selección de datos	26
3.2.1. Análisis conjunto de datos de agenda	26
3.2.2. Análisis conjunto de datos de concreción de cirugías	29
3.2.3. Análisis conjunto de datos de entrenamiento	30
3.3. Modelamiento y evaluación	35
3.4. Creación de las rutas quirúrgicas por cirugía	46
3.4.1. Definición de Ruta Quirúrgica para Cirugía Bariátrica	48
3.4.2. Definición de Ruta Quirúrgica para Cirugía Urológica	49
3.4.3. Definición de Ruta Quirúrgica para Colectomía	50
3.4.4. Definición de Ruta Quirúrgica para Hernia Abdominal	52
3.4.5. Definición de Ruta Quirúrgica para Meniscectomía	53
3.4.6. Definición de Ruta Quirúrgica para Parto	54
3.4.7. Definición de Ruta Quirúrgica para Rinoplastía	55
3.4.8. Definición de Ruta Quirúrgica para Ruptura Manguito Rotador	56
4. Prueba de Campo	59

4.1. Cirugía elegida e implementación	59
4.2. Obtención de los RUTs a partir del modelo	61
4.3. Situación actual de acompañamiento quirúrgico	63
4.4. Situación TO-BE de acompañamiento quirúrgico	64
4.4.1. Formulación piloto para validación de hipótesis	67
4.5. Evaluación piloto	70
5. Conclusiones	73
Bibliografía	76
Anexos	79
Anexo A. Planteamiento del Problema	79
A.1. Problema identificado	79
Anexo B. Modelamiento y Evaluación de Resultados	80
B.1. Análisis y selección de datos	80
B.1.1. Análisis conjunto de datos de entrenamiento	80
B.2. Modelamiento y evaluación	83
B.3. Creación de las rutas quirúrgicas por cirugía	86
Anexo C. Prueba de Campo	87
C.1. Obtención de los RUTs a partir del modelo	87
C.2. Evaluación piloto	89

Índice de Tablas

3.1.	Variables conjunto de datos de agenda	24
3.2.	Variables conjunto de datos de concreción de cirugías	25
3.3.	Variables categóricas del conjunto de datos horizontalizado.	25
3.4.	Variables numéricas del conjunto de datos horizontalizado.	25
3.5.	Ejemplo de cómo se verían los datos de cinco pacientes en el conjunto de datos de entrenamiento con datos falsos.	25
3.6.	Agrupación de las cirugías	30
3.7.	División de los datos para el entrenamiento.	36
3.8.	Rangos de hiperparámetros introducidos a optuna.	37
3.9.	Descripción de la configuración de los entrenamientos.	38
3.10.	Resultados entrenamientos [1-49].	39
3.11.	Resultados entrenamientos [50-98].	40
3.12.	Resultados entrenamientos [99-140].	41
3.13.	Resultados entrenamientos [141-154].	42
3.14.	Reporte de clasificación por cirugía modelo datos crudos <i>LightGBM</i>	43
3.15.	Ejemplo de cómo se verían los datos de cinco pacientes en el conjunto de datos de entrenamiento con datos falsos.	44
3.16.	Nueva agrupación de las cirugías	44
3.17.	Nueva división de los datos	45
3.18.	Resultados entrenamiento final	45
3.19.	Reporte de clasificación por cirugía nuevo modelo datos crudos <i>LightGBM</i> y validación cruzada estratificada.	46
3.20.	Servicios de Ruta Quirúrgica por cirugía.	57
4.1.	Resultados simulación modelo por días ($P(x) > 0.5$).	61
4.2.	Búsquedas de servicios semanas siguientes a obtención de datos.	62
4.3.	Información del paciente entregada a ejecutivo/a de <i>Call Center</i> [1/2]	68
4.4.	Información del paciente entregada a ejecutivo/a de <i>Call Center</i> [2/2]	68
4.5.	Datos evaluados previos al piloto.	70
4.6.	Resultados piloto	71
4.7.	Resultados piloto agrupados	71
C.1.	Resultados simulación modelo por días ($P(x) > 0.6$).	87
C.2.	Resultados simulación modelo por días ($P(x) > 0.7$).	87
C.3.	Resultados simulación modelo por días ($P(x) > 0.8$).	88
C.4.	Resultados simulación modelo por días ($P(x) > 0.9$).	88
C.5.	Agrupaciones de las conclusiones del piloto	90

Índice de Ilustraciones

1.1.	Organigrama prestador de salud. Elaboración propia.	2
2.1.	Función sigmoide que toma el valor de un número real y lo convierte al rango (0, 1). Obtenido de [8].	8
2.2.	Diagrama de un árbol de decisión para la clasificación de un animal como mamífero. Obtenido de [9].	8
2.3.	Diagrama explicativo del funcionamiento de <i>Random Forests</i> . Obtenido de [9].	9
2.4.	Diagrama de <i>Adaboost</i> . Obtenido de [10].	10
2.5.	Diagrama comparativo de árboles “ <i>level-wise</i> ” y “ <i>leaf-wise</i> ”. Obtenido de [13].	11
2.6.	Modelo no lineal de una neurona. Obtenido de [15].	13
2.7.	Funciones de activación. Obtenido de [16].	13
2.8.	Arquitectura general de un <i>Multilayer Perceptron</i> . Obtenido de [17].	14
2.9.	Validación Cruzada de conjunto de datos dividido en 10 subconjuntos. Obtenido de [19].	15
2.10.	Matriz de Confusión. Obtenido de [23].	18
2.11.	Proceso de CRISP-DM para minería de datos. Obtenido de [27].	22
2.12.	Flujo Solicitudes Web. Obtenido de [28].	23
3.1.	Histograma servicios. Muestra el nivel de demanda que han tenido los 10 servicios más populares dentro del prestador de salud. Elaboración propia.	27
3.2.	Histograma citas asistidas. Indica a cuántas citas asistió cada paciente en el lapso de tiempo que abarca el conjunto de datos. Elaboración propia.	28
3.3.	Histograma Área. Indica el número de atenciones realizadas por área en el prestador de salud. Elaboración propia.	28
3.4.	Histograma de las cirugías con los nombres más comunes por código. Elaboración Propia.	29
3.5.	Histograma con las cirugías agrupadas según lo presentado en la tabla 3.6. Elaboración propia.	30
3.6.	Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Colectectomía. Elaboración propia.	32
3.7.	Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Cirugía Bariátrica. Elaboración propia.	32
3.8.	Gráficos reducción de dimensionalidad. Elaboración propia.	34
3.9.	Gráfico de resumen Shap para Cirugía Bariátrica. Elaboración propia.	48
3.10.	Gráfico de resumen Shap para Cirugía Urológica. Elaboración propia.	49
3.11.	Gráfico de resumen Shap para Colectectomía. Elaboración propia.	50
3.12.	Gráfico de resumen Shap para Hernia Abdominal. Elaboración propia.	52
3.13.	Gráfico de resumen Shap para Meniscectomía. Elaboración propia.	53
3.14.	Gráfico de resumen Shap para Parto. Elaboración propia.	54
3.15.	Gráfico de resumen Shap para Rinoplastía. Elaboración propia.	55

3.16.	Gráfico de resumen Shap para Ruptura Manguito Rotador. Elaboración propia.	56
4.1.	Gráfico histograma y acumulado de la cantidad de días que se tarda un paciente desde su primera cita hasta llegar a su cirugía de Colectectomía. Elaboración propia.	60
4.2.	BPMN situación AS-IS del acompañamiento a pacientes que ingresan solicitud de presupuesto web. Elaboración propia.	64
4.3.	Diagrama validación de hipótesis. Elaboración propia.	65
4.4.	Cascada de pacientes entregados por el modelo. Elaboración propia.	66
4.5.	BPMN situación TO-BE del acompañamiento a pacientes que ingresan solicitud de presupuesto web e ingresados por el modelo. Elaboración propia.	67
4.6.	Flujo para ejecutivo/a de <i>Call Center</i> en piloto. Elaboración propia.	69
A.1.	Árbol de problemas. Elaboración propia.	79
B.1.	Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Hernia abdominal. Elaboración propia.	80
B.2.	Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Cesárea. Elaboración propia.	81
B.3.	Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Vasectomía. Elaboración propia.	81
B.4.	Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Meniscectomía. Elaboración propia.	82
B.5.	Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Circuncisión. Elaboración propia.	82
B.6.	Matriz de confusión modelo final. Elaboración propia.	85
C.1.	Gráfico cascada de la obtención de datos para el piloto. Elaboración propia. . .	89

Capítulo 1

Planteamiento del Problema

1.1. Características del prestador

El presente trabajo de tesis, se realiza en conjunto con un prestador de salud privado. Este cuenta con más de 15 años de trayectoria y presencia a lo largo de todo Chile, desde la Región de Arica y Parinacota, hasta la Región de Magallanes y la Antártica Chilena, donde a través de 9 clínicas (3 ubicadas en la Región Metropolitana y 6 en otras regiones), 26 centros de atención médica ambulatoria, y 45 centros de atención dental, ofrece una variedad de servicios a precios accesibles para sus pacientes.

Misión

“Contribuir a que las personas vivan más y mejor, a través de una atención integral de salud médica y dental, centrada en el paciente y su familia. Nuestro compromiso es entregar amplio acceso, con altos estándares de calidad, según los principios y valores de la Cámara Chilena de la Construcción.”

Visión

“Ser un referente en salud médica y dental privada del país, reconocida por su amplia accesibilidad y prestigio; ser un buen lugar para trabajar, y resolver las necesidades de salud de gran parte de la población.”

Organigrama

El organigrama del prestador se presenta en la figura 1.1

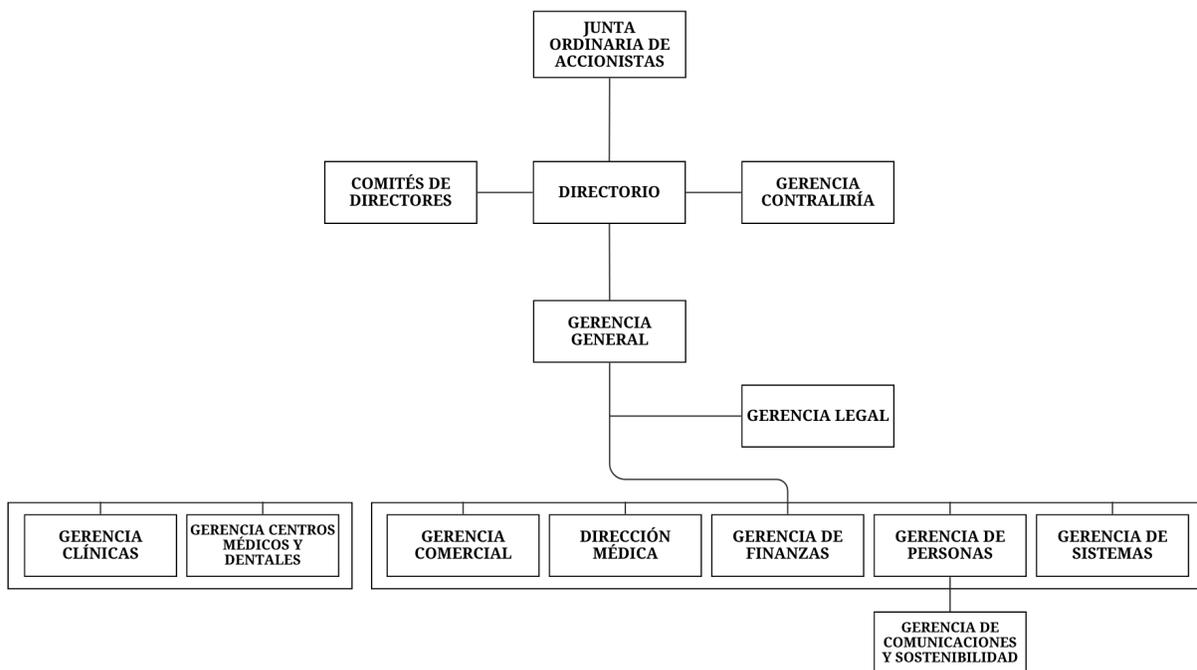


Figura 1.1: Organigrama prestador de salud. Elaboración propia.

En particular, el trabajo se realiza en conjunto con los equipos de la Gerencia Comercial, que se desempeñan en la Casa Matriz de la institución. Específicamente, primero se trabaja con el equipo de analistas de demanda, el cual tiene como función principal, tomar los datos respectivos a la oferta y la demanda de los distintos servicios médicos que se ofrecen, para luego crear indicadores que permitan medir y mejorar los distintos servicios. Junto con esto, se crean proyectos que tienen como objetivo mejorar los indicadores, como lo puede ser la implementación de la venta de sobrecupos para aumentar la demanda, o la priorización de pacientes para la reserva de servicios.

Del mismo modo, se trabaja con el equipo de “*Call Center*”, el cual se encarga de crear, coordinar, y realizar los llamados programados por parte del prestador, al paciente. En este sentido, son los encargados de crear los diálogos con las posibles respuestas a los pacientes, responsabilizándose también de capacitar a los ejecutivos que realizan los llamados, entre otras funciones relacionadas.

1.2. Características del sistema de salud de Chile

A nivel mundial, el funcionamiento de los sistemas de salud en distintos países, es un tema recurrente de conversación, debido a la complejidad de su funcionamiento y financiamiento. Los avances tecnológicos en esta disciplina han logrado que la tasa de mortalidad de las personas tenga un constante decrecimiento desde que existe registro en 1960 [1]. En este sentido, las cirugías juegan un papel fundamental al momento de salvar vidas y aumentar la expectativa de vida de las personas.

En Chile, la salud se basa en un sistema mixto de atención, integrado por un seguro público denominado Fondo Nacional de Salud (FONASA), y por un seguro privado denominado Instituciones de Salud Previsional (ISAPRE). En FONASA, los usuarios deben cotizar el 7% de sus ingresos mensuales para aportar en el financiamiento de este fondo. En él, existe un sistema de clasificación en el que, dependiendo del nivel de vulnerabilidad social del usuario, es cómo se entregan los beneficios del sistema. Este se divide en cuatro grupos. Primero está el tramo A, correspondiente a personas carentes de recursos que pueden atenderse de manera gratuita en Hospitales y Consultorios Públicos. Luego, está el tramo B, donde se encuentra a personas cuyo ingreso imponible mensual sea menor o igual a 276,000 pesos, quienes también pueden atenderse de manera gratuita en Hospitales y Consultorios Públicos. En tercer lugar está el tramo C, donde se encuentra a aquellas personas con un sueldo imponible mensual entre 276,001 y 402,960 pesos, quienes pagan el 10% del arancel en Hospitales Públicos. Finalmente está el tramo D, donde se encuentra a aquellas personas con un sueldo mensual imponible mayor a 402,961 pesos, quienes pagan el 20% del arancel en Hospitales Públicos.

Además, FONASA incluye dos modalidades de atención. Por un lado, está la Modalidad de Atención Institucional (MAI), que corresponde a la atención de los beneficiarios de los distintos tramos en los establecimientos públicos de la red asistencial. Por otro lado está la Modalidad de Libre Elección (MLE), donde pueden atenderse beneficiarios de los grupos B, C y D comprando el bono de atención correspondiente.

Por otro lado, en ISAPRE están aquellas personas que cotizan el 7% o un monto superior de sus ingresos mensuales. Aquí, la cobertura depende del plan elegido en conjunto con el asegurador. En este sentido, primero está la modalidad de prestadores en convenio o preferentes, la cual cubre en mayor magnitud si las prestaciones atendidas son en uno de los centros en convenio con la ISAPRE, y cubre menos si se atiende en un centro que no tiene convenio con la misma. En segundo lugar, está la modalidad de libre elección, donde la ISAPRE bonifica cualquier centro de salud en el que la persona se atiende. Por último, esta la modalidad cerrada, en la cual la ISAPRE bonifica solamente si el beneficiario se atiende en el centro de salud que su plan le señala [2].

A finales del año 2023, el porcentaje de chilenos adscritos a FONASA fue del 81%, lo cual corresponde a más de 16 millones de usuarios, mientras que los usuarios adscritos a alguna ISAPRE fueron alrededor de 2.7 millones.

1.3. Problema identificado

Para que un paciente quirúrgico pueda efectivamente realizarse su cirugía en una de las clínicas habilitadas por el prestador, este debe atravesar un proceso que difiere según la cirugía, e incluso entre pacientes, para una misma intervención quirúrgica. En general, este consiste de una serie de consultas médicas y exámenes, que permiten al médico tratante diagnosticar al paciente, y prepararse para la realización de la cirugía. Además, existe un proceso administrativo, en el cual el paciente, una vez en posesión de la orden médica para su cirugía, debe acercarse al prestador, a través de la web, o presencialmente, para solicitar el presupuesto de su cirugía. Una vez emitido el presupuesto, se debe reservar el pabellón, y coordinar los últimos detalles con el cirujano, para finalmente llegar a la realización de la intervención quirúrgica.

En particular, la solicitud de presupuesto, es un proceso determinante para la realización de una cirugía con el prestador. A través del formulario web disponible en su página, el prestador recibe solicitudes de pacientes que adjuntan, o no, su orden médica. De esta manera, en promedio se reciben alrededor de 14,000 solicitudes de presupuesto quirúrgico al mes. De estas, aproximadamente 2,400 corresponden a una solicitud con orden médica, de las cuales el 75 % terminan siendo correctamente validadas en su posterior revisión.

Los pacientes que ingresan su solicitud sin orden médica, o que en su defecto, adjuntan orden médica inválida, reciben un acompañamiento, de manera que se les pueda guiar hasta la obtención de orden médica de cirugía, y la posterior emisión de presupuesto.

El proceso anterior espera a que el paciente acuda al prestador, y este le entregue el presupuesto de una cirugía que requiere realizarse. Sin embargo, la mayoría de estos pacientes, al contrario de sus expectativas, reciben un llamado para ayudarles a agendar una cita médica que les ayude a obtener su orden médica. Esto es algo que puede llegar a causar frustración en los pacientes, haciendo que su viaje quirúrgico sea insatisfactorio.

Con esto, se define como problemática que, a los pacientes se les dificulta la completitud de sus viajes quirúrgicos, debido a que no saben qué es todo lo que involucra este viaje.

Dentro de las posibles causas para este problema, se encuentra que en ningún momento se le explica al paciente acerca del funcionamiento del proceso administrativo, para la realización de una cirugía con el prestador. Es por esta razón que tantos pacientes ingresan al formulario web esperando conocer el valor de su cirugía, pero se terminan encontrando con una realidad diferente a lo esperado.

Adicionalmente, no existe un acompañamiento que ayude al paciente en el viaje hacia su cirugía, en el sentido de la reserva de citas médicas y exámenes, por lo que el paciente debe gestionar por su propia cuenta estas reservas.

Junto con lo anterior, los pacientes tienen la posibilidad de asistir a cualquier prestador de su preferencia. Esto puede causar que se entrapen en los distintos procesos administrativos de cada uno de ellos, complejizando aún más el viaje del paciente.

La principal consecuencia que tiene el problema antes mencionado, corresponde a que, cuando un paciente ingresa a su viaje quirúrgico, este posee expectativas sobre los plazos a seguir, de manera que su problema de salud se pueda solucionar lo más pronto posible. Sin embargo, dada la situación actual, es difícil que estas expectativas se cumplan satisfactoriamente.

El árbol de problemas creado se encuentra en el anexo A.1.

1.4. Objetivos

1.4.1. Objetivo general

Diseñar un modelo de aprendizaje automático, que permita identificar pacientes con una mayor predisposición quirúrgica, para acompañarlos y acercarlos al pabellón.

1.4.2. Objetivos específicos

1. Diagnosticar la situación actual del problema.
2. Explorar y analizar los datos disponibles, generando gráficos que permitan entender la situación, y dar una idea de cómo abordar la problemática.
3. Diseñar un modelo de clasificación multiclase, que identifique a aquellos pacientes con mayor probabilidad de requerir una intervención quirúrgica.
4. Evaluar el modelo entrenado, contrastándolo con datos fuera del conjunto de entrenamiento.
5. Crear una implementación piloto, que permita al prestador acompañar a los pacientes seleccionados.
6. Generar indicadores a partir del piloto, y discutir sobre los mismos, evidenciando formas en las que se podría mejorar, para un eventual trabajo futuro.

Capítulo 2

Marco Teórico, Conceptual y Metodología

En este capítulo, se presenta el Marco teórico y conceptual del trabajo realizado, con el objetivo de exponer los fundamentos teóricos que sustentan esta investigación. Estos son esenciales para comprender el contexto y la base científica sobre la cual se desarrolla el proyecto. Además, se presenta la metodología en la que se basa la realización del trabajo, enfocándose en el carácter iterativo de la misma.

2.1. Aprendizaje Automático

El Aprendizaje Automático, o *Machine Learning* [5], es una rama de la inteligencia artificial que, según Samuel Arthur, uno de los pioneros de esta área de estudio en 1950, se define como “El campo de estudio que le da a las computadoras la habilidad de aprender sin ser explícitamente programadas”. Así, se reconoce que al entregarle datos que pueden ser palabras, números, imágenes, entre otros, preparados de la manera correcta a una computadora, entonces esta puede aprender sobre los mismos, sin necesidad de mayor intervención humana. En este sentido, se identifican tres tipos de aprendizajes derivados del Aprendizaje Automático:

Aprendizaje Supervisado

El Aprendizaje Supervisado [6], es el tipo de aprendizaje que le entrega a la computadora datos etiquetados con el objetivo de que esta pueda aprender correctamente cuáles son las características que definen a la etiqueta como tal. Por ejemplo, se le puede entregar imágenes de frutas junto con la etiqueta de a qué fruta corresponde la imagen. De esta manera, al entrenar un modelo con estas imágenes, la máquina aprende, por ejemplo, cuáles son las características específicas de las manzanas para clasificarlas correctamente.

Aprendizaje No Supervisado

El Aprendizaje No Supervisado [6], a diferencia del supervisado, no se le entregan los datos etiquetados. De esta manera, su objetivo es encontrar patrones o tendencias en los datos que no necesariamente sean conocidos previamente. Por ejemplo, en *marketing* se puede utilizar este tipo de aprendizaje para encontrar segmentos de compradores en función de cómo interactúan con la página web de una tienda en particular.

Aprendizaje Reforzado

El Aprendizaje Reforzado [7] se basa en un aprendizaje de prueba y error, con la idea de un sistema de recompensas. Por ejemplo, se puede entrenar una máquina para que juegue al ajedrez haciéndole saber que hizo un movimiento erróneo o uno acertado. Así, a medida que vaya jugando, este aprenderá cuáles son los mejores movimientos para cada situación.

2.2. Algoritmos de Clasificación

La clasificación es una de las aristas que involucra el aprendizaje supervisado. Esta, de igual manera que se explicó en el ejemplo anterior en 2.1, corresponde al tipo de aprendizaje que recibe una cantidad de etiquetas discreta sobre los datos que se utilizan para su entrenamiento.

Existen dos tipos de clasificación, la binaria y la multiclase. La primera corresponde a cuando existen sólo dos etiquetas diferentes, mientras que la segunda corresponde a cuando existen más de dos etiquetas, aunque manteniéndose en una cantidad finita.

A continuación se presentan los algoritmos de clasificación a considerar para el trabajo:

Regresión Logística

La Regresión Logística [8], también conocida como *Logit*, es un método estadístico probabilístico utilizado para modelar la relación entre una variable dependiente binaria y una o más variables independientes. Su objetivo es predecir la probabilidad de que una instancia pertenezca a una determinada clase, utilizando la función sigmoide para transformar la salida en un valor entre 0 y 1. Si la probabilidad predicha es mayor a 0.5, se asigna la clase positiva; de lo contrario, se asigna la clase negativa.

De forma más teórica, la Regresión Logística busca encontrar las probabilidades:

$$\mathbb{P}(y = 1|x)$$

$$\mathbb{P}(y = 0|x)$$

Donde el primero indica la probabilidad de que se tome la clase positiva dados los parámetros \mathbf{x} , y el segundo indica la probabilidad de que se tome la clase negativa dados los parámetros \mathbf{x} . Luego a partir de los datos de entrenamiento se definen los pesos w_i , asignándose a cada uno de los x_i de entrada, y junto con esto también se define el valor del intercepto b que corresponde al término del sesgo. Estos parámetros w_i y b se estiman utilizando métodos de optimización como el descenso del gradiente, donde se ajustan iterativamente para minimizar el error de predicción del modelo.

Así, se define el valor de z , que expresa la suma de las ponderaciones de los pesos encontrados anteriormente, sumándole el valor del intercepto:

$$z = \left(\sum_{i=1}^n w_i \cdot x_i \right) + b$$

Luego, de manera que la probabilidad se encuentre entre 0 y 1, se introduce la función sigmoide, llamada así por la similitud que tiene con la letra **S**, se describe como:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

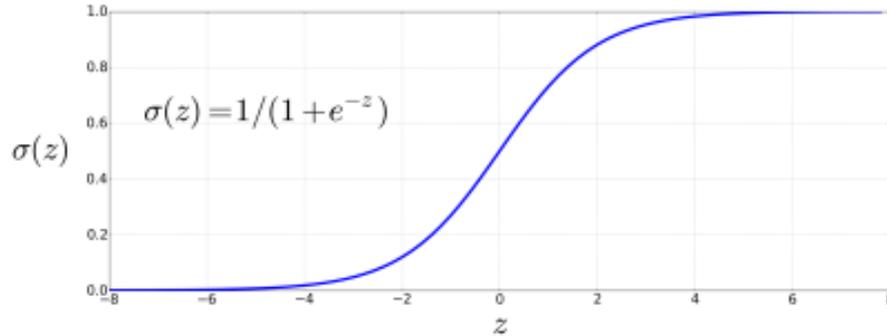


Figura 2.1: Función sigmoide que toma el valor de un número real y lo convierte al rango (0, 1). Obtenido de [8].

Árboles de Decisión

Los Árboles de Decisión [9], también conocidos por su nombre en inglés *Decision Trees*, es un método basado en una estructura tipo árbol donde cada nodo corresponde a un atributo, cada rama al valor del atributo, y cada hoja a la predicción final del modelo.

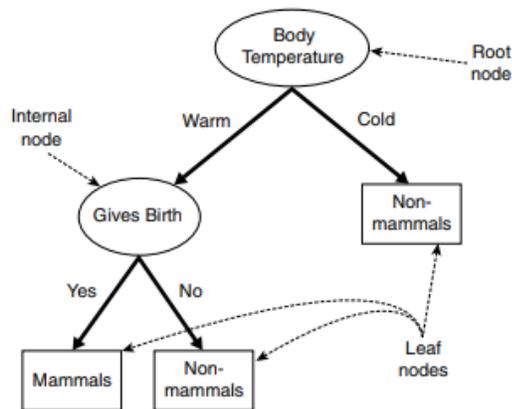


Figura 2.2: Diagrama de un árbol de decisión para la clasificación de un animal como mamífero. Obtenido de [9].

En el caso de la figura 2.2, se puede ver que el nodo raíz es la temperatura del cuerpo (el atributo), de donde sale una rama para temperatura caliente, y otra para temperatura fría (los valores del atributo). Finalmente en los nodos hoja se establece si el animal es mamífero o no.

A partir de un conjunto de atributos se puede construir una cantidad exponencial de árboles de decisión. Es por esto que se han creado algoritmos que, si bien no entregan el mejor resultado posible, sí son lo suficientemente precisos.

El Algoritmo de *Hunt* es uno de ellos. Este es un algoritmo recursivo que en cada nodo selecciona el atributo que mejor separa los datos en subconjuntos más puros, a través de medidas como la ganancia de información o el índice de Gini. El proceso es repetido hasta que se cumple un criterio de detención, los cuales pueden ser que no se puedan realizar más divisiones que mejoren la homogeneidad de los datos, que el árbol haya alcanzado un nivel máximo de profundidad preestablecido, haber alcanzado un número mínimo de muestras en un nodo, entre otros.

Bosques Aleatorios

Bosques Aleatorios [9], también conocidos como *Random Forests*, son un tipo de método de ensamblaje, es decir que combina resultados para obtener un resultado final, diseñado específicamente para los árboles de decisión. Este método toma los datos de entrenamiento, dividiéndolos primeramente en diferentes muestras aleatorias de los datos, y posteriormente seleccionando atributos aleatorios para cada árbol de decisión a entrenar.

Una vez entrenado, cuando se usa el modelo para hacer una predicción, se pasan los datos a través de todos los árboles obteniendo la predicción de cada uno, para luego, a partir de un sistema de votación obtener la predicción final del Bosque Aleatorio. Un ejemplo de este método de ensamblaje se puede ver a continuación:

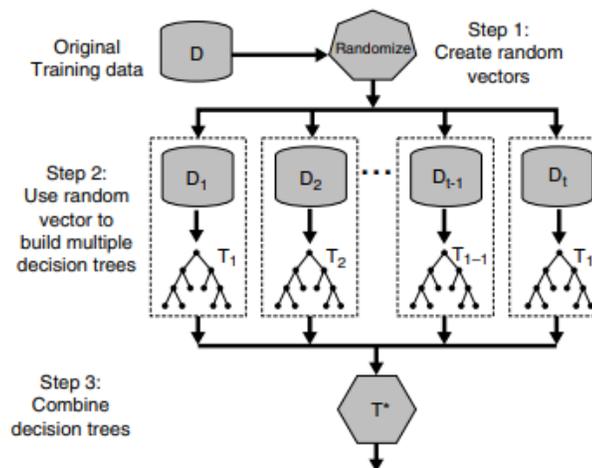


Figura 2.3: Diagrama explicativo del funcionamiento de *Random Forests*. Obtenido de [9].

En particular, *Random Forests* es un método robusto que reduce el riesgo de sobreajuste presentado por los Árboles de Decisión. Además, tiene una gran capacidad para manejar grandes conjuntos de datos con alta dimensionalidad, aunque puede llegar a ser lento para el procesamiento de los mismos.

Boosting

Los algoritmos de *Boosting* trabajan bajo el fundamento de ensamblaje. En general, estos consisten en entrenar distintos modelos secuencialmente, donde el siguiente aprende sobre los errores del anterior. En este sentido, los modelos que componen un algoritmo de *Boosting* ya entrenado corresponden a varios modelos con capacidades de predicción débil, pero que en conjunto generan un clasificador robusto. Esto se puede observar en la figura 2.4 donde se muestra un diagrama del funcionamiento de *Adaboost* que es uno de los algoritmos que utiliza este concepto para el entrenamiento de modelos.

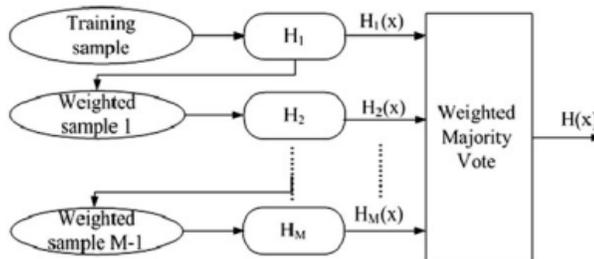


Figura 2.4: Diagrama de *Adaboost*. Obtenido de [10].

A continuación se describen tres algoritmos basados en *Boosting*:

Extreme Gradient Boosting

XGBoosting o *Extreme Gradient Boosting* [11], es un método de ensamblaje que se basa en la minimización de la función de pérdida a través del método del gradiente, el cual toma los gradientes de los errores de los árboles anteriores para ajustar los parámetros del siguiente árbol a entrenar. De esta manera, el algoritmo continúa hasta que logra predecir correctamente todos los datos del conjunto de entrenamiento, o hasta que se cumple con una condición de detención previamente establecida.

Más en específico, para un conjunto de datos $D = \{(x_i, y_i)\} (|D| = n, x_i \in \mathbb{R}^m, y_i \in \mathbb{R})$ con n instancias y m atributos, el modelo de árboles ensamblados utiliza K funciones aditivas para calcular la salida:

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^K f_k(x_i), \quad f_k \in F$$

Donde $F = \{f(x) = w_{q(x)}\} (q : \mathbb{R}^m \rightarrow T, w \in \mathbb{R}^T)$ es el espacio de árboles de decisión con los que se entrena el modelo. Aquí, q corresponde a la representación de la estructura de cada árbol, y T es el número de hojas que tiene cada árbol. Así, cada f_k corresponde a un árbol independiente de estructura q y ponderaciones de hojas w .

Luego, para aprender el conjunto de funciones a utilizar en el modelo se minimiza

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \omega(f_k)$$

$$\text{con } \omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2$$

De donde se puede notar que l es la función de pérdida que indica la diferencia entre el valor predicho \hat{y}_i y el valor real y_i . Este valor es acompañado por factor de regularización ω que penaliza la complejidad del modelo.

AdaBoost

Adaboost, acrónimo de *Adaptive Boosting* [10, 12], es un algoritmo de aprendizaje automático que se destaca por su capacidad para mejorar la precisión de los modelos de clasificación mediante la combinación de múltiples clasificadores débiles. Este método se basa en la idea de asignar pesos diferenciados a cada instancia de datos durante el proceso de entrenamiento, otorgando mayor importancia a aquellas que han sido clasificadas incorrectamente en las iteraciones previas.

En cada iteración, los pesos de los datos se actualizan para centrarse en las instancias más desafiantes, lo que permite que los modelos subsiguientes se enfoquen en corregir los errores más significativos. Esta adaptación progresiva suele conducir a una mejora significativa en la precisión del modelo final.

LightGBM

LightGBM, acrónimo de *Light Gradient Boosting* [13, 14], es un algoritmo desarrollado por científicos de *Microsoft* que se basa en el concepto *Gradient Boosting*, de forma similar a *XGBoost*, pero con la diferencia principal de que el primero busca ser más eficiente y escalable.

Esto lo logra cambiando el enfoque de crecimiento de los árboles desde uno “*level-wise*” hacia uno “*leaf-wise*”. Esto es, que los árboles de decisión entrenados en este método, en vez crecer nivel por nivel, crecen en función de las hojas que entregan una mayor reducción del error. De este modo, los árboles generados por este enfoque tienden a ser más profundos como se puede ver en la figura 2.5.

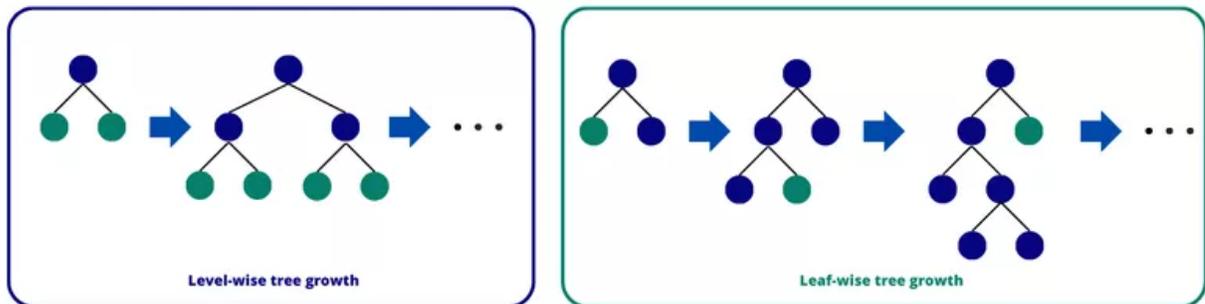


Figura 2.5: Diagrama comparativo de árboles “*level-wise*” y “*leaf-wise*”. Obtenido de [13].

Además, *LightGBM* utiliza una técnica basada en histogramas para la división de nodos. Esto significa que al momento de decidir la división de un nodo, en vez de analizar cada instancia, agrupa las variables descriptivas en histogramas, donde cada barra corresponde a un rango de valores para esa característica en particular.

Dos conceptos introducidos por *LightGBM*, son *Gradient-Based One Side Sampling* y *Exclusive Feature Bundling*. Estos son el fundamento de la eficiencia de este método y se definen como:

1. *Gradient-Based One Side Sampling* (GOSS): introduce un método de muestreo sobre los datos basado en el gradiente. Este toma todos los datos con gradientes más altos (los que necesitan más entrenamiento), mientras que los datos con menor gradiente (los que ya están bien entrenados) son muestreados aleatoriamente.
2. *Exclusive Feature Bundling* (EBF): la construcción de histogramas sobre las variables descriptivas puede ser altamente costosa en términos computacionales. Es por esto que se introduce un método que permite agrupar aquellas variables que son mutuamente excluyentes en una sola variable antes de crear los histogramas, lo cual ayuda a simplificar el proceso de creación de estos últimos.

Multilayer Perceptron

Las Redes Neuronales Artificiales nacen a partir del creciente interés por intentar desarrollar sistemas capaces de emular la capacidad humana para el aprendizaje y la toma de decisiones.

A grandes rasgos, estas son un conjunto de “neuronas” que intentan imitar el aprendizaje biológico, donde si bien para ciertas tareas incluso logran un mejor desempeño que una persona, también existen otros tipos de tareas en las cuales no pueden acercarse a algo que una persona realiza con facilidad.

Al estar imitando la estructura y el funcionamiento de las redes neuronales biológicas, las redes neuronales artificiales están formadas por nodos conectados entre sí. A estos se les llama neuronas, las cuales son capaces de recibir una entrada, realizar un procesamiento, y entregar una señal de salida.

En este sentido, la red neuronal más básica posible consiste de una sola neurona. Esta recibe un input $x_1, x_2, x_3, \dots, x_n$, les asigna pesos a estos valores $w_1, w_2, w_3, \dots, w_n$, y finalmente entrega un output $h_{w,b}(x) = f(w^T x) = f(\sum_{i=1}^n w_i x_i + b)$. Todo esto se puede evidenciar gráficamente en la figura 2.6.

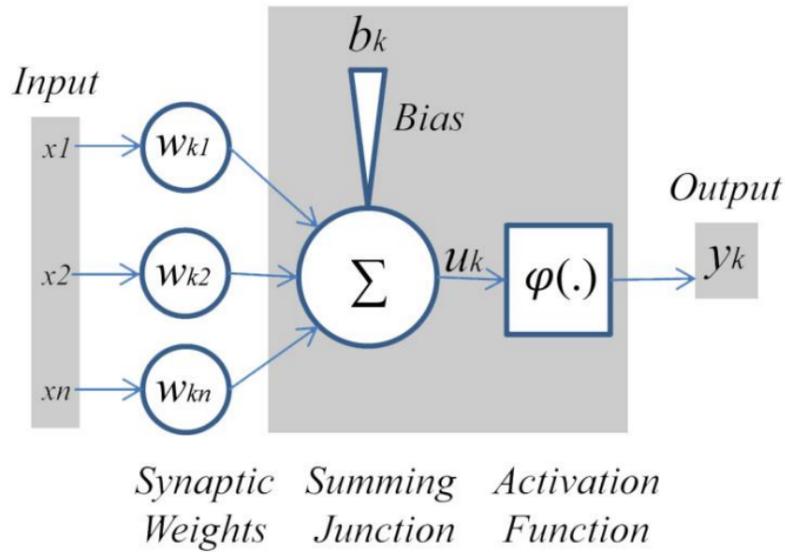


Figura 2.6: Modelo no lineal de una neurona. Obtenido de [15].

Aquí, $f : \mathbb{R} \rightarrow \mathbb{R}$ es la función de activación, que tiene como finalidad determinar si la neurona de la red debe activarse o no en función del *input* que recibe. Esta suele ser no lineal en problemas de clasificación, y permite transmitir información a las neuronas de la siguiente capa de la red.

La función de activación se puede modificar en función del tipo de problema a resolver, pero algunas de ellas son la función Sigmoide $\sigma(x) = \frac{1}{1+e^{-x}}$, *Rectified Linear Unit* (ReLU) $f(x) = \max(0, x)$ y la tangente hiperbólica $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ presentes en la figura 2.7.

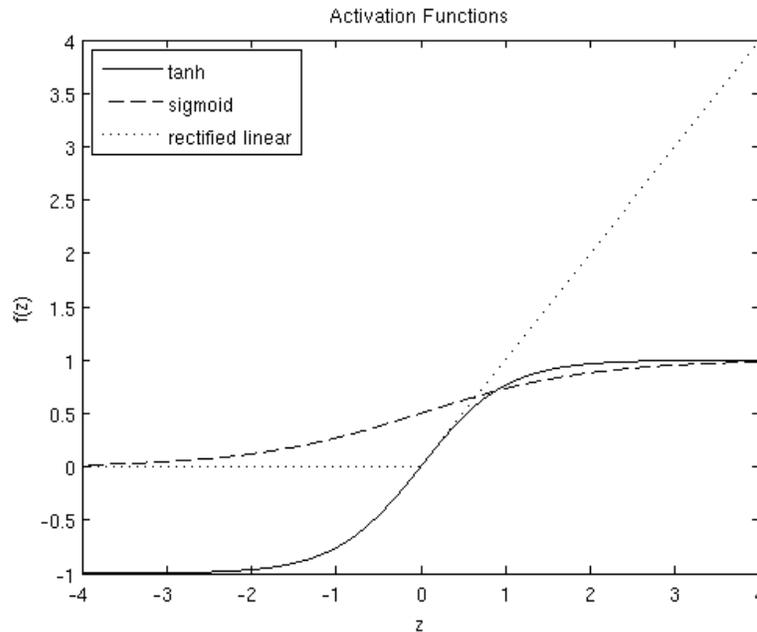


Figura 2.7: Funciones de activación. Obtenido de [16].

Al unir varias neuronas simples en diferentes capas, se forma una red neuronal, como se ilustra en la figura 2.8, que muestra la arquitectura general de un *Multilayer Perceptron* (MLP). El MLP es una arquitectura específica de redes neuronales artificiales en la cual se interconectan distintas capas de neuronas. Esta estructura se compone típicamente de una capa de entrada, seguida de una o más capas ocultas y finalmente una capa de salida.

En cada capa, las neuronas realizan transformaciones no lineales de las entradas ponderadas, lo que permite al MLP aprender y modelar relaciones no lineales complejas entre las variables de entrada y salida. Esta flexibilidad lo hace especialmente adecuado para una amplia gama de tareas en campos como la visión por computadora, el procesamiento de lenguaje natural y muchos otros.

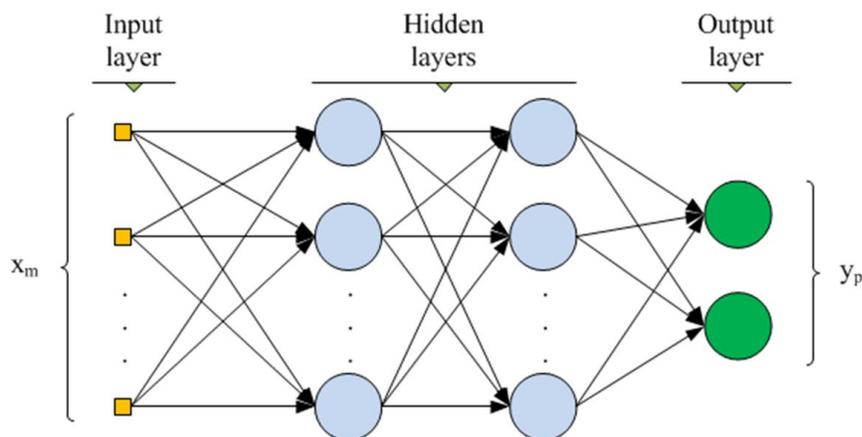


Figura 2.8: Arquitectura general de un *Multilayer Perceptron*. Obtenido de [17].

En el proceso de entrenamiento de un MLP, primero se inicializan los pesos de las conexiones entre las neuronas aleatoriamente, para luego pasar a la propagación hacia adelante. Aquí, los datos de entrada avanzan a través de la red neuronal capa por capa, de manera que cada neurona realiza una transformación no lineal de la suma ponderada de las entradas, seguida de la función de activación. Este proceso continúa hasta llegar a la capa de salida, donde se realiza la predicción del modelo.

Una vez teniendo esta predicción, se compara la salida predicha por el modelo con el valor real a través de una función de pérdida (algunas de ellas son la entropía cruzada, o el error cuadrático, su uso dependerá del problema a resolver) que mide la discrepancia entre ambos valores. Junto con esto, comienza el paso de *backpropagation*, el cual se encarga de propagar inversamente, es decir desde la salida hasta la entrada, el valor calculado, de manera que se pueda calcular el gradiente de la función de pérdida con respecto a los pesos ya definidos y llegar al último paso en el que se actualizan los pesos a través de un algoritmo de optimización (nuevamente depende del problema el tipo de algoritmo que obtiene los mejores resultados, pero algunos de ellos son descenso del gradiente y *Adam*) que permite ajustar los pesos de manera que se minimice la función de pérdida.

Este proceso se repite por un determinado número de “épocas” o hasta que se alcanza un criterio de convergencia predefinido.

2.3. Optuna

Optuna [18] es una estructura de software dedicada a la optimización de hiperparámetros diseñada para implementar una variedad de algoritmos de optimización del estado del arte, con el objetivo de lograr una convergencia eficiente y un buen desempeño en la búsqueda de los mejores hiperparámetros para un modelo de aprendizaje automático.

A diferencia de enfoques más simples como Grid Search, que examina un conjunto predefinido de combinaciones de hiperparámetros, Optuna recibe rangos de valores para los hiperparámetros, ajustando dinámicamente su búsqueda en función de los resultados observados. Esto significa que Optuna identifica y se enfoca en las áreas del espacio de búsqueda que tienen el mayor potencial para mejorar los resultados del modelo.

En el proceso de optimización, se define un número de *trials* en los cuales el algoritmo de optimización entrena modelos con diferentes combinaciones de hiperparámetros. Durante cada *trial*, Optuna ajusta los hiperparámetros con el objetivo de minimizar o maximizar una función objetivo definida por el usuario. Esta función objetivo suele basarse en métricas de rendimiento del modelo, como el área bajo la curva (AUC), la precisión o cualquier otra métrica relevante para el problema en cuestión.

2.4. Validación Cruzada

La Validación Cruzada, [19] es una técnica ampliamente utilizada en el Aprendizaje Automático para prevenir el sobreajuste del modelo a los datos. Esta consiste en evaluar el rendimiento del modelo en diferentes conjuntos de datos no utilizados en el entrenamiento, lo que permite una evaluación más precisa del rendimiento del modelo sobre datos “desconocidos”.

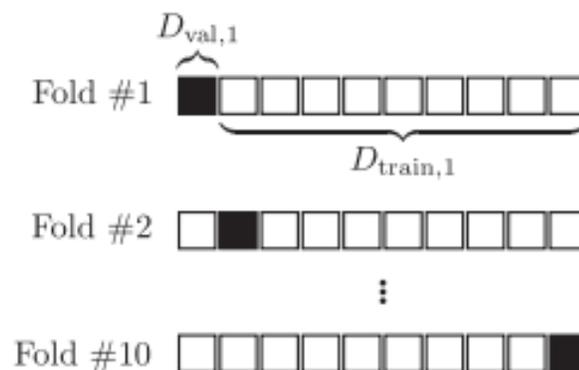


Figura 2.9: Validación Cruzada de conjunto de datos dividido en 10 subconjuntos. Obtenido de [19].

En la práctica, la validación cruzada implica dividir el conjunto de datos en k subconjuntos llamados *folds*, donde k es un número predefinido. Luego, se entrena el modelo k veces, utilizando un subconjunto diferente como conjunto de validación en cada iteración, y los restantes subconjuntos como conjunto de entrenamiento. Esto garantiza que cada instancia de datos se utilice tanto para entrenamiento como para validación en alguna de las iteraciones, como se puede ver en la figura 2.9, donde se muestra una aplicación de Validación Cruzada para $k = 10$.

Existen varios tipos de Validación Cruzada, entre los cuales se incluyen:

1. *K-fold Cross Validation*: Es el método más común, donde se dividen los datos en k subconjuntos aleatoriamente y se reserva uno para la validación en cada iteración de entrenamiento.
2. *Leave One Out Cross Validation (LOOCV)*: Corresponde a un caso especial de *K-fold Cross Validation* donde $k = N^{\circ} \text{ total de datos} - 1$. De esta manera sólo se deja una instancia fuera del entrenamiento para validación, realizando n entrenamientos.
3. *Stratified Cross Validation*: Esta técnica utiliza la misma metodología que *K-fold Cross Validation*, pero con la diferencia de que mantiene las proporciones de los datos para cada clase entre el conjunto de entrenamiento y validación, lo que es útil cuando se trabaja con conjuntos desbalanceados.

2.5. Reducción de Dimensionalidad

La Reducción de Dimensionalidad, [20] es un concepto con variadas aplicaciones en el campo del Aprendizaje Automático, que consiste en representar un conjunto de datos en un espacio de menor dimensión que el original.

A diferencia de simplemente eliminar columnas de datos, lo que podría provocar la pérdida de información importante, la reducción de dimensionalidad implica la creación de nuevas características que son combinaciones de las existentes, permitiendo así una representación más compacta de los datos. Esto es ampliamente utilizado en el contexto del Aprendizaje Automático, pues permite visualizar y entender los datos de mejor manera.

Además, al reducir el tamaño de los datos, también puede mejorar la eficiencia computacional y reducir la tendencia al sobreajuste de los modelos de Aprendizaje Automático.

Existen variados métodos de reducción de dimensionalidad, sin embargo, en el presente trabajo de tesis se hace uso de *Principal Component Analysis (PCA)*, *Uniform Manifold Approximation and Projection (UMAP)* y *Pairwise Controlled Manifold Approximation Projection (PaCMAP)*.

Principal Component Analysis (PCA)

PCA, [20] es una técnica de reducción de dimensionalidad que busca nuevos ejes llamados “Componentes Principales”. Estos son combinaciones lineales de las variables originales que son ortogonales entre sí, lo que significa que son independientes y no correlacionados.

De esta manera, el primer componente principal es el que captura la mayor variabilidad de los datos, el segundo componente principal es el segundo que captura la mayor variabilidad, y así sucesivamente.

Uniform Manifold Approximation and Projection (UMAP)

UMAP [21] es una técnica de reducción de dimensionalidad no lineal que, a diferencia de PCA, no busca mantener la estructura global de los datos, sino que tiene un enfoque más local.

Para lograr esto, el algoritmo utiliza una combinación de técnicas geométricas y topológicas que permiten una representación eficiente y escalable de los datos en un espacio de baja dimensionalidad. Esto hace que UMAP sea especialmente útil para visualizar conjuntos de datos complejos y para descubrir patrones subyacentes que pueden no ser evidentes en dimensiones más altas.

Pairwise Controlled Manifold Approximation Projection (PaC-MAP)

PaCMAP, [22] es una técnica de reducción de dimensionalidad que se distingue por su capacidad para preservar tanto la estructura global como local de los datos en el espacio original.

Para lograr esto, PaCMAP utiliza un enfoque que combina técnicas avanzadas de mapeo y proyección que permiten la conservación de las relaciones entre los puntos de datos a diferentes escalas. Esto se traduce en la capacidad de preservar tanto las relaciones locales entre puntos cercanos, como las relaciones globales entre puntos más distantes en el espacio de alta dimensionalidad.

2.6. Métricas de evaluación

Las métricas de evaluación son herramientas que permiten medir el desempeño de los modelos de aprendizaje automático y comprender su capacidad para realizar predicciones precisas. Gracias a estas es posible medir qué tan bien o que tan mal predice un modelo, lo cual es relevante para notar dónde falla el mismo, para mejorar así su rendimiento.

Algunos conceptos importantes para entender las métricas de evaluación son los Verdaderos Positivos (VP), Verdaderos Negativos (VN), Falsos Positivos (FP) y Falsos Negativos (FN), que se utilizan para construir la matriz de confusión, una herramienta visual que muestra el desempeño del modelo en la clasificación de datos.

Estos se definen como:

- Verdaderos Positivos (VP): Corresponden a aquellas predicciones que el modelo predice como positivas y se encuentran etiquetadas como positivas.
- Verdaderos Negativos (VN): Corresponden a aquellas predicciones que el modelo predice como negativas y se encuentran etiquetadas como negativas.
- Falsos Positivos (FP): Corresponden a aquellas predicciones que el modelo predice como positivas, pero están etiquetadas como negativas.
- Falsos Negativos (FN): Corresponden a aquellas predicciones que el modelo predice como negativas, pero están etiquetadas como positivas.

La matriz de confusión de la figura 2.10 permite entender de mejor manera el significado de estos valores.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	True Positive (TP)	False Positive (FP)
	Negative (0)	False Negative (FN)	True Negative (TN)

Figura 2.10: Matriz de Confusión. Obtenido de [23].

De esta manera, algunas métricas de evaluación son las siguientes:

1. *Precision*: Mide la proporción de predicciones positivas que fueron correctas entre todas las predicciones positivas realizadas por el modelo.

$$Precision = \frac{TP}{TP + FP}$$

2. *Recall*: Indica la habilidad del modelo de predecir verdaderos positivos y qué tan efectiva es esta predicción.

$$Recall = \frac{TP}{TP + FN}$$

3. *Accuracy*: Indica el porcentaje de predicciones correctas que realiza el modelo.

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

4. *F1-Score*: Indica la media armónica entre *Precision* y *Recall*.

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Es importante destacar que estas métricas se calculan para cada clase en problemas de clasificación, lo que permite evaluar el desempeño del modelo en la predicción de cada categoría por separado.

Para generalizar las métricas anteriores, se utilizan técnicas como *macro average* y *weighted average*. Estas entregan un único valor para todas las clases del modelo en las métricas de *precision*, *recall* y *F1-score*, permitiendo evaluar el modelo completo sintetizado en esas métricas. Estas se definen como:

1. *Macro average*: calcula la métrica para cada clase, y luego obtiene la media aritmética a través de todas ellas. Específicamente, para las métricas mencionadas anteriormente, su valor *macro average* se calcula como:

$$Precision_{macro} = \frac{1}{n} \sum_{i=1}^n Precision_i$$

$$Recall_{macro} = \frac{1}{n} \sum_{i=1}^n Recall_i$$

$$F1Score_{macro} = \frac{1}{n} \sum_{i=1}^n F1Score_i$$

2. *Weighted average*: calcula la métrica para cada clase, multiplicando cada una por un ponderador equivalente a su importancia en el conjunto de datos, para posteriormente sumarlas. Específicamente, para las métricas mencionadas anteriormente, su valor *weighted average* se calcula como:

$$Precision_{weighted} = \sum_{i=1}^n w_i * Precision_i$$

$$Recall_{weighted} = \sum_{i=1}^n w_i * Recall_i$$

$$F1Score_{weighted} = \sum_{i=1}^n w_i * F1Score_i$$

Interpretar los resultados de estas métricas es crucial para comprender las fortalezas y debilidades del modelo, para así poder guiar el proceso de ajuste y mejorar su desempeño en problemas específicos.

Es por lo anterior que, para la evaluación y comparación de modelos, se utilizan las métricas macro de *precision*, *recall* y *f1-score*, las cuales entregan una visión general de cómo se comportan los modelos. Posteriormente, y una vez elegido el modelo a utilizar, se utilizan las mismas métricas de *precision*, *recall* y *f1-score*, orientadas a cada una de las clases presentes en la evaluación.

2.7. *Shap Values*

SHapley Additive exPlanations (SHAP) [24] es una herramienta ampliamente utilizada en la ciencia de datos, con la finalidad de explicar las salidas de los modelos de Aprendizaje Automático. Utiliza la teoría de juegos, para asignar crédito a cada variable descriptiva por la contribución que hace a la predicción del modelo.

El principal uso que se le da a esta herramienta es para el entendimiento de cómo afectan las variables descriptivas a los modelos y sus resultados, de manera que se pueda mejorar el rendimiento del mismo.

2.8. Estado del arte

En la literatura investigada previa a la realización del presente trabajo, no se encuentra ninguna investigación similar en el contexto estudiado. Es por esto que a continuación, se presentan estudios del estado del arte en otros contextos, en los cuales se apoya el trabajo actual:

Big Data Preprocessing

El libro [25], se refiere al aumento de la cantidad de los datos a través de los años, haciendo énfasis en que el beneficio del *Big Data* no viene de los datos como tal, sino de la habilidad de poder descubrir patrones inesperados en ellos, permitiendo generar conocimiento a través de las técnicas de ciencia de datos apropiadas. En este sentido, el libro presenta diferentes técnicas y metodologías para lograr el objetivo de sacar conocimiento a los datos.

En particular, para la reducción de los datos presenta dos metodologías. En primer lugar, presenta la posibilidad de aplicar selección de atributos. Esto se refiere al proceso de identificar y eliminar la mayor cantidad de información irrelevante e innecesaria posible.

Por otro lado, está la aplicación de extracción de atributos. Esta se refiere al uso de diferentes algoritmos que permiten generar un nuevo conjunto de atributos, combinando los originales de distintas formas. En este libro se plantea el uso de métodos lineales y no lineales para esta reducción de dimensionalidad.

Classification and authentication of materials using prompt gamma ray activation analysis

El artículo [26] utiliza los espectros de *Prompt Gamma Ray Activation Analysis* (PGAA) para entrenar diferentes modelos y evaluar cuán discriminativos son estos espectros para distintos tipos de materiales. Para ello, se plantea un pipeline en el que se aplican diferentes preprocesamientos a los datos. En primer lugar, los conjuntos de entrenamiento y de prueba se someten a un filtro de umbral de varianza, de manera que si el valor sobrepasa este umbral, el dato se elimina. En segundo lugar, se realiza un balanceo de los datos. En tercer lugar, se aplica un escalamiento a los datos. Finalmente, se realiza una reducción de dimensionalidad, probando tanto un método lineal, como PCA, y un método no lineal, como PaCMAP. En este último paso, se reducen los 4,056 atributos a 2 componentes principales y se visualiza la distribución de las clases en distintos gráficos variando el umbral del procesamiento anterior. A partir de estos gráficos, se concluye que aunque ambos métodos se desempeñan adecuadamente, se decide proceder con PCA debido a que es una alternativa más simple que PaCMAP, a pesar de separar los datos de una forma ligeramente menos efectiva.

2.9. Metodología

CRISP-DM

En el contexto de que el trabajo presentado por el objetivo general corresponde a un proyecto de ciencia de datos, la metodología *Cross-Industry Standard Process for Data Mining* (CRISP-DM) ofrece un marco estructurado y ampliamente reconocido para guiar el proceso de la presente tesis. Este enfoque iterativo presenta seis fases principales que abarcan desde la comprensión del negocio hasta el despliegue del modelo.

- **Comprensión del negocio:** Corresponde al estudio y entendimiento del negocio. En esta fase se debe definir el objetivo del proyecto de minería de datos. Además se debe establecer la métrica a utilizar que medirá el éxito del proyecto.
- **Comprensión de los datos:** Corresponde a la recolección y exploración de los datos. En esta fase se deben explorar los datos de manera que se puedan encontrar posibles errores en los mismos, para procesarlos en la siguiente fase. Por otro lado, resulta relevante la creación de gráficos para comprender cómo se comportan los datos de forma estadística.
- **Preparación de los datos:** en esta fase se definen criterios de inclusión y exclusión de datos. Aquellos datos con errores identificados en la fase anterior se procesan a través de la limpieza de los mismos, y se puede crear nuevos atributos derivados de los ya existentes.
- **Modelado:** Aquí es donde se seleccionan y aplican técnicas de modelado de datos para construir y entrenar modelos predictivos o descriptivos. Se pueden explorar diferentes algoritmos y enfoques de modelado para encontrar la mejor solución al problema específico en cuestión.
- **Evaluación:** En esta fase se evalúan los modelos desarrollados utilizando métricas definidas previamente en la fase de comprensión del negocio. Se analizan los resultados obtenidos y se determina si los modelos cumplen con los objetivos establecidos.

- Despliegue: Consiste en la planificación de la implementación, el seguimiento, y el mantenimiento del modelo en el uso que se le dará dentro de la organización. Además, se establecen procesos para monitorear y mantener el modelo en funcionamiento asegurando su eficacia continua en el entorno operativo.

Una representación gráfica de esta metodología se puede ver en la figura 2.11, donde se muestran las seis fases de la misma, y se destaca el carácter iterativo de CRISP-DM.

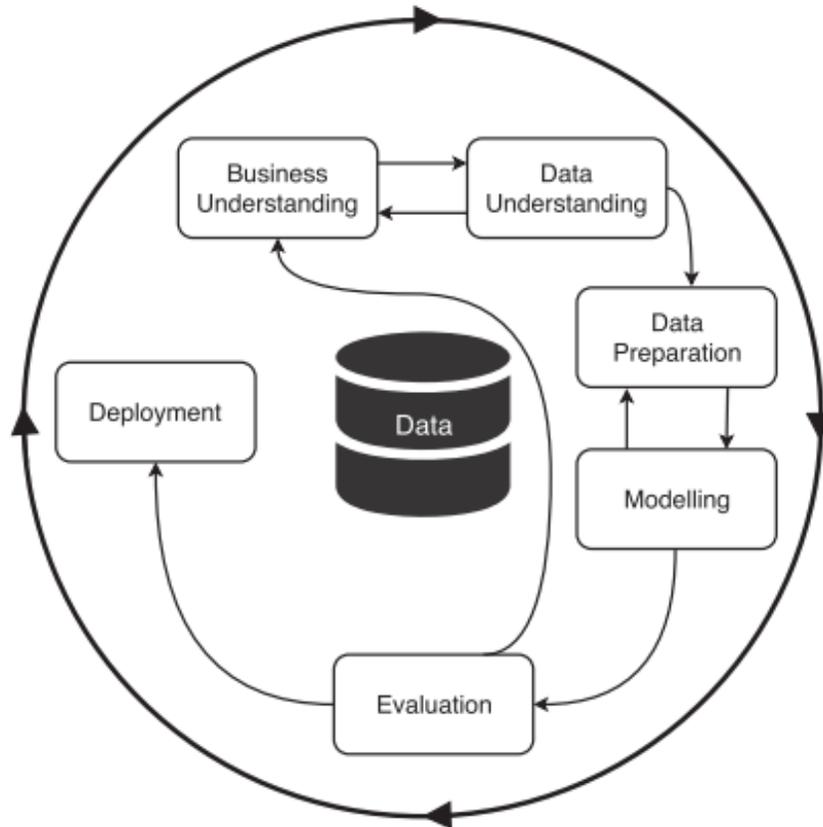


Figura 2.11: Proceso de CRISP-DM para minería de datos. Obtenido de [27].

2.10. Antecedentes

Actualmente, para que un paciente pueda realizarse una cirugía con el prestador, este debe estar en posesión de una orden médica válida con indicación quirúrgica, otorgada por un médico acreditado. De esta manera, debe realizar una solicitud de presupuesto a través del sitio web, o de forma presencial, adjuntando la respectiva orden médica. Estas solicitudes luego son validadas por un ejecutivo de call center, quien determina si se cumple con el estándar mínimo requerido para la emisión del presupuesto. Luego, el proceso continúa a través del área de presupuestos de la clínica correspondiente, que se encarga de evaluar y finalmente emitir el presupuesto de la intervención quirúrgica solicitada.

Por otro lado, dentro del formulario web para la solicitud de presupuesto, también se encuentra la opción de no entregar orden médica. Estos casos son más complejos, porque no es posible entregar un presupuesto sin que el paciente tenga su orden médica válida. Es por esto que aquí se realiza un seguimiento a estos pacientes, efectuando un acompañamiento para que quienes no se encuentren en posesión de su orden médica, puedan obtenerla y realicen su solicitud correctamente.

Así, de acuerdo con Espiñeira [28], a lo largo del año 2023 se registró un promedio mensual de aproximadamente 8,400 solicitudes, de las cuales un 26 % contaba con su respectiva orden médica, y de estas el 71 % fueron consideradas como válidas para la emisión del presupuesto. Mientras que del otro 74 % que no contaba con su orden médica, se logra contactar al 70 %, y un 12 % de ellos agenda realmente una cita ambulatoria al ser contactado.

De esta manera, y como se puede ver en la figura 2.12, el prestador tiene definidos ciertos contactos con los pacientes que identifica con intención quirúrgica. Aquí, se considera a todos los pacientes que han rellenado el formulario web con o sin orden médica. Del primero, una vez que se valida la orden médica, esta se envía a la clínica correspondiente para que se haga la emisión del presupuesto y finalmente se envíe a través del correo electrónico.

Por otro lado, cuando el paciente no entrega una orden médica, se aplica un protocolo de seguimiento en el cual se llama al paciente hasta tres veces. Si después de los tres llamados este no contesta, entonces se le elimina del flujo. Si el paciente contesta, pero decide no agendar ninguna cita para avanzar con su viaje quirúrgico, también se le elimina del flujo. Por último, si el paciente contesta la llamada y decide agendar, entonces se continúa con el seguimiento de ese paciente en el futuro.

Esto se puede evidenciar gráficamente en la figura 2.12.

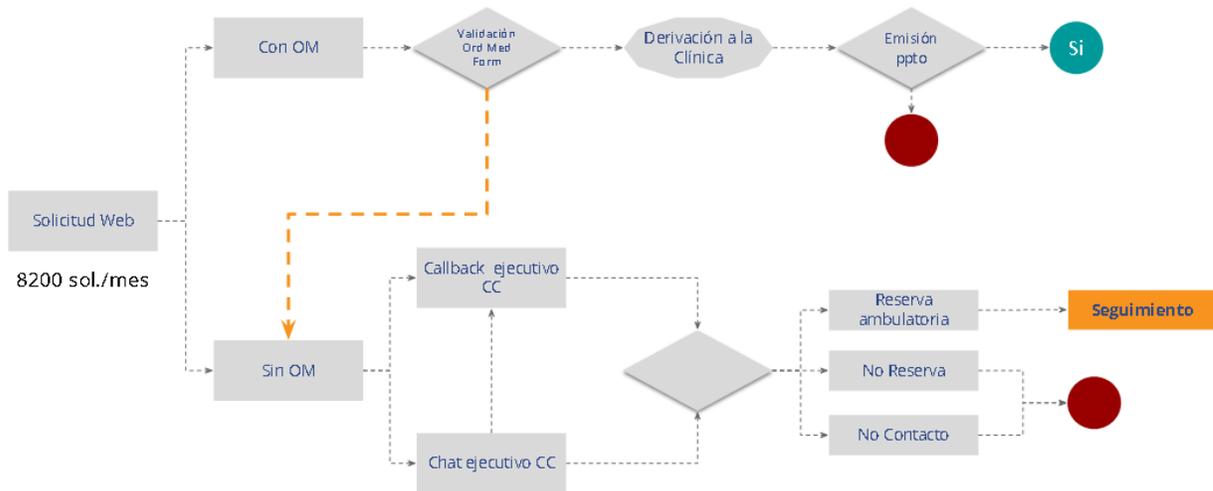


Figura 2.12: Flujo Solicitudes Web. Obtenido de [28].

Capítulo 3

Modelamiento y Evaluación de Resultados

3.1. Comprensión de los datos

Conjunto de datos de agenda

El primer conjunto de datos corresponde a los de agenda, con información sobre las citas asistidas por los pacientes, en los distintos servicios médicos ofrecidos por el prestador de salud. En este, se pueden encontrar las variables de la tabla 3.1.

Tabla 3.1: Variables conjunto de datos de agenda

Conjunto de datos de agenda		
Variables sobre la cita	Variables sobre el paciente	Variables sobre el médico
<ul style="list-style-type: none"> • Fecha • Nombre del servicio • Estado de la cita • Área médica • Nombre del centro médico 	<ul style="list-style-type: none"> • RUT del paciente • Fecha de nacimiento • Sexo • N° de citas atendidas^a 	<ul style="list-style-type: none"> • Nombre del médico • RUT del médico

^a Indica la cantidad de servicios a los que atendió en el periodo de tiempo que abarca el conjunto de datos

De esta manera, se tiene información sobre cada una de las citas asistidas por los pacientes dentro de los centros de salud provistos por el prestador.

Conjunto de datos de concreción de cirugías

Por otro lado, el segundo conjunto de datos corresponde a los de concreción de intervenciones quirúrgicas, con información sobre las cirugías realizadas en las distintas clínicas del prestador. En este, se pueden encontrar las variables de la tabla 3.2.

Tabla 3.2: Variables conjunto de datos de concreción de cirugías

Conjunto de datos de concreción de cirugías	
Variables sobre el paciente	Variables sobre la cirugía
<ul style="list-style-type: none"> RUT del paciente 	<ul style="list-style-type: none"> Código identificador Nombre de la cirugía Clínica donde se realiza la intervención quirúrgica

Conjunto de datos de entrenamiento

A partir de los datos de agenda y de concreción de cirugías, se crea el conjunto fundamental que se utiliza para el entrenamiento de los modelos. Este corresponde a la horizontalización de los datos de agenda, de manera que las filas entreguen información de cada paciente, y cada columna los servicios disponibles. Así, para cada paciente, se puede saber cuántas veces y a qué servicio asistió.

Además, se agrega una columna con las clases del modelo, correspondientes a las cirugías asistidas por los pacientes, o en su defecto, si no se realizaron ninguna cirugía. Así el conjunto de datos contiene 2 variables categóricas, y 1,541 variables numéricas, descritas en las tablas 3.3 y 3.4.

Tabla 3.3: Variables categóricas del conjunto de datos horizontalizado.

Variables categóricas		
Nombre variables	Descripción	Cantidad de columnas
rut_paciente	Identificador del paciente	1
cirugia	Nombre de la cirugía que se realiza. Si no se realiza cirugía este valor es 'No cirugía'	1

Tabla 3.4: Variables numéricas del conjunto de datos horizontalizado.

Variables Numéricas		
Nombre variables	Descripción	Cantidad de columnas
Servicios	Columnas de servicios ofrecidos por el prestador de salud	1541

De esta forma, un ejemplo de cómo se vería este conjunto de datos, se observa en la tabla 3.5.

Tabla 3.5: Ejemplo de cómo se verían los datos de cinco pacientes en el conjunto de datos de entrenamiento con datos falsos.

rut_paciente	cirugía	Medicina_General_Adulto	Examen_de_Sangre_Orina_Otros	...	N21_Cardiologia_Hipertension
123456789	No cirugía	1	0	...	0
987654321	Cirugia Bariatrica	2	1	...	0
314159265	No cirugía	2	0	...	0
358979323	Colecistectomía	0	1	...	0
846264338	No cirugía	0	0	...	0

En esta última, se muestra a modo de ejemplo, cómo se ven los datos de entrenamiento, a través de datos creados al azar. En esta, se ve que hay datos de cinco pacientes distintos distribuidos en cinco filas. Luego, se encuentra la columna que indica el nombre de la cirugía asistida por el paciente. Por último, están las columnas que indican los servicios a los que asiste cada paciente, con un número positivo entero cuando el paciente asiste a ese servicio en particular. Aquí, se ejemplifican las 1,541 columnas de servicios con tres ejemplos, y un espacio “...” indicando dónde se encontrarían las demás columnas.

Así, se obtienen los tres principales conjuntos de datos a utilizar durante el presente trabajo, (1) el **conjunto de datos de agenda**, (2) el **conjunto de datos de concreción de intervenciones quirúrgicas**, y (3) el **conjunto de datos de entrenamiento**.

3.2. Análisis y selección de datos

En esta sección se presenta un análisis de datos realizado con respecto a los tres conjuntos previamente mencionados. Aquí, se busca contextualizar el comportamiento que tienen los pacientes con respecto a los servicios que asisten, incluyendo cómo se relaciona esto con la cirugía que se realiza.

3.2.1. Análisis conjunto de datos de agenda

En el conjunto de datos de agenda se estudia el comportamiento general de los pacientes con respecto a sus citas médicas. Cabe destacar que para este análisis se utilizan datos de fechas ubicadas entre julio de 2023 y el 15 de enero de 2024.

En la figura 3.1, donde se puede ver un histograma de los distintos servicios ofrecidos por el prestador en el periodo mencionado, es posible notar que el servicio más frecuente corresponde al de “Medicina General Adulto”, con aproximadamente 800,000 citas asistidas. Este posee más del doble de la demanda del segundo servicio más común, correspondiente a “Exámenes de Sangre/Orina/Otros”. Luego los siguientes servicios más comunes son “Tratamiento Kinesiología”, “Ginecología General Adulto”, “Traumatología Adulto”, “Diagnóstico Dental”, “Pediatría”, “Tratamiento de Odontología General”, “Urología Adulto” y “Otorrinolaringología Adulto”.

Histograma de servicios

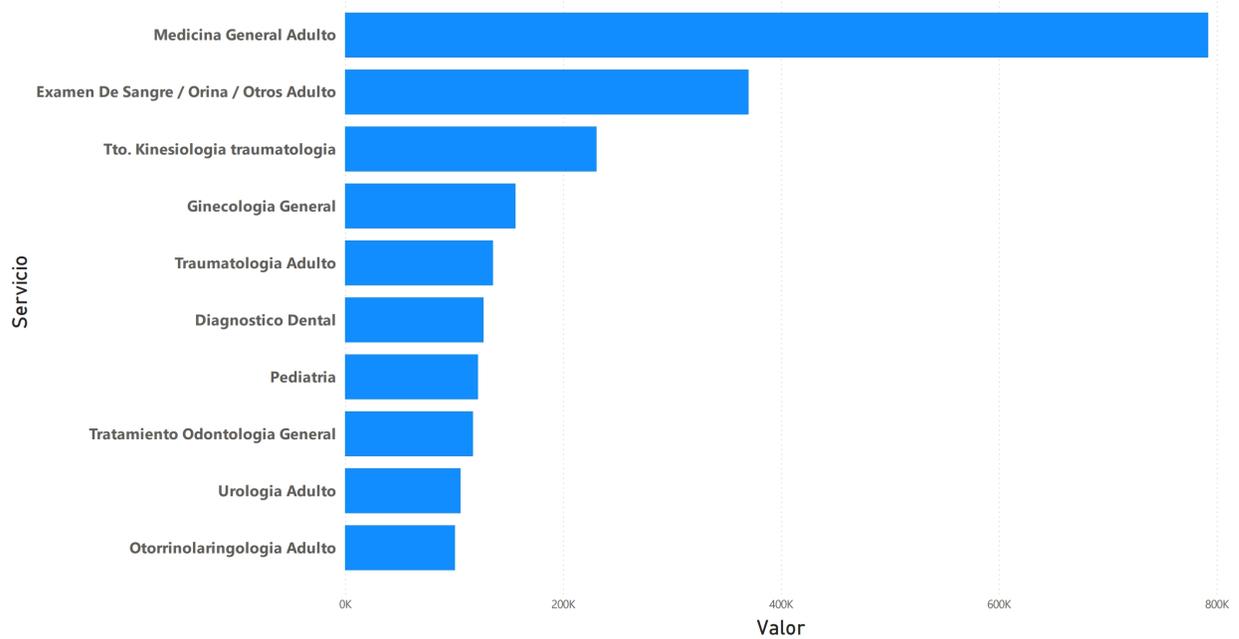


Figura 3.1: Histograma servicios. Muestra el nivel de demanda que han tenido los 10 servicios más populares dentro del prestador de salud. Elaboración propia.

En la figura 3.2, que describe un histograma de la variable “Citas Asistidas”, se puede notar que a medida que aumenta el número de citas asistidas, decrece el número de pacientes que efectivamente las asisten. Así, se infiere que la mayoría de los pacientes tienen menos de tres asistencias en los seis meses abarcados por el conjunto de datos. En este sentido, la mediana de citas asistidas por los pacientes es 5, lo cual permite concluir que, del total de pacientes que han asistido a algún servicio ofrecido por el prestador, al menos la mitad fue a 5 o menos citas entre julio de 2023 y el 15 de enero de 2024, y la otra mitad fue a 5 o más citas en el mismo periodo de tiempo.

Histograma Citas Atendidas

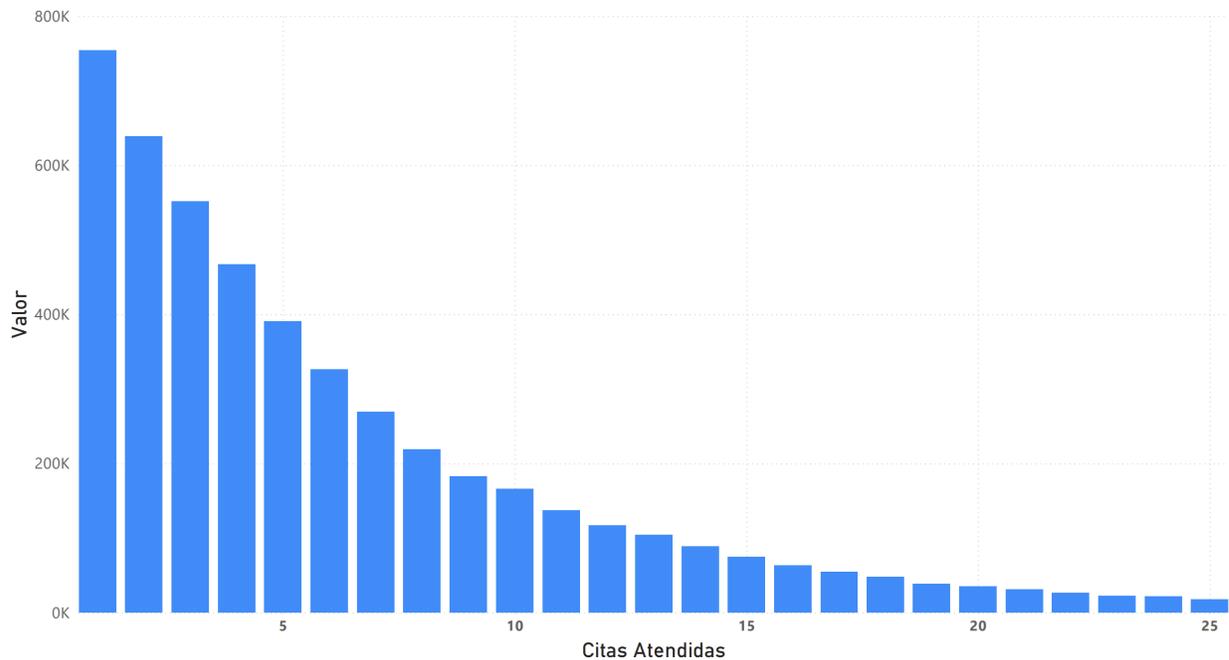


Figura 3.2: Histograma citas asistidas. Indica a cuántas citas asistió cada paciente en el lapso de tiempo que abarca el conjunto de datos. Elaboración propia.

En la figura 3.3, que muestra un histograma de la variable que describe las distintas áreas correspondientes a los servicios asistidos, es posible notar que, como se puede deducir de la figura 3.1, la mayoría de los servicios asistidos corresponden al área médica. Luego en menor medida le siguen las áreas dental, imágenes, laboratorio y kinesiología.

Histograma Área

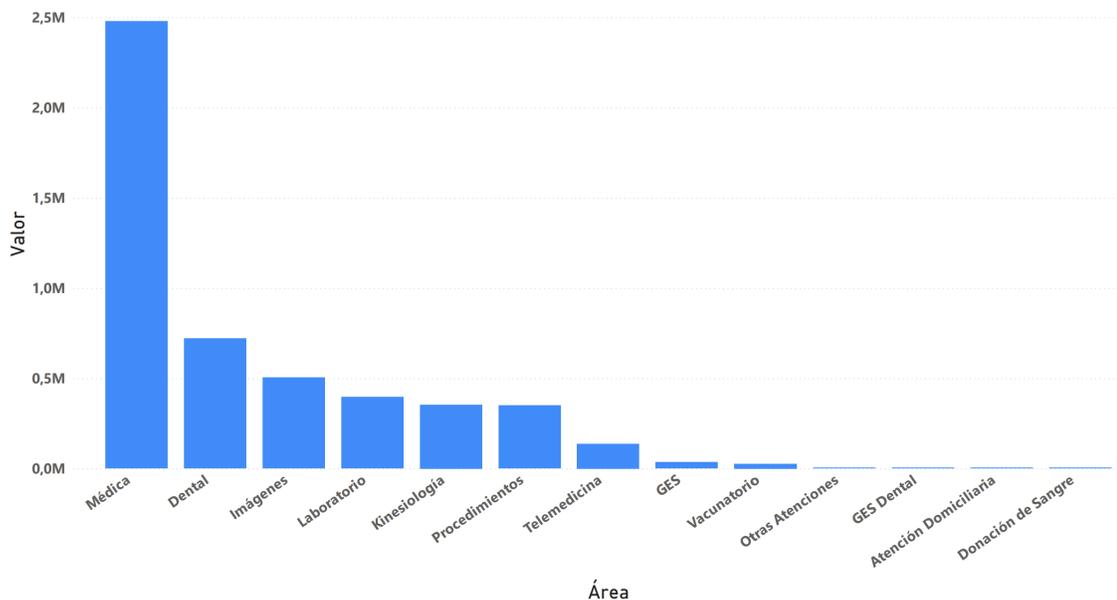


Figura 3.3: Histograma Área. Indica el número de atenciones realizadas por área en el prestador de salud. Elaboración propia.

3.2.2. Análisis conjunto de datos de concreción de cirugías

El conjunto de datos de cirugías, se construye incluyendo todas las cirugías que ofrece el prestador de salud. Además, se incluye una muestra de 20,000 instancias de pacientes que no se realizaron, o no se han realizado una intervención quirúrgica en el intervalo de tiempo abarcado.

En este conjunto de datos, existen nombres de cirugías que son distintos, pero que se refieren a lo mismo. En este caso, la problemática se aborda utilizando los códigos de las cirugías, manteniendo así el nombre más común por cada código presente en los datos. De esta manera, la distribución de cirugías realizadas se ve como en la figura 3.4.

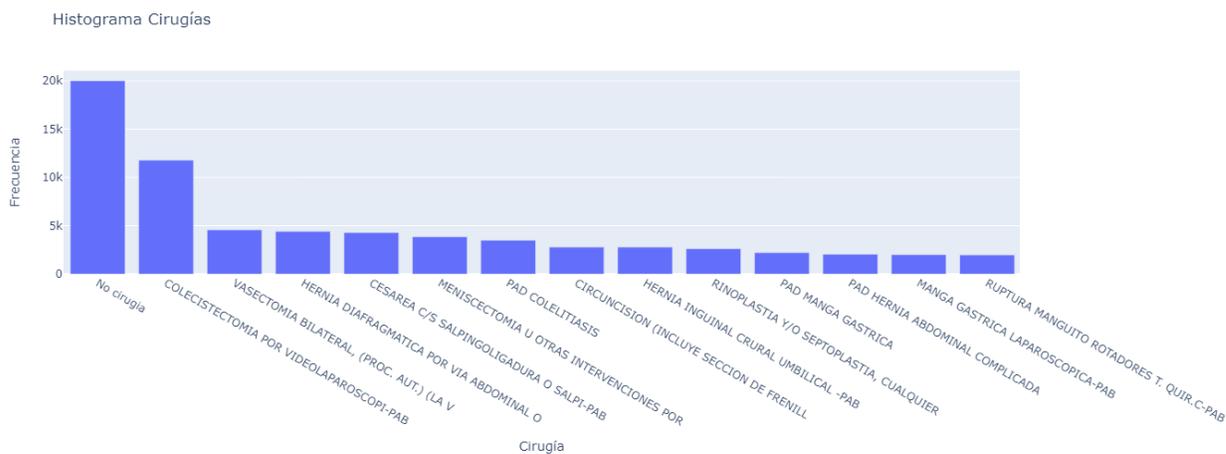


Figura 3.4: Histograma de las cirugías con los nombres más comunes por código. Elaboración Propia.

En la figura 3.4, se puede notar cuáles son las cirugías más frecuentes realizadas por el prestador una vez agrupadas por código. Sin embargo, es necesario realizar una nueva agrupación, pues, existen cirugías que si bien tienen nombre y códigos diferentes, para los fines de este trabajo de tesis y los intereses del prestador, estas cirugías pueden ser agrupadas como una sola dada la equivalencia entre ellas, aumentando así la cantidad de datos para esas cirugías en particular, lo cual también ayuda a mejorar la representatividad de cada clase en el entrenamiento del modelo. De esta manera, se decide agrupar las cirugías como se ve en la tabla 3.6.

Tabla 3.6: Agrupación de las cirugías

Primera agrupación de cirugías	
Nombre agrupado	Nombre cirugía en conjunto de datos
Colecistectomía	<ul style="list-style-type: none"> • PAD COLELITIASIS • COLECISTECTOMIA POR VIDEOLAPAROSCOPI-PAB
Meniscectomía	<ul style="list-style-type: none"> • MENISCECTOMIA U OTRAS INTERVENCIONES POR
Hernia Abdominal	<ul style="list-style-type: none"> • HERNIA INGUINAL CRURAL UMBILICAL -PAB • PAD HERNIA ABDOMINAL COMPLICADA • HERNIA DIAFRAGMATICA POR VIA ABDOMINAL O
Cirugía Bariátrica	<ul style="list-style-type: none"> • PAD MANGA GASTRICA • MANGA GASTRICA LAPAROSCOPICA-PAB
Circuncisión	<ul style="list-style-type: none"> • CIRCUNCISION (INCLUYE SECCION DE FRENILL
Vasectomía	<ul style="list-style-type: none"> • VASECTOMIA BILATERAL, (PROC. AUT.) (LA V
Ruptura Manguito Rotador	<ul style="list-style-type: none"> • RUPTURA MANGUITO ROTADORES T. QUIR.C-PAB
Cesárea	<ul style="list-style-type: none"> • CESAREA C/S SALPINGOLIGADURA O SALPI-PAB
Rinoplastía	<ul style="list-style-type: none"> • RINOPLASTIA Y/O SEPTOPLASTIA, CUALQUIER

Luego, el gráfico con los nombres de las cirugías agrupadas, queda como se aprecia en la figura 3.5.

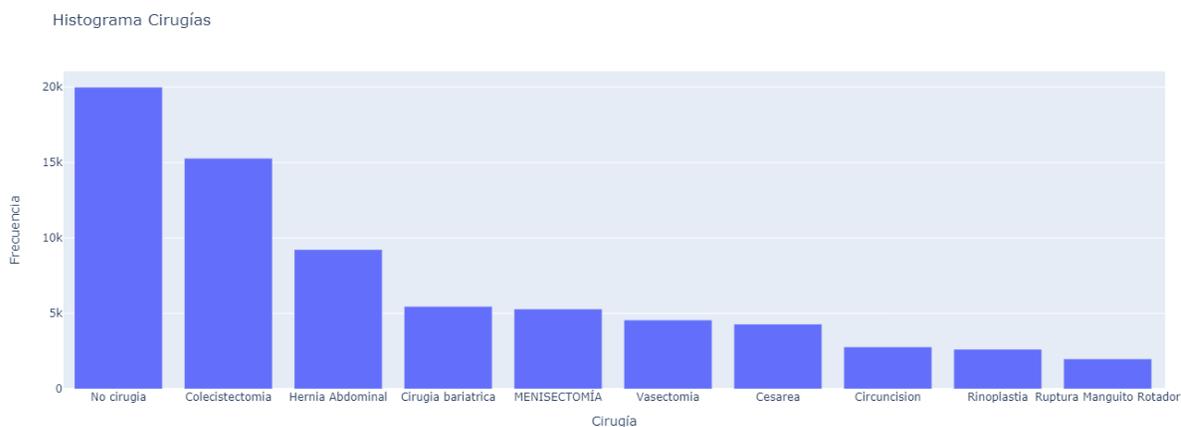


Figura 3.5: Histograma con las cirugías agrupadas según lo presentado en la tabla 3.6. Elaboración propia.

A partir de la figura 3.5, se puede apreciar cuáles son los 10 tipos de cirugías más comunes, y los que, debido a la cantidad de datos disponibles, son las clases a predecir por el modelo.

3.2.3. Análisis conjunto de datos de entrenamiento

Para la realización del siguiente análisis, se utiliza un conjunto de datos previo a la agrupación realizada en la tabla 3.6. Esto implica que los gráficos de las figuras 3.6, 3.7, y las figuras B.1, B.2, B.3 y B.4, disponibles en el anexo, solamente abarcan las cirugías de Colecistectomía, Cirugía Bariátrica, Hernia Abdominal, Cesárea, Vasectomía y Meniscectomía.

Sin embargo, la función de estos gráficos es ejemplificar el comportamiento general de los pacientes que se realizan ciertas cirugías, con respecto a sus servicios más frecuentes, por lo que no es necesario incluir las agrupaciones en su totalidad.

Con este conjunto, se analiza la posibilidad de definir cuáles son los servicios a los que asisten los pacientes antes de llegar a cada cirugía. Al mismo tiempo, se busca definir una lista que incluya los servicios a los que acuden los pacientes antes de llegar a sus respectivas cirugías, a la cual de ahora en adelante se denomina como “Ruta Quirúrgica”. Esto en consideración de que el paciente se encuentra en un viaje quirúrgico, en el cual la cirugía es el destino. De esta manera, la “Ruta Quirúrgica” corresponde a los distintos caminos que puede elegir el paciente para llegar a su destino correctamente.

La función de esta “Ruta Quirúrgica” es asegurar que, para cada cirugía, exista una lista con todos los servicios que el paciente debe, o potencialmente podría, asistir antes de someterse a dicha cirugía. Esto se logra mediante el análisis del porcentaje de pacientes que efectivamente cumplen con cierto conjunto de citas médicas antes de su intervención.

Así, en las figuras 3.6, 3.7, y las figuras B.1, B.2, B.3 y B.4 disponibles en el anexo, se puede notar seis gráficos basados en algunas de las cirugías más comunes. En cada uno de estos, se ordenan los servicios desde el más frecuente (arriba), al menos frecuente (abajo), para esa cirugía. Las barras grises, indican el porcentaje de pacientes que asistieron a esos servicios, antes de la realización de su cirugía.

La línea roja, indica un acumulado de pacientes, sumando aquellos que acudieron a un servicio en particular, pero no a los servicios de arriba. Esto se ejemplifica en la figura 3.6, donde se ve que cerca del 60 % de los pacientes asisten a Cirugía General Adulto. Luego, del 40 % que acude al Electrocardiograma de Reposo, solamente el 10 % asistió a este último y no a Cirugía General Adulto, por lo que la línea roja solamente suma esa diferencia al pasar del primer al segundo servicio.

Finalmente, la línea azul indica el porcentaje de pacientes que han asistido a todos los servicios de arriba al que se está viendo. Por ejemplo, en la figura 3.7, se ve que cerca del 90 % de los pacientes asistieron al servicio de Cirugía Bariátrica, pero menos del 60 % asistió a este último, y a Nutricionista Bariátrica. Luego, menos del 20 % de los pacientes asistió a la Endoscopia, al Control Nutriología Bariátrica, al Control Psicología Bariátrica, al Nutricionista Bariátrica y a Cirugía Bariátrica. Los gráficos de las demás cirugías se pueden encontrar en el anexo B.1, B.2, B.3 y B.4.

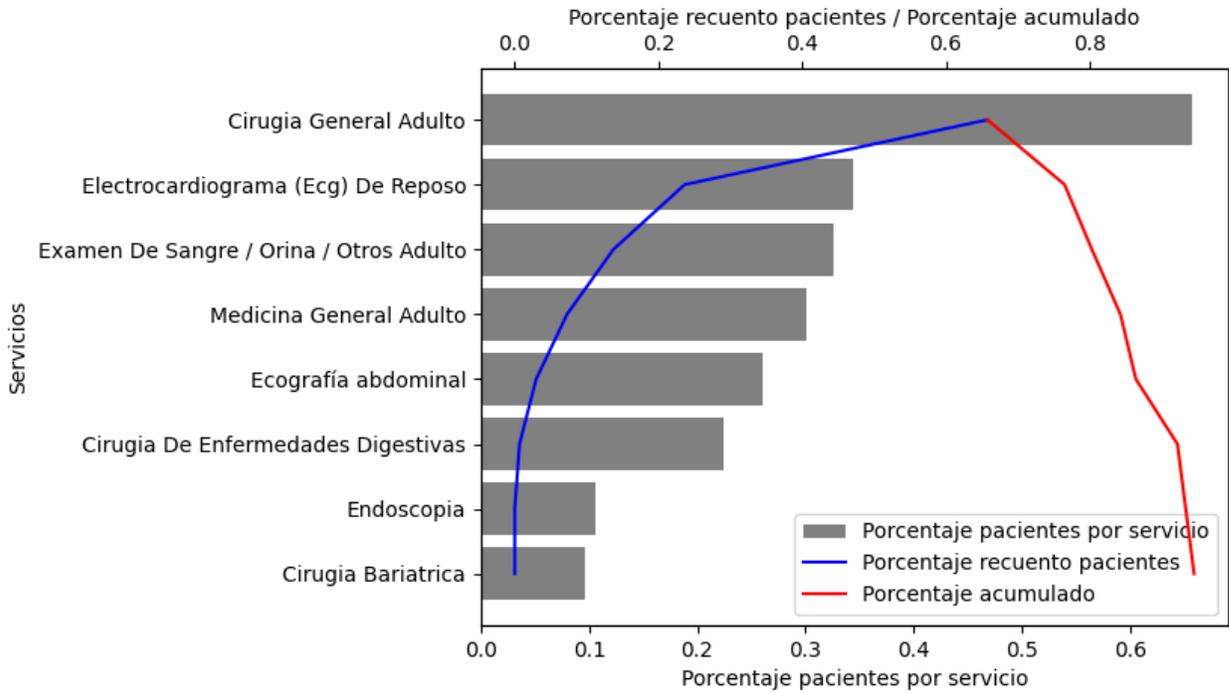


Figura 3.6: Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Colecistectomía. Elaboración propia.

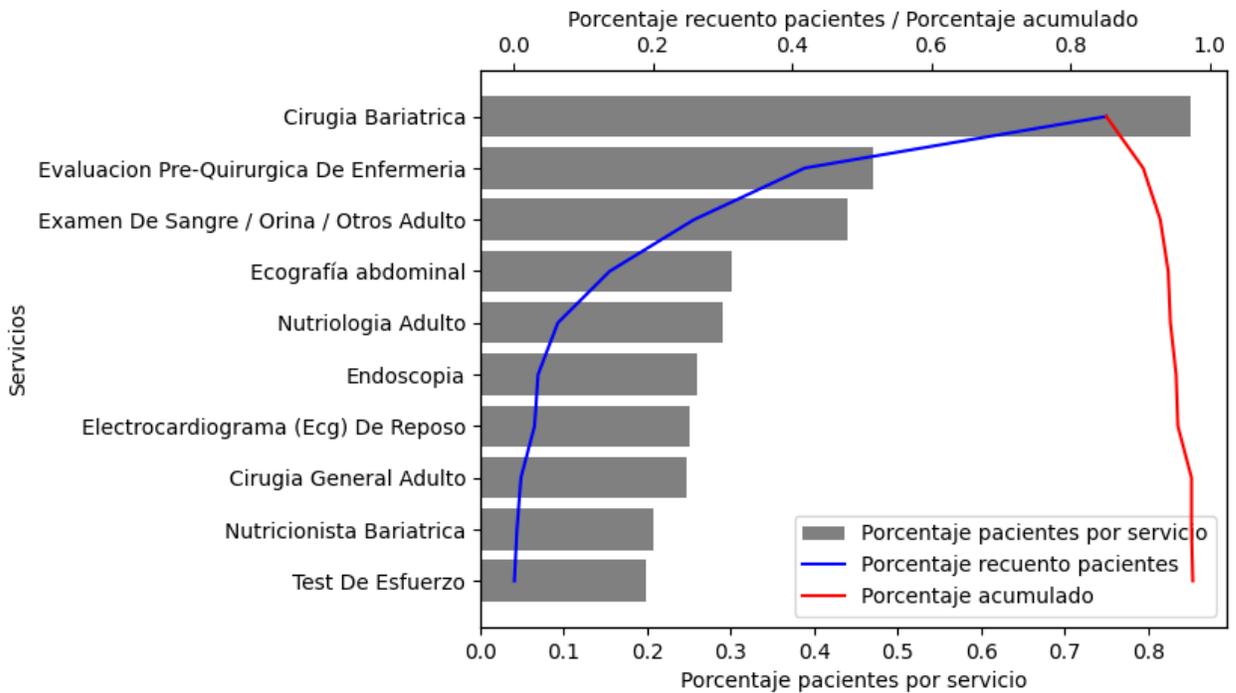


Figura 3.7: Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Cirugía Bariátrica. Elaboración propia.

De esta forma, se infiere que es un porcentaje muy bajo de pacientes el que cumple con asistir, al menos, a los tres servicios más comunes de cada cirugía. Es por esto que posterior al entrenamiento de los modelos, se plantea la utilización de Shap Values para encontrar aquellos servicios que son mejores predictores de la realización de cada cirugía, y junto con esto, encontrar los servicios a los que realmente asisten para llegar bien preparados a sus cirugías.

Por otro lado, se plantea la visualización del conjunto de datos en un formato reducido, a través de las herramientas de reducción de dimensionalidad propuestas en el Marco Teórico. De esta manera, se reducen los datos desde las 1,541 variables descriptivas, a 2 componentes principales, resultantes de la aplicación de los algoritmos de PCA, UMAP y PACMAP.



(a) PCA



(b) UMAP



(c) PACMAP

Figura 3.8: Gráficos reducción de dimensionalidad. Elaboración propia.

Con respecto a lo observado en 3.8, se puede notar que, en primer lugar, PCA (3.8.a) no separa claramente las distintas clases, con un primer componente principal disperso en toda su extensión. Por otro lado, al observar el segundo componente principal, se puede notar que este sí logra separar en cierta medida algunas de las clases. En esta, la mayoría de los datos de “Cirugía Bariátrica” se alejan hacia arriba dentro del gráfico, mientras que los de “Vasectomía” tienden a irse hacia abajo. Además, se logra ver ciertas agrupaciones para las otras cirugías, pero sobrepuestas unas con otras.

UMAP (3.8.b), por otro lado, muestra *clusters* más claros, en los que se pueden diferenciar ciertos segmentos de algunas cirugías, como lo puede ser (1) “cesárea” en la esquina inferior derecha, (2) “ruptura manguito rotador” en el centro, y (3) “vasectomía” en el centro arriba.

Finalmente, PACMAP (3.8.c) muestra un *cluster* grande en el que se logra apreciar la sobreposición de algunas cirugías, y además muestra dos clusters más pequeños en los que se nota una preponderancia de “vasectomía” en el de la izquierda, y de “cesárea” en el central.

3.3. Modelamiento y evaluación

En esta sección, el objetivo es diseñar un modelo de clasificación multiclase, que permita identificar si un paciente se encuentra o no en el viaje quirúrgico de una de las cirugías en estudio. En este modelado de clasificación, se presentan diez clases entre las que están (1) “No cirugía”, (2) “Colecistectomía”, (3) “Hernia Abdominal”, (4) “Cirugía Bariátrica”, (5) “Meniscectomía”, (6) “Vasectomía”, (7) “Cesárea”, (8) “Circuncisión”, (9) “Ruptura Manguito Rotador” y (10) “Rinoplastía”. Luego, como se revisó anteriormente, las variables dependientes del modelo son las que se pueden ver en la tabla 3.7.

Los datos se dividen en un conjunto de entrenamiento (80 %) y uno de prueba (20 %), aplicando la estrategia de muestreo estratificado, pues debido al desbalanceo de los datos, que las clases se encuentren de manera proporcional, ayuda a que cada una tenga la misma representatividad en los dos conjuntos a trabajar.

Así, el número total de datos en el conjunto de entrenamiento alcanza las 59,623 filas, siendo mayoritaria la clase que describe a datos de pacientes “sin cirugías” con 16,000 instancias. La segunda clase con mayor representación es la “Colecistectomía”, con 12,225 instancias, y luego están las demás clases con una representación entre las 7,000 y las 2,000 instancias.

Del mismo modo, y en sintonía con las proporciones de la división antes mencionada en el conjunto de prueba, la clase “No cirugía” posee 4,000 instancias, mientras que la “Colecistectomía” le sigue con 3,056 instancias. Luego, las demás clases poseen entre 2,000 y 500 instancias. Esto se puede ver en detalle en la tabla 3.7, donde se muestra el número exacto de valores que posee cada clase en ambos conjuntos.

Tabla 3.7: División de los datos para el entrenamiento.

Clases	Conjunto de entrenamiento	Conjunto test
No cirugía	16,000	4,000
Colecistectomía	12,225	3,056
Hernia abdominal	7,386	1,847
Cirugía bariátrica	6,264	1,566
Meniscectomía	4,233	1,058
Vasectomía	3,655	914
Cesárea	3,430	858
Circuncisión	2,226	556
Ruptura Manguito Rotador	2,106	527
Rinoplastía	2,098	524
Total	59,623	14,906

Para el entrenamiento de los modelos, se utiliza Optuna con 100 *trials* en la optimización de los hiperparámetros, en conjunto con una grilla particular para cada algoritmo. La definición de cada uno de los hiperparámetros a optimizar se encuentra en el anexo B.2.

De esta manera, se crean las grillas de hiperparámetros presentadas en la tabla 3.8. Estos rangos buscan que los algoritmos funcionen correctamente para grandes cantidades de variables descriptivas, como lo es el caso actual.

Tabla 3.8: Rangos de hiperparámetros introducidos a optuna.

Algoritmo de Clasificación	Hiperparámetros numéricos	Hiperparámetros Categóricos
Regresión Logística	<ul style="list-style-type: none"> 'C': [0.01, 10.0] [Valor decimal] 	<ul style="list-style-type: none"> 'penalty': ['l1', 'l2']
Árboles de Decisión	<ul style="list-style-type: none"> 'max_depth': [2, 32] [Valor entero] 'min_samples_split': [2, 20] [Valor entero] 'min_samples_leaf': [1, 20] [Valor entero] 	<ul style="list-style-type: none"> 'criterion': ['gini', 'entropy']
Random Forests	<ul style="list-style-type: none"> 'n_estimators': [10, 100] [Valor entero] 'max_depth': [2, 32] [Valor entero] 'min_samples_split': [2, 20] [Valor entero] 'min_samples_leaf': [1, 20] [Valor entero] 	<ul style="list-style-type: none"> 'max_features': ['sqrt', 'log2']
XGBoost	<ul style="list-style-type: none"> 'learning_rate': [1e-4, 1e-1] [Valor decimal] 'n_estimators': [10, 100] [Valor entero] 'lambda': [1e-8, 1.0] [Valor decimal] 'alpha': [1e-8, 1.0] [Valor decimal] 'max_depth': [1, 8] [Valor entero] 'gamma': [1e-8, 1.0] [Valor decimal] 'min_child_weight': [1e-8, 1.0] [Valor decimal] 'subsample': [0.5, 1.0] [Valor decimal] 	<ul style="list-style-type: none"> 'objective': ['multi:softmax', 'multi:softprob']
AdaBoost	<ul style="list-style-type: none"> 'n_estimators': [10, 100] [Valor entero] 'learning_rate': [0.01, 1] [Valor decimal] 	<ul style="list-style-type: none"> 'algorithm': ['SAMME', 'SAMME.R']
LightGBM	<ul style="list-style-type: none"> 'learning_rate': [1.e-4, 1.e-1] [Valor decimal] 'n_estimators': [10, 100] [Valor entero] 'lambda_l1': [1e-8, 10] [Valor decimal] 'lambda_l2': [1e-8, 10] [Valor decimal] 'max_depth': [1, 8] [Valor entero] 'min_child_samples': [5, 100] [Valor entero] 'subsample': [0.5, 1.0] [Valor decimal] 'colsample_bytree': [0.5, 1.0] [Valor decimal] 	<ul style="list-style-type: none"> 'boosting_type': ['gbdt', 'rf', 'dart']
Multi-Layer Perceptron	<ul style="list-style-type: none"> 'num_hidden_layers': [1, 10] [Valor entero] 'hidden_layer_units': [16, 256] [Valor entero] 'dropout_rate': [0.0, 0.5] [Valor decimal] 'learning_rate': [1e-5, 1e-1] [Valor decimal] 'batch_size': [16, 1024] [Valor entero] 	<ul style="list-style-type: none"> 'activation': ['relu', 'sigmoid', 'tanh']

Finalmente, se muestran diferentes representaciones del conjunto de datos para su posterior entrenamiento. Primero, se utilizan los datos sin ningún tratamiento (Datos crudos), y luego, se aplican las técnicas de reducción de dimensionalidad mencionadas en el Marco Teórico, PCA, UMAP y PACMAP, representando los datos en diferentes cantidades de componentes principales, para su posterior entrenamiento.

De esta manera, se entrenan siete modelos por cada representación de datos diferente. El detalle de cómo se configura cada representación de los datos, el número de componentes principales utilizadas, y el número total de entrenamientos se puede apreciar en la tabla 3.9.

Tabla 3.9: Descripción de la configuración de los entrenamientos.

Reducción de dimensionalidad	Nº de componentes principales	Modelos entrenados en total
Ninguna	No aplica	7
PCA	[50, 100, 200, 300, 500, 1000, 1500]	49
UMAP	[50, 100, 200, 300, 500, 1000, 1500]	49
PACMAP	[50, 100, 200, 300, 500, 1000, 1500]	49

Para la evaluación de los modelos, se utilizan las métricas *Precision macro*, *Recall macro* y *F1-Score macro*, sobre el conjunto de datos de *test*, de manera que se pueda poner a prueba los distintos modelos con datos distintos a los del entrenamiento.

Así, se procede con la fase de entrenamiento y evaluación. La configuración de cada uno de los entrenamientos, junto con las métricas obtenidas sobre los datos de *test*, se presentan en las tablas 3.10, 3.11, 3.12 y 3.13.

Tabla 3.10: Resultados entrenamientos [1-49].

Nº de modelo	Reducción de dimensionalidad	Nº de componentes principales	Modelo	Métricas		
				Precision	Recall	F1-Score
1	Ninguna	No aplica	Regresión Logística	0.76	0.71	0.72
2	Ninguna	No aplica	Árbol de decisión	0.74	0.685	0.71
3	Ninguna	No aplica	Random Forest	0.74	0.671	0.709
4	Ninguna	No aplica	XGBoost	0.76	0.707	0.728
5	Ninguna	No aplica	AdaBoost	0.74	0.685	0.711
6	Ninguna	No aplica	LightGBM	0.76	0.71	0.73
7	Ninguna	No aplica	Multilayer Perceptron	0.74	0.693	0.71
8	PCA	50	Regresión Logística	0.73	0.668	0.676
9	PCA	50	Árbol de decisión	0.71	0.657	0.677
10	PCA	50	Random Forest	0.73	0.67	0.699
11	PCA	50	XGBoost	0.74	0.688	0.711
12	PCA	50	AdaBoost	0.69	0.623	0.637
13	PCA	50	LightGBM	0.74	0.69	0.71
14	PCA	50	Multilayer Perceptron	0.74	0.686	0.713
15	UMAP	50	Regresión Logística	0.66	0.574	0.598
16	UMAP	50	Árbol de decisión	0.7	0.636	0.657
17	UMAP	50	Random Forest	0.72	0.681	0.657
18	UMAP	50	XGBoost	0.72	0.665	0.685
19	UMAP	50	AdaBoost	0.64	0.553	0.561
20	UMAP	50	LightGBM	0.72	0.66	0.68
21	UMAP	50	Multilayer Perceptron	0.65	0.6	0.63
22	PACMAP	50	Regresión Logística	0.64	0.542	0.563
23	PACMAP	50	Árbol de decisión	0.69	0.634	0.656
24	PACMAP	50	Random Forest	0.71	0.65	0.673
25	PACMAP	50	XGBoost	0.71	0.654	0.675
26	PACMAP	50	AdaBoost	0.63	0.514	0.521
27	PACMAP	50	LightGBM	0.71	0.66	0.68
28	PACMAP	50	Multilayer Perceptron	0.58	0.462	0.484
29	PCA	100	Regresión Logística	0.75	0.697	0.715
30	PCA	100	Árbol de decisión	0.71	0.657	0.674
31	PCA	100	Random Forest	0.74	0.68	0.706
32	PCA	100	XGBoost	0.75	0.7	0.722
33	PCA	100	AdaBoost	0.71	0.65	0.68
34	PCA	100	LightGBM	0.75	0.7	0.72
35	PCA	100	Multilayer Perceptron	0.73	0.7	0.7
36	UMAP	100	Regresión Logística	0.66	0.59	0.62
37	UMAP	100	Árbol de decisión	0.7	0.63	0.67
38	UMAP	100	Random Forest	0.72	0.66	0.68
39	UMAP	100	XGBoost	0.72	0.66	0.68
40	UMAP	100	AdaBoost	0.64	0.55	0.56
41	UMAP	100	LightGBM	0.72	0.67	0.69
42	UMAP	100	Multilayer Perceptron	0.68	0.64	0.65
43	PACMAP	100	Regresión Logística	0.66	0.57	0.59
44	PACMAP	100	Árbol de decisión	0.7	0.63	0.66
45	PACMAP	100	Random Forest	0.71	0.65	0.68
46	PACMAP	100	XGBoost	0.71	0.66	0.68
47	PACMAP	100	AdaBoost	0.64	0.53	0.55
48	PACMAP	100	LightGBM	0.71	0.66	0.68
49	PACMAP	100	Multilayer Perceptron	0.67	0.59	0.61

Tabla 3.11: Resultados entrenamientos [50-98].

Nº de modelo	Reducción de dimensionalidad	Nº de componentes principales	Modelo	Métricas		
				Precision	Recall	F1-Score
50	PCA	200	Regresión Logística	0.75	0.71	0.72
51	PCA	200	Árbol de decisión	0.71	0.66	0.68
52	PCA	200	Random Forest	0.74	0.68	0.71
53	PCA	200	XGBoost	0.75	0.7	0.72
54	PCA	200	AdaBoost	0.71	0.65	0.68
55	PCA	200	LightGBM	0.75	0.7	0.72
56	PCA	200	Multilayer Perceptron	0.75	0.71	0.72
57	UMAP	200	Regresión Logística	0.67	0.6	0.62
58	UMAP	200	Árbol de decisión	0.7	0.64	0.66
59	UMAP	200	Random Forest	0.72	0.66	0.68
60	UMAP	200	XGBoost	0.72	0.66	0.68
61	UMAP	200	AdaBoost	0.64	0.52	0.53
62	UMAP	200	LightGBM	0.72	0.66	0.68
63	UMAP	200	Multilayer Perceptron	0.7	0.64	0.66
64	PACMAP	200	Regresión Logística	0.65	0.56	0.58
65	PACMAP	200	Árbol de decisión	0.69	0.64	0.65
66	PACMAP	200	Random Forest	0.71	0.65	0.67
67	PACMAP	200	XGBoost	0.71	0.66	0.68
68	PACMAP	200	AdaBoost	0.65	0.58	0.59
69	PACMAP	200	LightGBM	0.71	0.66	0.68
70	PACMAP	200	Multilayer Perceptron	0.7	0.64	0.66
71	PCA	300	Regresión Logística	0.75	0.71	0.72
72	PCA	300	Árbol de decisión	0.71	0.64	0.67
73	PCA	300	Random Forest	0.73	0.67	0.7
74	PCA	300	XGBoost	0.75	0.7	0.72
75	PCA	300	AdaBoost	0.71	0.65	0.68
76	PCA	300	LightGBM	0.75	0.7	0.72
77	PCA	300	Multilayer Perceptron	0.74	0.69	0.71
78	UMAP	300	Regresión Logística	0.67	0.6	0.62
79	UMAP	300	Árbol de decisión	0.7	0.64	0.66
80	UMAP	300	Random Forest	0.72	0.66	0.68
81	UMAP	300	XGBoost	0.72	0.67	0.69
82	UMAP	300	AdaBoost	0.64	0.54	0.56
83	UMAP	300	LightGBM	0.72	0.66	0.68
84	UMAP	300	Multilayer Perceptron	0.69	0.62	0.64
85	PACMAP	300	Regresión Logística	0.65	0.55	0.57
86	PACMAP	300	Árbol de decisión	0.69	0.63	0.65
87	PACMAP	300	Random Forest	0.71	0.66	0.68
88	PACMAP	300	XGBoost	0.72	0.66	0.68
89	PACMAP	300	AdaBoost	0.65	0.57	0.59
90	PACMAP	300	LightGBM	0.71	0.65	0.67
91	PACMAP	300	Multilayer Perceptron	0.69	0.62	0.64
92	PCA	500	Regresión Logística	0.76	0.71	0.72
93	PCA	500	Árbol de decisión	0.71	0.65	0.68
94	PCA	500	Random Forest	0.73	0.66	0.69
95	PCA	500	XGBoost	0.75	0.7	0.72
96	PCA	500	AdaBoost	0.71	0.65	0.68
97	PCA	500	LightGBM	0.75	0.7	0.72
98	PCA	500	Multilayer Perceptron	0.75	0.69	0.71

Tabla 3.12: Resultados entrenamientos [99-140].

Nº de modelo	Reducción de dimensionalidad	Nº de componentes principales	Modelo	Métricas		
				Precision	Recall	F1-Score
99	UMAP	500	Regresión Logística	0.68	0.61	0.63
100	UMAP	500	Árbol de decisión	0.7	0.64	0.66
101	UMAP	500	Random Forest	0.72	0.66	0.68
102	UMAP	500	XGBoost	0.72	0.66	0.68
103	UMAP	500	AdaBoost	0.65	0.55	0.56
104	UMAP	500	LightGBM	0.72	0.66	0.68
105	UMAP	500	Multilayer Perceptron	0.63	0.53	0.53
106	PACMAP	500	Regresión Logística	0.65	0.56	0.58
107	PACMAP	500	Árbol de decisión	0.7	0.64	0.66
108	PACMAP	500	Random Forest	0.71	0.65	0.67
109	PACMAP	500	XGBoost	0.71	0.66	0.68
110	PACMAP	500	AdaBoost	0.63	0.52	0.52
111	PACMAP	500	LightGBM	0.71	0.65	0.68
112	PACMAP	500	Multilayer Perceptron	0.7	0.64	0.65
113	PCA	1.000	Regresión Logística	0.75	0.71	0.72
114	PCA	1.000	Árbol de decisión	0.71	0.65	0.67
115	PCA	1.000	Random Forest	0.72	0.64	0.68
116	PCA	1.000	XGBoost	0.75	0.7	0.72
117	PCA	1.000	AdaBoost	0.71	0.65	0.67
118	PCA	1.000	LightGBM	0.75	0.7	0.72
119	PCA	1.000	Multilayer Perceptron	0.63	0.46	0.44
120	UMAP	1.000	Regresión Logística	0.66	0.58	0.6
121	UMAP	1.000	Árbol de decisión	0.7	0.64	0.66
122	UMAP	1.000	Random Forest	0.72	0.66	0.68
123	UMAP	1.000	XGBoost	0.74	0.67	0.69
124	UMAP	1.000	AdaBoost	0.58	0.55	0.56
125	UMAP	1.000	LightGBM	0.75	0.67	0.69
126	UMAP	1.000	Multilayer Perceptron	0.64	0.54	0.56
127	PACMAP	1.000	Regresión Logística	0.64	0.54	0.56
128	PACMAP	1.000	Árbol de decisión	0.69	0.63	0.65
129	PACMAP	1.000	Random Forest	0.71	0.65	0.67
130	PACMAP	1.000	XGBoost	0.71	0.66	0.68
131	PACMAP	1.000	AdaBoost	0.64	0.52	0.54
132	PACMAP	1.000	LightGBM	0.71	0.65	0.67
133	PACMAP	1.000	Multilayer Perceptron	0.64	0.57	0.58
134	PCA	1.500	Regresión Logística	0.76	0.71	0.72
135	PCA	1.500	Árbol de decisión	0.74	0.64	0.67
136	PCA	1.500	Random Forest	0.77	0.65	0.68
137	PCA	1.500	XGBoost	0.77	0.7	0.72
138	PCA	1.500	AdaBoost	0.74	0.65	0.68
139	PCA	1.500	LightGBM	0.78	0.7	0.72
140	PCA	1.500	Multilayer Perceptron	0.75	0.69	0.71

Tabla 3.13: Resultados entrenamientos [141-154].

N° de modelo	Reducción de dimensionalidad	N° de componentes principales	Modelo	Métricas		
				Precision	Recall	F1-Score
141	UMAP	1.500	Regresión Logística	0.7	0.61	0.63
142	UMAP	1.500	Árbol de decisión	0.7	0.64	0.66
143	UMAP	1.500	Random Forest	0.75	0.66	0.68
144	UMAP	1.500	XGBoost	0.74	0.67	0.69
145	UMAP	1.500	AdaBoost	0.58	0.55	0.56
146	UMAP	1.500	LightGBM	0.75	0.66	0.68
147	UMAP	1.500	Multilayer Perceptron	0.72	0.64	0.65
148	PACMAP	1.500	Regresión Logística	0.69	0.55	0.56
149	PACMAP	1.500	Árbol de decisión	0.71	0.63	0.65
150	PACMAP	1.500	Random Forest	0.73	0.66	0.68
151	PACMAP	1.500	XGBoost	0.72	0.66	0.68
152	PACMAP	1.500	AdaBoost	0.65	0.53	0.53
153	PACMAP	1.500	LightGBM	0.73	0.66	0.68
154	PACMAP	1.500	Multilayer Perceptron	0.7	0.64	0.65

A partir de los resultados presentados, es posible notar que, de todas las métricas, los modelos entrenados con los datos crudos, y con PCA, tienen resultados levemente superiores a los obtenidos por UMAP y PACMAP, para todas las cantidades de componentes principales incluidas en los entrenamientos. De esta manera, no se puede definir claramente cuál de los entrenamientos con PCA fue el que obtuvo mejores resultados, debido a que las diferencias entre ellos son mínimas y variantes entre los diferentes modelos.

Sin embargo, sí se puede notar que el mejor resultado lo obtiene el modelo de *LightGBM* que fue entrenado con datos crudos, obteniendo un *precision* de 0.76, un *recall* de 0.71, y un *F1-Score* de 0.73. Esto se acerca bastante a los demás modelos entrenados con PCA, donde *LightGBM* obtuvo resultados de 0.75 en *precision*, 0.7 en *recall* y 0.72 en *F1-Score*.

Luego, y además de los mejores resultados obtenidos por la configuración mencionada, se decide ahondar en la investigación con los datos crudos sobre el algoritmo de *LightGBM*. Esto, debido a que los tiempos de ejecución de *LightGBM* son menores que los de los demás algoritmos, y además, los datos requieren de un menor procesamiento para su posterior entrenamiento, por lo que se determina como la mejor opción para el futuro uso del modelo por la contraparte.

De esta manera se presenta el reporte de clasificación, de cada una de las clases, obtenido del entrenamiento antes mencionado en la tabla 3.14.

Tabla 3.14: Reporte de clasificación por cirugía modelo datos crudos *LightGBM*.

Nombre Cirugía	Métricas			
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
Cesárea	0.87	0.78	0.82	858
Circuncisión	0.78	0.37	0.50	556
Colecistectomía	0.69	0.81	0.75	3056
Cirugía Bariátrica	0.83	0.86	0.84	1566
Hernia Abdominal	0.83	0.49	0.62	1847
Meniscectomía	0.74	0.79	0.77	1058
No cirugía	0.78	0.85	0.81	4000
Rinoplastía	0.79	0.57	0.66	524
Ruptura Manguito Rotador	0.86	0.77	0.81	527
Vasectomía	0.61	0.80	0.69	914

De esta tabla 3.14, en primer lugar, se infiere que las clases de Cesárea, Cirugía Bariátrica, No cirugía y Ruptura Manguito Rotador son las mejor clasificadas con *F1-Scores* sobre 0.8. Luego, se nota que las clases de Circuncisión, Hernia Abdominal, Rinoplastía, y Vasectomía obtienen los peores resultados, con *F1-Scores* menores a 0.7.

Además, la Circuncisión y la Vasectomía, al ser ambas cirugías que intuitivamente, requieren que el paciente asista a servicios urológicos, se infiere que el modelo no logra distinguir correctamente una de la otra.

Del mismo modo, se infiere que la Hernia Abdominal se puede estar confundiendo con la Colecistectomía, ya que ambas son cirugías que se realizan en el abdomen del paciente, lo que hace que los servicios necesarios para llegar a estas cirugías sean similares.

Así, se decide realizar una segunda iteración del modelo, integrando cambios para mejorar los resultados obtenidos.

En primer lugar, se plantea incluir dos nuevas variables a los datos de entrenamiento. Primero, la variable “edad”, para ayudar al modelo a identificar aquellas cirugías que podrían llegar a tener patrones con ciertos grupos etarios. Además, se incluye la variable sexo, para ayudar a distinguir aquellas cirugías que son estrictamente de ciertos géneros, como lo puede ser la Circuncisión y la Vasectomía del sexo masculino, o la Cesárea del sexo femenino.

Así, un ejemplo de como se vería el conjunto de datos, incluyendo las variables descritas, es lo que se puede ver en la tabla 3.15, donde se aprecia que las últimas dos columnas de la derecha corresponden a la inclusión de las columnas anteriormente mencionadas.

Tabla 3.15: Ejemplo de cómo se verían los datos de cinco pacientes en el conjunto de datos de entrenamiento con datos falsos.

rut_paciente	cirugía	Medicina_General_Adulto	...	N21_Cardiología_Hipertension	edad	sexo
123456789	No cirugía	1	...	0	27	Femenino
987654321	Cirugía Bariátrica	2	...	0	25	Masculino
314159265	No cirugía	2	...	0	57	Masculino
358979323	Colecistectomía	0	...	0	57	Femenino
846264338	No cirugía	0	...	0	27	Femenino

Luego, se realizan modificaciones en cómo se agrupan las cirugías para la creación de las clases. En este sentido, el cambio principal se realiza en las cirugías de Circuncisión y Vasectomía, ya que se decide unir ambas clases, generando una clase única llamada Cirugía Urológica.

A continuación, se cambia el nombre de la clase Cesárea a Parto, ya que se agregan dos grupos de cirugías relacionadas a la realización del parto natural, por lo que resulta más general la utilización de este nuevo nombre.

Finalmente, se agregan a ciertas agrupaciones otras cirugías que difieren de la cirugía general como para tener otro nombre/código, pero no lo suficiente como para ser descartadas. El detalle de la nueva definición de las clases se puede notar en la tabla 3.16.

Tabla 3.16: Nueva agrupación de las cirugías

Nueva agrupación de cirugías	
Nombre agrupado	Nombre cirugía en conjunto de datos
Colecistectomía	<ul style="list-style-type: none"> • PAD COLELITIASIS • COLECISTECTOMIA POR VIDEOLAPAROSCOPI-PAB
Meniscectomía	<ul style="list-style-type: none"> • MENISCECTOMIA U OTRAS INTERVENCIONES POR • MENISCECTOMIA
Hernia Abdominal	<ul style="list-style-type: none"> • HERNIA INGUINAL CRURAL UMBILICAL -PAB • PAD HERNIA ABDOMINAL COMPLICADA • HERNIA DIAFRAGMATICA POR VIA ABDOMINAL O
Cirugía Bariátrica	<ul style="list-style-type: none"> • PAD MANGA GASTRICA • MANGA GASTRICA LAPAROSCOPICA-PAB • PAD BY PASS GASTRICO • BY PASS GASTRICO LAPAROSCOPICO-PAB • CIRUGÍA BARIÁTRICA POR BY PASS GASTRICO
Cirugía Urológica	<ul style="list-style-type: none"> • CIRCUNCISION (INCLUYE SECCION DE FRENILL • VASECTOMIA BILATERAL, (PROC. AUT.) (LA V
Ruptura Manguito Rotador	<ul style="list-style-type: none"> • RUPTURA MANGUITO ROTADORES T. QUIR.C-PAB • RUPTURA MANGUITO ROTADOR
Parto	<ul style="list-style-type: none"> • CESAREA C/S SALPINGOLIGADURA O SALPI-PAB • PARTO PRESENTA.CEFALICA O PODALICA C-PAB • PARTO
Rinoplastía	<ul style="list-style-type: none"> • RINOPLASTIA Y/O SEPTOPLASTIA, CUALQUIER

Así, la nueva distribución de los datos queda como se puede apreciar en la tabla 3.17, dada por la división del conjunto de datos en un 80 % para el de entrenamiento, y un 20 % para el de *test* con la técnica de división de datos estratificada. Aquí, se puede ver que la mayor diferencia se encuentra en los datos agregados por la inclusión de los nuevos casos en la clase de Parto, mientras que las demás clases se mantienen constantes con respecto a la primera división presentada en la tabla 3.7.

Tabla 3.17: Nueva división de los datos

Clases	Conjunto de entrenamiento	Conjunto test
No cirugía	16,000	4,000
Colecistectomía	12,224	3,056
Hernia Abdominal	7,386	1,847
Cirugía Bariátrica	6,262	1,566
Cirugía Urológica	6,255	1,564
Parto	5,702	1,426
Meniscectomía	4,233	1,058
Ruptura Manguito Rotador	2,107	526
Rinoplastía	2,098	524
Total	62,267	15,567

Una vez teniendo los datos listos, se plantea, de la misma forma que se mencionó anteriormente, la utilización de los datos crudos en conjunto con el algoritmo de LightGBM. Esto nuevamente es optimizado en 100 *trials* con Optuna, pero con la novedad de que para cada uno de los entrenamientos realizados por el optimizador, las métricas de evaluación a optimizar, son obtenidas a partir del método de validación cruzada estratificada con 100 particiones sobre los datos. De esta manera, el resultado obtenido es más robusto y tiene una menor tendencia al sobreajuste de los datos. Así, el resultado del nuevo entrenamiento se puede ver en la tabla 3.18.

Tabla 3.18: Resultados entrenamiento final

N° de modelo	Reducción de dimensionalidad	N° de componentes principales	Modelo	Métricas		
				Precision	Recall	F1-Score
155	Ninguna	No aplica	LightGBM + Validación Cruzada Estratificada	0.81	0.79	0.80

Estos resultados superan los del mejor entrenamiento previo en todas las métricas utilizadas. En particular, *precision* aumenta de 0.76 a 0.81, *recall* aumenta de 0.71 a 0.79, y *F1-Score* aumenta de 0.73 a 0.80.

El detalle de estos resultados se puede ver en el reporte de clasificación de la tabla 3.19.

Tabla 3.19: Reporte de clasificación por cirugía nuevo modelo datos crudos *LightGBM* y validación cruzada estratificada.

Nombre Cirugía	Métricas			
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
Parto	0.87	0.89	0.88	1426
Cirugía Urológica	0.81	0.86	0.83	1564
Colecistectomía	0.77	0.76	0.76	3056
Cirugía Bariátrica	0.83	0.87	0.85	1566
Hernia Abdominal	0.74	0.65	0.69	1847
Meniscectomía	0.81	0.77	0.79	1058
No cirugía	0.80	0.84	0.82	4000
Rinoplastía	0.78	0.68	0.73	524
Ruptura Manguito Rotador	0.87	0.77	0.82	526

En esta tabla 3.19, todos los resultados mejoran con respecto a los del mejor entrenamiento anterior, disponibles en la tabla 3.14. Primero, el resultado de la clase Cirugía Urológica, creada a partir de la Vasectomía y la Circuncisión, logra un *precision* de 0.81, un *recall* de 0.86 y un *F1-Score* de 0.83.

Segundo, el resultado de la clase Parto, también creada a partir de la combinación de otras cirugías, logra un *precision* de 0.87, un *recall* de 0.89, y un *F1-Score* de 0.88 mejorando aún más los resultados anteriores de la clase Cesárea.

De esta manera, los *F1-Scores* de todas las clases se encuentran sobre 0.7, a excepción de Hernia Abdominal, que a pesar de mejorar con respecto al entrenamiento anterior, solamente obtiene un *F1-Score* de 0.69. El detalle de la matriz de confusión del modelo se puede ver en el anexo B.6.

3.4. Creación de las rutas quirúrgicas por cirugía

En esta sección, se presenta cómo se utilizan los resultados de los modelos, para ayudar a definir la “Ruta Quirúrgica” de servicios necesarios que debe asistir un paciente, y así poder llegar a su cirugía con una preparación correcta. Para esta parte, resulta relevante el entendimiento general de cada una de las cirugías en análisis. De esta manera, se puede entender a qué parte del cuerpo se refiere, y en qué consiste cada cirugía, dándole contexto a los servicios de las distintas “Rutas Quirúrgicas” generadas. Así, la definición de cada una de las cirugías en cuestión, se puede encontrar en el anexo B.3.

Con esto, y como se mencionó anteriormente, además de poder identificar a los pacientes con una mayor probabilidad quirúrgica, se quiere definir la “Ruta Quirúrgica” con los servicios necesarios para que cada paciente pueda llegar correctamente a su cirugía. El primer acercamiento a la construcción de este listado, se basa en la frecuencia de los servicios asistidos por pacientes de cada cirugía, en la sección 3.2.3.

Ahora, se busca poder definir este listado a través de aquellos servicios que se comportan como predictores positivos para las distintas cirugías. De esta manera, se crean los gráficos de *Shap Values*.

En estos, se puede apreciar que existen dos ejes. En primer lugar, en el eje horizontal se encuentra el *Shap Value* de la característica en particular, mientras que en el eje vertical se encuentran las 20 características más importantes para la predicción de esa clase. Además, los colores de los puntos indican el impacto que tiene la característica específica en la predicción del modelo, siendo de color rojo las que tienen una mayor relevancia, y de color azul las que tienen una menor relevancia.

De esta manera, los servicios seleccionados para la “Ruta Quirúrgica” a través de estos gráficos, son aquellos servicios que, al presentarse positivamente, también tienen un gran impacto en la predicción. Estos serían aquellos servicios en que los puntos rojos se encuentran al lado derecho de la vertical en cero.

Para esta sección, es relevante mencionar que los servicios de “Evaluación Pre Quirúrgica de Enfermería”, “Examen de Sangre/Orina/Otros Adulto” y “Electrocardiograma de Reposo” se realizan necesariamente a todos los pacientes que se someten a una cirugía, de forma preventiva para verificar que el paciente no presentará complicaciones, fuera de las estipuladas, durante la cirugía. Los gráficos de *Shap Values* para todas las clases del último modelo entrenado se presentan a continuación.

3.4.1. Definición de Ruta Quirúrgica para Cirugía Bariátrica

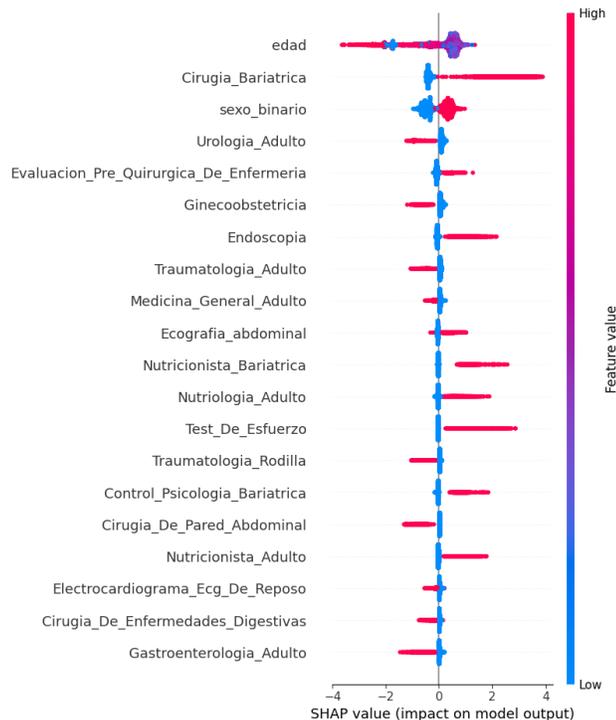


Figura 3.9: Gráfico de resumen Shap para Cirugía Bariátrica. Elaboración propia.

En la figura 3.9, que describe las características más relevantes para “Cirugía Bariátrica”, se puede observar que los servicios con mayor impacto positivo para la predicción son:

1. Cirugía Bariátrica
2. Evaluación Pre Quirúrgica de Enfermería
3. Endoscopia
4. Ecografía Abdominal
5. Nutricionista Bariátrica
6. Nutriología Adulto
7. Test de Esfuerzo
8. Control Psicología Bariátrica
9. Nutricionista Adulto

En general, todos los servicios mencionados se relacionan directamente con la Cirugía Bariátrica, considerando que este procedimiento implica un cambio significativo en el estilo de vida del paciente, y requiere una extensa preparación que puede durar varios meses.

En este sentido, dentro del listado se incluyen servicios que corresponden al proceso quirúrgico, como “Cirugía Bariátrica”, “Endoscopia”, y “Ecografía Abdominal”, que son servicios diagnósticos. Además, se incluyen servicios enfocados en la preparación del paciente, como lo son “Nutricionista Bariátrica/Adulto”, “Nutriología Adulto”, “Test de esfuerzo” y “Control Psicología Bariátrica”, los cuales buscan mejorar ciertos hábitos del paciente para asegurar que la cirugía sea lo más favorable posible.

3.4.2. Definición de Ruta Quirúrgica para Cirugía Urológica

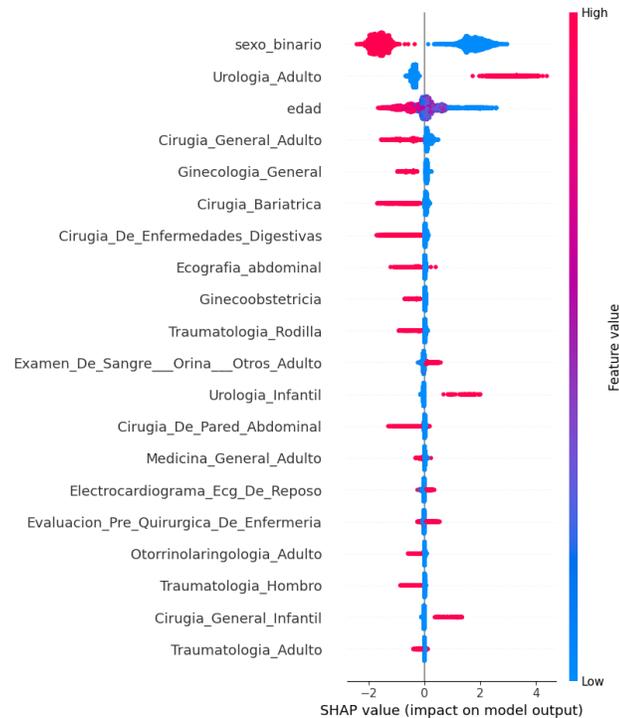


Figura 3.10: Gráfico de resumen Shap para Cirugía Urológica. Elaboración propia.

En la figura 3.10, que describe las características más relevantes para “Cirugía Urológica”, se puede observar que los servicios con mayor impacto positivo son:

1. Urología Adulto
2. Urología Infantil
3. Cirugía General Infantil
4. Examen De Sangre/Orina/Otros Adulto
5. Evaluación Pre Quirúrgica de Enfermería
6. Electrocardiograma de Reposo

Los servicios de la lista para Cirugía Urológica son pocos y se diferencian entre sí, principalmente por la edad del paciente que lo asiste. Esto es comprensible, ya que una de las cirugías, la Circuncisión, puede realizarse en cualquier etapa de la vida del hombre, mientras que la otra, la Vasectomía, se realiza únicamente en adultos.

De esta manera, los servicios incluidos en este listado son específicamente servicios urológicos, y aquellos que los pacientes deben someterse antes de cualquier intervención quirúrgica.

3.4.3. Definición de Ruta Quirúrgica para Colectectomía

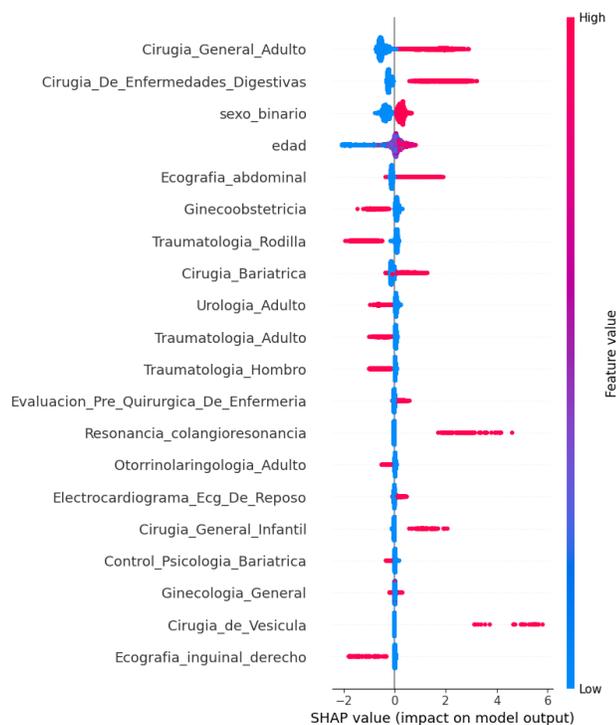


Figura 3.11: Gráfico de resumen Shap para Colectectomía. Elaboración propia.

En la figura 3.11, que describe las características más relevantes para “Colecistectomía”, se puede observar que los servicios con mayor impacto positivo son:

1. Cirugía General Adulto
2. Cirugía de Enfermedades Digestivas
3. Ecografía Abdominal
4. Cirugía Bariátrica
5. Resonancia Colangioreonancia
6. Cirugía General Infantil
7. Cirugía de Vesícula
8. Electrocardiograma de Reposo
9. Evaluación Pre Quirúrgica de Enfermería

Los servicios de la lista para Colecistectomía corresponden, en su mayoría, a distintos servicios diagnósticos de cirugía, incluyendo algunos que incluso no se relacionan directamente con esta intervención específica. Esto se debe a que hay médicos cirujanos que atienden múltiples servicios, por lo que en ciertas situaciones, atienden a pacientes de Colecistectomía en el servicio de “Cirugía Bariátrica”, por ejemplo.

Además, se incluye “Resonancia Colangioreonancia”, un examen no invasivo que permite evaluar, entre otros órganos, la vesícula biliar.

3.4.4. Definición de Ruta Quirúrgica para Hernia Abdominal

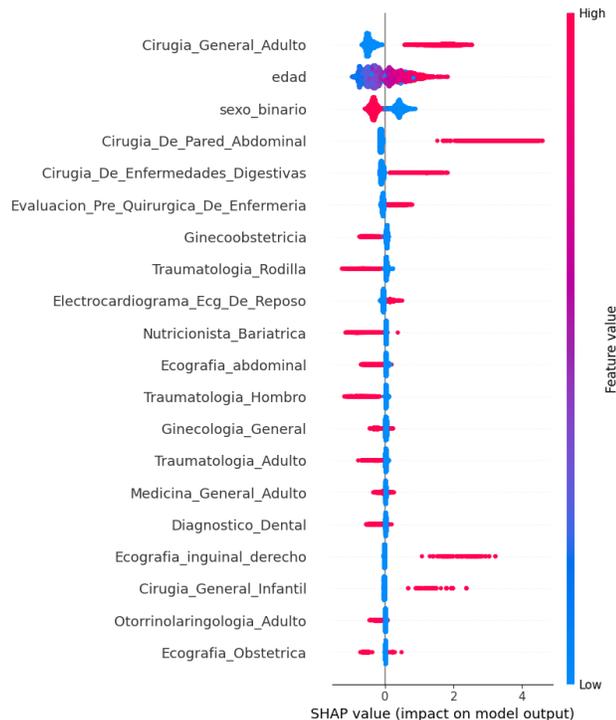


Figura 3.12: Gráfico de resumen Shap para Hernia Abdominal. Elaboración propia.

En la figura 3.12, que describe las características más relevantes para “Hernia Abdominal”, se puede observar que los servicios con mayor impacto positivo son:

1. Cirugía General Adulto
2. Cirugía de Pared Abdominal
3. Cirugía de Enfermedades Digestivas
4. Evaluación Pre Quirúrgica de Enfermería
5. Electrocardiograma de Reposo
6. Ecografía Inguinal Derecho
7. Cirugía General Infantil

Los servicios de la lista para Hernia Abdominal corresponden a aquellos que permiten diagnosticar la necesidad de extraer la hernia del abdomen del paciente. En particular, se incluyen cuatro servicios de cirugía, lo cual puede deberse a que los médicos cirujanos atienden en más de un servicio específico, similar a lo que ocurre en la Colectectomía.

Además, se incluye el servicio de “Ecografía Inguinal”, que se encarga específicamente de detectar la presencia de hernias en la zona inguinal del cuerpo.

3.4.5. Definición de Ruta Quirúrgica para Meniscectomía

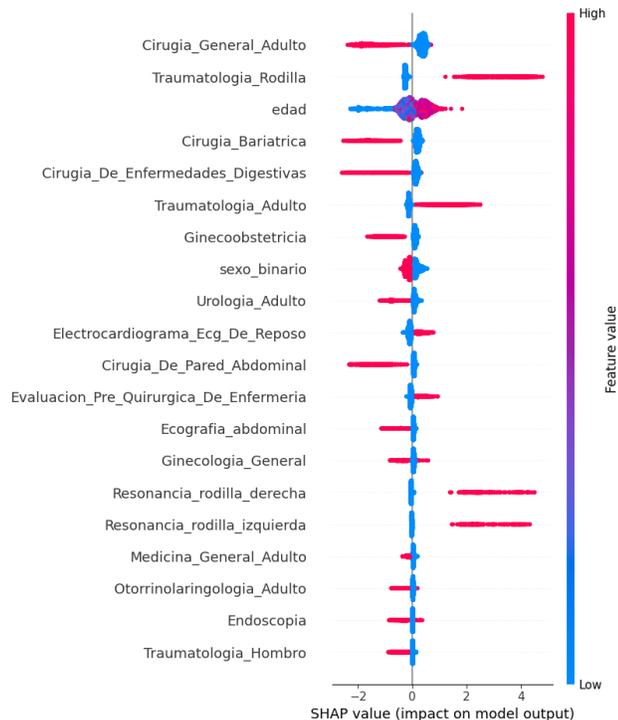


Figura 3.13: Gráfico de resumen Shap para Meniscectomía. Elaboración propia.

En la figura 3.13 que describe las características más relevantes para “Meniscectomía”, se puede observar que los servicios con mayor impacto positivo son:

1. Traumatología Rodilla
2. Traumatología Adulto
3. Electrocardiograma de Reposo
4. Evaluación Pre Quirúrgica de Enfermería
5. Resonancia de Rodilla Derecha
6. Resonancia de Rodilla Izquierda

Los servicios de la lista para Meniscectomía, son aquellos que se encargan de tratar una lesión traumática en la rodilla de una persona. Además, se incluye la Resonancia de Rodilla de ambos lados como servicio diagnóstico de la lesión y su gravedad.

3.4.6. Definición de Ruta Quirúrgica para Parto

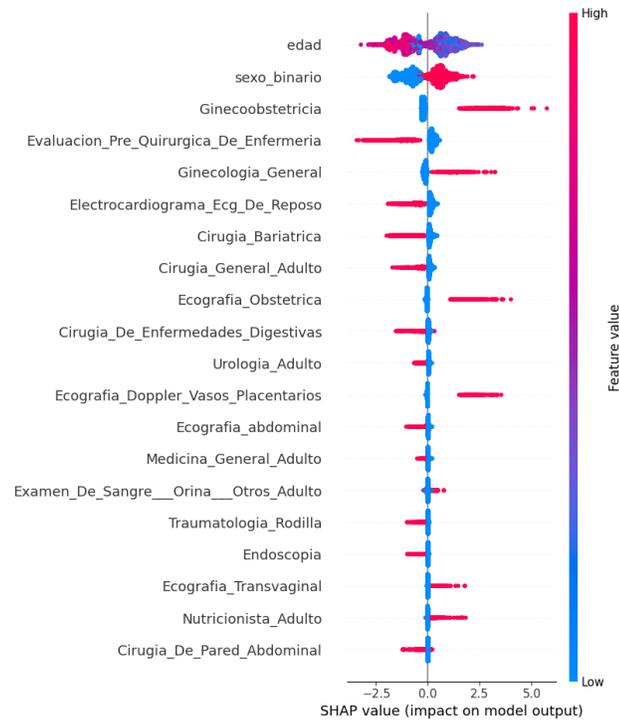


Figura 3.14: Gráfico de resumen Shap para Parto. Elaboración propia.

En la figura 3.14, que describe las características más relevantes para “Parto”, se puede observar que los servicios con mayor impacto positivo son:

1. Ginecoobstetricia
2. Ginecología General
3. Ecografía Obstétrica
4. Ecografía Doppler Vasos Placentarios
5. Examen de Sangre/Orina/Otros Adulto
6. Ecografía Transvaginal
7. Nutricionista Adulto

Los servicios de la lista para Parto, se encargan de supervisar el proceso prenatal de una mujer embarazada. En particular, todos los servicios incluidos corresponden a exámenes que permiten evaluar el estado del futuro bebé desde diferentes perspectivas, garantizando así que el parto se desarrolle de la mejor manera posible.

3.4.7. Definición de Ruta Quirúrgica para Rinoplastía

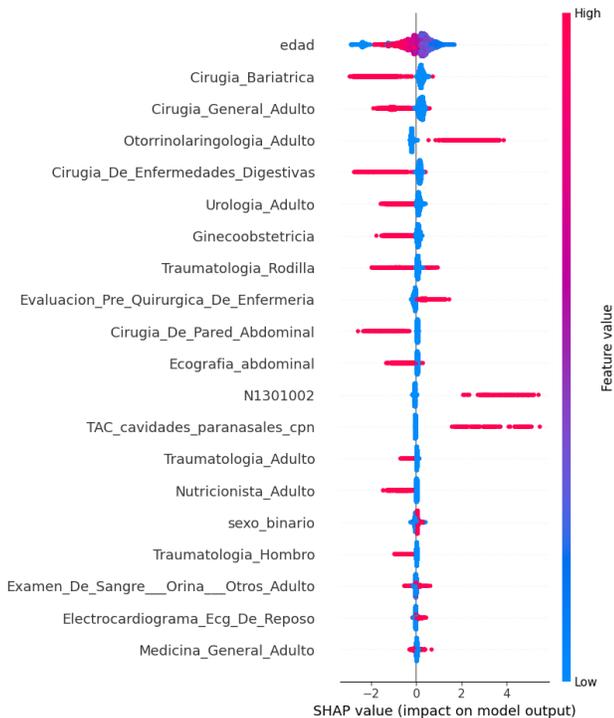


Figura 3.15: Gráfico de resumen Shap para Rinoplastía. Elaboración propia.

En la figura 3.15, que describe las características más relevantes para “Rinoplastía”, se puede observar que los servicios con mayor impacto positivo son:

1. Otorrinolaringología Adulto
2. Evaluación Pre Quirúrgica de Enfermería
3. TAC Cavidades Paranasales

Los servicios de la lista para Rinoplastía, se relacionan directamente con el tipo de médico especialista del área nasal, incluyendo un examen que permite al médico identificar posibles anomalías que podrían tenerse en cuenta, o incluso corregirse durante la cirugía.

3.4.8. Definición de Ruta Quirúrgica para Ruptura Manguito Rotador

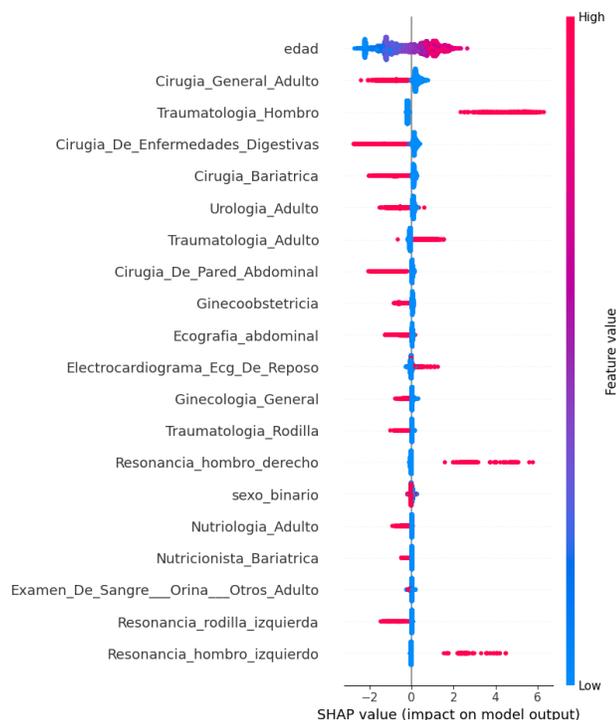


Figura 3.16: Gráfico de resumen Shap para Ruptura Manguito Rotador. Elaboración propia.

En la figura 3.16, que describe las características más relevantes para “Ruptura Manguito Rotador”, se puede observar que los servicios con mayor impacto positivo son:

1. Traumatología Hombro
2. Traumatología Adulto
3. Electrocardiograma de Reposo
4. Resonancia Hombro Derecho
5. Resonancia Hombro Izquierdo

Los servicios seleccionados son análogos a los de Meniscectomía, pero orientados a posibles traumas en los hombros del paciente. En este, se presenta inicialmente una consulta con un médico especialista, junto con un examen que permite diagnosticar la ruptura del manguito rotador.

A continuación se presenta una tabla resumen con las “Rutas Quirúrgicas” de cada una de las cirugías bajo estudio:

Tabla 3.20: Servicios de Ruta Quirúrgica por cirugía.

Cirugía	Servicios
Cirugía Bariátrica	<ul style="list-style-type: none"> • Cirugía Bariátrica • Evaluación Pre Quirúrgica de Enfermería • Endoscopia • Ecografía Abdominal • Nutricionista Bariátrica • Nutriología Adulto • Test de Esfuerzo • Control Psicología Bariátrica • Nutricionista Adulto
Cirugía Urológica	<ul style="list-style-type: none"> • Urología Adulto • Urología Infantil • Cirugía General Infantil • Examen de Sangre Orina Otros Adulto • Evaluación Pre Quirúrgica de Enfermería • Electrocardiograma de Reposo
Colecistectomía	<ul style="list-style-type: none"> • Cirugía General Adulto • Cirugía de Enfermedades Digestivas • Ecografía Abdominal • Cirugía Bariátrica • Resonancia Colangioresonancia • Cirugía General Infantil • Cirugía de Vesícula • Electrocardiograma de Reposo • Evaluación Pre Quirúrgica de Enfermería
Hernia Abdominal	<ul style="list-style-type: none"> • Cirugía General Adulto • Cirugía de Pared Abdominal • Cirugía de Enfermedades Digestivas • Evaluación Pre Quirúrgica de Enfermería • Electrocardiograma de Reposo • Ecografía Inguinal Derecho • Cirugía General Infantil

Meniscectomía	<ul style="list-style-type: none"> • Traumatología Rodilla • Traumatología Adulto • Electrocardiograma de Reposo • Evaluación Pre Quirúrgica de Enfermería • Resonancia Rodilla Derecha • Resonancia Rodilla Izquierda
Parto	<ul style="list-style-type: none"> • Ginecoobstetricia • Ginegología General • Ecografía Obstétrica • Ecografía Doppler Vasos Placentarios • Examen de Sangre Orina Otros • Ecografía Transvaginal • Nutricionista Adulto
Rinoplastía	<ul style="list-style-type: none"> • Otorrinolaringología Adulto • Evaluación Pre Quirúrgica de Enfermería • TAC Cavidades Paranasales
Ruptura Manguito Rotador	<ul style="list-style-type: none"> • Traumatología Hombro • Traumatología Adulto • Electrocardiograma de Reposo • Resonancia hombro Derecho • Resonancia hombro Izquierdo

Capítulo 4

Prueba de Campo

Para poner a prueba del modelo, se propone la elección de una cirugía en particular, con el fin de poder analizar los servicios de su “Ruta Quirúrgica” y personalizar el accionar en función de la cirugía en particular, de manera que el piloto se pueda enfocar en ella.

De esta forma, se plantea la creación de un piloto que consista en la obtención de los RUTs entregados por el modelo, para la posterior aplicación de un acompañamiento a los pacientes seleccionados a través de los ejecutivos de call center. De esta manera, se planea validar que la predicción del modelo sea correcta, para posteriormente agendar a aquellos que necesitan servicios adicionales porque aún no han completado los requisitos para su obtención de presupuesto, y orientar a aquellos que ya han recibido todos los servicios necesarios para que reciban su presupuesto.

Cabe destacar que para el prestador de salud, el objetivo frente a posibles pacientes quirúrgicos, es que estos puedan recibir el presupuesto de su cirugía, ya que una vez conociendo este valor, es el paciente quien decide si se realiza o no la cirugía, dados otros factores como el financiamiento, convenios con previsión de salud, etc.

Para emitir este presupuesto al paciente, se deben cumplir ciertos requisitos. Entre ellos está el cumplir con la asistencia a un cierto listado de servicios (Ruta Quirúrgica) y estar en posesión de una orden médica válida para la cirugía en cuestión.

4.1. Cirugía elegida e implementación

La elección de una única cirugía para la realización del piloto considera diferentes aspectos. En primer lugar, se descarta Cirugía Bariátrica debido a que esta ya se encuentra en un flujo de acompañamiento por parte del prestador, entonces si bien es de las clases que mejor se predicen, se prefiere no intervenir con el trabajo en proceso por otros equipos del prestador.

Luego, se descartan Rinoplastía, Ruptura Manguito Rotador y Meniscectomía, debido a la baja cantidad de datos disponibles relacionados a estas cirugías. Esto se traduce en que la frecuencia de realización de las cirugías mencionadas es muy baja, haciendo que no sean la elección óptima para el despliegue de un piloto.

Posteriormente, se descarta Parto, ya que, dada la naturaleza del mismo, deben transcurrir aproximadamente 9 meses desde la primera atención hasta el momento de la cirugía, la cual, además no se considera como una cirugía “tradicional”. Por otro lado, se espera que el piloto se pueda realizar en un espacio acotado de tiempo, siendo 9 meses un periodo muy largo para la implementación de este piloto.

También se descarta Hernia Abdominal. Esto se debe a que es la cirugía que peor predice el modelo, obteniendo los peores resultados de las métricas en la mayoría de los modelos entrenados.

Finalmente, entre Cirugía Urológica y Colectomía, se decide por esta última, debido a que es más sencilla que la primera, en el sentido de que Cirugía Urológica contiene dos tipos de intervenciones diferentes dentro de la misma clase, y además, la Colectomía es la cirugía más frecuente que se realiza en las instalaciones del prestador.

Adicionalmente, como se puede ver en la figura 4.1, cerca del 50% de los pacientes que se realizaron esta cirugía entre enero de 2023 y marzo de 2024, se tardaron menos de 50 días entre su primera cita y la realización de su cirugía. Esto hace que la Colectomía sea la mejor opción para la realización de un piloto en el presente trabajo de tesis.

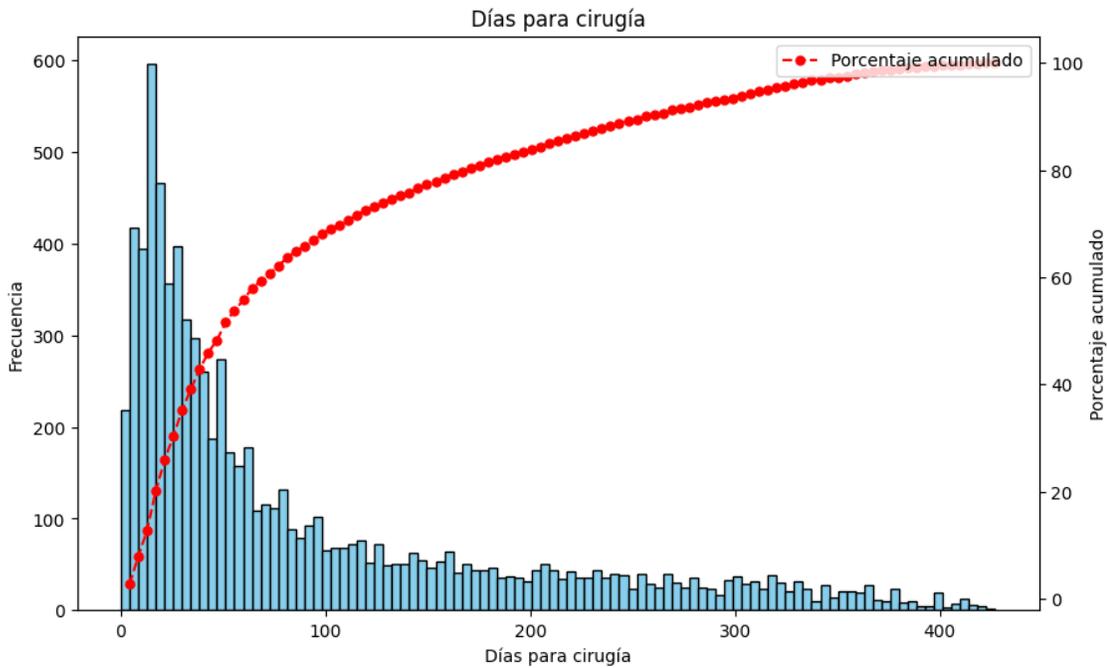


Figura 4.1: Gráfico histograma y acumulado de la cantidad de días que se tarda un paciente desde su primera cita hasta llegar a su cirugía de Colectomía. Elaboración propia.

4.2. Obtención de los RUTs a partir del modelo

Una vez elegida la cirugía a trabajar, y para llevar el modelo a la práctica, es necesario obtener los datos de los pacientes que se requieren para poder introducirlos en el mismo, y obtener las predicciones sobre las cirugías. Sin embargo, se debe definir un intervalo de tiempo hacia el pasado para extraer los servicios atendidos por cada paciente, pues no se puede predecir la cirugía que requiere un paciente en la actualidad, con la totalidad de servicios a los que ha atendido ni con los servicios que asistió cinco años en el pasado.

Es por esto que, dado el contexto de la Colectomía, pues puede ser distinto para otra cirugía, se decide tomar datos de 30, 60 y 90 días en el pasado. Además se define un umbral variante de probabilidad, entregado por la función *predict_proba*¹ de la librería *Scikit-Learn* de Python, para clasificaciones con probabilidades mayores a 50 %, 60 %, 70 %, 80 % y 90 %, para simular las predicciones del modelo en un caso más “realista”, y elegir cuál es la combinación que entrega mejores resultados en función de los pacientes que, dentro de ese mismo periodo de tiempo, se encuentran dentro del proceso de solicitud de presupuesto.

De esta manera, se presenta la tabla 4.1. En esta, se pueden ver los resultados para los RUTs que obtienen una probabilidad mayor a 50 % para datos de 30, 60 y 90 días previos.

Tabla 4.1: Resultados simulación modelo por días ($P(x) > 0.5$).

Nº de días	Nº de pacientes	Nº total de predicciones	Nº de predicciones con presupuesto emitido	Nº de predicciones con presupuesto emitido correctamente clasificado	Porcentaje de predicciones realizadas correctamente	Pacientes restantes por abordar
30	471,431	4,808	734	408	55.59%	4,074
60	748,532	7,832	1,404	750	53.42%	6,428
90	975,074	10,578	2,057	1073	52.16%	8,521

Aquí, se puede notar que a medida que aumenta la cantidad de días, también aumenta el número de pacientes que se ingresan al modelo, y en consecuencia, también la cantidad de pacientes clasificados hacia la cirugía en estudio.

Adicionalmente, del total de pacientes clasificados, se muestra el número que, además, dentro de ese mismo periodo de tiempo (30, 60 y 90 días) ya poseen su presupuesto emitido. Del total de presupuestos emitidos en cada periodo, se calcula el porcentaje de estos que se encuentran correctamente clasificados. Siendo el mejor de estos, el correspondiente a 30 días de datos con un 55.59 % de las predicciones correctamente realizadas.

¹ *Predict_proba*, es una función de *Scikit Learn* que se utiliza para predecir las probabilidades de todas las clases para las instancias introducidas en un modelo de aprendizaje de máquinas. A diferencia de la función *predict*, que entrega la clase predicha más probable, *predict_proba* entrega una matriz con las probabilidades estimadas de pertenencia a cada clase.

Aquí, es importante resaltar que este último número difiere de lo entregado por las métricas del modelo. Esto se debe a que, en la realidad, existen más cirugías que las incluidas en el entrenamiento del mismo. De esta manera, además de equivocarse lo esperable en función de las métricas entregadas, también está el factor del error por sesgo de omisión, el cual no se puede evitar a menos que se entrene un nuevo modelo con la cantidad de datos suficientes para cada cirugía.

De todas maneras, el poder encontrar pacientes en camino a una cirugía de Colectomía con una probabilidad mayor al 50 % es valioso para el prestador de salud. Además, de que, para un trabajo futuro, se podría utilizar los distintos umbrales probabilísticos presentados, y establecer un acompañamiento diferente para cada intervalo considerando que en algunos se tiene mayor seguridad que el paciente efectivamente se realizará la cirugía.

De esta manera, como este umbral (50 %), es el que entrega la mayor cantidad de pacientes para acompañar, se decide continuar el despliegue con la configuración de 30 días presente en la tabla 4.1.

De todas formas, las tablas con las simulaciones más estrictas en el umbral de probabilidad (60 %, 70 %, 80 % y 90 %) se encuentran en el anexo C.1, C.2, C.3 y C.4.

Continuando con esta simulación, se realiza un análisis de las búsquedas de servicios realizadas por los pacientes seleccionados en la página web del prestador. En particular, esto se realiza en dos oportunidades, para los primeros 7 días siguientes a la fecha de la extracción de datos, y también para los primeros 14 días a partir de esta misma fecha. De esta manera, se obtiene que de los 4,808 pacientes seleccionados, en la semana siguiente a la obtención de datos, 701 pacientes buscó para agendar un servicio en la web, y de estos, 291 buscó alguno de los servicios presentes en la Ruta Quirúrgica de la Colectomía.

Del mismo modo, en las dos semanas siguientes a la obtención de datos, 1,071 pacientes buscó para agendar un servicio en la web, y de estos, 472 buscó alguno de los servicios presentes en la Ruta Quirúrgica de la Colectomía. De esta manera, el número de búsquedas por cada uno de los servicios son los que se pueden ver en la tabla 4.2.

Tabla 4.2: Búsquedas de servicios semanas siguientes a obtención de datos.

Nombre Servicio	N° de Búsquedas	
	7 días	14 días
Cirugía General Adulto	312	532
Cirugía de Enfermedades Digestivas	83	124
Cirugía de Vesícula	64	113
Examen de Sangre / Orina / Otros Adulto	59	125
Cirugía Bariátrica	46	81
Electrocardiograma de Reposo	24	37
Ecografía Abdominal	17	50
Cirugía General Infantil	2	4
Resonancia Colangiografía	1	1

En estas, se puede notar que, los pacientes que están buscando atenderse con el prestador, corresponden a un bajo porcentaje del total seleccionado por el modelo, en una y dos semanas después de la obtención de datos. De esto se infieren dos casos. Primero, que la mayoría de los pacientes no se están atendiendo siempre con el prestador, lo cual presenta una oportunidad para buscarlos y que se atiendan en el mismo. Y segundo, que los pacientes que buscan y se agendan por su cuenta, podrían no llegar a necesitar un acompañamiento.

Con respecto a la cantidad de búsquedas por cada servicio en la tabla 4.2, en general, se mantiene el orden de los servicios más buscados entre ambos intervalos de tiempo, siendo el más frecuente Cirugía General Adulto. Aquí, es importante recalcar que existen médicos que atienden varios servicios. Es decir, un mismo médico podría tener consultas de Cirugía Bariátrica y de Cirugía de Enfermedades Digestivas, por ejemplo. Esto permite entender por qué aparece Cirugía Bariátrica como buen predictor de Colecistectomía, cuando a partir del mismo nombre del servicio se puede notar que su objetivo es la derivación hacia otro tipo de cirugía.

Por otro lado, con respecto a los servicios presentes en la Ruta Quirúrgica, se agrega el servicio de “Examen de Sangre / Orina / Otros Adulto”. Esto se debe a que al validar los servicios en conjunto con la contraparte, esta menciona que también se debe incluir este último servicio, puesto que es un requisito para que el paciente obtenga el presupuesto de la operación.

Además, se decide eliminar “Cirugía General Infantil” del listado, debido a que, en conjunto con el prestador, se decide incluir dentro del piloto solamente a pacientes mayores de edad.

4.3. Situación actual de acompañamiento quirúrgico

Actualmente, el prestador no presenta un proceso en el cual se pueda encontrar tempranamente a aquellos pacientes que presentan una mayor predisposición quirúrgica. En particular, el principal proceso de acompañamiento que se realiza en el ámbito quirúrgico, nace a partir de las solicitudes de presupuestos que se realizan a través de la página web del prestador. En esta, existe un formulario en el cual el paciente que quiere cotizar su cirugía debe ingresar su nombre, su RUT, su previsión de salud, su edad, el nombre de la cirugía a cotizar, el nombre de la clínica en la cual se quiere realizar la operación, el nombre de la clínica en la cual atiende el cirujano que le indicó la cirugía y finalmente se consulta si se cuenta, o no, con la orden médica para la cirugía.

A partir de este formulario, se define un proceso representado en el BPMN de la figura 4.2. En ella, se puede notar que el acompañamiento se divide en dos partes, aquellos pacientes que adjuntan su orden médica (COM), y aquellos que no (SOM).

Del primer grupo, se debe validar la orden médica, la cual en caso de ser correcta, se envía a la clínica correspondiente al caso, para que se emita el presupuesto y se le envíe mediante correo electrónico al paciente. En caso de que la orden médica no sea válida, se introduce al paciente al conjunto sin orden médica. Aquí, se contacta al paciente mediante llamado telefónico. Si el paciente no contesta en tres oportunidades, entonces se le saca del proceso.

Si contesta, pero comunica no querer el acompañamiento, también se le saca del proceso. Si contesta, y reserva una hora ambulatoria, entonces se le hace un seguimiento hasta que obtenga su orden médica y se le pueda ingresar al conjunto COM. En caso de no obtener su orden médica, también se le saca del proceso.

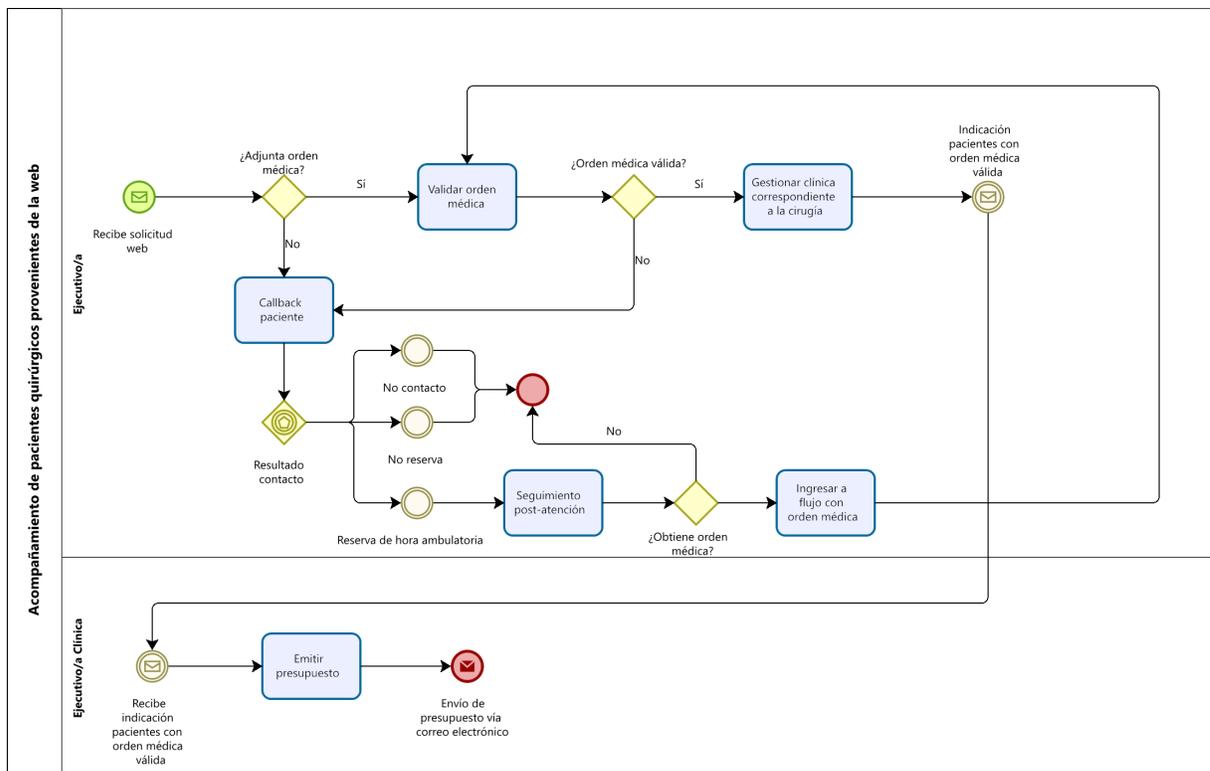


Figura 4.2: BPMN situación AS-IS del acompañamiento a pacientes que ingresan solicitud de presupuesto web. Elaboración propia.

4.4. Situación TO-BE de acompañamiento quirúrgico

Para la situación TO-BE del proceso, se propone la creación de un proceso que pueda trabajar en paralelo al presentado en la situación AS-IS, de manera que este implique el menor esfuerzo posible por parte del prestador. Esto debido a que, como se puede ver en la figura 4.3, a partir de los RUTs que entrega el modelo, se obtienen pacientes que ya se encuentran en el conjunto SOM y COM, pero falta incluir a aquellos pacientes que no se encuentran en ninguno de estos, impulsándolos a entrar en este flujo.

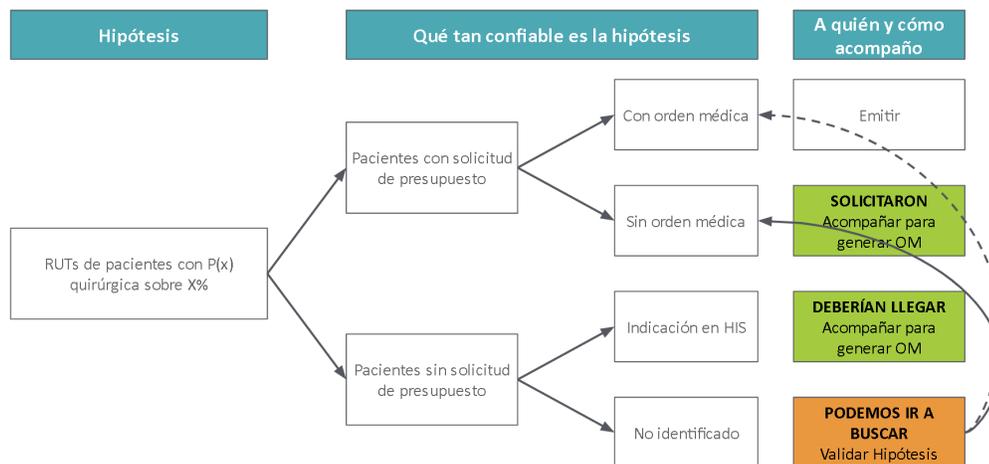


Figura 4.3: Diagrama validación de hipótesis. Elaboración propia.

De esta manera, se plantea un flujo para poder introducir a estos pacientes a los conjuntos COM y SOM. Sin embargo, antes de poder realizar este acompañamiento, es necesario restar a aquellos pacientes seleccionados que no se están atendiendo con un médico staff² del prestador.

Así, al filtrar por médicos staff, se asegura que el acompañamiento a realizar sea para pacientes que efectivamente se intervendrán en las clínicas del prestador, y el médico no se los llevará a otro centro de su preferencia para la cirugía. Con esto, y a partir de la simulación realizada en la sección 4.2, se realiza el gráfico de cascada de la figura 4.4.

² Los médicos staff son aquellos que el prestador está seguro que atienden consultas, y realizan las cirugías dentro de las mismas clínicas del prestador.

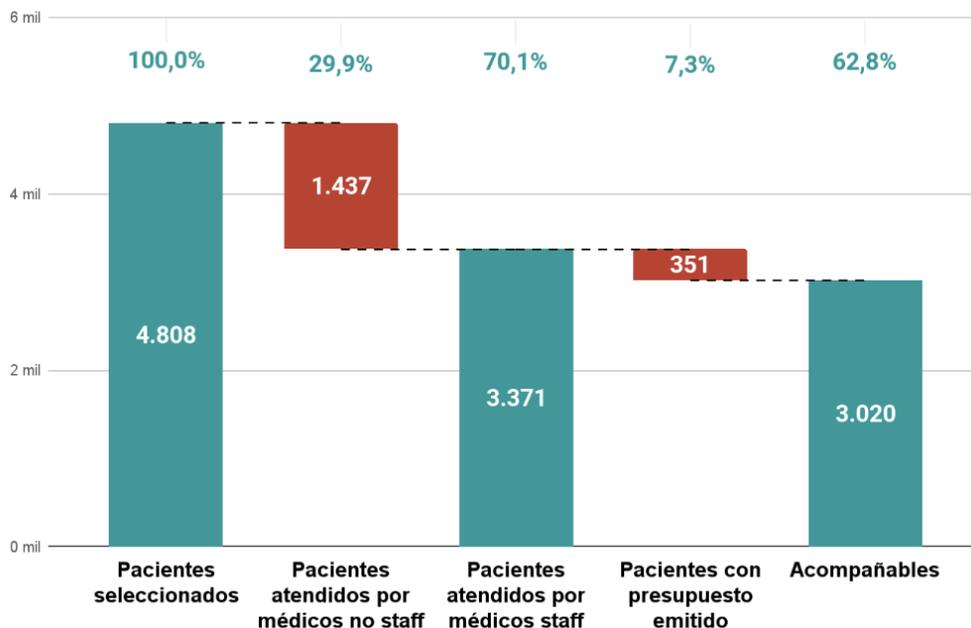


Figura 4.4: Cascada de pacientes entregados por el modelo. Elaboración propia.

En ella, se puede ver que a partir de los 4,808 pacientes seleccionados por el modelo, se resta un 29.9% de ellos por no estar siendo atendidos por médicos staff, quedando un total de 3,371 pacientes por acompañar. A este número se le termina restando 351 pacientes que ya se encuentran con su presupuesto emitido, quedando así un total de 3,020 pacientes por acompañar, lo cual es un 62.8% del total inicial de pacientes entregados por el modelo.

Dado lo anterior, se propone una situación TO-BE, en la cual, a partir de los pacientes acompañables de la figura, un ejecutivo de call center realice un llamado en el cual se pueda validar la hipótesis de que efectivamente ese paciente se encuentra en camino a la realización de una Colectectomía. Para esto, es sumamente importante que al momento del llamado, se le aclare al paciente que no se le está haciendo un diagnóstico, si no que solamente los médicos están calificados para hacerlo.

Junto con esto, se le debe preguntar al paciente si efectivamente un médico le ha comentado que debe realizarse una Colectectomía. Una vez validada la hipótesis, se le debe consultar al paciente si ya posee su orden médica para la cirugía, de manera que, en caso positivo, se le derive al conjunto COM, y en caso negativo, se le pueda agendar su siguiente servicio, ingresándolo al conjunto SOM para la continuación de su acompañamiento. El detalle de cómo funcionaría este nuevo proceso se puede ver en la figura 4.5.

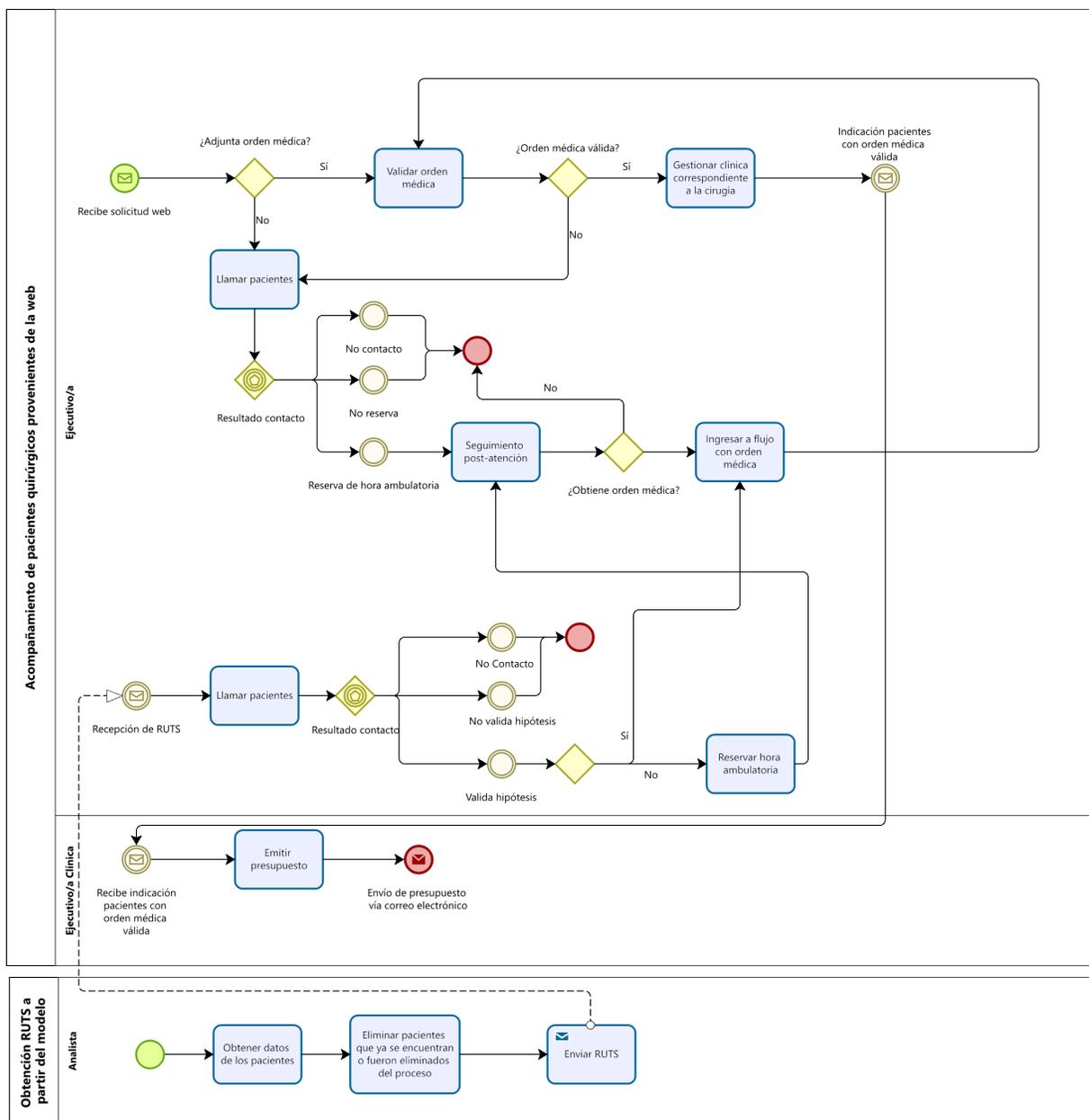


Figura 4.5: BPMN situación TO-BE del acompañamiento a pacientes que ingresan solicitud de presupuesto web e ingresados por el modelo. Elaboración propia.

4.4.1. Formulación piloto para validación de hipótesis

Como una primera aproximación a la situación TO-BE presentada anteriormente, se plantea un piloto que tiene como principal objetivo, validar el que un médico le haya indicado a los pacientes seleccionados, que deben realizarse una Colectectomía, y en caso positivo, ayudarles a agendar su siguiente paso a seguir. Esto se realiza a través de un llamado por parte de un ejecutivo/a de *Call Center*, capacitado/a para este tipo de gestiones.

Para este llamado, es relevante para el prestador que al momento de validar la hipótesis con el paciente, se le indique que no se le está diagnosticando una enfermedad, ya que los médicos son los únicos con las aptitudes necesarias para hacerlo. De esta manera, el diálogo entregado a los ejecutivos de *Call Center* es meticulosamente creado para evitar este problema.

En primer lugar, el ejecutivo/a tiene a su disposición dos herramientas, la información del paciente, y el flujo del diálogo a seguir. El primero se presenta de forma tabular, como se puede ver a continuación en las tablas 4.3 y 4.4.

Tabla 4.3: Información del paciente entregada a ejecutivo/a de *Call Center*
[1/2]

Rut paciente	Nombre paciente	Teléfono de contacto	Correo electrónico	Plan de salud	Fecha última atención
12.345.678-9	Tomás Aguirre	+56912345678	tomas@correo.cl	Fonasa	17-07-2021

Tabla 4.4: Información del paciente entregada a ejecutivo/a de *Call Center*
[2/2]

Servicios agendados	Servicios por agendar	Cirugía
Cirugía de Vesícula, Electrocardiograma (Ecg) De Reposo, Resonancia colangiorensonancia, Cirugía De Enfermedades Digestivas, Ecografía abdominal	Examen De Sangre / Orina / Otros Adulto, Cirugía General Adulto, Cirugía Bariátrica	Colecistectomía

En ellas, se puede ver datos de identificación del paciente en las columnas “Rut paciente” y “Nombre paciente”. Luego, sus datos de contacto en las columnas “Teléfono de contacto” y “Correo electrónico”. Además, se encuentra su previsión en la columna “Plan de salud”, la fecha de la última vez que se atendió en algún servicio ofrecido por el prestador en “Fecha última cita”, los servicios de la Ruta Quirúrgica de Colecistectomía que el paciente ya atendió en los últimos 30 días, en “Servicios agendados”, los servicios de la Ruta Quirúrgica de Colecistectomía que el paciente no ha atendido en los últimos 30 días, en la columna “Servicios por agendar”, y finalmente el nombre de la cirugía por validar en la columna “Cirugía”.

Luego, el flujo del diálogo a seguir por el ejecutivo/a es el que se puede ver en la figura 4.6. En este, se puede notar la intención de validar el que los pacientes seleccionados se encuentren en el camino a la realización de una Colecistectomía, y en caso positivo, agendarle el siguiente paso a seguir según lo indicado por el mismo paciente.

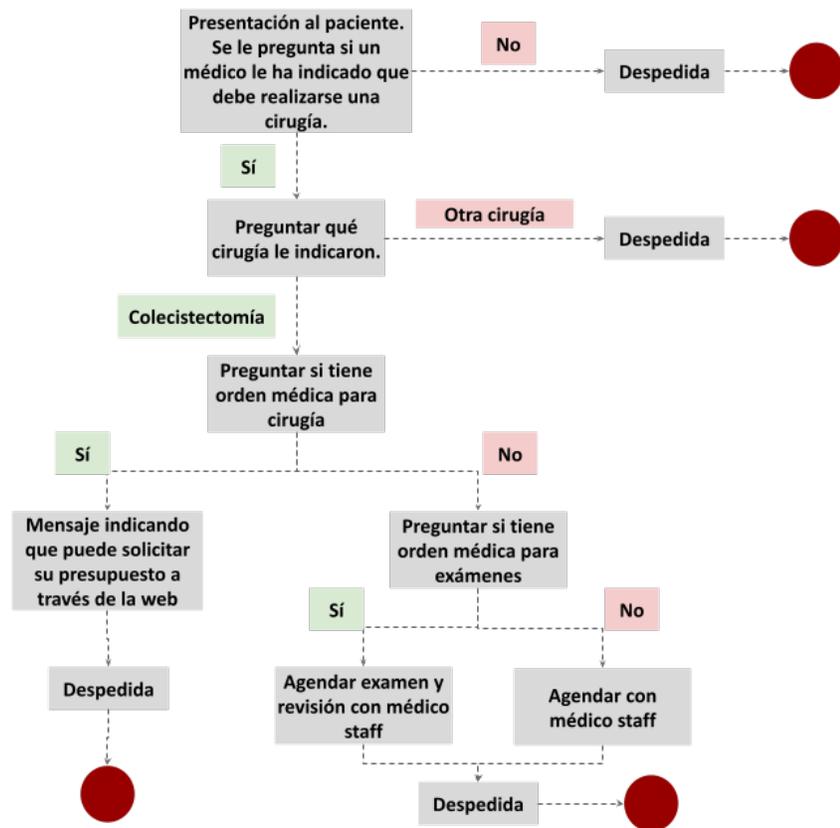


Figura 4.6: Flujo para ejecutivo/a de *Call Center* en piloto. Elaboración propia.

Así, una vez finalizado el piloto, se obtiene un conjunto de datos con los resultados por paciente, con variables que informan sobre la fecha y la hora a la que se realizó el último llamado, y el resultado conclusión del mismo. En este sentido, los posibles resultados son:

1. No está en un viaje Quirúrgico
2. Ya se realizó la cirugía
3. No contesta
4. No interesado, ya se lo realizó con el prestador
5. En viaje Quirúrgico / Otra cirugía / No agenda
6. No interesado
7. Se corta la llamada
8. En viaje Quirúrgico / Otra cirugía / Agenda
9. En viaje Quirúrgico / Colecistectomía / No agenda
10. En viaje Quirúrgico / Otra cirugía / COM

11. En viaje Quirúrgico / Colectomía / COM
12. En viaje Quirúrgico / Colectomía / Agenda
13. Buzón de voz
14. Volver a llamar
15. Recado con tercero

4.5. Evaluación piloto

El piloto se inicia con un total de 2,527 pacientes que fueron seleccionados por el modelo, restando aquellos que no se atienden con médicos staff, y aquellos que ya tienen su presupuesto emitido, según lo anteriormente estipulado.

Con esto, se obtiene la siguiente tabla 4.5 con los resultados de la obtención de datos. En esta, del mismo modo que en la simulación realizada anteriormente, se puede notar que el número total de pacientes ingresados al modelo es de 469,659. De estos, el modelo predice para Colectomía a 4,529 pacientes, de los cuales 723 ya tenían una solicitud de presupuesto emitida, siendo 424 de Colectomía. Luego, una vez restados aquellos pacientes con presupuesto emitido, y que no se atienden con médicos staff, quedan 2,535 pacientes por llamar en el piloto. Se puede encontrar un gráfico de cascada con estos datos en el anexo C.1.

Tabla 4.5: Datos evaluados previos al piloto.

Nº total de pacientes	Nº de predicciones	Nº de predicciones con presupuesto emitido	Nº de predicciones con presupuesto emitido de Colectomía	Nº de pacientes por llamar
469,659	4,529	723	424	2,527

Luego, se presentan los resultados del piloto en la tabla 4.6.

Tabla 4.6: Resultados piloto

Código conclusión	Nº de pacientes	Porcentaje de pacientes
No está en un viaje Quirúrgico	776	30,71 %
No contesta	717	28,37 %
Ya se realizó Cirugía	533	21,09 %
En viaje Quirúrgico - Otra Cirugía - No Agenda	135	5,34 %
No interesado - No decide/ no quiere	99	3,92 %
En viaje Quirúrgico - Otra Cirugía - COM	79	3,13 %
Se corta la llamada	61	2,41 %
En viaje Quirúrgico - Colectectomía - No Agenda	39	1,54 %
En viaje Quirúrgico - Colectectomía - COM	33	1,31 %
En viaje Quirúrgico - Otra Cirugía - Agenda	23	0,91 %
Volver a llamar	18	0,71 %
Recado con tercero	7	0,28 %
En viaje Quirúrgico - Colectectomía - Agenda	4	0,16 %
Buzón de voz	2	0,08 %
No interesado - Ya se lo realizó con el prestador	1	0,04 %

A partir de esta tabla, se puede notar que el código de conclusión con mayor cantidad de pacientes, es el que indica que el paciente no se encuentra en un viaje quirúrgico, seguido por aquellos que no contestan al llamado realizado. Luego, se encuentra a pacientes que ya se realizaron la cirugía. En este último punto, si bien en un principio se descarta a aquellos pacientes con solicitud de presupuesto emitida, es relevante indicar que, desde la obtención de datos, hasta el inicio del piloto, transcurren alrededor de 12 días, y la duración del piloto es de 6 días. De esta manera, entre la obtención de datos y el último llamado pasan alrededor de 18 días, creando un desfase relevante entre la obtención y filtrado de los datos, y el inicio de los llamados, lo cual entrega espacio para que un porcentaje de los pacientes se realice sus cirugías, tanto dentro, como fuera de las instalaciones del prestador.

Posteriormente, se realiza una agrupación de las conclusiones, para la realización de un análisis de estos resultados. Estos se pueden ver en la tabla 4.7. La forma en que se agrupan las conclusiones se puede encontrar en el anexo C.5.

Tabla 4.7: Resultados piloto agrupados

Grupo	Nº de pacientes	Porcentaje de pacientes
Falla contacto	805	31.86 %
No Quirúrgico	776	30.71 %
Acompañamiento tardío	534	21.13 %
Quirúrgico	313	12.39 %
No interesado	99	3.92 %

A partir de esta última tabla, se infiere que hay un punto positivo y tres puntos principales por mejorar a partir del piloto realizado.

Lo positivo que se infiere de estos resultados, es que el 33 % de los pacientes contactados resultaron ser efectivamente quirúrgicos (Grupos de acompañamiento tardío y quirúrgico), lo cual a pesar de no ser un número cercano a lo esperado dadas las evaluaciones anteriores, de todas maneras corresponden a 847 pacientes, que si se les realiza un acompañamiento a tiempo, se convierten en potenciales pacientes que se pueden intervenir quirúrgicamente con el prestador. Esto incluso si no corresponden necesariamente a la cirugía en estudio, pues de todas maneras se les puede acompañar a los pacientes mal clasificados.

Para los puntos a mejorar, en primer lugar, es necesario realizar cambios para mejorar los contactos con el paciente. En este sentido, para un nuevo piloto, se podría plantear que el contacto con el paciente se realice a través de la aplicación “WhatsApp” mediante un *chatbot*, o se haga una mayor cantidad de intentos de llamados, en horarios diferentes, a aquellos que no contestan en la primera ocasión.

Por otro lado, es necesario que exista coordinación entre la obtención y filtrado de los datos, y el inicio del piloto, de manera que se reduzca la tasa de pacientes a los que se les aplica el acompañamiento de forma tardía. De esta manera, el caso ideal es que cada día se realice la obtención y el filtrado de pacientes, esperando que se pueda mejorar la correcta utilización de los recursos disponibles, no llamando a aquellos pacientes que se encuentran en una etapa más avanzada del proceso al que apunta la propuesta de acompañamiento.

Finalmente, si bien el modelo obtiene buenas métricas en su evaluación inicial, al momento de llevarlo a la realidad, su desempeño empeora debido al sesgo por omisión presentado anteriormente. En este sentido, el óptimo es poder incluir la mayor cantidad de cirugías posibles, de manera que solamente se excluyan aquellas que rara vez se realizan.

Capítulo 5

Conclusiones

En el presente trabajo de tesis, se desarrolla y presenta, el diseño de una herramienta que permite identificar y acompañar a aquellos pacientes que, según sus servicios médicos atendidos, tienen una mayor probabilidad de requerir una intervención quirúrgica. En particular, para la creación de la herramienta se utiliza un enfoque de aprendizaje automático supervisado, planteándolo como un problema de clasificación multiclase. Además, se presenta la implementación de un piloto que prueba esta herramienta, realizando un primer acercamiento a los pacientes seleccionados por la misma.

El modelo final mencionado, se entrena utilizando el algoritmo *LightGBM*, optimizado con *Optuna* para la selección de hiperparámetros, y evaluado mediante *Cross Validation*. Los resultados obtenidos muestran un *Precision macro* de 0.81, un *Recall macro* de 0.79 y un *F1-Score macro* de 0.8, prediciendo satisfactoriamente las nueve clases planteadas en su entrenamiento.

Junto con esto, se desarrolla la implementación de un piloto para acompañar a los pacientes identificados por el modelo en el proceso de realización de una Colectomía. Así, se plantea que a través de un llamado telefónico, se pueda validar la hipótesis postulada, y junto con esto, se acompañe al paciente para que reserve su siguiente servicio a asistir.

Una vez finalizado el piloto de acompañamiento, se obtienen los resultados de los llamados, donde falla el contacto en el 32 % de los casos, un 31 % son no quirúrgicos, se contacta tardíamente al 21 % de los pacientes, y el 12 % de los casos resulta ser quirúrgico.

De esta manera, se cumple el objetivo general enunciado como “Diseñar un modelo de aprendizaje automático, que permita identificar pacientes con una mayor predisposición quirúrgica, para acompañarlos y acercarlos al pabellón”. Además, se alcanzan los objetivos específicos (1-6) establecidos en la sección 1.4.

Así, la herramienta resulta relevante para el prestador, como una primera aproximación a la identificación temprana de pacientes con intención quirúrgica, mediante el uso de aprendizaje automático, antes de que el paciente busque cotizar su cirugía dentro de la institución. Además, al utilizar el planteamiento del modelo con los datos crudos, y el algoritmo de *LightGBM*, se hace posible que el prestador siga trabajando a futuro con el proyecto, dado que el entrenamiento con mejores resultados, también termina siendo el más eficiente y sencillo de trabajar.

Con esto, el prestador de salud tiene la oportunidad de mejorar la relación con sus pacientes quirúrgicos, personalizando el acompañamiento de cada cirugía, y de esta manera, ayudando a que el viaje quirúrgico de sus pacientes sea más sencillo e informado.

Con respecto al modelo seleccionado para la simulación y el piloto, se observa una disminución en el rendimiento al contrastarlo con datos reales, en comparación con la evaluación en la etapa de entrenamiento. Esto se debe principalmente a que el modelo está sesgado hacia las cirugías proporcionadas durante el entrenamiento, cuando en la práctica, existen más cirugías, que si bien son menos frecuentes, no dejan de depreciar el rendimiento del modelo. A pesar de esto, el modelo logra predecir correctamente alrededor del 33 % de los casos, lo cual si bien no es lo esperado inicialmente, de todas maneras son aproximadamente 800 pacientes que, de ser acompañados en el momento correcto, son relevantes para el prestador.

Para un eventual trabajo futuro, se plantea una evaluación exhaustiva de las cirugías a incluir dentro de los datos de entrenamiento, agrupando aquellas con nombres diferentes, pero que sean equivalentes o muy similares entre sí. Además, con acceso a mayor capacidad de cómputo y memoria, se puede plantear la inclusión de datos previos al año 2022, aumentando la cantidad de datos disponibles para el entrenamiento y evaluación del modelo.

Con los dos puntos anteriores, se podría maximizar el número de cirugías presentes en los datos de entrenamiento, reduciendo el sesgo del modelo, y excluyendo solamente aquellas cirugías con una muy baja representatividad, o en otras palabras, aquellas que rara vez se realizan. En este caso, es esperable que aumente la cantidad de datos, y junto con esto, el tiempo de procesamiento de la metodología propuesta. Es por esto que se propone, además, ahondar en la reducción de dimensionalidad, buscando mejorar los resultados al modificar los hiperparámetros de los métodos mencionados, o incluyendo nuevos métodos que logren una mejor aproximación.

Por otro lado, también se encuentra la posibilidad de incluir un conjunto específico de cirugías al entrenamiento del modelo, de manera tal que a pesar de no estar conteniendo todas las cirugías, de todas maneras se obtengan buenos resultados, dada la especificidad de estas cirugías.

En el mediano/largo plazo, se plantea diseñar un acompañamiento específico para cada tipo de cirugía, atendiendo las necesidades particulares de cada una. Además, se podría segmentar el acompañamiento según la probabilidad indicada por el modelo para cada paciente, considerando que un paciente con probabilidad entre 50 % y 60 %, tiene un camino recorrido distinto al de un paciente entre el 90 % y 100 %.

Finalmente, una vez teniendo un planteamiento de piloto más avanzado, se debe probar la efectividad de los acompañamientos tomando un grupo de control similar en características a los pacientes acompañados, y evaluando métricas relevantes para el prestador, como lo podrían ser la emisión de los presupuestos y la concreción de la cirugía específica.

En resumen, este trabajo de tesis demuestra la viabilidad y efectividad de una herramienta basada en aprendizaje automático, para identificar y acompañar a pacientes que probablemente requieren realizarse una intervención quirúrgica. La validación realizada en el piloto, sugiere que, de ser implementada correctamente en las operaciones diarias del prestador, se podría mejorar la relación con los pacientes, haciendo que sus viajes quirúrgicos tengan menos obstáculos, y ayudándoles a seguir la ruta correcta, para que tengan una llegada satisfactoria a su cirugía. Así, esta tesis cumple con los objetivos planteados, estableciendo una base sólida para la continuidad del desarrollo de esta herramienta, esperando que el prestador logre entregar una atención personalizada a cada uno de sus pacientes quirúrgicos y facilitando sus viajes de salud.

Bibliografía

- [1] Mundial, G. B., “World bank open data”, <https://datos.bancomundial.org/indicador/S.P.DYN.CDRT.IN?end=2021&start=1960&view=chart&year=1960>.
- [2] Superintendencia de Salud, “Cómo funciona el sistema de salud en Chile - cómo funciona el sistema de salud en Chile”, 2023, https://www.superdesalud.gob.cl/difusion/665/w3-article-17328.html#accordion_0.
- [3] Aguilar-Barrientos, R. y Velasco, C., “¿cómo se usan los pabellones quirúrgicos en Chile? radiografía al sector estatal de salud”, Documentos de trabajo, no. 1, 2022.
- [4] Instituto de Políticas Públicas en Salud, “Uso de pabellones quirúrgicos: hasta 540 mil cirugías mayores más podrían realizarse cada año”, 2022, <https://ipsuss.cl/actualidad/uso-de-pabellones-quirurgicos-hasta-540-mil-cirugias-mayores-mas>.
- [5] Zhou, Z.-H., Machine learning. Springer nature, 2021.
- [6] Cunningham, P., Cord, M., y Delany, S. J., “Supervised learning”, en Machine learning techniques for multimedia: case studies on organization and retrieval, pp. 21–49, Springer, 2008.
- [7] van Otterlo, M. y Wiering, M., Reinforcement Learning and Markov Decision Processes, pp. 3–42. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, [doi:10.1007/978-3-642-27645-3_1](https://doi.org/10.1007/978-3-642-27645-3_1).
- [8] Daniel, J. y Martin, J., “Speech and language processing”, 2024, <https://web.stanford.edu/~jurfafsky/slp3/5.pdf> (visitado el 2024-03-20).
- [9] Pang-Ning, T., Steinbach, M., y Kumar, V., Introduction to Data Mining. Pearson Education, updated edition ed., 2013.
- [10] Ferreira, A. J. y Figueiredo, M. A. T., Boosting Algorithms: A Review of Methods, Theory, and Applications, pp. 35–85. New York, NY: Springer New York, 2012, [doi:10.1007/978-1-4419-9326-7_2](https://doi.org/10.1007/978-1-4419-9326-7_2).
- [11] Chen, T. y Guestrin, C., “Xgboost: a scalable tree boosting system”, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16, p. 785–794, 2016, [doi:10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785).
- [12] Schapire, R., Explaining adaboost, pp. 37–52. Springer Berlin Heidelberg, 2013, [doi:10.1007/978-3-642-41136-6_5](https://doi.org/10.1007/978-3-642-41136-6_5). Publisher Copyright: © Springer-Verlag Berlin Heidelberg 2013.
- [13] Minaro, A., “Mastering lightgbm: Unravelling the magic behind gradient boosting”, 2023, <https://www.sicara.fr/blog-technique/mastering-lightgbm-unravelling-the-magic-behind-gradient-boosting>.

- [14] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., y Liu, T.-Y., LightGBM: A Highly Efficient Gradient Boosting Decision Tree. 2017, https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.
- [15] Taravat, A., Proud, S., Peronaci, S., Del Frate, F., y Oppelt, N., “Multilayer perceptron neural networks model for meteosat second generation seviri daytime cloud masking”, *Remote Sensing*, vol. 7, no. 2, pp. 1529–1539, 2015, [doi:10.3390/rs70201529](https://doi.org/10.3390/rs70201529).
- [16] Maas, A. y Tandon, S., “Unsupervised feature learning and deep learning tutorial”, 2019, <http://deeplearning.stanford.edu/tutorial/supervised/MultiLayerNeuralNetworks/>.
- [17] Geronimo, T. M., Cruz, C. E. D., de Souza Campos, F., Aguiar, P. R., y Bianchi, E. C., “Mlp and anfis applied to the prediction of hole diameters in the drilling process”, en *Artificial Neural Networks (Suzuki, K., ed.)*, cap. 7, Rijeka: IntechOpen, 2013, [doi:10.5772/51629](https://doi.org/10.5772/51629).
- [18] Akiba, T., Sano, S., Yanase, T., Ohta, T., y Koyama, M., “Optuna: A next-generation hyperparameter optimization framework”, en *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.
- [19] Berrar, D., “Cross-validation”, en *Encyclopedia of Bioinformatics and Computational Biology (Ranganathan, S., Gribskov, M., Nakai, K., y Schönbach, C., eds.)*, pp. 542–545, Oxford: Academic Press, 2019, [doi:https://doi.org/10.1016/B978-0-12-809633-8.20349-X](https://doi.org/10.1016/B978-0-12-809633-8.20349-X).
- [20] Guestrin, C., CS229: Machine Learning Dimensionality Reduction Principal Component Analysis (PCA) CS229: Machine Learning Slides include content developed by and co-developed with Emily Fox. 2021, <https://cs229.stanford.edu/notes2021fall/lecture14-pca.pdf>.
- [21] McInnes, L., Healy, J., y Melville, J., “Umap: Uniform manifold approximation and projection for dimension reduction”, 2020.
- [22] Wang, Y., Huang, H., Rudin, C., y Shaposhnik, Y., “Understanding how dimension reduction tools work: An empirical approach to deciphering t-sne, umap, trimap, and pacmap for data visualization”, *Journal of Machine Learning Research*, vol. 22, no. 201, pp. 1–73, 2021, <http://jmlr.org/papers/v22/20-1061.html>.
- [23] Naidu, G., Zuva, T., y Sibanda, E. M., “A review of evaluation metrics in machine learning algorithms”, en *Artificial Intelligence Application in Networks and Systems (Silhavy, R. y Silhavy, P., eds.)*, (Cham), pp. 15–25, Springer International Publishing, 2023.
- [24] Lundberg, S. M. y Lee, S.-I., “A unified approach to interpreting model predictions”, en *Advances in Neural Information Processing Systems (Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., y Garnett, R., eds.)*, vol. 30, Curran Associates, Inc., 2017, https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.
- [25] Luengo, J., García-Gil, D., Ramírez-Gallego, S., García, S., y Herrera, F., *Big Data Preprocessing*. Cham: Springer International Publishing, 2020, [doi:https://doi.org/10.1007/978-3-030-39105-8](https://doi.org/10.1007/978-3-030-39105-8).
- [26] Mahynski, N. A., Monroe, J. I., Sheen, D. A., Paul, R. L., Chen-Mayer, H. H., y Shen,

- V. K., “Classification and authentication of materials using prompt gamma ray activation analysis”, *Journal of Radioanalytical and Nuclear Chemistry*, vol. 332, p. 3259–3271, 2023, doi:<https://doi.org/10.1007/s10967-023-09024-x>.
- [27] Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., Lachiche, N., Ramírez-Quintana, M. J., y Flach, P., “Crisp-dm twenty years later: From data mining processes to data science trajectories”, *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 8, pp. 3048–3061, 2021, doi:[10.1109/TKDE.2019.2962680](https://doi.org/10.1109/TKDE.2019.2962680).
- [28] Espiñeira, J., Rediseño del Proceso de Solicitudes Web de Presupuesto Quirúrgico para un Prestador Privado de Salud. Memoria, Universidad de Chile, 2024.
- [29] Mayo Clinic, “Stages of labor and birth: Baby, it’s time!”, 2022, <https://www.mayoclinic.org/healthy-lifestyle/labor-and-delivery/in-depth/stages-of-labor/art-20046545>.
- [30] Mayo Clinic, “Circumcision (male) - mayo clinic”, 2022, <https://www.mayoclinic.org/tests-procedures/circumcision/about/pac-20393550>.
- [31] Mayo Clinic, “Vasectomy - mayo clinic”, 2023, <https://www.mayoclinic.org/tests-procedures/vasectomy/about/pac-20384580#:~:text=Overview>.
- [32] Mayo Clinic, “Cholecystectomy”, 2018, <https://www.mayoclinic.org/tests-procedures/cholecystectomy/about/pac-20384818>.
- [33] Mayo Clinic, “Bariatric surgery”, 2019, <https://www.mayoclinic.org/tests-procedures/bariatric-surgery/about/pac-20394258>.
- [34] Cleveland Clinic, “Hernia surgery”, 2023, <https://my.clevelandclinic.org/health/procedures/17967-hernia-repair-surgery>.
- [35] Cleveland Clinic, “Meniscus surgery: Who needs it, what to expect before after”, 2021, <https://my.clevelandclinic.org/health/treatments/21508-meniscus-surgery>.
- [36] Mayo Clinic, “Rhinoplasty - mayo clinic”, 2019, <https://www.mayoclinic.org/tests-procedures/rhinoplasty/about/pac-20384532>.
- [37] Medicine, J. H., “Rotator cuff repair”, <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/rotator-cuff-repair>.

Anexo A

Planteamiento del Problema

A.1. Problema identificado

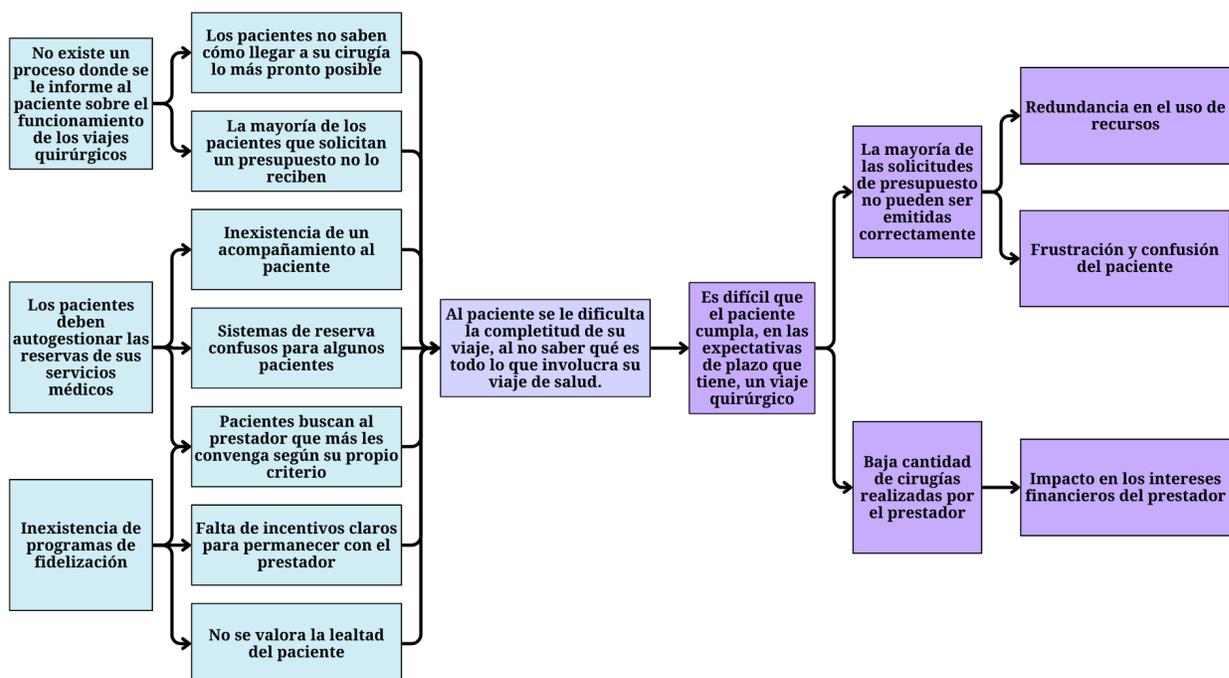


Figura A.1: Árbol de problemas. Elaboración propia.

Anexo B

Modelamiento y Evaluación de Resultados

B.1. Análisis y selección de datos

B.1.1. Análisis conjunto de datos de entrenamiento

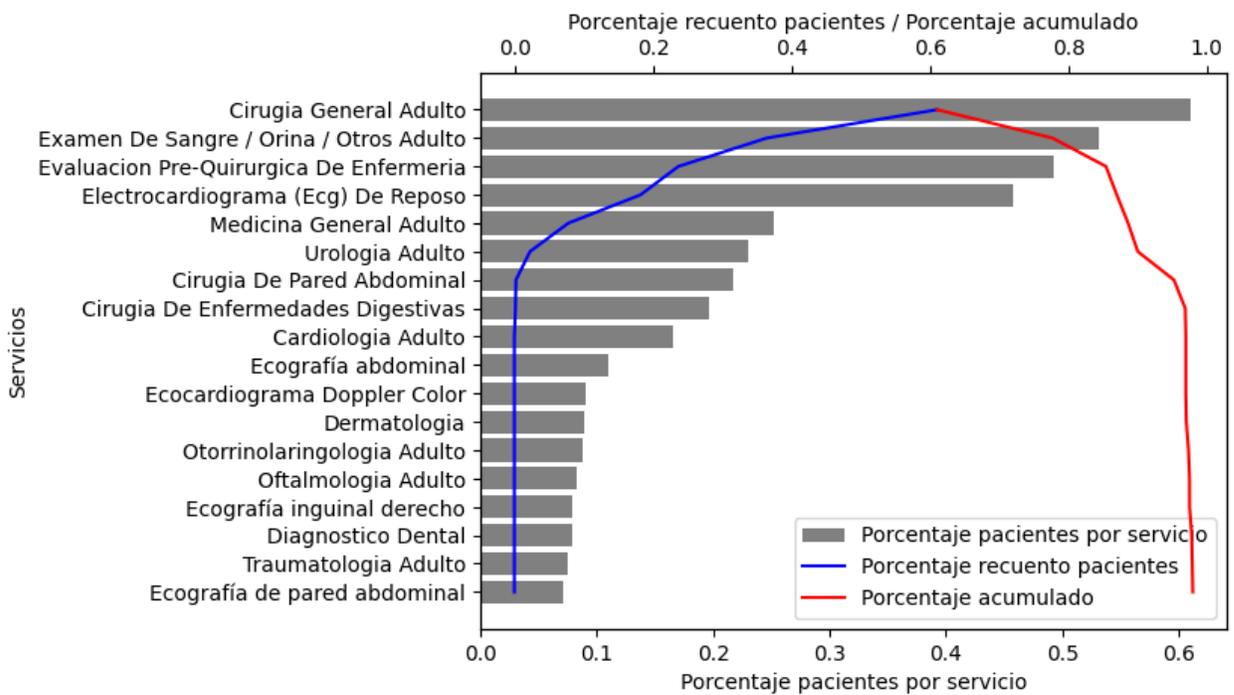


Figura B.1: Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Hernia abdominal. Elaboración propia.

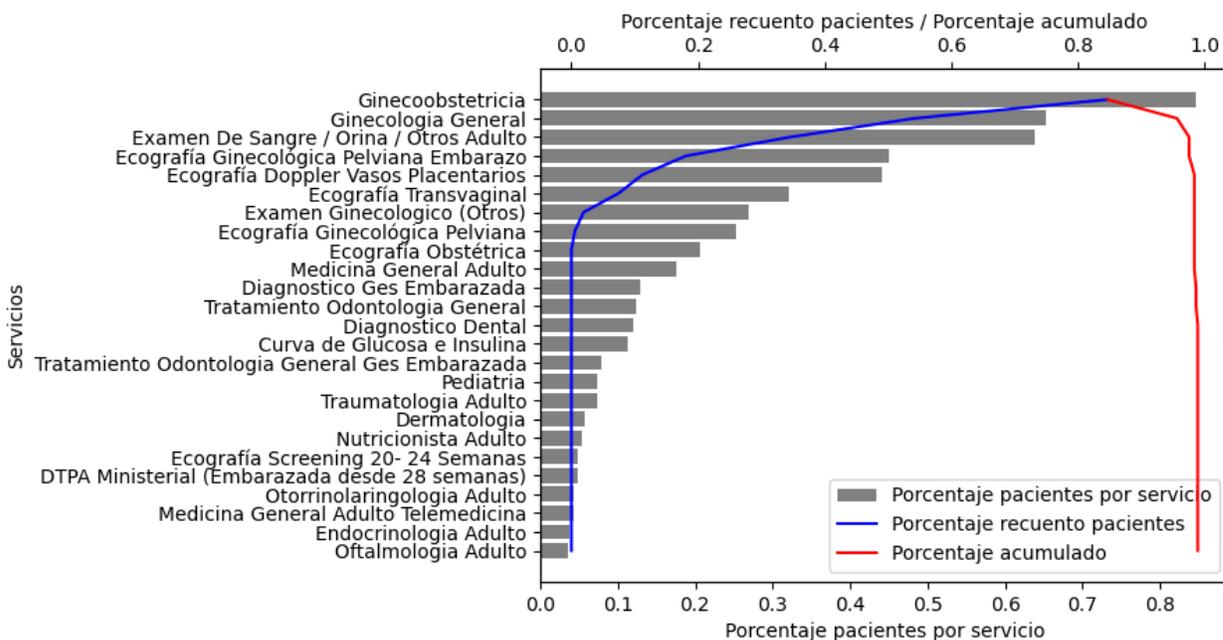


Figura B.2: Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Cesárea. Elaboración propia.

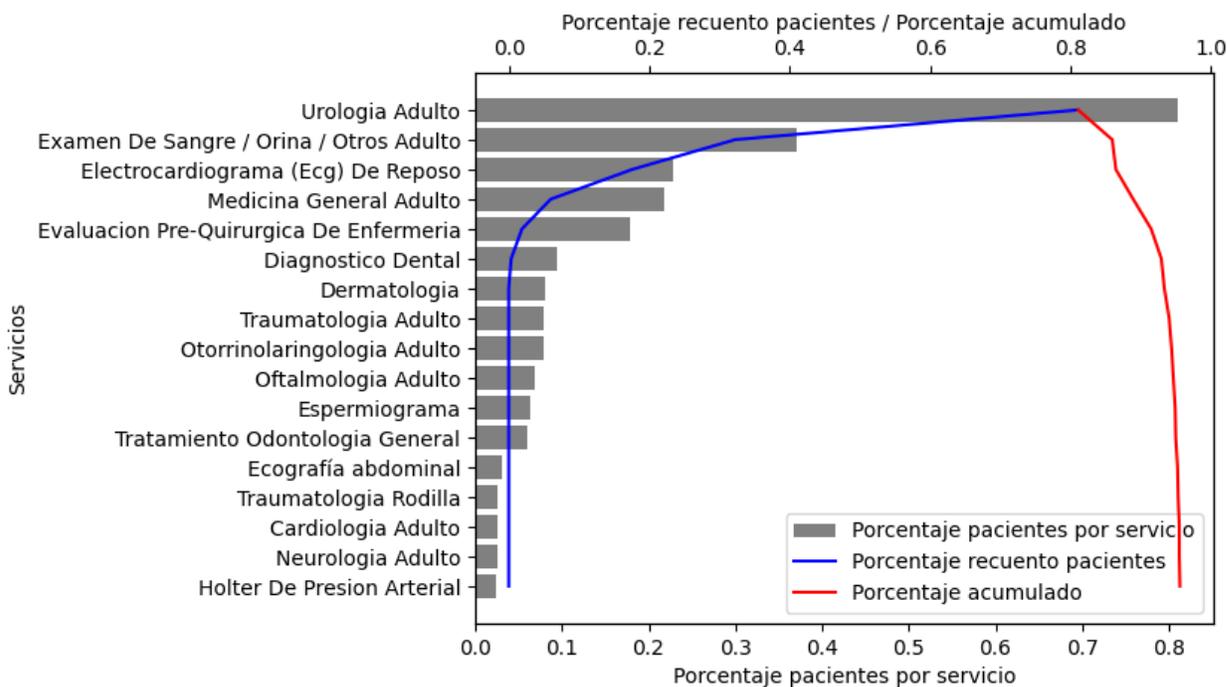


Figura B.3: Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Vasectomía. Elaboración propia.

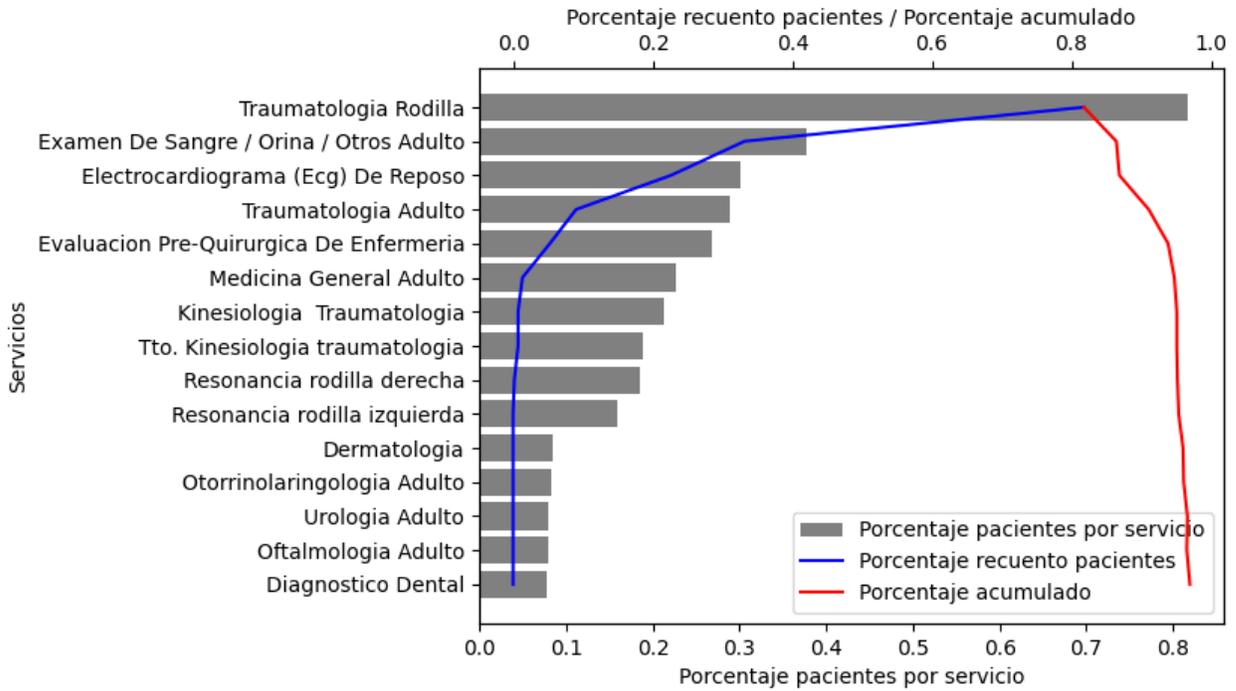


Figura B.4: Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Meniscectomía. Elaboración propia.

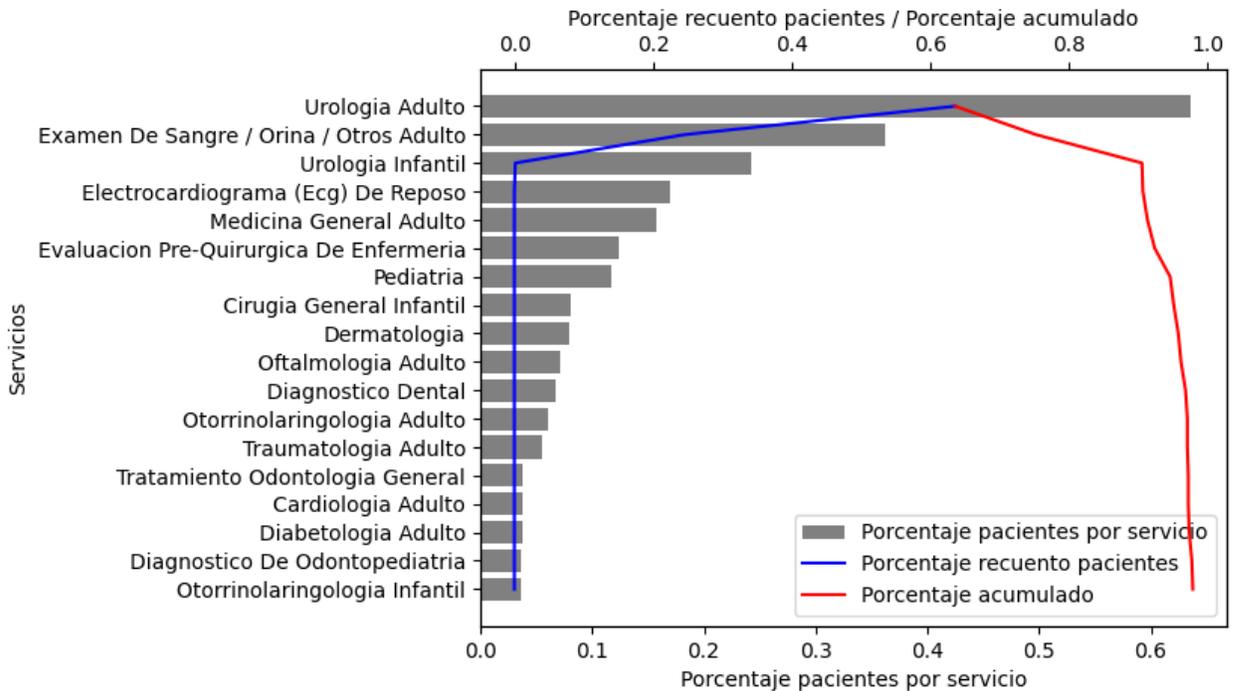


Figura B.5: Gráficos análisis del nivel de cumplimiento de los pacientes en los servicios para Circuncisión. Elaboración propia.

B.2. Modelamiento y evaluación

1. Regresión logística:

- *penalty*: Indica la norma de penalización a aplicar en la optimización, ayudando a evitar el sobreajuste.
- *C*: Controla la fuerza de la regularización. Un valor más bajo de este hiperparámetro conlleva una regularización más fuerte, haciendo que el modelo sea más simple. Por otro lado, un valor más alto de este hiperparámetro, conlleva una regularización más débil, haciendo que el modelo se ajuste más a los datos de entrenamiento.

2. Árboles de Decisión:

- *criterion*: Especifica la función a utilizar para medir la impureza de las particiones dentro del árbol.
- *max_depth*: Controla la profundidad máxima del árbol a construir, considerando que una profundidad más alta indica un mayor ajuste a los datos de entrenamiento, logrando encontrar relaciones más complejas. Por el contrario, una menor profundidad, tiende a una mayor generalización sobre los mismos.
- *min_samples_split*: Establece el número mínimo de muestras requeridas para dividir un nodo. Si el número de muestras de un nodo es menor a este valor, entonces no se dividirá más.
- *min_samples_leaf*: Establece el número mínimo de muestras requeridas para que un nodo sea terminal. Esto es, si la hoja resultante de una división resulta tener menos muestras que el valor establecido, entonces se descarta la división, y se convierte al nodo padre en terminal.

3. Bosques Aleatorios:

- *n_estimators*: Indica el número de árboles de decisión a construir en el bosque. Así, mientras más árboles se incluyan, mejor será la generalización del modelo, y junto con esto, también aumentará el costo computacional.
- *max_depth*: Controla la profundidad máxima de cada árbol generado dentro del bosque aleatorio.
- *min_samples_split*: Controla la cantidad mínima de muestras requeridas, para la división de un nodo en los árboles generados en el bosque.
- *min_samples_leaf*: Controla la cantidad mínima de muestras en una hoja para que esta sea considerada como tal. Si no, de igual manera que en los árboles de decisión, se elimina la división y se convierte al nodo padre en terminal.
- *max_features*: Determina el número máximo de características a considerar para buscar la mejor división de cada nodo.

4. XGBoost:

- *learning_rate*: Corresponde al factor de reducción aplicado a las actualizaciones de los pesos, después de cada paso del árbol. Una menor tasa de aprendizaje hará que los árboles crezcan más lentamente, lo que puede hacer que se requiera de una mayor cantidad de estimadores, si se quiere lograr un buen desempeño.

- *n_estimators*: Corresponde al número de árboles que se construirán en el modelo.
- *objective*: Es la función de pérdida a utilizar.
- *eval_metric*: Define la métrica de evaluación a utilizar para medir el rendimiento del modelo durante el entrenamiento.
- *lambda*: Corresponde al término de la regularización Ridge, controlando su magnitud aplicada a los pesos de las características.
- *alpha*: Corresponde al término de la regularización Lasso, controlando su magnitud aplicada a los pesos de las características.
- *max_depth*: Controla la profundidad máxima de cada árbol generado.
- *gamma*: También conocido como *min_split_loss*, es el valor mínimo de la pérdida que se produciría de generar cierta división. Si esta división no genera una pérdida mayor a gamma, entonces no se realiza la división y el nodo se convierte en terminal.
- *min_child_weight*: Corresponde al peso mínimo necesario para crear una partición en un nodo del árbol.
- *subsample*: Determina la proporción de muestras utilizadas para el entrenamiento de cada árbol.

5. *AdaBoost*:

- *n_estimators*: Determina el número de estimadores débiles que se utilizarán en el ensamblaje.
- *learning_rate*: Corresponde al peso que se le aplica a cada clasificador en cada iteración de *boosting*.
- *algorithm*: Determina el algoritmo base a utilizar para ajustar los estimadores.

6. *LightGBM*:

- *learning_rate*: Corresponde al factor de reducción aplicado a las actualizaciones de los pesos después de cada paso del árbol.
- *n_estimators*: Indica el número de árboles a construir en el modelo.
- *lambda_l1*: Corresponde al término de la regularización Lasso, controlando su magnitud aplicada a los pesos de las características.
- *lambda_l2*: Corresponde al término de la regularización Ridge, controlando su magnitud aplicada a los pesos de las características.
- *max_depth*: Controla la profundidad máxima de cada árbol generado.
- *min_child_samples*: Es el número mínimo de muestras necesarias en cada nodo terminal del árbol.
- *subsample*: Determina la proporción de muestras utilizadas para el entrenamiento de cada árbol.
- *colsample_bytree*: Es la proporción de características utilizadas para entrenar cada árbol.
- *boosting_type*: Determina el tipo de algoritmo de refuerzo que se utilizará durante el entrenamiento del modelo. Específicamente, es el método que se utilizará para la forma en que se construyen los árboles en el proceso de entrenamiento.

7. Multi-Layer Perceptron:

- *num_hidden_layers*: Indica el número de capas ocultas en la red neuronal.
- *hidden_layer_units*: Determina el número de unidades en cada capa oculta de la red neuronal.
- *dropout_rate*: Es un tipo de regularización, que desactiva ciertas unidades de una capa con una cierta probabilidad. Un mayor valor de este hiperparámetro, implica una mayor probabilidad de desactivar unidades durante la fase de entrenamiento.
- *learning_rate*: Controla la magnitud de las actualizaciones de los pesos durante el entrenamiento.
- *batch_size*: Determina el número de ejemplos de entrenamiento que se utilizan en cada iteración de entrenamiento.

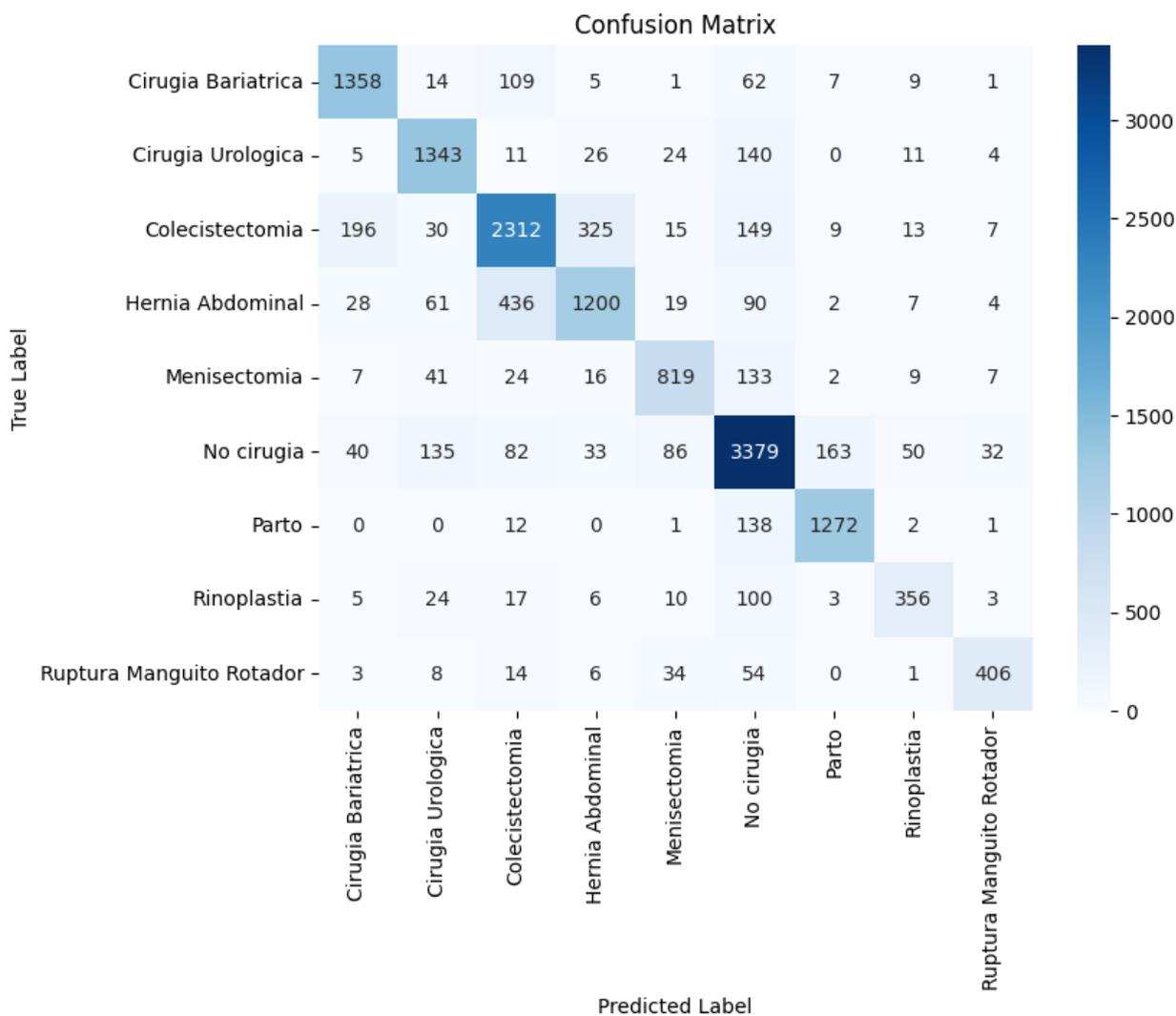


Figura B.6: Matriz de confusión modelo final. Elaboración propia.

B.3. Creación de las rutas quirúrgicas por cirugía

- Parto [29]: Corresponde al proceso mediante el cual una mujer embarazada expulsa al feto del útero al final del embarazo. Esto puede ser de forma vaginal o por cesárea.
- Cirugía Urológica [30, 31]: En primer lugar está la circuncisión, la cual corresponde al retiro total o parcial del prepucio del pene masculino. En segundo lugar, la vasectomía es un procedimiento quirúrgico que consiste en el bloqueo de los conductos deferentes, impidiendo el flujo de espermatozoides.
- Colectectomía [32]: Procedimiento quirúrgico mediante el cual se extirpa la vesícula biliar. Esta, es un órgano que se encuentra debajo del hígado, y es el encargado de almacenar la bilis, que ayuda a la digestión de grasas.
- Cirugía Bariátrica [33]: Corresponde a un conjunto de procedimientos quirúrgicos realizados a pacientes con obesidad para facilitar la pérdida de peso y mejorar su salud.
- Hernia Abdominal [34]: Tratamiento quirúrgico a los bultos que aparecen a través de los músculos del abdomen, o de una cicatriz de una intervención quirúrgica.
- Meniscectomía [35]: Es el procedimiento quirúrgico que se encarga de extirpar una parte, o la totalidad de un menisco, el cual es un disco de cartílago ubicado en la rodilla.
- Rinoplastía [36]: Cirugía que modifica la forma o la función de la nariz. Esta se realiza por fines estéticos, o también para corregir problemas respiratorios.
- Ruptura Manguito Rotador [37]: Procedimiento quirúrgico encargado de reparar los tendones dañados del manguito de los músculos del hombro. Esto se realiza cuando los tendones se encuentran desgarrados o separados de su inserción en el hueso del húmero.

Anexo C

Prueba de Campo

C.1. Obtención de los RUTs a partir del modelo

Tabla C.1: Resultados simulación modelo por días ($P(x) > 0.6$).

Nº de días	Nº de pacientes	Nº total de predicciones	Nº de predicciones con presupuesto emitido	Nº de predicciones con presupuesto emitido correctamente clasificado	Porcentaje de predicciones realizadas correctamente	Pacientes restantes por abordar
30	471.431	3,877	635	390	61.42 %	3,242
60	748.532	6,254	1,203	707	58.77 %	5,051
90	975.074	8,427	1,754	1,010	57.58 %	6,673

Tabla C.2: Resultados simulación modelo por días ($P(x) > 0.7$).

Nº de días	Nº de pacientes	Nº total de predicciones	Nº de predicciones con presupuesto emitido	Nº de predicciones con presupuesto emitido correctamente clasificado	Porcentaje de predicciones realizadas correctamente	Pacientes restantes por abordar
30	471.431	3,307	584	374	64.04 %	2,723
60	748.532	5,325	1,069	661	61.83 %	4,256
90	975.074	7,114	1,567	950	60.63 %	5,547

Tabla C.3: Resultados simulación modelo por días ($P(x) > 0.8$).

Nº de días	Nº de pacientes	Nº total de predicciones	Nº de predicciones con presupuesto emitido	Nº de predicciones con presupuesto emitido correctamente clasificado	Porcentaje de predicciones realizadas correctamente	Pacientes restantes por abordar
30	471.431	1,762	371	283	76.28 %	1,391
60	748.532	2,921	707	515	72.84 %	2,214
90	975.074	3,927	1,040	729	70.10 %	2,887

Tabla C.4: Resultados simulación modelo por días ($P(x) > 0.9$).

Nº de días	Nº de pacientes	Nº total de predicciones	Nº de predicciones con presupuesto emitido	Nº de predicciones con presupuesto emitido correctamente clasificado	Porcentaje de predicciones realizadas correctamente	Pacientes restantes por abordar
30	471.431	751	230	206	89.57 %	521
60	748.532	1,227	432	380	87.96 %	795
90	975.074	1,596	609	510	83.74 %	987

C.2. Evaluación piloto

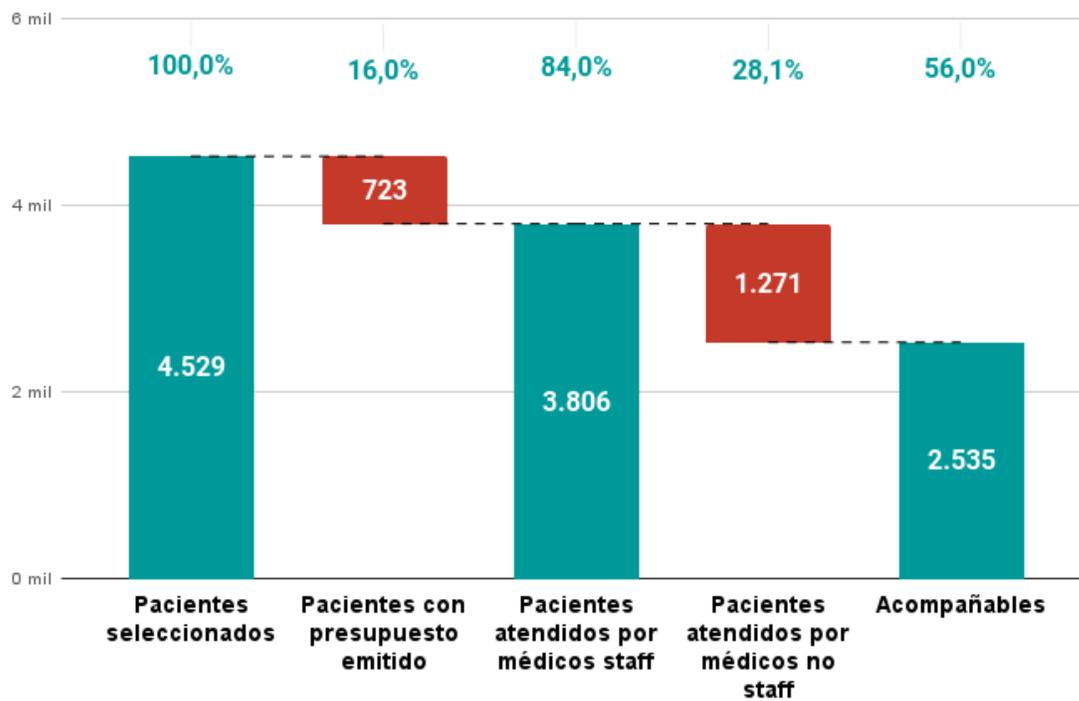


Figura C.1: Gráfico cascada de la obtención de datos para el piloto. Elaboración propia.

Tabla C.5: Agrupaciones de las conclusiones del piloto

Grupo	Código conclusión
Falla contacto	<ul style="list-style-type: none"> • No contesta • Se corta la llamada • Buzón de voz • Volver a llamar • Recado con tercero
No Quirúrgico	<ul style="list-style-type: none"> • No está en un viaje Quirúrgico
Acompañamiento tardío	<ul style="list-style-type: none"> • Ya se realizó cirugía • No interesado - Ya se lo realizó en con el prestador
Quirúrgico	<ul style="list-style-type: none"> • En viaje Quirúrgico - Otra Cirugía - No Agenda • En viaje Quirúrgico - Otra Cirugía - Agenda • En viaje Quirúrgico - Colectomía - No Agenda • En viaje Quirúrgico - Otra Cirugía - COM • En viaje Quirúrgico - Colectomía - COM • En viaje Quirúrgico - Colectomía - Agenda
No interesado	<ul style="list-style-type: none"> • No interesado - No decide/ no quiere