



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

PREDICCIÓN DE EFECTIVIDAD DE TRATAMIENTOS NO CONVENCIONALES EN  
PACIENTES CON CÁNCER DE MAMA HER2

MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL EN COMPUTACIÓN

ESTEBAN JAVIER RAMÍREZ FUENZALIDA

PROFESOR GUÍA:  
NELSON BALOIAN TATARYAN

PROFESOR CO-GUÍA:  
SERGIO PEÑAFIEL AGUILERA

MIEMBROS DE LA COMISIÓN:  
NANCY HITSCHFELD KAHLER  
JAVIER BUSTOS JIMÉNEZ

SANTIAGO DE CHILE  
2023

# Resumen

Este trabajo presenta el desarrollo de una herramienta computacional innovadora para apoyar la toma de decisiones en el tratamiento del cáncer de mama HER2-positivo. El cáncer de mama femenino es una preocupación mundial debido a su alta incidencia y mortalidad. En particular, el subtipo HER2-positivo representa un desafío significativo en la práctica clínica.

El objetivo principal del trabajo es aplicar un modelo de clasificación interpretable basado en la teoría de Dempster-Shafer. Este modelo utiliza características clínicas e histopatológicas del paciente para predecir la respuesta al tratamiento neoadyuvante. Se busca identificar qué variables influyen en la obtención de respuestas completas patológicas, lo que permitiría tomar decisiones informadas en el tratamiento.

El enfoque interpretable del modelo brinda la oportunidad de discutir y validar los resultados con especialistas en oncología. Además, el modelo proporciona, en fase de entrenamiento, más de 20 reglas, las que contienen información relevante sobre la brecha que existe en algunos pacientes usando un esquema u otro en el tratamiento.

En términos de proyección, se destaca el potencial del trabajo para influir en políticas públicas de salud. La colaboración con instituciones médicas permitirá validar su utilidad en situaciones clínicas reales y respaldar decisiones. El enfoque de investigación continua y el análisis prospectivo impulsarán la transformación en el tratamiento del cáncer de mama HER2-positivo.

En resumen, este trabajo ha logrado un avance en el desarrollo de una herramienta computacional con alta proyección en la atención oncológica. Su buen desempeño en comparación con otros métodos interpretables, su interpretabilidad y su potencial para mejorar la toma de decisiones clínicas y políticas en el tratamiento del cáncer de mama HER2 abren una nueva perspectiva en el avance de la medicina oncológica.

*A mi madre y padre; Julia y Javier.  
A mis hermanas; Daniela y Mariana*

# Agradecimientos

Quiero comenzar agradeciendo a mis padres, Julia y Javier, quienes me han criado y educado en un hogar lleno de amor y alegría. Admiro profundamente su historia de esfuerzo y su noble corazón. Gracias por su amor y aliento, les debo todo lo que soy y lo que seré.

También quiero agradecer a mis hermanas mayores, Daniela y Mariana. Por sus buenos consejos y cariño fraterno. Elegí este camino porque fui inspirado por dos hermanas muy esforzadas, cuya constancia admiro. Siempre serán mi ejemplo.

Por supuesto, también quiero agradecer mi profesor guía Nelson Baloian, quien me recibió en su equipo y depositó fe en mí. Como parte de su equipo he aprendido a hacer ciencia y a encantarme con ella.

Además, un especial agradecimiento a Sergio Peñafiel, mi coguía en este proceso y en cuyo brillante trabajo se basa mi memoria. Muchas gracias compartir tu conocimiento, por tu paciencia y diligencia con mi proyecto. También quiero agradecer a los profesores que integran la comisión que revisa mi trabajo. Gracias a su retroalimentación y comentarios mi trabajo ha mejorado considerablemente.

Quiero agradecer además al equipo de investigación de la Fundación Arturo López Pérez, quienes han mostrado gran interés en colaborar con nosotros, tanto haciendo ciencia como en el desarrollo de herramienta, poniendo a nuestra disposición todo lo necesario. Fue en una reunión con ellos que surgió el tema de esta memoria. Admiro la labor de la institución y estoy muy feliz de poder aportar en su trabajo.

Finalmente, quiero dar las gracias a Natalia, mi polola, amiga y compañera de estos últimos e intensos años. Gracias por tu cariño, contención y apoyo en todo este proceso. Te admiro y te amo muchísimo.

# Tabla de Contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Objetivo General . . . . .	3
1.2. Objetivos Específicos . . . . .	3
1.3. Estructura del Informe . . . . .	3
<b>2. Marco teórico</b>	<b>5</b>
2.1. Marco Conceptual . . . . .	5
2.1.1. El cáncer de mama HER2-positivo . . . . .	5
2.1.2. Tratamientos de cáncer de mama . . . . .	6
2.1.3. Situación actual del tratamiento del cáncer . . . . .	7
2.1.4. Aprendizaje de máquinas . . . . .	9
2.1.5. Clasificación interpretable . . . . .	10
2.1.6. Teoría de Dempster-Shafer . . . . .	10
2.1.7. Clasificador basado en la teoría de Dempster-Shafer . . . . .	12
2.1.8. Desarrollo de <i>software</i> . . . . .	12
2.2. Estado del arte . . . . .	15
2.2.1. Predictores de Respuesta Completa Patológica . . . . .	15
2.2.2. Herramientas predictivas de efectividad . . . . .	16
<b>3. Descripción e implementación de la solución</b>	<b>17</b>
3.1. Clasificador interpretable en pacientes con esquema completo . . . . .	17
3.1.1. Descripción de los datos . . . . .	18

3.1.2.	Planteamiento del problema . . . . .	18
3.1.3.	Modelo . . . . .	18
3.1.4.	Resultados . . . . .	19
3.2.	Clasificador interpretable en pacientes con esquemas diferentes . . . . .	25
3.2.1.	Descripción de los datos . . . . .	25
3.2.2.	Planteamiento del problema . . . . .	26
3.2.3.	Modelo . . . . .	27
3.2.4.	Resultados . . . . .	27
3.2.5.	Conclusiones . . . . .	33
3.3.	Desarrollo del <i>software</i> . . . . .	33
3.3.1.	Arquitectura del <i>software</i> . . . . .	34
3.3.2.	Modelo de datos . . . . .	36
3.3.3.	API REST para administrar modelos . . . . .	37
3.3.4.	API FALP . . . . .	38
3.3.5.	Interfaz de usuario . . . . .	39
3.3.6.	Contenedor . . . . .	41
<b>4.</b>	<b>Evaluación de la Herramienta Predictiva</b>	<b>43</b>
4.1.	Observaciones Importantes sobre el Modelo Predictivo . . . . .	43
4.2.	Observaciones Importantes sobre la Herramienta . . . . .	44
4.3.	Ideas de Trabajo Futuro . . . . .	44
4.4.	Conclusión . . . . .	44
<b>5.</b>	<b>Conclusión</b>	<b>45</b>
5.1.	Trabajos futuros . . . . .	46
5.2.	Contribuciones . . . . .	47
	<b>Siglas</b>	<b>48</b>
	<b>Glosario</b>	<b>49</b>

<b>Bibliografía</b>	<b>53</b>
<b>A. Datos clasificador interpretable en pacientes con esquema completo</b>	<b>54</b>
<b>B. Clasificador interpretable en pacientes tratados con Trastuzumab</b>	<b>56</b>
<b>C. Diseño previo de la aplicación</b>	<b>62</b>
<b>D. Aplicación implementada</b>	<b>64</b>

# Índice de Tablas

3.1. Desempeño del modelo DSGD. . . . .	19
3.2. Rendimiento de todos los métodos en la predicción de la respuesta de los pacientes. . . . .	20
3.3. Reglas más importantes para No pCR y su incertidumbre (Inc). Ver glosario para entender las siglas. . . . .	22
3.4. Reglas más importantes para pCR y su incertidumbre (Inc). Ver glosario para entender las siglas. . . . .	22
3.5. Desempeño del modelo DSGD. . . . .	27
3.6. Rendimiento de todos los métodos en la predicción de la respuesta de los pacientes. . . . .	29
3.7. Reglas más importantes para No pCR y su incertidumbre (Inc). Ver glosario para entender las siglas. . . . .	30
3.8. Reglas más importantes para pCR y su incertidumbre (Inc). Ver glosario para entender las siglas. . . . .	30
3.9. Detalle de cambio de clasificación al cambiar esquema. Donde VP: Verdadero Positivo, FN: Falso Negativo, VN: Verdadero Negativo, FP: Falso Positivo, T-P: Trastuzumab - Pertuzumab y T: Trastuzumab . . . . .	32



# Índice de Ilustraciones

2.1. Extracto del flujo a seguir en el tratamiento de pacientes con cáncer de mama HER2. Cortesía de la FALP. . . . .	8
3.1. Matriz de confusión después de validación cruzada. . . . .	20
3.2. Curvas ROC promedio para los métodos utilizados en el problema. . . . .	21
3.3. Árbol de decisión resultante. Ver glosario para entender las siglas. . . . .	23
3.4. Distribución de esquema administrado en pacientes que obtuvieron respuesta. . . . .	26
3.5. Matriz de confusión después de validación cruzada. . . . .	28
3.6. Curvas ROC promedio para los métodos utilizados en el problema. . . . .	29
3.7. Matriz de confusión anterior luego de cambiar el esquema de tratamiento. . . . .	32
3.8. Primera arquitectura propuesta. . . . .	35
3.9. Arquitectura implementada. . . . .	35
3.10. UML del modelo de datos de la herramienta. . . . .	37
3.11. Esto es lo que se encuentra el usuario el ingresar. . . . .	40
3.12. Vista principal de la aplicación. . . . .	41
A.1. Distribución de variables usadas para entrenar el modelo. . . . .	55
B.1. Distribución de pCR de pacientes. . . . .	56
B.2. Distribución de edad, estado menopáusico e IMC en pacientes de la muestra. . . . .	57
B.3. Distribución de cT (tamaño), cN (linfonodos comprometidos) y etapa del cáncer en pacientes de la muestra. . . . .	58
B.4. Distribución de involucramiento axilar, tipo histológico y nivel receptores de estrógeno en pacientes de la muestra. . . . .	59

B.5. Distribución de receptores de progesterona, índice Ki67 y relación del gen HER/CENTROMERO en pacientes de la muestra. . . . .	60
B.6. Distribución de amplificación de HER2, perfil molecular y esquema de tratamiento en pacientes de la muestra. . . . .	61
C.1. El usuario llena el formulario con datos de un paciente. Envía el formulario. .	62
C.2. La solicitud es enviada a la API, realiza el cálculo de predicción y responde con los resultados. . . . .	63
C.3. Su última predicción se puede visualizar en la barra lateral. . . . .	63
D.1. Una vez ingresado un rut, la aplicación realiza una consulta directamente a la API de la FALP y obtiene los datos del paciente. Es posible rellenar con un solo clic los valores que están bien estructurados e individualizados. . . . .	64
D.2. Además, se pueden ver informes más completos sobre el paciente para extraer mas información. Por ejemplo, el tipo histológico. . . . .	65
D.3. Al enviar el formulario, se muestra un modal con los resultados de la predicción.	65
D.4. Los resultados se muestran en la barra lateral derecha. . . . .	66
D.5. Es posible realizar otro cálculo para el mismo paciente, agregando algunas de las variables ausentes. . . . .	66
D.6. Es posible comparar el nuevo output con los resultados anteriores. . . . .	67

# Capítulo 1

## Introducción

El cáncer de mama femenino es el cáncer más comúnmente diagnosticado a nivel mundial, superando incluso al cáncer de pulmón. Es el quinto cáncer más mortal, siendo la primera causa de muerte por cáncer en mujeres [1]. Lo anterior lo convierte en un problema de salud mundial, por lo que los avances científicos relacionados a la prevención y al tratamiento son muy importantes.

Existen distintos tipos de cáncer de mama. Para este trabajo es de particular interés el tipo HER2-positivo, clasificado así por la alta presencia de la proteína HER2 en las células. Este representa cerca del 20% de los cánceres de mama [2]. Diversos estudios indican que un tratamiento exitoso en este tipo de cáncer, es la terapia neoadyuvante con anticuerpos monoclonales tales como Pertuzumab y Trastuzumab, en conjunto con quimioterapia [3] convencional. Este tipo de tratamiento ha demostrado una respuesta positiva en los pacientes, disminuyendo las células cancerosas, lo cual clínicamente se conoce como una alta respuesta completa patológica (rcp) o, por sus siglas en inglés, pCR (pathological complete response).

Para el tratamiento mencionado, hay distintas formas en que los pacientes pueden costearlo: sistema de salud público o programas gubernamentales, seguros oncológicos, entre otros. En Chile, el Trastuzumab está asegurado para pacientes con diagnóstico de cáncer de mama HER2-positivo, gracias a la Ley N° 20.850, más conocida como Ley Ricarte Soto. Sin embargo, es otro el panorama con el Pertuzumab, el cual al día de hoy es un tratamiento de difícil acceso y con un alto costo monetario.

Los antecedentes indican que pacientes que son tratados mediante ambos anticuerpos monoclonales tienen una mayor tasa de respuesta completa patológica, que pacientes que reciben solo un anticuerpo. Por esto, es lógico deducir que siempre busca tratar a los pacientes con ambos fármacos para asegurar el éxito. Sin embargo, además del difícil acceso mencionado anteriormente, un factor importante a considerar es la calidad de vida de una persona bajo este tratamiento. Pues, sin lugar a dudas, administrar dos fármacos sumamente invasivos como lo son el Trastuzumab y Pertuzumab deteriora el estado del paciente.

En este contexto, el Instituto Oncológico Fundación Arturo López Pérez (FALP), ha llevado a cabo estudios con el objetivo de identificar, mediante métodos estadísticos, qué características clínicas hacen que un paciente sea más propenso a responder al tratamiento

neoadyuvante (Saffie et al. [4]).

En colaboración con la FALP, surge la idea de utilizar herramientas de aprendizaje de máquinas y datos de pacientes que ya han recibido este tratamiento, para desarrollar un modelo que permita predecir si un paciente tendrá una respuesta al tratamiento neoadyuvante. El modelo se basa en las características clínicas e histopatológicas del paciente, y puede aplicarse tanto a un esquema que incluya ambos fármacos, como a aquellos que solo utilicen Trastuzumab.

No solo se busca abordar el problema de esta manera por el desempeño de los algoritmos de aprendizaje de máquinas, sino también por la oportunidad de utilizar métodos interpretables. En este tipo de problemas, la falta de interpretabilidad plantea un desafío, ya que basar decisiones de tratamiento en algoritmos de tipo ‘caja negra’, impide discutir los resultados con especialistas en oncología. Esto, debido a que no se dispone de información sobre cómo el algoritmo llegó a una decisión, lo cual va en contra del principio de la medicina basada en la evidencia [5].

El modelo en cuestión se construye a partir de un algoritmo de clasificación interpretable basado en la Teoría de Dempster-Shafer, la cual es un marco matemático utilizado para modelar y combinar la incertidumbre y la evidencia en la toma de decisiones [6]. El método de clasificación basado en esta teoría fue desarrollado por Sergio Peñafiel, quien es co-guía en este trabajo [7]. Además, este método ha sido utilizado en problemas de naturaleza clínica, demostrando un alto desempeño en comparación con algoritmos tan potentes como Support Vector Machine o Multilayer Perceptron [8].

La aplicación del algoritmo mencionado, ofrece un doble beneficio, ya que este tipo de modelo proporcionaría información relevante a los médicos sobre las características de los pacientes que influyen en el éxito de un tratamiento en comparación con otro. Esto permitiría la discusión y validación del modelo, fomentando así la investigación en el campo. Además, el modelo serviría como indicador para determinar si un paciente podría prescindir de Pertuzumab en su tratamiento.

El objetivo principal de esta memoria es desarrollar el modelo mencionado previamente y crear una herramienta computacional que, utilizando dicho modelo, brinde apoyo a los profesionales de la salud al momento de decidir qué tratamiento es más adecuado en función de las características clínicas del paciente. Con el fin de que esta herramienta sea práctica para los médicos, se desarrollará una aplicación web que se integre con la ficha médica utilizada en la página de la FALP. Esta aplicación web deberá consumir las APIs de la institución para obtener los datos de los pacientes y generar predicciones al evaluar un perfil.

El trabajo consta principalmente 3 etapas para llegar al objetivo principal. En primera instancia, como un acercamiento al clasificador basado en Dempster-Shafer, se resuelve lo que puede considerarse un sub problema dentro del problema principal. Se entrena el clasificador para estimar qué pacientes tienen o no respuesta al tratamiento, identificando así las variables predictoras de respuesta completa patológica.

Posteriormente, una vez validado el modelo mencionado anteriormente con los profesionales, se procede a entrenar el algoritmo incluyendo como variable el esquema de tratamiento: trastuzumab y pertuzumab o solo trastuzumab.

A continuación, se procede a la etapa de desarrollo de *software*, en la cual se crea una interfaz de usuario simple en forma de formulario de una sola vista. Los usuarios pueden optar por ingresar manualmente los datos o permitir que el formulario se complete automáticamente con la información disponible en la base de datos de la FALP. La interfaz envía solicitudes a una API que utiliza el modelo previamente desarrollado, para realizar una predicción con ambos tratamientos. Finalmente, se proporciona al usuario la probabilidad de tener una respuesta al tratamiento con los diferentes esquemas.

## 1.1. Objetivo General

Obtener un modelo, en base a un algoritmo de clasificación interpretable, que sea capaz de predecir la efectividad de un tratamiento en base a la información de los pacientes. Entregando información sobre qué variables son importantes en esta decisión.

Además, desarrollar una herramienta computacional que proporcione apoyo a los especialistas en oncología de la FALP en la toma de decisiones para determinar el tratamiento más adecuado para un paciente con diagnóstico de cáncer de mama HER2-positivo.

## 1.2. Objetivos Específicos

1. Entrenar el clasificador en un sub problema y validarlo con los especialistas.
2. Entrenar el clasificador con una base de datos más grande, agregando el esquema de tratamiento como variable adicional.
3. Realizar un análisis estadístico del trabajo de las secciones anteriores y validarlo con expertos en oncología.
4. Desarrollar el software que haga accesible el modelo entrenado.

## 1.3. Estructura del Informe

En el siguiente capítulo se presenta el marco teórico que sustenta y justifica este trabajo. Explicando conceptos clínicos claves para entender el trabajo, tales como los tipos de cáncer de mama y los tipos de tratamientos existentes. Además, se presentan algunos antecedentes del problema, cuál es la situación país y cómo soluciones como la propuesta en este trabajo son un potencial aporte.

Además, se desarrolla un breve repaso sobre conceptos de aprendizaje de máquinas, algoritmos interpretables y la teoría de Dempster - Shafer, con el fin de sentar las bases necesarias para entender el modelo utilizado. Luego, se realiza una breve descripción de conceptos de desarrollo de *software*, abarcando algunos conceptos, patrones de diseño y tecnologías.

Luego, en el tercer capítulo se muestra la metodología seguida para entrenar los modelos clasificadores a partir del algoritmo. Desde la descripción y tratamientos de los datos, descripción y construcción del modelo clasificador, presentación y análisis de los resultados obtenidos.

La metodología presenta dos casos de estudio, es decir, dos modelos con grupos de estudio distintos. Por un lado, se describe el procedimiento aplicado a un grupo de pacientes que recibió el esquema completo (Trastuzumab y Pertuzumab). Este grupo proviene de diversos centros médicos, donde los exámenes difieren y solo algunos atributos pudieron ser utilizados para obtener experimentos representativos y confiables. Por otro lado, se analiza un grupo de pacientes provenientes de centros médicos chilenos, que en algunos casos recibieron un esquema parcial (solo Trastuzumab) y en otros, el esquema completo. Aunque este último grupo es más reducido, al tener un origen común, se trata de pacientes con los mismos exámenes realizados en el momento del diagnóstico, lo que permitió trabajar con datos de gran relevancia médica en este tipo de investigación.

La última parte del tercer capítulo muestra el desarrollo del software en que se despliega el modelo entrenado para predecir la respuesta de los pacientes al tratamiento. En esta parte se describe la arquitectura de los elementos como bases de datos y APIs involucradas, el modelo de datos para administrar los modelos y peticiones, y el diseño e implementación de la interfaz.

El cuarto capítulo describe la evaluación de la herramienta predictiva con algunos especialistas de la FALP en una sesión grupal. Aquí se comentan algunas observaciones hechas respecto a la herramienta, tanto valorando su aporte y valor, como entregando retroalimentación para mejorarla y darle más valor.

En el quinto y último capítulo, se concluye de forma general el trabajo presentado, poniendo énfasis en el valor de obtener información interesante a partir del método interpretable de clasificación y en el valor de la herramienta en el contexto nacional. Además proponiendo trabajo futuro para mejorar el modelo y avanzar en el desarrollo del uso de algoritmos de aprendizaje de máquinas en este tipo de problemas.

# Capítulo 2

## Marco teórico

En este capítulo se presenta el marco teórico de las distintas herramientas usadas para cumplir el objetivo de este trabajo. En primer lugar, se describe el cáncer de mama HER2-positivo y los métodos que existen para tratar este tipo de cáncer. Luego, se explican las distintas técnicas y métodos de aprendizaje de máquinas, poniendo énfasis en los métodos interpretables, con el fin de identificar el método de clasificación interpretable basado en la teoría de Dempster-Shafer.

### 2.1. Marco Conceptual

#### 2.1.1. El cáncer de mama HER2-positivo

El cáncer es una enfermedad en la cual las células del cuerpo crecen sin control y se propagan por todo el organismo. Se trata de una enfermedad de naturaleza genética, provocada por alteraciones en los genes responsables de regular el crecimiento celular. Estos cambios genéticos pueden ser hereditarios o estar influenciados por factores externos, como la exposición a sustancias químicas o radiación.

El funcionamiento normal de las células se basa en un ciclo de vida en el cual se forman y se multiplican a través de la división celular, en respuesta a las necesidades del cuerpo. Cuando una célula envejece o se daña, se muere y se produce un reemplazo que reinicia el ciclo. No obstante, este proceso no siempre se desarrolla de esa manera, ya que pueden existir células dañadas o anormales, que se resisten a morir y se multiplican por el cuerpo. Estas células pueden formar tumores (bultos de tejido), los que pueden ser cancerígenos (malignos) o no cancerígenos (benignos) [9]. Los tumores malignos se propagan a los tejidos cercanos, esparciéndose por distintas zonas del cuerpo pudiendo formar nuevos tumores. Por otro lado, los tumores benignos no se propagan por el cuerpo, una vez extirpados no vuelven a aparecer.

El cáncer de mama se caracteriza por la formación mayoritariamente de tumores malignos en las glándulas mamarias, afectando principalmente a las mujeres. Para evaluar el estado de un cáncer de mama, es fundamental comprender la presencia y el nivel de expresión de

la proteína HER2 (siglas que corresponden a Human Epidermal growth factor Receptor-2 en inglés). La proteína HER2 está presente en la superficie de todas las células mamarias que promueve el crecimiento celular. Las células cancerosas que presentan altos niveles de HER2 se denominan HER2-positivas o HER2+, y son el foco de atención en este trabajo [10].

### 2.1.2. Tratamientos de cáncer de mama

Una vez se diagnostica el cáncer, un equipo de profesionales decide el tratamiento a seguir. El tratamiento se decide en función de factores como el tipo de cáncer, el grado, tamaño, etc. Además, se considera la opinión del paciente al decidir el plan a seguir. Los pacientes usualmente son sometidos a más de uno de los siguientes tratamientos:

1. **Quimioterapia:** La quimioterapia es el uso de fármacos con el fin de eliminar las células cancerosas o detener su propagación. Lamentablemente este tipo de tratamiento no solo ataca las células cancerígenas, si no que también afecta las células sanas, deviniendo en efectos secundarios como nauseas, fatiga o pérdida de cabello [11].
2. **Cirugía:** En una cirugía los doctores remueven o extirpan el tumor. La cirugía puede ser conservativa, es decir, solo extirpar el tumor y zonas comprometidas al rededor de este, pero sin remover el seno completo. O puede ser una cirugía mas radical, conocida como Mastectomía, en la cual sí es necesario remover el seno completo. Es usual realizar quimioterapia a los pacientes antes de ser sometidos a una cirugía, en la cual se busca reducir el tamaño de la zona comprometida, buscando una cirugía menos radical.
3. **Radioterapia:** En este tipo de terapia se usa radiación electromagnética de alta frecuencia con el fin de eliminar las células cancerosas del paciente o detener su esparcimiento. La radioterapia externa es la que se aplica en cáncer de seno, enfocándose en una parte específica del cuerpo (el seno en este caso). Lamentablemente, al igual que la quimioterapia, no diferencia entre células dañadas y células sanas, trayendo efectos secundarios para el paciente [12].
4. **Terapia hormonal:** Los tumores sensibles a hormonas suelen crecer por la acción de estas. Esta terapia detiene el crecimiento de tumores hormono-sensibles, limitando la producción de hormonas del cuerpo y interfiriendo sus efectos en el tumor [13].
5. **Inmunoterapia:** Es un tipo de terapia biológica, que usa sustancias producidas por el cuerpo para tratar una enfermedad. Esta terapia fortalece el sistema inmunitario de un paciente para combatir el cáncer [14].
6. **Terapia dirigida:** La terapia dirigida usa fármacos u otras sustancias para atacar células cancerosas específicas. Qué sustancias usar en el tratamiento se decide, en primera instancia, en base a los resultados de un examen de biomarcadores [15]. Existen distintos tipos de terapias dirigidas:
  - (a) Anticuerpos monoclonales (Margetuximab, Pertuzumab, Sacituzumab govitecán, Trastuzumab y Trastuzumab deruxtecán).
  - (b) Inhibidores de tirosina-cinasas (Lapatinib, Neratinib y Tucatinib).



- (c) Inhibidores de las cinasas dependientes de ciclinas (Abemaciclib, Alpelisib, Palbociclib y Ribociclib).
- (d) Inhibidores de la diana de la rapamicina en mamíferos (Everólimus).
- (e) Inhibidores de PARP (Olaparib y Talazoparib).

7. **Terapia neoadyuvante:** La terapia neoadyuvante en el tratamiento del cáncer es un enfoque en el que se administra tratamiento (como quimioterapia, terapia hormonal u otros) antes de la cirugía o de cualquier tratamiento local, como la radioterapia. El objetivo principal de la terapia neoadyuvante es reducir el tamaño del tumor en la mama y, en algunos casos, en los ganglios linfáticos cercanos antes de la cirugía. Esto puede facilitar la extirpación completa del tumor durante la cirugía y permitir la conservación de la mama en lugar de realizar una mastectomía. También puede ayudar a evaluar la respuesta del tumor al tratamiento y ajustar la estrategia terapéutica en función de esa respuesta.

En el contexto del cáncer de mama, la terapia neoadyuvante puede incluir quimioterapia, terapia hormonal (especialmente para tumores sensibles a hormonas, como los receptores de estrógeno o progesterona), terapia dirigida (como anticuerpos monoclonales contra el receptor del factor de crecimiento epidérmico humano 2 o HER2) o una combinación de estos tratamientos. La administración de tratamiento antes de la cirugía puede ayudar a reducir la extensión del tumor, lo que puede hacer posible una cirugía menos invasiva.

La evaluación de la respuesta al tratamiento neoadyuvante es fundamental para determinar la efectividad del tratamiento y planificar el tratamiento adicional posterior a la cirugía. Esto se suele hacer mediante imágenes médicas y análisis patológicos de muestras de tejido tomadas durante la cirugía. Dependiendo de la respuesta al tratamiento neoadyuvante, se decidirá el curso del tratamiento adicional, que puede incluir más cirugía, radioterapia o terapia adyuvante (tratamiento después de la cirugía) [16].

Este trabajo está enfocado en la efectividad de un tratamiento neoadyuvante que combina quimioterapia con terapia dirigida basada en anticuerpos monoclonales (Trastuzumab y Pertuzumab).

### 2.1.3. Situación actual del tratamiento del cáncer

El cáncer es una enfermedad que implica un proceso largo y costoso a quienes lo padecen. Son pocas las familias que pueden solventar los gastos asociados al tratamiento y más aún considerando la situación económica que enfrenta Chile actualmente [17]. Por ello instituciones como la FALP son de las primeras alternativas recomendadas a pacientes para comenzar el largo proceso.

Las pacientes con cáncer de mama HER2-positivo que son sometidas al tratamiento neoadyuvante que combina quimioterapia con anticuerpos monoclonales (Trastuzumab y Pertuzumab), tienen diversas formas de costear su tratamiento, como lo son el apoyo del sistema

de salud público o programas de gobierno, seguros oncológicos, entre otros. En Chile a día de hoy, el Trastuzumab está asegurado para pacientes con diagnóstico de cáncer de mama HER2-positivo, gracias a la Ley N° 20.850, más conocida como Ley Ricarte Soto. Por el contrario, el Pertuzumab es un fármaco costoso y de difícil acceso. Por lo tanto, poder caracterizar e identificar el perfil de pacientes que puedan prescindir de este segundo fármaco, es un aporte a mejorar la calidad de vida del paciente.

Para entender en qué parte del tratamiento se integra la solución trabajada en esta memoria, se presenta en forma breve el flujo por el que pasa un paciente con sospechas de cáncer de mama que llega por primera vez a la FALP.

Cuando un paciente presenta síntomas que sugieren un cáncer de mama se le realizan exámenes imagenológicos para determinar si el cáncer es benigno o maligno. Si los resultados indican que puede ser maligno, se realiza una biopsia para confirmar las sospechas. En el caso de confirmar las sospechas y si el cáncer de mama es invasor, a partir de los resultados histopatológico se procede a clasificar el tipo de cáncer:

1. Cáncer de mama triple negativo
2. Cáncer de mama HER2 enriquecido
3. Cáncer de mama luminal

El segundo cáncer de mama es el abarcado en este proyecto. Un paciente diagnosticado con cáncer de mama HER2 es sometido a una serie de exámenes adicionales a partir de los cuales el comité oncológico especializado definirá qué tipo de tratamiento es conveniente para el paciente. Como se puede apreciar en la Figura 2.1.

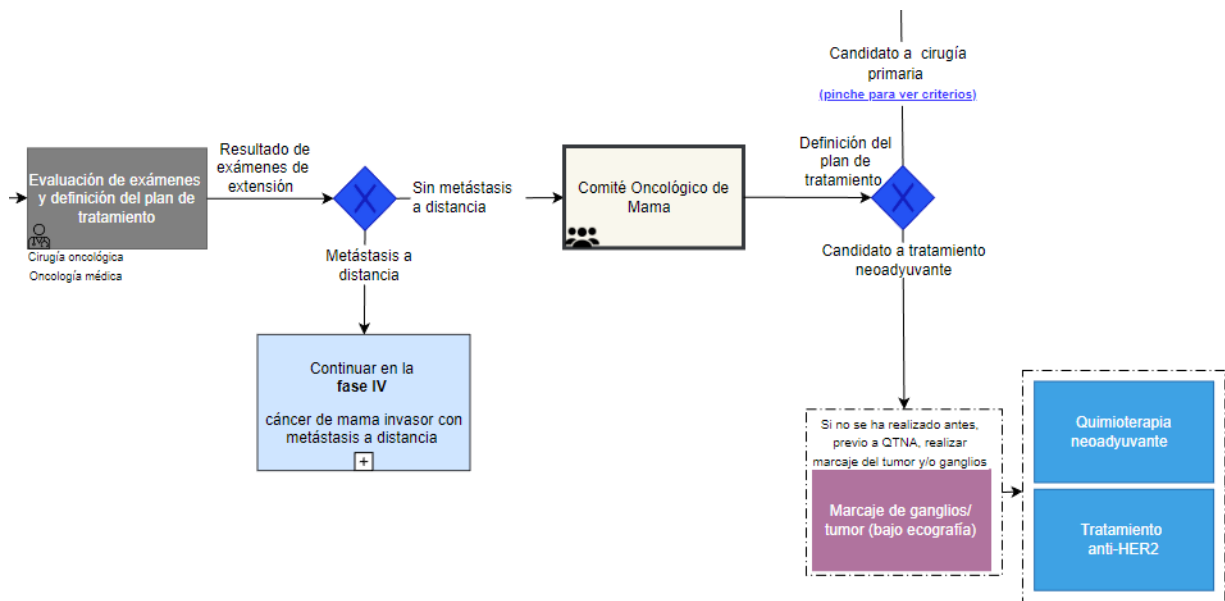


Figura 2.1: Extracto del flujo a seguir en el tratamiento de pacientes con cáncer de mama HER2. Cortesía de la FALP.

Si el paciente es candidato a tratamiento neoadyuvante, lo que quiere decir que es posible disminuir las zonas comprometidas por el cáncer, entonces el comité debe decidir qué tipo de tratamiento va a seguir y es en ese punto donde el software entrega valor. Pues actualmente no se cuenta con un método establecido para determinar si el paciente que se somete al tratamiento anti-HER2 le es conveniente usar ambos anticuerpos (Trastuzumab y Pertuzumab) o, en base a sus características y a la experiencia con otros pacientes, podría ser suficiente con solo un medicamento.

#### **2.1.4. Aprendizaje de máquinas**

Es un campo dentro de la inteligencia artificial que se centra en el desarrollo de algoritmos y herramientas estadísticas capaces de discernir, a partir de conjuntos de datos, patrones o tomar decisiones. Estos algoritmos adquieren conocimientos a través de un proceso de aprendizaje y pueden emplearse para realizar predicciones, clasificaciones y otras tareas de mayor complejidad. Existen principalmente dos enfoques en aprendizaje de máquinas: aprendizaje supervisado y no supervisado [18].

##### **Supervisado**

En el aprendizaje supervisado, los algoritmos se entrenan utilizando conjuntos de datos etiquetados, es decir, se disponen de datos de entrada junto con sus correspondientes salidas. Por ejemplo, consideremos un algoritmo al cual le proporcionamos numerosas imágenes de perros y gatos. El algoritmo utiliza diversas técnicas para reconocer los patrones distintivos de perros y gatos en diferentes situaciones, razas, colores, etc. Luego, se espera que el modelo pueda identificar imágenes que no fueron incluidas durante la fase de entrenamiento, y su desempeño se evalúa mediante diversas métricas, determinando cuántos gatos y perros clasifica correctamente. Un algoritmo como el descrito anteriormente puede ser aplicado en problemas como reconocimiento facial, diagnóstico médico en base a radiografías, entre otros.

El aprendizaje supervisado se divide en dos grupos, según el tipo de problema que abordan. El primer grupo se enfoca en la clasificación, donde el algoritmo asigna a los datos de entrada una clase o categoría específica. Por otro lado, el segundo grupo se centra en problemas de regresión, donde la salida consiste en un valor numérico continuo. En este caso, el objetivo es comprender la relación de dependencia entre las variables de entrada y salida.

##### **No supervisado**

Los algoritmos de aprendizaje no supervisado, al contrario del enfoque anterior, aprenden patrones, analizan y agrupan en base a un conjunto de datos no etiquetados. Por ejemplo, consideremos un conjunto de datos con la información de los movimientos de usuarios en una plataforma web de compra. En base a dichos datos, usando algoritmos de aprendizaje no supervisado, es posible segmentar a los clientes en base a su comportamiento y preferencias. Lo anterior permite implementar un sistema de recomendaciones o extraer información relevante para aplicar estrategias de marketing personalizadas.

Al igual que el aprendizaje supervisado mencionado anteriormente, este tipo de aprendizaje no supervisado se divide en dos grupos. Por un lado, está el *clustering* o agrupamiento, en el cual los modelos buscan agrupar datos en base a sus similitudes. Por otro lado, se encuentran los problemas de asociación, en los cuales los algoritmos utilizan diversas reglas para descubrir relaciones entre las variables en el conjunto de datos.

### 2.1.5. Clasificación interpretable

Los algoritmos de clasificación son ampliamente utilizados en el aprendizaje supervisado, donde se asigna una clase a un vector de datos. Este vector puede representar una imagen, audio, datos de un paciente, entre otros. Existen diversos métodos de clasificación, tales como Árboles de Clasificación, Support Vector Machine (SVM), Redes Neuronales, etc. Además, estos algoritmos mencionados pueden ser categorizados como interpretables o no interpretables (también conocidos como cajas negras).

Se dice que un algoritmo es una caja negra o no es interpretable, cuando no entrega información sobre cómo y por qué llegó a clasificar cierto dato entrante con una etiqueta. Es decir, se tiene un resultado y una clasificación pero no una justificación de ella. Un algoritmo interpretable es todo lo contrario, ya que entregan información sobre cómo una decisión fue tomada y en qué medida cada variable fue importante en la tarea.

La falta de interpretabilidad es un problema en el campo de la salud. Usar algoritmos de este tipo para determinar diagnósticos, o tratamientos, sin entender cómo se llegó al resultado, viola el principio de medicina basada en la evidencia [5], donde se requiere una comprensión clara y una justificación de las decisiones médicas. Por ejemplo, es posible que una red neuronal profunda consiga diagnosticar enfermedades en base a radiografías, con una gran precisión. Sin embargo, los médicos no pueden entender cómo el algoritmo llega a una conclusión, entonces podría estar tomando decisiones sin sentido alguno, teniendo un modelo poco confiable a la hora de usarlo para tomar decisiones.

La confianza y aceptación de un modelo se basan en entender el razonamiento detrás de un diagnóstico, para tomar decisiones informadas y tratar de forma adecuada a los pacientes.

Sin embargo, modelos interpretables suelen mostrar un menor desempeño que algoritmos de tipo caja negra, como las redes neuronales. Como regla general, mientras más preciso es el modelo, menor es su interpretabilidad [19]. Por lo tanto, generar algoritmos con alto desempeño e interpretabilidad es un desafío en ciencias de la computación.

En este trabajo, se usa un clasificador interpretable basado en la teoría de Dempster-Shafer. Este modelo tiene la ventaja de poder trabajar con la incertidumbre y valores nulos, además permite agregar conocimiento experto para mejorar sus predicciones.

### 2.1.6. Teoría de Dempster-Shafer

La teoría de Dempster Shafer [6] es un método matemático que permite razonar bajo la incerteza y datos incompletos. Usualmente es considerada una generalización de la teoría

Bayesiana porque incluye incerteza a los clásicos modelos Bayesianos. Para entender mejor este método, digamos que  $X$  es un conjunto de todos los estados del sistema, que llamaremos marco de discernimiento. Una función de asignación de masa que llamaremos  $m$ :

$$m : 2^X \rightarrow [0, 1], \quad m(\phi) = 0, \quad \sum_{A \subseteq X} m(A) = 1 \quad (2.1)$$

donde  $A$  es un sub conjunto de  $X$  y  $\phi$  es un conjunto vacío. La expresión  $m(A)$  puede ser interpretada como la probabilidad de obtener los resultados esperados exactamente de  $A$  y no de un subconjunto del mismo. Las masas asignadas a los distintos sub conjuntos agrupan el conocimiento y valor del modelo.

Por ejemplo, podemos tomar el problema de predecir la respuesta a un tratamiento del paciente, donde se consideran dos clases: tiene respuesta (R) o no tiene respuesta (NR). Así, la función de asignación de masa sobre el marco de discernimiento asigna valores a el conjunto vacío, al singleton R, al singleton NR y al conjunto  $\{R, NR\}$ . Por ejemplo, podría tener las siguientes masas asignadas:

$$\begin{aligned} m(\phi) &= 0 \\ m(\{R\}) &= 0,2 \\ m(\{NR\}) &= 0,6 \\ m(\{R, NR\}) &= 0,2 \end{aligned}$$

La función de masa es equivalente a una distribución de probabilidad en la teoría Bayesiana.

La teoría de Dempster-Shafer define dos indicadores para medir la cantidad de evidencia que soporta una clase resultante. Estos son la creencia (*Belief*) y la plausibilidad (*Plausibility*).

La creencia es la evidencia total que soporta una clase y es definida así:

$$Bel(A) = \sum_{X \subseteq A} m(X) \quad (2.2)$$

La métrica de plausibilidad define la cantidad de evidencia que puede soportar una clase resultante:

$$Pl_m(A) = \sum_{B \cap A \neq \phi} m(B) \quad (2.3)$$

La creencia, considera la evidencia que respalda a la misma clase y a sus subconjuntos, definiendo un límite inferior para la probabilidad del conjunto. Mientras que la plausibilidad, que considera las intersecciones no nulas con el conjunto, define un límite superior para la probabilidad. Esta es la forma en que la teoría puede entregar una estimación sobre la probabilidad de una clase, mediante métodos como la propuesta por Smets [20] o por Cobb y Shenoy [21].

La Regla de Dempster [22] permiten combinar múltiples fuentes de evidencia representadas por su función de asignación de masa respectiva para el mismo marco de discernimiento.

De acuerdo a la Regla de Shafer, dadas dos masas  $m_1$  y  $m_2$ , la masa que las combine a partir de la siguiente fórmula:

$$\begin{aligned} m_c(A) &= m_1(A) \oplus m_2(A) \\ &= \frac{1}{1-K} \sum_{B \cap C = A \neq \phi} m_1(B)m_2(C) \end{aligned} \quad (2.4)$$

donde  $K$  es una constante que representa el grado de conflicto entre las masas a combinar, dado por la siguiente expresión:

$$K = \sum_{B \cap C \neq \phi} m_1(B)m_2(C) \quad (2.5)$$

### 2.1.7. Clasificador basado en la teoría de Dempster-Shafer

El algoritmo está basado en reglas, las que son compuestas por aseveraciones respecto a los datos. Estas pueden ser resultado de la exploración de los datos o a partir de conocimiento de expertos en el área, por ejemplo, pacientes con sobrepeso o un tumor mayor a cierto tamaño.

Para clasificar, el método busca la clase a la cual una observación pertenece. Una observación, en el problema antes mencionado, pueden ser las características de un paciente, las cuales se representan como un vector de características. Así, una regla está relacionada a una función de masa, entonces el que un paciente tenga sobrepeso va a tener asignada una función de asignación de masa que expresa cuan propenso es o no un paciente a tener respuesta a un tratamiento [7]. Peñafiel comenta que las reglas son definidas por métodos de generación de reglas que implementa en su algoritmo, pero además permite ingresar reglas personalizadas para incluir la opinión experta, por ejemplo.

Entonces, el modelo sigue los siguientes pasos para realizar una predicción:

1. A partir de un vector de características, filtra las reglas que no son satisfechas.
2. Combina las funciones de masa asignadas a las reglas seleccionadas usando la Regla de Dempster.
3. A partir de la función de asignación de masa obtenida en la parte anterior, estima una probabilidad y elige la clase con mayor probabilidad (la que tiene mayor creencia).

Considerando lo anterior, el proceso de entrenamiento del modelo consta de buscar las masas para cada regla que produzcan el mínimo error en la clasificación de un conjunto de datos, optimizando sus valores mediante el método del Gradiente Descendente.

### 2.1.8. Desarrollo de *software*

El desarrollo de *software* es un área de ciencias de la computación, en la cual se crea, diseña, despliega y mantiene *software*. Donde un *software* se entiende como un conjunto de

algoritmos o instrucciones que le dicen a un computador qué hacer [23]. Existen, a grandes rasgos, tres tipos de desarrollos:

1. *Software* de sistema: implementación de sistemas operativos, manejo de discos y hardware (drivers, por ejemplo), etc.
2. *Software* de programación: desarrollo de herramientas para programadores, como editores de textos, compiladores, depuradores y otras herramientas para crear código.
3. *Software* de aplicación: crear herramientas que ayudan a usuarios a resolver tareas, por ejemplo: manejo y visualización de datos, sistemas de gestión, programas de monitoreo, seguridad, etc.

Existen diversos patrones de diseño, arquitecturas y herramientas usadas en el ámbito del desarrollo de *software*. A continuación se describen las más relevantes para entender este trabajo.

### **Patrón Modelo Vista Controlador**

El Modelo-Vista-Controlador (MVC) es un patrón de diseño arquitectónico ampliamente utilizado en el desarrollo de software. Su objetivo principal es separar la lógica de negocio y la interacción con el usuario en tres componentes distintos: el Modelo, la Vista y el Controlador. Esta separación facilita la modularidad, el mantenimiento y la escalabilidad de las aplicaciones [24].

1. **Modelo:** El Modelo representa la capa de datos y la lógica de negocio de la aplicación. Es responsable de gestionar y mantener los datos, así como de implementar las operaciones y reglas de negocio. El Modelo no tiene conocimiento de la interfaz de usuario o de cómo se presentan los datos, lo que lo hace independiente de la Vista y el Controlador.
2. **Vista:** La Vista representa la interfaz de usuario y es responsable de presentar los datos al usuario en un formato comprensible. Se encarga de la visualización y la presentación de los datos que provienen del Modelo. La Vista no realiza operaciones de negocio, sino que simplemente muestra la información y responde a las acciones del usuario, transmitiéndolas al Controlador.
3. **Controlador:** El Controlador actúa como intermediario entre el Modelo y la Vista. Es responsable de recibir las interacciones del usuario desde la Vista, procesarlas y actualizar el Modelo en consecuencia. También se encarga de actualizar la Vista con los cambios realizados en el Modelo, asegurando que la interfaz de usuario refleje siempre el estado actual de los datos.

El flujo de trabajo en el patrón MVC es el siguiente: cuando el usuario interactúa con la interfaz de usuario (vista), la acción se comunica al controlador. El controlador procesa la acción y actualiza el modelo en consecuencia. A su vez, el modelo notifica al controlador sobre los cambios, y este actualiza la vista para reflejar los nuevos datos.

El patrón MVC ofrece varios beneficios, incluyendo:

1. Separación de responsabilidades: Permite que cada componente se centre en su tarea específica, lo que facilita el mantenimiento y la reutilización del código.
2. Facilidad de colaboración: Dado que cada componente tiene su función bien definida, diferentes desarrolladores pueden trabajar en paralelo en diferentes partes de la aplicación.
3. Flexibilidad y escalabilidad: Al estar bien organizada, la arquitectura MVC permite que la aplicación crezca y evolucione sin comprometer su estructura.

El patrón MVC es ampliamente utilizado en el desarrollo de aplicaciones web y de escritorio, así como en otros contextos de desarrollo de *software*, y ha demostrado ser una estrategia efectiva para crear aplicaciones bien estructuradas y mantenibles.

## Tecnología de contenedores

La tecnología de contenedores es una forma de virtualización a nivel de sistema operativo que permite empaquetar y distribuir aplicaciones junto con sus dependencias y configuraciones en un entorno aislado y portátil. Cada contenedor se ejecuta como un proceso independiente y aislado, compartiendo el mismo kernel del sistema operativo del host. Esto ofrece una mayor eficiencia y rendimiento en comparación con la virtualización tradicional basada en hipervisores, ya que no es necesario ejecutar múltiples sistemas operativos completos [25].

Docker es una de las herramientas más populares para implementar y administrar contenedores. Proporciona una plataforma abierta que permite a los desarrolladores y administradores crear, distribuir y ejecutar aplicaciones dentro de contenedores de manera sencilla. Algunas de las características clave de Docker incluyen:

1. Portabilidad y Consistencia: Los contenedores Docker son altamente portátiles, lo que significa que una vez que una aplicación se empaqueta en un contenedor, puede ejecutarse en cualquier entorno que tenga Docker instalado, independientemente de las diferencias en la infraestructura subyacente.
2. Aislamiento y Seguridad: Los contenedores Docker están aislados unos de otros y del sistema operativo host, lo que proporciona un nivel de seguridad adicional. Cada contenedor tiene su propio sistema de archivos, procesos y recursos asignados, lo que evita conflictos y garantiza que los contenedores no interfieran entre sí.
3. Escalabilidad y Eficiencia: Docker facilita el despliegue de aplicaciones en entornos de nube o en infraestructuras de servidores, permitiendo una rápida y fácil escalabilidad tanto horizontal como vertical.
4. Gestión de Imágenes: Docker utiliza imágenes para construir contenedores. Las imágenes son plantillas de solo lectura que contienen el sistema operativo, el código de la aplicación y las dependencias necesarias para ejecutarla. Esto facilita la replicación y distribución de las aplicaciones en diferentes entornos.



5. Orquestación de Contenedores: Además de Docker, existen herramientas como Kubernetes que facilitan la gestión, el equilibrio de carga y la escalabilidad de contenedores en clústeres y entornos de producción a gran escala.

En resumen, la tecnología de contenedores, con Docker como una de sus principales representantes, ha revolucionado la forma en que se desarrollan, distribuyen y gestionan aplicaciones en la actualidad. Proporciona una mayor eficiencia, portabilidad y facilidad de gestión, lo que la convierte en una solución atractiva para el desarrollo y despliegue de aplicaciones en diversos escenarios.

## 2.2. Estado del arte

### 2.2.1. Predictores de Respuesta Completa Patológica

En la literatura ya existen estudios respecto a la efectividad de la combinación de Trastuzumab y Pertuzumab con otros componentes en el tratamiento de cáncer de seno [26]. Sin embargo, son escasos los estudios hechos en casos reales. Un ejemplo de un estudio hecho con casos reales se tiene en Díaz-Acedo et al [27], trabajo en el cual se contrasta el tratamiento con Trastuzumab-Pertuzumab-Quimioterapia y Trastuzumab-Quimioterapia, que es justamente lo que se busca hacer en este trabajo y es uno de los pocos estudios hechos en base a casos reales. A pesar de la poca cantidad de datos usados, lograron inferir que en pacientes con cáncer de alto riesgo se tiene una mayor respuesta al tratamiento usando esquema completo y sin que suponga incrementar la toxicidad cardiaca.

Otro trabajo, en base a casos reales, relacionado con la predicción de la respuesta de pacientes con cáncer de mamas con sobre expresión HER2 a este tratamiento neoadyuvante que combina anticuerpos con quimioterapia es una investigación de la FALP (*Saffie* et al. [4]) a partir de los datos de 94 pacientes con cáncer de mamá evaluó el efecto de las variables clínicas e histopatológicas en el nivel de respuesta al tratamiento neoadyuvante ya mencionado.

El análisis estadístico del estudio mencionado fue hecho mediante un estudio de control de casos [28], donde los *casos* fueron las mujeres que consiguieron una respuesta al tratamiento y el *control* se dio sobre mujeres que no lo consiguieron. Las variables expuestas fueron las características clínicas de los pacientes y la variable de respuesta fue justamente si se alcanzó o no la respuesta completa patológica completa.

El efecto de cada variable de exposición se midió con diferentes métodos, las variables continuas fueron evaluadas mediante la prueba de los rangos con signo de Wilcoxon [29], las variables nominales mediante la prueba exacta de Fisher [30] y las variables ordinales con la prueba de tendencias Cochran-Armitage [31]. Luego se evaluó el efecto combinado de las variables de exposición usando una regresión logística múltiple, determinando las variables que mejor predicen la respuesta al tratamiento con un algoritmo de eliminación hacia atrás.

A partir del trabajo de *Saffie* et al. se puede desprender que pacientes con tumores que presentan una alta expresión de receptores de progesterona y testosterona, tienen una menor

tendencia a tener una respuesta al tratamiento. Mientras que los pacientes con un alto índice ki-67 tienen mayor probabilidad de mostrar una respuesta al tratamiento que quienes tienen un bajo índice. Además, también se menciona que la etapa del cáncer también es una variable importante a considerar en el tratamiento neoadyuvante, donde los pacientes en etapas más avanzadas, sorprendentemente, suelen mostrar una mayor tasa de respuesta. Entre otras variables que dan cuenta de la pureza del cáncer, también se obtiene que un tumor con mayor amplificación de HER2 tiene una mayor tendencia a una respuesta al tratamiento.

Los resultados del estudio de Saffie et al. se resumen en la pureza del cáncer en que mientras más puro en la proteína HER2 sea un cáncer, más probabilidad de respuesta tiene al tratamiento, y por otro lado, mientras más heterogéneo o más Luminal sea el tumor, menor probabilidad de respuesta tendrá el paciente.

Variables como el tamaño del tumor también han sido evaluadas como predictoras, sin embargo, hasta el momento los estudios indican que no es una característica importante al someterse a este tipo de tratamiento [32, 33].

### **2.2.2. Herramientas predictivas de efectividad**

Existen variadas herramientas disponibles con las que se puede proyectar el éxito de un tratamiento en base a las características de un paciente. Los especialistas de la FALP usan una plataforma llamada Predict, desarrollada por investigadores de la Universidad de Cambridge, la cual es específica para pacientes con cáncer de mama HER2-positivo. Esta herramienta recibe las características de la paciente, muchas de las cuales también son usadas en este trabajo, y en base a eso calcula la probabilidad de sobrevida del paciente.

Esta herramienta, además permite realizar el cálculo para distintas opciones de tratamiento, e incluso permite visualizar cuánto mejora la probabilidad de sobrevida si al tratamiento se le agrega Trastuzumab previo a la cirugía [34]. Este método matemático fue desarrollado a partir de los datos de 5000 pacientes con cáncer de mama diagnosticado, y validado con más de 23000 pacientes alrededor del mundo.

Con el tiempo se ha ido extendiendo la herramienta, por ejemplo, agregando nuevas variables del paciente como entrada o nuevas opciones para realizar las proyecciones, como la cantidad de años desde la cirugía en la que se quiere la probabilidad de sobrevida del paciente.

# Capítulo 3

## Descripción e implementación de la solución

En este capítulo se presenta la metodología, los datos, los modelos y herramientas usadas en cada una de las etapas comprendidas para concluir en la herramienta de apoyo de decisiones médicas. En la primera parte se desarrolla un acercamiento al algoritmo interpretable elegido para resolver el problema de clasificación. Se trabaja con un conjunto de datos de pacientes sometidos al tratamiento con esquema completo, se entrena el modelo y se obtienen variables predictoras de una respuesta al tratamiento gracias a la característica interpretable del algoritmo.

La segunda sección se trabaja con un subconjunto del conjunto de datos anterior, el cual está constituido por pacientes de Chile, con variables cuyo análisis es de particular interés y además, estas pacientes fueron sometidas a distintos esquemas monoclonales.

Por último, en la tercera sección se describe la estructura de la aplicación mediante la cual los doctores tendrán acceso al modelo entrenado en la parte anterior. Se describe la arquitectura, tecnologías e interfaz de la herramienta creada.

### 3.1. Clasificador interpretable en pacientes con esquema completo

A continuación se presenta un modelo para predecir la respuesta completa patológica de pacientes a un tratamiento neoadyuvante basado en fármacos monoclonales Trastuzumab y Pertuzumab, combinados con quimioterapia convencional, antes de que los pacientes se sometan a algún tratamiento. El modelo es en base a un algoritmo de clasificación interpretable. Esto permite determinar qué características clínicas del paciente podrían afectar o determinar su respuesta al tratamiento.

Se utiliza un método interpretable basado en reglas, basado en la teoría de Dempster-Shafer que ya ha sido usado previamente en problemas del área médica [8]. La motivación

principal para utilizar este modelo es que al ser interpretable permite discutir los resultados obtenidos de forma transparente con los profesionales y, además, al estar basado en reglas genera conocimiento en base a la evidencia o importancia que da el modelo a cierta regla al clasificar un paciente.

El modelo fue entrenado a partir de un conjunto de datos de 390 pacientes, donde 57% obtuvo respuesta patológica completa después del tratamiento.

### 3.1.1. Descripción de los datos

El conjunto de datos contiene información de pacientes con cáncer de mama HER2-positivo, tratados en 8 hospitales de distintos países (Portugal, España y Chile), entre Enero del 2018 y Diciembre del 2021.

Los pacientes recibieron una terapia neoadyuvante con doble bloqueo hormonal. Las características clínicas e histopatológicas consideradas en este estudio son: edad, estado menopáusico, categorías T y N (describen tamaño y propagación a ganglios, respectivamente), etapa inicial del cáncer, grado, presencia de receptores de estrógeno y progesterona expresión, y port último, el índice Ki-67. Finalmente, la variable a predecir es si el paciente muestra o no una respuesta luego de ser sometido al tratamiento. Ver en Anexo A la distribución de los datos.

Esta parte del trabajo se realiza con datos de 390 pacientes de 8 hospitales, con una edad aproximada de 53 años. De estos, el 65,1% tienen receptores hormonales positivos y el 54,9% presentan afectación de los ganglios linfáticos regionales. En cuanto al esquema de quimioterapia, el 59,7% de los pacientes habían recibido un régimen basado en antraciclinas. Por último, se logró una respuesta patológica completa en el 56,7% de los pacientes.

### 3.1.2. Planteamiento del problema

El desafío presentado en esta parte es determinar si un paciente con cáncer de mama HER2-positivo, que fue tratado con la terapia neoadyuvante ya descrita, mostrará una respuesta al tratamiento, basado en sus características clínicas e histopatológicas. El problema puede ser entendido como una clasificación binaria, donde la entrada es la información del paciente descrita en la parte anterior y la salida puede ser: sí muestra respuesta o no muestra respuesta.

### 3.1.3. Modelo

El clasificador basado en Dempster-Shafer, optimizado con Descenso de Gradiente (DSGD, por sus siglas en inglés), de Sergio Peñafiel [7] muestra un buen desempeño, comparable con otros modelos como *Support Vector Machine* o el algoritmo *K-nearest neighbors*. Además, el algoritmo DSGD entrega la libertad de definir reglas a mano o dejar que se generen automáticamente. En esta sección, algunas reglas básicas fueron configuradas de forma manual

para evaluar dependencias cruzadas, pero la mayoría fueron determinadas en el proceso de entrenamiento por el mismo algoritmo. El algoritmo fue entrenado usando el error cuadrático medio como función de pérdida y el algoritmo ADAM [35] como optimizador, con una tasa de aprendizaje de 0,0015. Estos hiperparámetros fueron determinados eligiendo valores de forma aleatoria, evaluando el desempeño en cada combinación mediante el método de validación cruzada.

### 3.1.4. Resultados

#### Desempeño del modelo

Para tener una evaluación más confiable del desempeño del modelo, evitar el sobreajuste de los datos y que la evaluación sea independiente de la partición de estos, se usa la técnica de validación cruzada [36]. Esta técnica implica entrenar y evaluar el modelo combinando diferentes sub conjuntos de los datos, para evaluar su capacidad de generalización. Es este estudio, se realiza una validación cruzada con 3 conjuntos, lo que significa que el 66.6 % de los datos es usado para entrenamiento y el resto para testeo. Luego, se repite el proceso 3 veces, cambiando cada vez el tercio que se usa para testeo.

La tabla 3.1 muestra el desempeño medio del modelo después de los tres experimentos, además la figura 3.1 muestra la matriz de confusión conjunta de las 3 iteraciones. Es posible notar una mayor certeza al clasificar correctamente a pacientes con respuesta al tratamiento, sin embargo, también se tiene una alta tasa de pacientes que el modelo dice que tendrán respuesta, sin embargo, no la tuvieron. Es tipo de error podría afectar la calidad de vida de un paciente, ya que implicaría someterlo a un tratamiento invasivo el cual no le servirá en para mitigar el cáncer. Por otro lado, cuando se toman estas importantes decisiones, los expertos consideran otros factores, además de la sugerencia de modelos matemáticos.

Indicador	Valor
Exactitud	0.61
AUC ROC	0.61
F1	0.69
Precisión	0.63
Exhaustividad	0.76

Tabla 3.1: Desempeño del modelo DSGD.

Para entender qué tan buen desempeño tiene el modelo anterior, se llevó a cabo el mismo proceso con otros algoritmos de clasificación, cuyos hiperparámetros son ajustados usando la técnica de *GridSearchCV* [37]. Como resultado de lo anterior, lo modelos usados son:

1. Árbol de decisión: criterio entropía y profundidad máxima 3.
2. *K-nearest neighbors (KNN)*:  $K = 21$ .

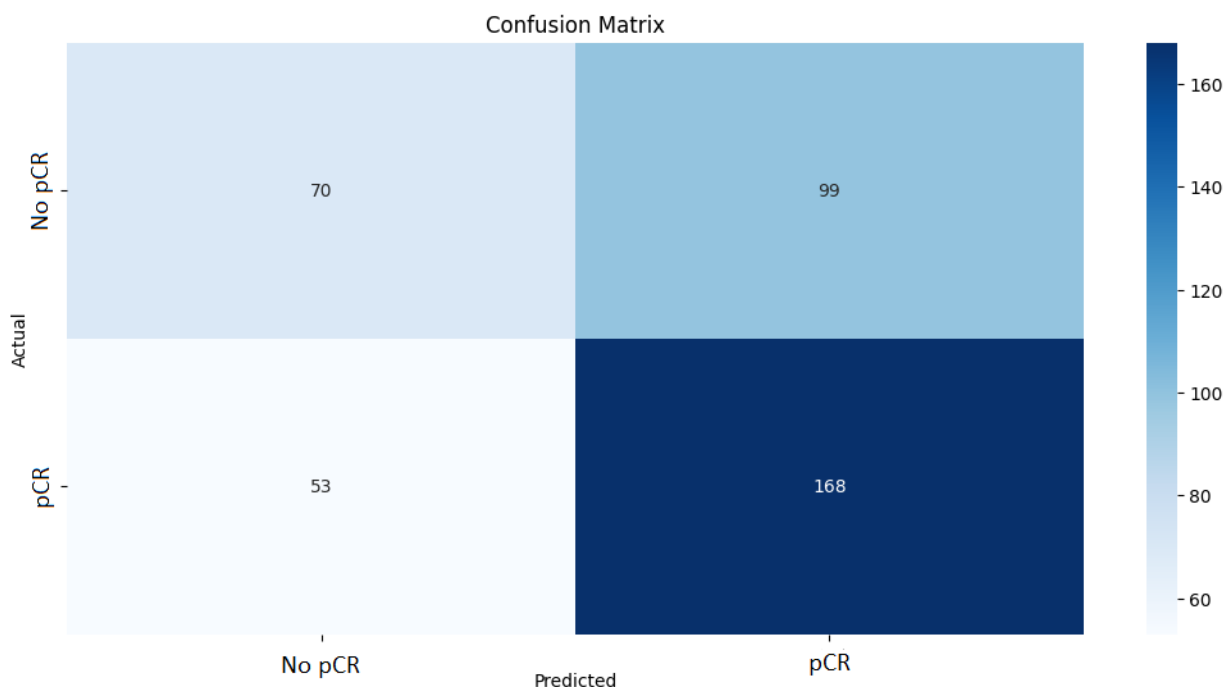


Figura 3.1: Matriz de confusión después de validación cruzada.

3. *Support Vector Machines (SVM): kernel* polinomial de grado 1.
4. *XGBoost: gmtree booster*.

La tabla 3.2 presenta el desempeño del modelo DSGD y los otros modelos evaluados, cuyas métricas son la media obtenida a partir de los experimentos de validación cruzada.

Model	Exactitud	AUC ROC	F1
DSGD	0.61	0.61	0.69
Árbol de decisión	0.52	0.53	0.54
KNN	0.52	0.49	0.64
SVM	0.56	0.56	0.72
XGBoost	0.56	0.59	0.60

Tabla 3.2: Rendimiento de todos los métodos en la predicción de la respuesta de los pacientes.

Además, la figura 3.2 muestra las curvas *ROC* de los modelos entrenados y evaluados, ilustrando el *trade-off* entre la tasa de verdaderos positivos (sensibilidad o recuerdo) y la tasa de falsos positivos (especificidad) en diferentes umbrales de clasificación.

A partir de los resultados presentados, es posible observar que el desempeño del modelo DSGD es moderado, probablemente debido a lo pequeño que es el dataset y en menor medida al desbalance en las clases. Sin embargo, para este problema, el modelo ha demostrado el mejor desempeño en precisión y la curva *ROC*, en comparación con los otros modelos. Presenta

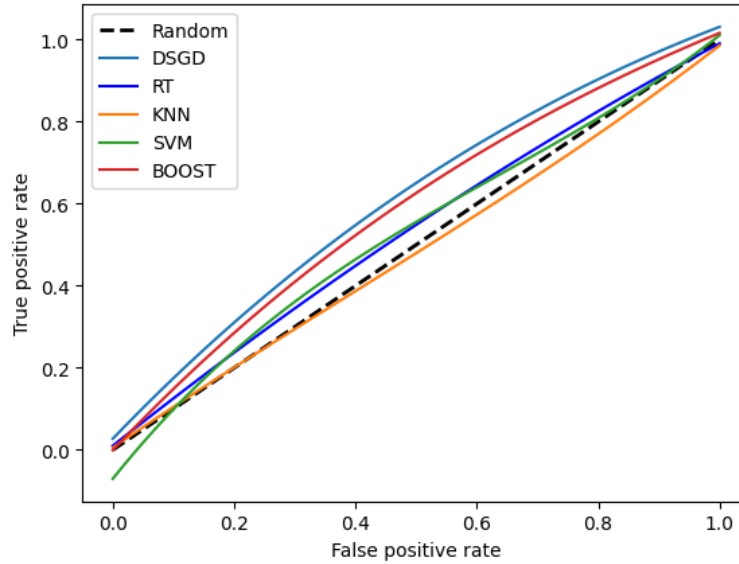


Figura 3.2: Curvas ROC promedio para los métodos utilizados en el problema.

métricas incluso mejores que las de un clasificador binario conocido por su alto rendimiento, como lo es SVM, pero con la gran ventaja de ser un algoritmo interpretable.

A partir de la matriz de confusión, es posible obtener varias métricas para evaluar el rendimiento del modelo en la clasificación de cada clase. Debido al desequilibrio de clases, es preferible utilizar la puntuación F1 para comparar los rendimientos. Por un lado, la puntuación F1 para *pCR* es aproximadamente 0.7, mientras que, por otro lado, la puntuación F1 para *No pCR* es aproximadamente 0.5. Por lo tanto, el modelo demuestra un mejor rendimiento de clasificación para los pacientes con *pCR*.

## Reglas generadas

Además de su desempeño, el modelo es capaz de identificar la importancia de distintas reglas al momento de predecir un resultado. Dichas reglas representan distintos valores que pueden tomar los datos de entrada, es decir, distintas características de los pacientes. Durante la fase de entrenamiento, las masas de las reglas son ajustadas con tal de reducir el error en la predicción. A cada regla se le asigna una masa para cada posible salida, y una incertidumbre, donde una masa mide la contribución de una regla a una salida.

Las tablas 3.3 y 3.4 muestran las reglas más importantes, según el modelo DSGD entrenado, al momento de predecir la respuesta al tratamiento de un paciente:

Los resultados anteriores sugieren que los pacientes con una alta expresión de receptores de estrógeno (ER) y progesterona (PR), bajos niveles de índice Ki-67 y un bajo grado de cáncer de mama, son menos propensos a mostrar una respuesta completa patológica (*pCR*) al tratamiento neoadyuvante. Esto se puede ver en la tabla 3.3, donde la masa de dichas reglas favorece clasificar a los pacientes sin una respuesta favorable al tratamiento.

Además, las reglas presentadas también sugieren que un estado premenopausico y una

Regla	Masa No pCR	Masa pCR	Inc
ER $\geq$ 20 % and PR $\geq$ 20 %	0.279	0.000	0.721
Ki-67 < 20 % and ER $\geq$ 20 %	0.259	0.000	0.741
Grado 1	0.231	0.000	0.769
ER $\geq$ 20 %	0.206	0.000	0.794

Tabla 3.3: Reglas más importantes para No pCR y su incertidumbre (Inc). Ver glosario para entender las siglas.

Regla	Masa No pCR	Masa pCR	Inc
Estado menopausico	0.000	0.244	0.756
Etapas III	0.000	0.206	0.794
cN = cN2	0.000	0.206	0.794
cT = cT1	0.000	0.204	0.796

Tabla 3.4: Reglas más importantes para pCR y su incertidumbre (Inc). Ver glosario para entender las siglas.

etapa avanzada de cáncer, tienden a alcanzar una pCR. Además, los resultados sugieren una tendencia hacia el logro de pCR después del tratamiento en pacientes con cN1 y cT2 en las categorías cN y cT, respectivamente.

### Comparación con árbol de decisión

El Árbol de decisión fue uno de los métodos con desempeño más bajo en el grupo, sin embargo, es otro método directamente interpretable, además del DSGD. La figura 3.3 muestra el mejor árbol de decisión obtenido en la validación cruzada (exactitud de 0.55). Cada nodo en el árbol muestra la regla de decisión del modelo, el valor de entropía, la proporción de las muestras utilizadas para generar la división, la proporción de la división ( $[p_{\text{No pCR}}, p_{\text{pCR}}]$ ), y la clase asignada si el algoritmo tiene que tomar una decisión en este nodo. La clase predicha en un nodo se representa mediante su color: naranja para *No pCR* y azul para *pCR*, donde colores más oscuros significan una predicción más segura. Si la condición en el nodo es verdadera (*true*), el algoritmo sigue el ramo de la izquierda; por otro lado, si es falsa (*false*), sigue el ramo de la derecha. Es importante considerar que algunos atributos se normalizaron utilizando la técnica de codificación one-hot; por lo tanto,  $PR > 0,5$  significa receptor de progesterona mayor o igual al 20 % (lo mismo ocurre con *ER* y *Ki - 67*).

El resultado muestra reglas similares que se pueden leer en las tablas 3.3 y 3.4. Por ejemplo, los pacientes con niveles bajos de progesterona, receptor de estrógeno y alta presencia de la proteína *Ki-67* tienden a mostrar *pCR*. Sin embargo, las reglas sobre la edad de los pacientes no son mostradas por el modelo DSGD.



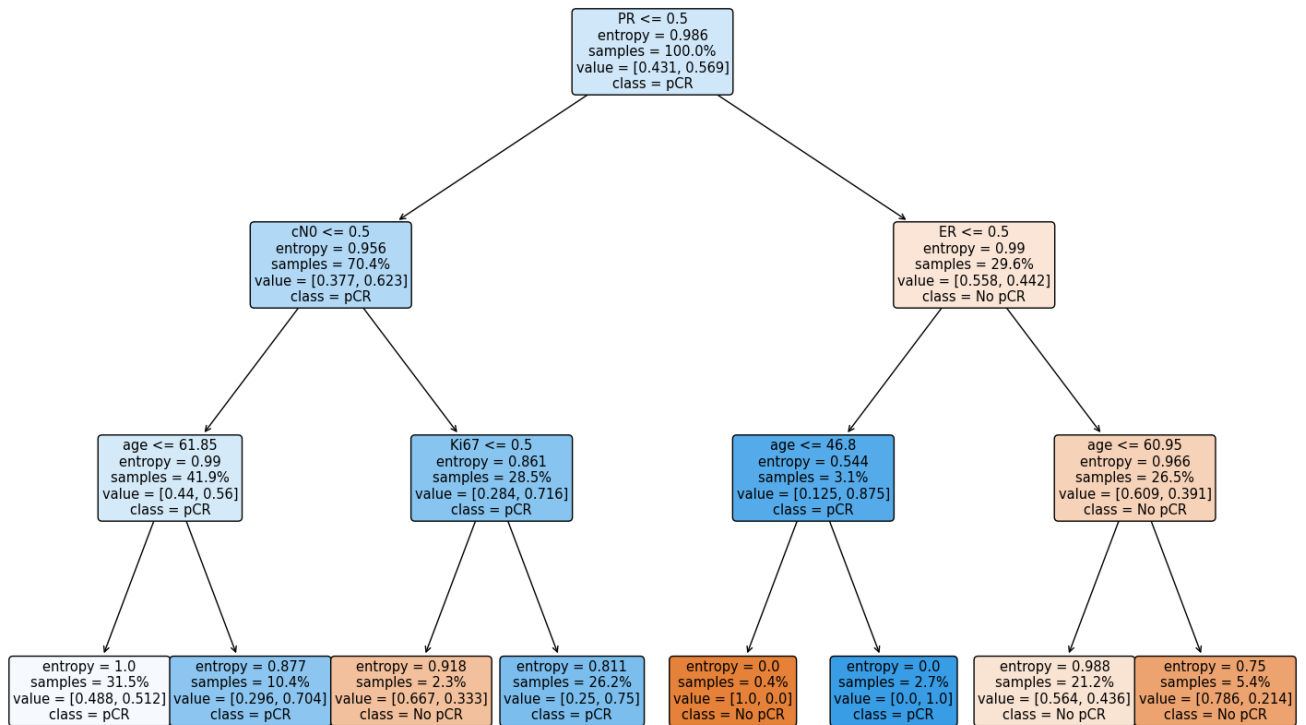


Figura 3.3: Árbol de decisión resultante. Ver glosario para entender las siglas.

## Discusión

El modelo DSGD supera a los otros modelos en términos de rendimiento. Sin embargo, aún hay margen de mejora. Una forma de mejorar su rendimiento es incorporar más datos. Se ha demostrado que más datos y más casos de estudio implican un mejor rendimiento en los algoritmos de aprendizaje automático [36], tener un conjunto de datos más robusto y diverso puede favorecer el encontrar patrones y mejorar las predicciones. Además, agregar más características, como aquellas que describen la amplificación del gen HER2 en el cáncer de mama, como la relación HER2/centrómero y el número de copias de HER2, que han demostrado un papel predictivo en la respuesta completa patológica (pCR) en estudios previos [38], podría mejorar aún más el rendimiento de nuestro modelo.

Una característica destacada del modelo presentado es su interpretabilidad, que es fundamental para problemas de esta naturaleza. La interpretabilidad del modelo permite una comprensión más fácil de las reglas subyacentes, facilitando a los médicos utilizar este conocimiento en la práctica.

Las reglas resultantes obtenidas después de entrenar el modelo ofrecen algunas ideas interesantes para la discusión. Uno de los hallazgos más significativos es que es menos probable que los pacientes con tumores positivos para los receptores de estrógeno y progesterona muestren pCR después del tratamiento, como se muestra en la tabla 3.3. Considerando la discusión

presentada en Saffie et al. [4], la presencia de positividad de los receptores hormonales indica un tumor más luminal. Además, la heterogeneidad del cáncer de mama HER2/luminal es un predictor de una menor respuesta a las terapias neoadyuvantes. Otro estudio relevante realizado por Schneeweiss et al. [39] respalda estos hallazgos al demostrar una mayor tasa de pCR en pacientes con tumores negativos para los receptores hormonales en comparación con aquellos con tumores positivos para los receptores hormonales. Por lo tanto, los resultados obtenidos son consistentes con investigaciones previas.

Otra regla interesante es que es menos probable que los pacientes con un valor de Ki-67  $< 20\%$ , además de ER  $\geq 20\%$ , logren una respuesta patológica completa (pCR). Esta característica también se discute en [4], donde se afirma que los pacientes con un valor de Ki-67  $> 20\%$  tienen más probabilidades de tener una pCR en comparación con aquellos con un valor de Ki-67  $< 20\%$ . Además, se ha establecido que los tumores con una alta expresión de Ki-67 son más agresivos y muestran una mayor respuesta tanto a tratamientos citotóxicos como dirigidos [40].

Además de las reglas mencionadas anteriormente, la tabla 3.3 indica que los tumores de grado bajo o bien diferenciados no muestran una respuesta patológica completa al tratamiento. Aunque hay evidencia limitada que respalde esta observación, se puede hipotetizar que la falta de pCR en los tumores de grado bajo puede atribuirse a su naturaleza menos agresiva, lo que resulta en una menor respuesta a este tratamiento en particular.

En la tabla 3.4, se muestra que un cáncer en etapas más avanzadas tiende a mostrar pCR, lo cual es consistente con los hallazgos de Saffie et al. [4]. Este resultado se puede explicar considerando que un cáncer más avanzado tiende a ser un tumor menos luminal, lo que lleva a una mayor probabilidad de lograr pCR.

Además, la última regla presentada sugiere que los tumores más pequeños tienden a mostrar una respuesta patológica completa después del tratamiento neoadyuvante. Es importante aclarar que si bien el tamaño del tumor puede ser un factor para determinar la etapa del cáncer, no es el único factor determinante. También se utilizan otras variables y criterios para evaluar la etapa de un tumor. A pesar de que el modelo indica que el tamaño del tumor es un factor importante en este tipo de tratamiento, la literatura existente contradice esta noción, afirmando que el tamaño del tumor no es un predictor confiable de pCR [32, 33].

En la misma tabla 3.4, hay otra regla importante a considerar que indica que las mujeres en estado premenopáusico tienden a tener una respuesta patológica completa al tratamiento neoadyuvante. Esta regla genera controversia, ya que actualmente no hay suficiente evidencia para respaldarla. Desafortunadamente, el conjunto de datos disponible no brinda la oportunidad de analizar el perfil molecular del cáncer y establecer una conexión entre los estados premenopáusicos y los tipos de tumores menos luminales. En consecuencia, se requieren estudios adicionales para investigar cualquier relación potencial entre los estados premenopáusicos y la pCR en los tratamientos neoadyuvantes.

## Conclusiones

En esta sección, se presenta un modelo novedoso para predecir la efectividad de una terapia neoadyuvante basado en variables clínicas e histopatológicas. El modelo supera a todos los demás modelos probados, lo que indica su superioridad en predecir con precisión la respuesta completa patológica. Sin embargo, aún existe potencial para mejorar el modelo mediante la adición de más pacientes e integrando nuevas características.

Las reglas derivadas del proceso de entrenamiento revelan el impacto de ciertas variables en la respuesta patológica completa (pCR) de pacientes con cáncer de mama. Algunas de estas variables, como la expresión de receptores hormonales y las etapas del cáncer, tienen respaldo directo en investigaciones anteriores como predictores de pCR. Otras reglas, como la asociación con el estado premenopáusico o los grados del cáncer, no tienen una explicación clara. Por lo tanto, se requieren más datos y análisis para establecer una correlación definitiva. No obstante, todas las reglas indican consistentemente que la heterogeneidad del cáncer de mama sirve como indicador de una menor probabilidad de lograr pCR.

### 3.2. Clasificador interpretable en pacientes con esquemas diferentes

En esta parte se presentan los pasos a seguir para obtener el clasificador interpretable que es usado en la herramienta final por los oncólogos. La idea es que al usar la herramienta, a partir de las características de un paciente, se obtengan dos resultados: la probabilidad de la respuesta usando el esquema simple y la probabilidad de respuesta usando el esquema doble.

Para esta parte se considera manejar el tipo de esquema monoclonal usado en el tratamiento como variable de entrada en el modelo, así es posible evaluar la influencia del esquema en la respuesta al tratamiento, con el fin de identificar y caracterizar a los pacientes que pueden prescindir del Pertuzumab.

Para esta parte, se usa un conjunto de datos más reducido que en la sección anterior, ya que los datos usados en la parte anterior eran pacientes que solo habían recibido esquema completo en su tratamiento neoadyuvante, en base a Trastuzumab y Pertuzumab, no hay pacientes con un solo fármaco en su tratamiento. Por cortesía de la FALP, es posible tener acceso a datos de sus pacientes, dentro de los cuales se especifica el tipo de esquema que usan en el tratamiento, algunas variables interesantes como las mencionadas en la parte anterior y otras que serán descritas en esta sección.

#### 3.2.1. Descripción de los datos

En esta parte del trabajo se usan 2 grupos de pacientes. Por un lado, los que fueron sometidos a un tratamiento neoadyuvante solo con Trastuzumab, y por otro, los que tuvieron

un tratamiento con ambos anticuerpos monoclonales: Trastuzumab y Pertuzumab. Ambos conjuntos de datos son de pacientes tratados en Chile y fueron facilitados por la FALP.

Las características clínicas e histopatológicas consideradas en esta parte son: edad, estado menopaúsico, IMC, categorías T y N (describen tamaño y propagación a ganglios, respectivamente), etapa, involucramiento axilar, tipo histológico, receptores de progesterona y estrógenos, índice Ki-67, relación de hormona HER2- centrómero, número de copias o amplificación de HER2 y perfil molecular. Además de las anteriores, se tiene qué tipo de esquema tuvo el paciente durante el tratamiento, y además si llegó a no a mostrar una respuesta completa patológica al tratamiento. Ver distribución de datos en Anexo B.

Luego de hacer una limpieza de los datos y filtrar los pacientes que aportan información para poder entrenar los modelos, se cuentan con 142 pacientes, con una edad media aproximada de 54 años. De estos pacientes, el 61% tuvieron respuesta completa patológica al tratamiento neoadyuvante. De ese 61%, un 76% son pacientes que fueron sometidos a un esquema con los dos fármacos y el otro 24% solo recibió uno (ver figura 3.4).

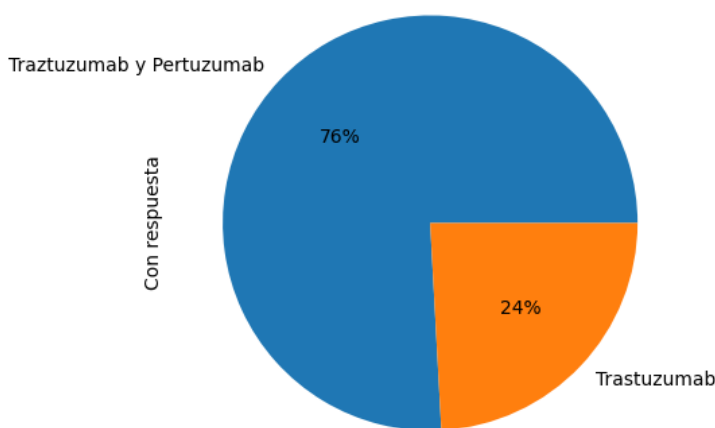


Figura 3.4: Distribución de esquema administrado en pacientes que obtuvieron respuesta.

### 3.2.2. Planteamiento del problema

Aquí, al igual que en la sección anterior, se busca crear un modelo que en base a las características de un paciente, pueda determinar si va o no a tener una respuesta al tratamiento. Pero en esta oportunidad se considera otra variable en la clasificación: el tipo de esquema monoclonal al cual es sometido un paciente. A partir de esta información, entrenando un modelo interpretable, se busca poder caracterizar a este 24% de pacientes que reciben solo Trastuzumab y obtienen una respuesta al tratamiento neoadyuvante.

### 3.2.3. Modelo

El modelo a usar es el mismo que en la sección anterior, el DSGD, llegando a los mismo hiperparámetros. En esta oportunidad se agregan más reglas personalizadas, donde se cruzan todos los posibles valores de las variables con las dos opciones de esquema posible. Por ejemplo, se agrega como regla que el paciente tenga un cáncer en cualquiera de las 3 etapas posibles y además reciba un esquema con Trastuzumab o con Pertuzumab, así se agregan 6 reglas (2 por cada una de las 3 etapas). Esto con el fin de encontrar una relación entre esquema y características clínicas que pueda ser determinando a la hora de identificar a los pacientes que pueden tener una respuesta al tratamiento solo con un fármaco.

### 3.2.4. Resultados

#### Desempeño del modelo

Se usa la técnica de validación cruzada para tener una evaluación más confiable y garantizar que los resultados son independientes de la partición de los datos. Para esta parte, se realiza la validación cruzada partiendo el conjunto de datos en 4 partes, por lo que el 75 % es usado para entrenamiento, y el 25 % para testear el modelo entrenado. Se realiza una partición diferente, ya que al realizar diversos experimentos con distintas particiones (3, 4 y 5 partes), dividir en 4 partes fue la que obtuvo mejores métricas promedio (mayor exactitud, precisión y F1).

En la tabla 3.5 se presenta el desempeño del modelo DSGD luego de realizar la validación cruzada.

Indicador	Valor
Exactitud	0.62
AUC ROC	0.66
F1	0.67
Precisión	0.71
Exhaustividad	0.65
Entropía cruzada	0.67

Tabla 3.5: Desempeño del modelo DSGD.

Además, la matriz de confusión de la figura 3.5 representa las clasificaciones acertadas hechas por el modelo en los conjuntos de testeo. Es posible notar que hay mayor precisión en predecir correctamente los pacientes que tienen respuesta al tratamiento, que al clasificar correctamente los que no tendrán respuesta (aunque ambos cuadrantes tienen la mayor certeza). Al clasificar con mayor certeza que los pacientes tienen respuesta, se puede decir que el clasificador es más confiable al decir que un paciente tendrá respuesta. Sin embargo, también se aprecia una tasa considerable de pacientes que se clasifican erróneamente como pacientes sin respuesta, lo cual presenta un caso peligroso, ya que si la decisión solo dependiera del modelo, hay chance de que un paciente no reciba el tratamiento adecuado. En estos casos

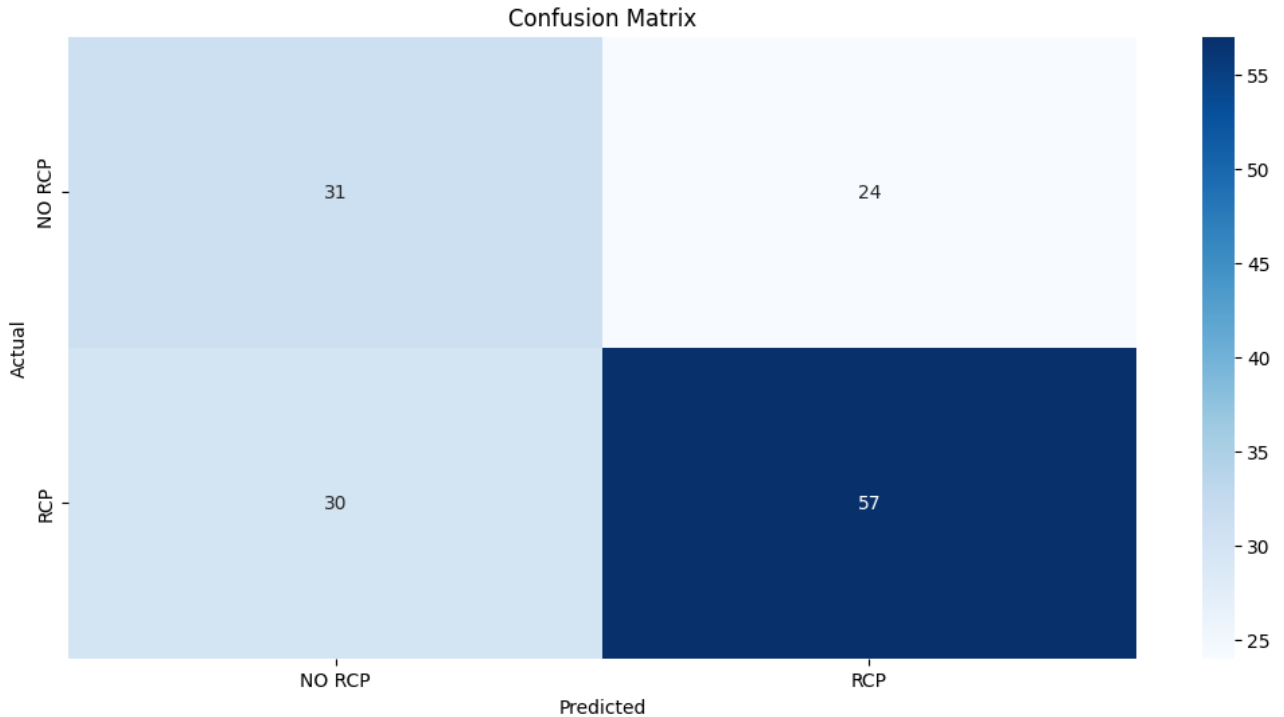


Figura 3.5: Matriz de confusión después de validación cruzada.

hay que tener en cuenta que al decidir qué tratamiento recibirá un paciente, se consideran muchos factores, no solo la estadística de una sola referencia.

Para tener un contraste, se realiza el mismo proceso de validación cruzada con otros modelos, al igual que en la sección anterior. Se ajustan los hiperparámetros mediante la técnica de *GridSearchCV*. Como resultado la configuración usada es:

1. *Árbol de decisión*: criterio entropía y profundidad máxima 3.
2. *K-nearest neighbors (KNN)*:  $K = 21$ .
3. *Support Vector Machines (SVM)*: *kernel* polinomial de grado 1.
4. *XGBoost*: *gbtree booster*.

La tabla 3.6 presenta el desempeño de todos los modelos evaluados, las métricas mostradas son el promedio de los 4 experimentos.

En la figura 3.6 se puede ver la curva ROC, la cual muestra que tan sensibles son los modelos a los falsos negativos. Nuevamente el modelo DSGD muestra ser el con mejor capacidad de discriminación entre verdaderos positivos y falsos positivos.

El modelo central de estudio es el que muestra el mejor desempeño en términos generales. En este caso, la cantidad de datos es menor que en la sección anterior, razón que explica los resultados. Es un problema complejo, ya que el modelo podría no tener suficiente información para poder establecer las relaciones y poder diferenciar las clases de mejor manera. Al realizar

Model	Exactitud	AUC ROC	F1
DSGD	0.62	0.66	0.67
Árbol de decisión	0.58	0.59	0.65
KNN	0.53	0.53	0.60
SVM	0.54	0.56	0.60
XGBoost	0.59	0.61	0.67

Tabla 3.6: Rendimiento de todos los métodos en la predicción de la respuesta de los pacientes.

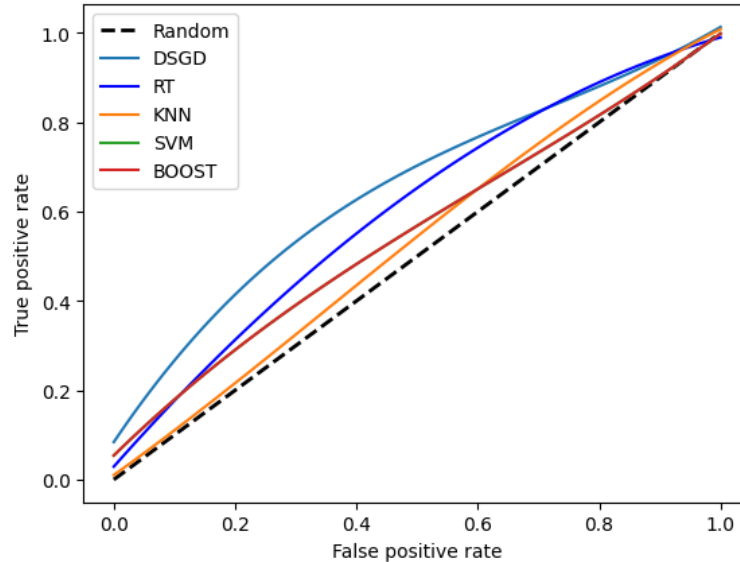


Figura 3.6: Curvas ROC promedio para los métodos utilizados en el problema.

la evaluación de diversos modelos y revisar los resultados, podemos notar que el desempeño no mejora demasiado si se usa un modelo u otro, por lo que muy probablemente se deba a la baja cantidad de muestras que se tienen para el entrenamiento.

Calculando la métrica F1-score para cada una de las clases, se obtiene un valor aproximado de 0.7 para la clase *pCR* y un 0.5 para *No pCR*. Es el mismo valor obtenido para los experimentos con el conjunto de datos de la sección anterior, mostrando además que este modelo entrenado también tiene un mejor desempeño en la clasificación de pacientes con *pCR*.

### Reglas generadas

Como ya ha sido explicado, el modelo entrega cierto peso o masa a las reglas, y en entrenamiento va a ajustando estos valores hasta alcanzar un óptimo. Posterior al proceso de entrenamiento, considerando las masas, se puede hacer un *ranking* de las masas más importantes. Las tablas 3.7 y 3.8 muestran dichas reglas.

La incertidumbre de las reglas resultantes es bastante menor que en la sección anterior,

Regla	Masa No pCR	Masa pCR	Inc
N° de copias de HER2 $\leq 10$	0.311	0.000	0.689
Etapa 3 y recibe esquema T	0.297	0.000	0.703
cN = cN2	0.284	0.000	0.716
Ki-67 $\geq 20\%$ y recibe esquema T	0.266	0.000	0.734
Etapa 2 y recibe esquema T-P	0.256	0.000	0.744
Tipo histológico CLI	0.253	0.000	0.747
El paciente tiene mas de 50 años y recibe esquema T	0.250	0.000	0.750
Relación HER2/CEN $< 6$	0.245	0.001	0.754
HER2 Puro y recibe esquema T	0.244	0.00	0.756
cT = cT4	0.235	0.000	0.765
tipo histologico = CDI y recibe esquema T	0.222	0.000	0.778
Ki-67 $\leq 20\%$	0.218	0.000	0.782

Tabla 3.7: Reglas más importantes para No pCR y su incertidumbre (Inc). Ver glosario para entender las siglas.

Regla	Masa No pCR	Masa pCR	Inc
cT = cT4 y recibe esquema T-P	0.000	0.382	0.618
ER $\leq 20\%$	0.000	0.272	0.728
cN = cN3 y recibe esquema T-P	0.000	0.272	0.728
HER2 Luminal y recibe esquema T	0.000	0.265	0.735
cN = cN3	0.000	0.253	0.747
cN = cN3 y recibe esquema T	0.000	0.252	0.748
Etapa 2 y recibe esquema T	0.000	0.247	0.753
cN = cN1 y recibe esquema T	0.000	0.240	0.760
Sin involucramiento axilar	0.000	0.223	0.777
N° de copias de HER2 $> 10$	0.000	0.225	0.775
Tipo histológico = CLI y recibe esquema T-P	0.000	0.220	0.780
HER2 Puro y recibe esquema T-P	0.000	0.216	0.784

Tabla 3.8: Reglas más importantes para pCR y su incertidumbre (Inc). Ver glosario para entender las siglas.

eso puede ser explicado porque se agregaron reglas más complejas, intersectando las características del paciente con el tratamiento al que fue sometido. Sin lugar a duda, el modelo indica que la respuesta de un paciente al tratamiento depende fuertemente del esquema al que fue sometido, además de algunas características del mismo.

En las tablas 3.7 y 3.8 se puede observar que los pacientes con n° de copias de HER2 bajo, es decir, con baja amplificación del gen HER2, no tienden a tener respuesta al tratamiento neoadyuvante. Lo mismo ocurre con la relación HER2/CEN baja. Estas reglas contienen información innovadora, ya que aportan datos que no estaban presentes en el estudio anterior (debido a que son variables que no existían en dicho conjunto de datos) y están respaldadas por un estudio previo realizado por Saffie et al. [4]. Saffie concluye en su estudio que los pa-



cientes con presencia de HER2 tienden a tener una mayor respuesta, ya que esto se relaciona con un cáncer más homogéneo o HER2 puro. No obstante, a pesar del argumento anterior, el modelo indica que los pacientes con un perfil molecular HER2 puro tienden a no tener una respuesta cuando son tratados solo con Trastuzumab, pero si reciben un esquema completo, son más propensos a mostrar una respuesta completa patológica a la terapia.

El modelo también sugiere que un cáncer Luminal, es decir, uno más heterogéneo, tiende a tener una respuesta al tratamiento si es el paciente es tratado solo con Trastuzumab.

Una regla interesante es que pacientes en Etapa 3 y con un esquema que comprende la administración solo de Trastuzumab, tienden a no tener respuesta completa patológica al tratamiento. En la experiencia anterior se obtuvo que pacientes en Etapa 3 mostraban respuesta, sin embargo eran pacientes que recibían el esquema completo. Esta regla sugiere que pacientes en etapa avanzada pueden no mostrar respuesta si es que son sometidas a un esquema solo con Trastuzumab. Por lo tanto, en estos casos la herramienta sugeriría si o si administrar un esquema monoclonal completo.

Además, el modelo señala que pacientes con índice Ki-67  $\geq 20\%$  y un tratamiento basado solo en Trastuzumab, no tienden a una respuesta al tratamiento, a pesar de ser más agresivos. Lo cual, al igual que la regla anterior, indica que es más probable que un paciente con Ki-67  $\geq 20\%$  muestre una respuesta si es que es sometido a un esquema completo. Pero, además se indica que pacientes con Ki-67  $< 20\%$ , independiente de recibir uno o los dos fármacos, es probable que no muestren una respuesta al tratamiento.

## **Análisis de cambio de esquema**

Es interesante evaluar la influencia del esquema empleado en la respuesta de los pacientes, según el modelo entrenado. Para esto, se realiza el ejercicio de cambiar el esquema empleado los pacientes del grupo de testeo, realizar la predicción y detectar cuales pacientes cambiaban su respuesta al tratamiento. La figura 3.7 muestra que un 29.58% de los casos evaluados cambian su respuesta al tratamiento al variar el esquema al que son sometidos.

En la tabla 3.9 se ve más en detalle cual es el cambio de esquema realizado en los grupos que cambiaron la respuesta, por ejemplo, qué porcentaje de verdaderos positivos (cuadrante inferior derecho) pasó a ser clasificado como paciente sin respuesta al cambiar el tratamiento de un esquema completo a un esquema parcial.

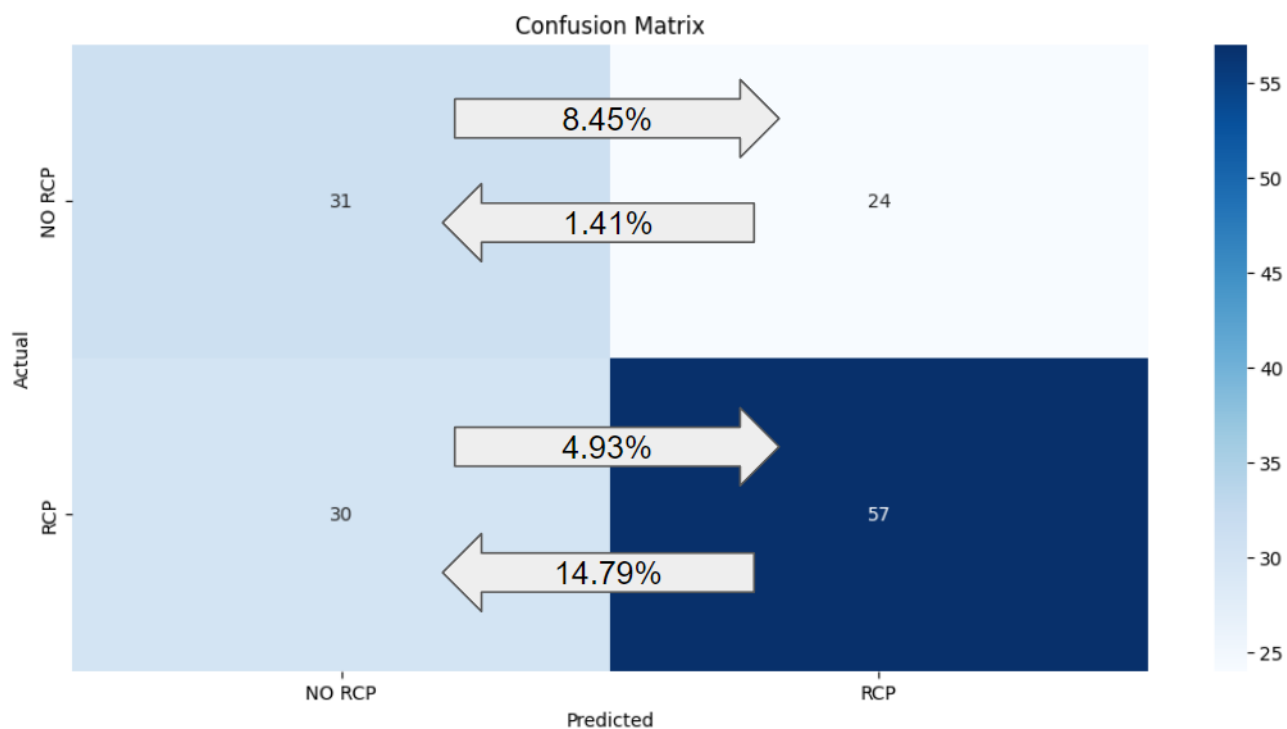


Figura 3.7: Matriz de confusión anterior luego de cambiar el esquema de tratamiento.

Porcentaje (%)	Movimiento	Cambio esquema
1.41	VP → FN	T → T-P
13.38	VP → FN	T-P → T
0.71	FP → VN	T → T-P
0.7	FP → VN	T-P → T
6.34	VN → FP	T → T-P
2.11	VN → FP	T-P → T
2.82	FN → VP	T → T-P
2.11	FN → VP	T-P → T

Tabla 3.9: Detalle de cambio de clasificación al cambiar esquema. Donde VP: Verdadero Positivo, FN: Falso Negativo, VN: Verdadero Negativo, FP: Falso Positivo, T-P: Trastuzumab - Pertuzumab y T: Trastuzumab

Uno de los cambios más con mayor tasa de pacientes representa uno de los casos más críticos, que son los pacientes que cambian de VP a FN al cambiar de esquema completo (T-P) a esquema parcial (T). Es decir, los pacientes que tienen efectivamente una respuesta al ser tratados con ambos fármacos, según el modelo, no tendrían una respuesta al tratamiento si es que le suspendo el Pertuzumab (13.38 %)

Otro caso interesante es el de los pacientes que son clasificados como pacientes sin respuesta completa al tratamiento con esquema parcial, pero al cambiar a esquema completo el modelo determina que podrían tener una respuesta (6.38 %).

### 3.2.5. Conclusiones

El modelo DSGD nuevamente es el con mejor desempeño en la tarea de clasificación binaria. Sin embargo, este también tiene métricas que abren la puerta a una mejoría, ya sea con más datos o mayor exploración por parte de expertos para ingresar reglas más complejas.

Por otro lado, el gran valor de usar DSGD es, además de su desempeño, el conocimiento que se puede extraer de los datos. Información que es relevante para poder caracterizar a los pacientes que podrían prescindir del uso de Pertuzumab y quienes no.

El modelo muestra de manera explícita que los pacientes que pueden prescindir de Pertuzumab para tener una respuesta al tratamiento son: pacientes con un cáncer luminal, con  $cN = cN3$  o  $cN = cN1$ , y pacientes en Etapa 2. Además se puede inferir que los pacientes con bajos niveles de receptores de estrógeno, sin involucramiento axilar y con el gen HER2 amplificado, también son más propensos a tener una respuesta al tratamiento.

Por otro lado, también es posible identificar características de los pacientes que indican que sí o sí deben ser sometidos a un esquema completo para tener respuesta: pacientes con tumor clasificado como  $cT = cT4$  o  $cN = cN3$ , tipo histológico CLI y perfil molecular HER2 puro.

Además, analizar la influencia del esquema suministrado a los pacientes en la respuesta del modelo es reveladora, indicado que en casi un tercio de los casos de prueba la respuesta de los pacientes es sensible al cambio de esquema. Donde hay una tendencia clara del modelo a que los pacientes que son sometidos a un tratamiento con ambos fármacos tienen mayor probabilidad de obtener una respuesta al tratamiento. Es menor la cantidad de casos en que el modelo indica que al quitarle el Pertuzumab a un paciente va a pasar de no tener una respuesta a tenerla. Lo cual se condice con la muestra utilizada y la experiencia misma de los especialistas de la FALP. Por lo que identificar las características de este pequeño grupo es sumamente desafiante e importante.

Este modelo encapsula información relevante para el estudio de la efectividad de este tipo de tratamientos en pacientes con cáncer de mama HER2-positivo. Al DSGD presentado en esta sección fue validado con profesionales de la FALP, donde la participación de Isabel Saffie fue sumamente importante para comprender el valor de las relaciones expresadas en las reglas más importantes. Por lo tanto, este es el modelo que se incluye en la herramienta computacional que servirá a los oncólogos como referencia al momento de decidir qué esquema seguirá el paciente.

## 3.3. Desarrollo del *software*

En esta sección, se detalla el proceso de construcción de la herramienta computacional, abordando los lenguajes y *frameworks* utilizados, la estructura y el diseño, el modelo de datos, la gestión de los datos y el acceso a los mismos.

Antes de describir el *software*, es fundamental definir los requisitos de la herramienta. Los requerimientos identificados son los siguientes:

1. Interfaz de tipo formulario para ingresar datos de pacientes.
2. Base de datos escalable que permita el registro de modelos y su expansión futura.
3. API para conectarse a la base de datos mencionada anteriormente.
4. Capacidad de conexión con la API de la FALP para obtener información de los pacientes

Con estas consideraciones en mente, se ha desarrollado la herramienta computacional que se describe a continuación.

### 3.3.1. Arquitectura del *software*

La arquitectura básica del software se basa en un enfoque cliente-servidor, en el cual el cliente corresponde a una aplicación desarrollada en VueJS, un framework de JavaScript ampliamente utilizado. Por su parte, el servidor está implementado como una API REST en Django.

Desde la perspectiva del cliente, la aplicación sirve como punto de entrada para los usuarios, quienes interactúan con ella mediante una interfaz intuitiva. A través de esta interfaz, los usuarios pueden ingresar datos en formularios y realizar acciones al hacer clic en botones. Por otro lado, el servidor despliega un papel crucial al recibir y procesar las solicitudes enviadas por el cliente (Vue). Entre las operaciones que realiza, se encuentra la de enviar datos a la base de datos, extraer información de PostgreSQL y ejecutar diversas operaciones según lo requerido.

Lo anterior describe las partes desarrolladas en este trabajo. Sin embargo, es importante destacar que una de estas partes necesita comunicarse con una API externa encargada de administrar los datos de pacientes de la FALP.

En primera instancia se decide resolver este problema siguiendo la arquitectura mostrado en la figura 3.8.

En la figura 3.8 se presentan 2 fuentes de datos: una es específica para el presente trabajo, y contiene información sobre los modelos de machine learning utilizados para la predicción; esta fuente es administrada por una API que será detallada en las próximas secciones. La otra fuente de datos es la información de los pacientes de la FALP, a la cual se puede acceder mediante peticiones a la API del servidor de la FALP.

El cliente establece comunicación con ambos servidores. Para obtener los datos de los pacientes, el cliente realiza consultas a la API de la FALP utilizando el rut como referencia, y rellena el formulario con esta información. Posteriormente, esta información se envía a la API encargada de administrar los modelos predictivos, y se recibe una respuesta (predicción) que se muestra al usuario.

Sin embargo, al utilizar esta estructura, surgieron algunos problemas relacionados con el Intercambio de Recursos de Origen Cruzado (CORS, por sus siglas en inglés). Como resultado, muchos navegadores restringen el acceso a las respuestas provenientes de un servidor

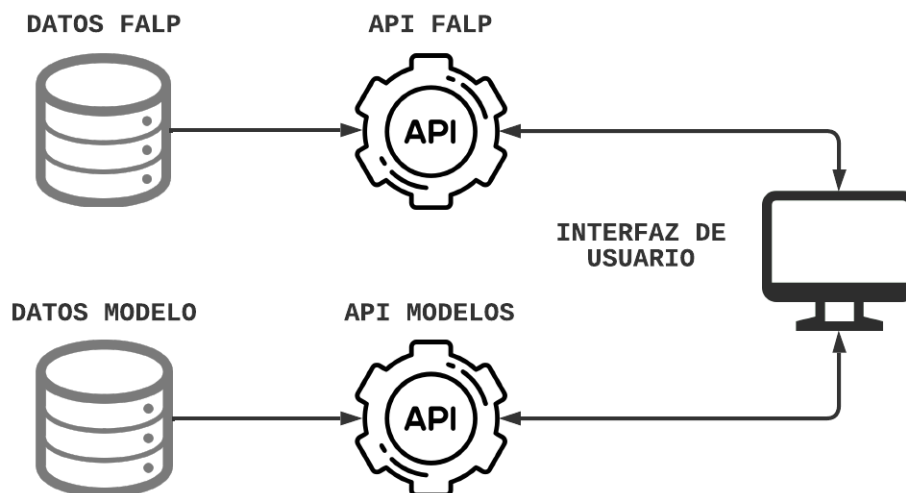


Figura 3.8: Primera arquitectura propuesta.

externo. Esto se debe a que la aplicación fue probada en servidores de prueba de la FALP, los cuales tienen un origen distinto al de la API consultada, la cual se encuentra en producción.

A pesar del esfuerzo realizado en colaboración con el equipo técnico de la FALP, no se logró encontrar una solución sencilla para este inconveniente. Por lo tanto, con el fin de poder presentar la aplicación a los médicos en un ambiente de prueba, se decide la decisión de reestructurar la comunicación con la API externa.

Finalmente, la estructura de la aplicación es la representada en la figura 3.9.

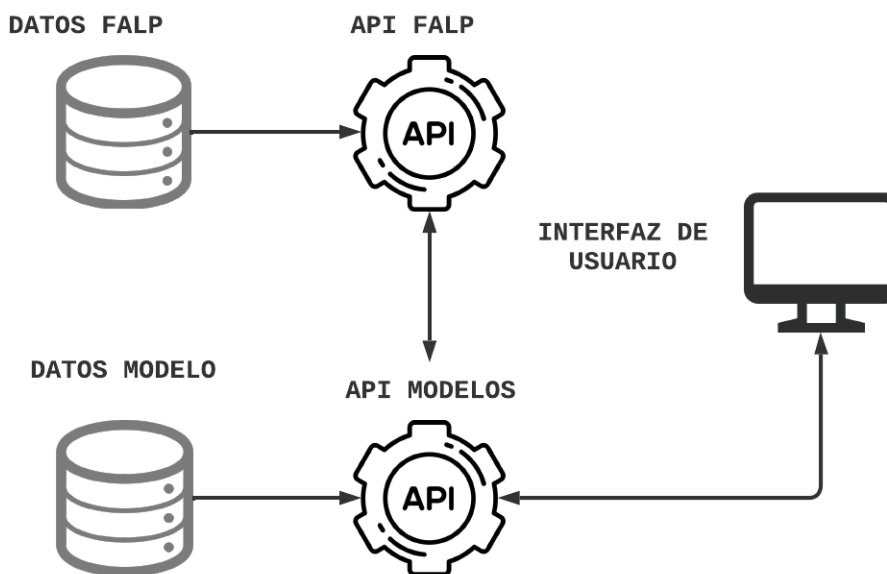


Figura 3.9: Arquitectura implementada.

En esta nueva configuración, todas las consultas del cliente pasan a través de la API creada para administrar los modelos. En esta arquitectura, cuando se realizan consultas sobre los datos de los pacientes, la API de los modelos actúa como un proxy. Esto implica que procesa la solicitud del cliente, la redirige a la API de la FALP y recibe su respuesta, para luego enviarla al cliente que hizo la solicitud.

Actualmente, esta es la arquitectura en uso para el software. Sin embargo, una vez que la aplicación esté en producción, el problema de CORS no existirá y será posible volver a la arquitectura propuesta inicialmente. La vuelta a la solución original es necesaria, ya que en dicha configuración se realizan menos llamadas para obtener la información de los pacientes, lo que genera un flujo de consultas más limpio y rápido.

### 3.3.2. Modelo de datos

La base de datos necesaria para resolver el problema tiene una estructura simple, ya que su tarea principal es registrar los modelos de clasificación entrenados. En primera instancia se requiere un esquema para registrar los modelos de clasificación en la base de datos. Además, considerando el requerimiento de escalabilidad, se requiere poder acceder a más de un modelo a través de la API.

Por lo tanto, se decide agregar un modelo cuya función es solo tener el registro de los distintos puntos de acceso para cada modelo en particular, para que sea simple y directo asociar, editar y consultar los distintos puntos de acceso a los modelos. Además, considerando que hay consultas que pueden ser hechas mas de una vez para un mismo paciente, ya sea para explorar distintas entradas o para probar las variables, surge la necesidad de tener un esquema de registro de las distintas consultas hechas a la API.

En resumen, se crean 3 modelos:

1. EndPoint: Registro de puntos de acceso asociado a un algoritmo de clasificación entrenado.
2. ClassifierAlgorithm: Representa los distintos algoritmos que se pueden agregar.
3. Request: Almacena toda la información de una consulta hecha. Tanto datos de entrada como de salida.

Las relaciones entre los modelos se pueden ver con en la figura 3.10. En ella también se representa la existencia de una clase abstracta que tiene todos los atributos básicos que debería tener un registro en la base de datos de la aplicación, de la cual extienden todas las clases que se crean, por lo que mas que tener representar un comportamiento común entre estas clases, tiene el objetivo de definir atributos comunes, lo cual la hace mas escalable. Además se muestran las multiplicidades de las relaciones, donde se representa que un modelo registrado está asociado a un *endpoint* y viceversa. Pero por otro lado, una instancia de *endpoint* puede estar asociada a muchas instancias de *request*.

Todo lo anterior es gestionado mediante el sistema de bases relacionales PostgreSQL.

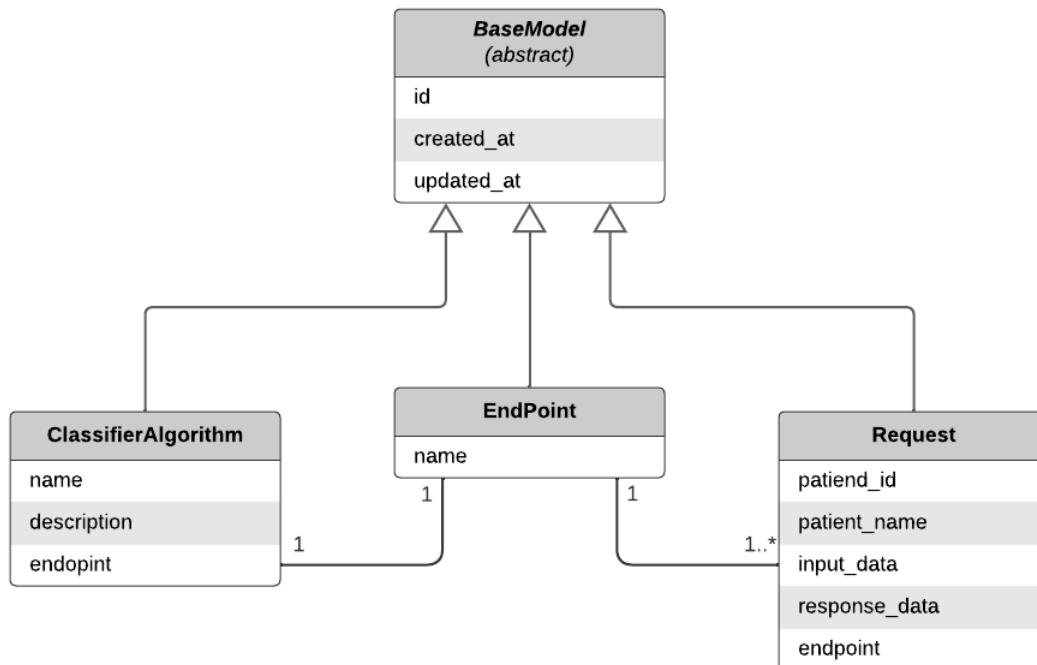


Figura 3.10: UML del modelo de datos de la herramienta.

### 3.3.3. API REST para administrar modelos

La API REST recibe y gestiona las peticiones del cliente, permite que éste tenga acceso a la base de datos descrita en la parte anterior.

Esta interfaz está desarrollada en Django REST framework (DRF), el cual sigue un patrón de diseño similar al de Modelo-Vista-Controlador, con ciertas diferencias que permiten una diferenciación clara entre el negocio (modelo) y la gestión de solicitudes.

Esta API tiene 4 principales *endpoints*:

1. *api/requests*: Una petición de tipo GET a esta ruta es recibida por la vista que maneja el modelo de solicitudes hechas a la API (*Request*), la cual retorna las últimas consultas hechas.
2. *api/endpoints*: Una petición de tipo GET a esta ruta es recibida por la vista que maneja el modelo de los *endpoints* hechas a la API (modelo *EndPoint*), la cual retorna todos los endpoints registrados y su información. Esta consulta es útil para obtener la documentación de los distintos modelos disponibles en caso de que luego existan mas modelos agregados.
3. *api/endpoint\_name/predict/*: Una solicitud con el método POST a esta ruta es recibida por una vista encargada de preprocesar los datos del paciente, determinar el algoritmo a usar en base al *endpoint\_name* entregado en la ruta (es el *endpoint* definido al agregar un modelo ya entrenado a la base de datos), realizar la predicción para ambos esquemas

y devolver el resultado.

4. *api/proxy/*: Una solicitud POST a esta ruta es procesada por una clase (o vista), la cual recibe información sobre la petición de la cual será intermediario hacer: la ruta, el método, encabezados y datos. Retornando la respuesta entregada, respetando el formato de entrega. Así este *endpoint* permite utilizar la API como proxy para llamadas a otras APIs, no solo la que se usa para este trabajo, en caso de que en un futuro exista información que accesible a través de otra API.

Es importante mencionar que las vistas descritas tienen sus respectivos serializadores que garantizan que la respuesta recibida por el cliente tenga una estructura definida y sea simple de cambiar.

Además de lo anterior, hay una parte importante del software que está contenido en la API REST, la cual contiene la información de los modelos. Esta es toda una sección, donde se guardan los *encoders* necesarios para procesar los datos de entrada antes de realizar las predicciones, las clases de los algoritmos

1. Codificadores (o *encoders*): necesarios para procesar los datos de entrada antes de realizar las predicciones.
2. Algoritmos como clase: permite instanciar algoritmos de clasificación, con distintas configuraciones, pudiendo registrarlos con distintos *endpoints*. Esto es similar a un patrón de diseño *factory*, ya que permite crear distintas instancias de un mismo algoritmo.
3. Registro: esta clase permite el registro de un algoritmo clasificador con sus respectivos datos: nombre, descripción, instancia de la clase del algoritmo y *endpoint*. Esta clase está inspirada en el patrón de diseño de registro [41]. Está pensada para ser usada desde la sección de administración de la API. Por ahora se usa la clase de registro para registrar el único modelo usado, una vez se levanta el servidor.

### 3.3.4. API FALP

De esta API solo se maneja la información necesaria para las tareas específicas que se hacen mediante esta comunicación: consultar datos de pacientes por rut y traer sus características clínicas e histopatológicas. En esta API se identifican 2 *endpoints* importantes:

1. *metadata*: Una consulta con el método POST a esta ruta del servidor de la FALP, entrega todas las variables clínicas que se pueden consultar de un paciente, con su respectivo indicador. Esto es fundamental en la exploración de la API de FALP, para determinar cómo se pueden solicitar los datos de los pacientes.
2. *data*: Una petición de tipo POST recibe el rut del paciente y ciertas características (los campos consultados con la ruta anterior), en un formato predeterminado (Ver 1). Se entrega más de un registro por campo, ya que hay ciertos exámenes que se repiten en la historia del paciente, por lo que se considera el más reciente.



El examen físico y el informe del comité tiene información valiosa, como la etapa o la amplificación HER2 del cáncer. Sin embargo, está en texto libre, por lo tanto se le muestra al usuario para que lo ingrese al formulario a mano.

```
1      {
2        id: rut,
3        fields: [
4          { field: "PACIENTE" },
5          { field: "FECHA_NACIMIENTO" },
6          { field: "FECHA_INGRESO" },
7          { field: "INT_mama_histologia" },
8          { field: "INT_mama_receptorEstrogeno" },
9          { field: "INT_mama_receptorProgesterona" },
10         { field: "INT_mama_Ki-67" },
11         { field: "INT_mama_fechaDgHistologico" },
12         { field: "INT_mama_lateralidad" },
13         { field: "INFORME_COMITE" },
14         { field: "INT_mama_examenFisico" },
15         { field: "INT_prostata_tnmT" },
16         { field: "INT_prostata_tnmN" },
17       ]
18     }
```

Listing 1: JSON petición a API FALP.

Para tener acceso a esta API, es necesario estar conectado en la red de la FALP. En etapas de desarrollo, se accede a la red mediante conexión VPN, con credenciales de invitado facilitadas por la institución.

### 3.3.5. Interfaz de usuario

La interfaz de usuario implementada en VueJS (versión 3), es sencilla y fácil de entender. En esta sección se describen los pasos seguidos en su desarrollo.

#### Mockups

Antes de comenzar a programar, es necesario tener un diseño de referencia a seguir, o definirlo en el camino. Para esto, se crean maquetas en una herramienta especializada y se presentan al equipo de la FALP. Luego de iterar maquetas con el equipo de la FALP, se define el diseño definitivo (ver figura 3.11)

Paralelo a la determinación del diseño, se define el flujo de uso del usuario (ver Anexo C).

Nombre completo <input type="text"/>	Rut <input type="text"/>	Edad <input type="text"/>
IMC <input type="text"/>	Etapa <input type="text"/>	Perfil Molecular <input type="text"/>
Estado menopausia <input type="text"/>	Receptor de Progesterona <input type="text"/>	Receptor de Estrógeno <input type="text"/>
Indice Ki67 <input type="text"/>	cN <input type="text"/>	cT <input type="text"/>
N° de copias <input type="text"/>	Relación cen <input type="text"/>	Tipo histológico <input type="text"/>

### Efectividad de tratamiento Neoadyuvante

No has realizado una predicción

### Últimas predicciones

Fecha	Nombre	T	TP
21-6-2023	Paciente 1	56%	67%
20-6-2023	Paciente 2	23%	88%
19-6-2023	Paciente 3	22%	88%
17-6-2023	Paciente 4	88%	56%
10-6-2023	Paciente 5	82%	44%

Figura 3.11: Esto es lo que se encuentra el usuario el ingresar.

## Desarrollo

Ya definido el diseño básico de la aplicación, se comienza la etapa de desarrollo. Para ello se usa el *framework* VueJS versión 3, el cual tiene una curva de aprendizaje suave, con un simple manejo de componentes y un flujo fácil de entender, ideal para lo simple que será la aplicación. Además, permite un diseño modular y escalable, donde cada parte de la aplicación puede ser tratado como un componente, y la asociación entre los mismos es de cómoda gestión. Lo anterior facilita cualquier cambio de diseño, agregar o quitar funcionalidades, conexiones a servidores y configuración.

Dado la simpleza de lo requerido, interacción del usuario se hace en una sola vista o página, es decir, el usuario no requiere navegar en la aplicación para acceder a las funcionalidades. Cada sección de la vista es un componente en Vue, por lo que el formulario es un componente, las barras laterales son un componente y así mismo lo es el modal (Ver anexo 3.12).

A continuación se describe cada una de estas componentes con las que interactúa el usuario:

1. Formulario de ingreso: formulario en el que el usuario agrega las características del paciente. La única entrada requerida es el rut del paciente, para registrarla en la base de datos y además consultar sus datos a la API de la FALP.
2. Barra lateral desplegable (izquierda): esta barra es desplegada automáticamente cuando se ingresa el rut de una paciente que existe en la base de datos de la FALP. Ésta contiene la información extraída desde la API de la FALP. Entrega la opción de rellenar automáticamente los datos que están individualizados en los registros de la institución,

Fecha	Nombre	T	TP
17-07-2023		30 %	36 %
17-07-2023		35 %	42 %
17-07-2023		35 %	42 %
17-07-2023		29 %	35 %
17-07-2023		21 %	29 %

Figura 3.12: Vista principal de la aplicación.

por ejemplo: edad, tipo histológico, Ki-67 y receptores hormonales.

3. Barra lateral (derecha): Es barra es fija, provee la información de la última consulta y hecha y el registro de las últimas 5 consultas hechas al modelo. Esto es útil para comparar los distintos resultados del modelo al cambiar los valores de las variables, para estudiar la influencia de cada una en la respuesta del paciente. Esto tiene fines científicos, ya que las características del paciente no pueden ser realmente modificadas.
4. Modal: Este modal se muestra al finalizar el cálculo de efectividad una vez enviado el formulario. Contiene la información resultante, mostrando la efectividad usando cada uno de los esquemas.

Es posible ver el flujo de uso de la aplicación implementada en el anexo D.

### 3.3.6. Contenedor

Durante la etapa de desarrollo, la API, la interfaz y la base de datos son montados de forma independiente. Así se puede aprovechar la característica reactiva a los cambios que poseen Django y Vue. Sin embargo, al momento de migrar el software a otro servidor, se decide hacerlo utilizando la tecnología de contenedores, en particular, con Docker.

Docker permite encapsular todos los servicios que necesarios para correr el software, incluso el sistema operativo, conteniendo además todas las dependencias necesarias.

Para este proyecto, solo se encapsula en un contenedor la API y la interfaz implementada en Vue. Porque por un lado, el sistema operativo del servidor de la FALP es una distribución de Linux, lo cual es compatible con las instrucciones y dependencias usadas en el desarrollo de la aplicación. Además, el servidor de la FALP cuenta con un servicio Postgres, el cual tiene un buen sistema de respaldos, muchos núcleos de CPU, entre otras características.

Todo el proyecto descrito anteriormente, incluido puede ser revisado en el repositorio [DidoTau/risk-calculator](#).

# Capítulo 4

## Evaluación de la Herramienta Predictiva

En este capítulo, se detalla la sesión de evaluación de la herramienta predictiva enfocada en pacientes con cáncer de mama HER2-positivo. La sesión contó con la participación de dos destacados miembros del equipo de investigación de la FALP: la cirujana oncológica Isabel Saffie y el experto en bioestadística Rodrigo Lagos. También estuvo presente Sergio Peñafiel, autor del método de clasificación interpretable utilizado para construir la herramienta y co-guía del proyecto.

### 4.1. Observaciones Importantes sobre el Modelo Predictivo

Durante la evaluación, se destacó la interpretabilidad del modelo, que proporciona información valiosa sobre las variables que influyen en la respuesta al tratamiento. Además, se elogió la rapidez con la que la herramienta calcula las predicciones.

Uno de los hallazgos más interesantes fue la tendencia a obtener una mayor respuesta al tratamiento con el esquema que incluye Trastuzumab y Pertuzumab. La posibilidad de comparar la efectividad de ambos esquemas es un enfoque novedoso que podría respaldar políticas públicas en cuanto al financiamiento de Pertuzumab. Se propone validar el modelo y publicar un estudio que compare ambos tratamientos.

Para mejorar el desempeño del modelo, se consideró crucial incorporar más datos de pacientes, incluyendo aquellos cuyas fichas aún no han sido ingresadas al sistema, lo que sumaría al menos 100 nuevos casos al análisis. Además, se sugirió agregar variables importantes, como el grado del tumor, para aumentar la precisión de las predicciones. También se planteó la posibilidad de utilizar ciertas variables como criterios de exclusión para evitar sesgos en los resultados de la predicción.

## 4.2. Observaciones Importantes sobre la Herramienta

Se valoró la limpieza y facilidad de uso de la herramienta, así como su conexión con la base de datos de la FALP, lo que agiliza el ingreso de datos y ofrece una experiencia rápida y expedita.

Los participantes destacaron la utilidad de la herramienta para realizar análisis interesantes al modificar las variables y ver cómo afecta la tasa de efectividad de los tratamientos. Se resaltó la importancia de usar la herramienta al momento del diagnóstico para informar a las pacientes sobre las probabilidades de respuesta con diferentes esquemas de tratamiento.

Para mejorar la usabilidad y experiencia del usuario, se propusieron algunas ideas, como presentar los resultados en gráficos circulares o diagramas de sectores para facilitar su comprensión. También se sugirió agregar descripciones en los distintos campos del formulario para que sea más claro para los usuarios que no estén familiarizados con ciertas variables. Además, se planteó la posibilidad de agregar el HER2 como otra variable en el formulario para ampliar el alcance de la herramienta.

## 4.3. Ideas de Trabajo Futuro

Entre las ideas de trabajo futuro, se propuso someter la herramienta a un análisis prospectivo, donde se guardarían las predicciones realizadas antes de someter a una paciente al tratamiento, y luego se compararía con la respuesta obtenida después de todo el proceso, incluida la operación.

Se consideró importante ampliar el modelo para determinar la probabilidad de sobrevida de los pacientes, además de la respuesta completa patológica, para tener un indicador más completo de un tratamiento exitoso.

También se sugirió realizar un estudio que incluya más pacientes con diferentes tipos de HER2, ya que esto mejoraría la aplicabilidad de la herramienta.

## 4.4. Conclusión

La evaluación de la herramienta predictiva en el focus group proporcionó información valiosa para mejorar su precisión y utilidad en el campo de la oncología. Las observaciones sobre el modelo y la herramienta, así como las ideas de trabajo futuro, ofrecen una guía clara para fortalecer la herramienta y aumentar su relevancia en la toma de decisiones clínicas para pacientes con cáncer de mama HER2-positivo. La colaboración con otros especialistas y la validación prospectiva son pasos fundamentales para el éxito y el potencial impacto positivo en la calidad de vida de los pacientes.

# Capítulo 5

## Conclusión

Este trabajo ha alcanzado con éxito el desarrollo de una herramienta computacional innovadora para apoyar a los especialistas en oncología en la toma de decisiones para el tratamiento del cáncer de mama HER2-positivo. Mediante el uso de un modelo de clasificación interpretable basado en la teoría de Dempster-Shafer, se logró predecir la respuesta al tratamiento neoadyuvante, identificando qué variables clínicas e histopatológicas influyen significativamente en la obtención de respuestas completas patológicas.

La generación de reglas con información relevante sobre las variables predictoras de respuesta ha sido un logro destacado. Estas reglas brindan una comprensión profunda y valiosa de los factores que pueden afectar la eficacia del tratamiento en cada paciente. La interpretabilidad del modelo permite a los especialistas en oncología discutir y validar los resultados, fomentando una toma de decisiones informada y basada en evidencia.

Es importante mencionar que, contrario a lo propuesto en un comienzo, más que destacar a la herramienta por ayudar a identificar qué pacientes pueden prescindir del Pertuzumab, se valora más el que la herramienta sirva para destacar la brecha que existe en la mayoría de pacientes entre su probabilidad de obtener respuesta con un esquema u otro. Dónde un esquema completo muestra en general una mayor tendencia a la respuesta.

Este trabajo representa solo el inicio de un proyecto con alta proyección. La herramienta desarrollada puede ser mejorada y enriquecida con datos adicionales, lo que la haría más robusta y confiable en la práctica clínica. Además, es posible ampliar su alcance para incluir otras características relevantes que puedan influir en la respuesta al tratamiento, lo que potencialmente la convertiría en una herramienta valiosa en el tratamiento de otros tipos de cáncer.

Además de su aplicación clínica, este proyecto podría tener un importante impacto en el ámbito de las políticas públicas de salud. La disponibilidad de una herramienta que permita una toma de decisiones personalizada y basada en evidencia para el tratamiento del cáncer de mama HER2-positivo, podría impulsar el desarrollo de políticas de acceso equitativo a tratamientos efectivos, especialmente considerando el alto costo y dificultades de acceso a Pertuzumab.

La investigación en paralelo es esencial para continuar mejorando y validando la herramienta, así como para explorar nuevas estrategias de tratamiento. La colaboración con instituciones de salud y organismos públicos permitirá aprovechar al máximo el potencial de este enfoque y contribuir al avance de la medicina oncológica.

En resumen, este trabajo ha logrado un hito significativo al desarrollar una herramienta computacional con potencial transformador en el tratamiento del cáncer de mama HER2. Las reglas identificadas proporcionan información crucial para decidir un tratamiento, lo que puede mejorar la calidad de vida y los resultados clínicos de las pacientes. Con miras al futuro, el enfoque investigativo y el esfuerzo por robustecer la herramienta y explorar políticas públicas en salud permitirán maximizar su impacto y contribuir al avance en el campo de la oncología.

## 5.1. Trabajos futuros

El desarrollo de la herramienta computacional para la predicción de la respuesta al tratamiento neoadyuvante ha sido un logro significativo. Sin embargo, hay varias oportunidades para mejorar y expandir este trabajo, lo que podría llevar a un mayor impacto en la práctica clínica y en las políticas de salud relacionadas con el cáncer de mama HER2.

Uno de los principales trabajos futuros es mejorar y validar el modelo actual. Esto implica incluir datos de más pacientes y considerar nuevas variables clínicas relevantes. Con una mayor cantidad de datos, se puede aumentar la robustez y precisión del modelo, lo que resultaría en predicciones más confiables para los especialistas en oncología.

Además, se sugiere ampliar el alcance de la herramienta a otros tipos de cáncer de mama HER2 y a diferentes fármacos utilizados en el tratamiento. Esto requeriría la exploración de nuevos subtipos de cáncer de mama HER2 y la evaluación de la efectividad de otros tratamientos, como la administración de antraciclina.

Incorporar la probabilidad de sobrevida de los pacientes es otro aspecto importante a considerar en futuros trabajos. Actualmente, existe un debate sobre el valor pronóstico de la respuesta completa patológica en términos de supervivencia. Un análisis más detallado y prospectivo permitiría determinar qué tan buen indicador de supervivencia es la respuesta al tratamiento.

Para validar la efectividad de la herramienta en la práctica clínica, se propone llevar a cabo un estudio prospectivo en colaboración con instituciones médicas y especialistas en oncología. Este estudio permitiría evaluar y validar la utilidad de la herramienta en situaciones clínicas reales, lo que respaldaría su adopción en la práctica médica.

Además de mejorar la herramienta en sí misma, es esencial realizar un análisis más profundo sobre la efectividad y el costo de los diferentes tratamientos. Se sugiere realizar un estudio detallado para comparar la efectividad entre el esquema de tratamiento con Trastuzumab solo y el esquema que incluye Trastuzumab y Pertuzumab. Además, se debe realizar un análisis de costo-efectividad que considere los tratamientos con Trastuzumab y Pertuzu-



mab, y el tratamiento alternativo con T-DM1, que se administra a pacientes que no alcanzan una respuesta completa patológica. Esto proporcionaría información valiosa para la toma de decisiones a nivel de políticas públicas de salud.

En resumen, aunque se ha logrado un avance significativo en el desarrollo de la herramienta computacional, hay una serie de trabajos futuros esenciales para fortalecer su desempeño y aplicabilidad en el campo de la oncología. La mejora y validación del modelo, la ampliación del alcance a otros tipos de cáncer de mama HER2 y fármacos, la incorporación de la probabilidad de sobrevida y el análisis de la efectividad y el costo son aspectos fundamentales para mejorar la toma de decisiones médicas y promover políticas de salud más efectivas en el tratamiento del cáncer de mama HER2. Estos trabajos futuros representan una oportunidad valiosa para seguir impulsando avances en la investigación y la atención oncológica.

## 5.2. Contribuciones

Es importante destacar que el trabajo presentado en la sección 3.1 es parte de un *paper* científico que fue enviado al *Journal of Clinical Oncology - Clinical Cancer Informatics* <https://ascopubs.org/journal/cci> y está en proceso de revisión. Cabe mencionar que en dicha publicación, el autor de esta memoria figura como primer autor y los resultados fueron validados tanto por los guías de esta memoria, como por expertos en oncología. Dentro de los expertos está Isabel Saffie, cuyo trabajo fue la inspiración y base de esta memoria.

# Siglas

**ER** Estrogen Receptor o Receptores de estrógeno. 22

**HER2** Human Epidermal Growth Factor Receptor 2. i

**pCR** Pathological Complete Response. 1

**PR** Progesterone Receptor o Receptores de progesterona. 22

# Glosario

**cN** La clasificación N describe la presencia y la extensión de la propagación del cáncer a los ganglios linfáticos regionales. Puede variar desde N0 (no hay ganglios linfáticos afectados) hasta NX (ganglios linfáticos regionales no evaluables) y luego de N1 a N3, donde N3 indica una mayor propagación a los ganglios linfáticos regionales. 22

**cT** La clasificación T describe el tamaño del tumor primario y su invasión a tejidos y estructuras cercanas. Puede variar desde T0 (sin evidencia de tumor primario) hasta TX (no se puede evaluar el tumor primario), y luego de T1 a T4, donde T4 indica un tumor más grande o que se ha extendido a estructuras cercanas. 22

**ki-67** Ki-67 es una proteína utilizada como marcador de proliferación celular en el diagnóstico de cáncer de seno y otras neoplasias. Indica la proporción de células en una muestra de tejido que están en la fase activa del ciclo celular. Cuanto mayor es la expresión de Ki-67, mayor es la actividad proliferativa de las células, lo que puede estar relacionado con un crecimiento tumoral más rápido y un peor pronóstico en algunos casos de cáncer de seno [42]. 16

**Pertuzumab** Medicamento que se usa con otros medicamentos para tratar el cáncer de mama positivo para HER2. Se usa para pacientes cuyo cáncer se diseminó a otras partes del cuerpo y todavía no se trató con otros medicamentos contra el cáncer. También se usa antes de la cirugía para pacientes con cáncer de mama localmente avanzado, inflamatorio o en estadio temprano, y se usa después de la cirugía para pacientes de cáncer de mama en estadio temprano con un riesgo alto de que el cáncer recidive (vuelva). Además, está en estudio para el tratamiento de otros tipos de cáncer. El pertuzumab se une a una proteína de la superficie de algunas células cancerosas que se llama HER2. Bloquear la acción de esta proteína tal vez ayude a destruir las células cancerosas. El pertuzumab es un tipo de anticuerpo monoclonal [43]. 1

**Trastuzumab** Medicamento que se usa solo o con otros medicamentos para tratar ciertos tipos de cáncer de mama, cáncer de estómago y cáncer de unión gastroesofágica positivos para HER2. También está en estudio para el tratamiento de otros tipos de cáncer. El trastuzumab se une a una proteína llamada HER2, que se encuentra en algunas células cancerosas, y ayuda al sistema inmunitario a destruir células cancerosas. El trastuzumab es un tipo de anticuerpo monoclonal y de antagonista del receptor de HER2 [44]..

# Bibliografía

- [1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, “Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [2] A. G. Waks and E. P. Winer, “Breast Cancer Treatment: A Review,” *JAMA*, vol. 321, pp. 288–300, 01 2019.
- [3] S. Loibl and L. Gianni, “Her2-positive breast cancer,” *The Lancet*, vol. 389, no. 10087, pp. 2415–2429, 2017.
- [4] I. Saffie Vega, J. Sapunar Zenteno, F. Buscaglia Fernandez, F. Reyes Cosmelli, R. Lagos Chavez, and B. Chahuán Manzur, “Predictors of pathologic complete response to neoadjuvant treatment in her2-overexpressing breast cancer: a retrospective analysis using real-world data,” *Ecancermedicalscience*, vol. 16, p. 1338, 2022.
- [5] A. J. London, “Artificial intelligence and black-box medical decisions: Accuracy versus explainability,” *Hastings Center Report*, vol. 49, no. 1, pp. 15–21, 2019.
- [6] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [7] S. Peñafiel, N. Baloian, H. Sanson, and J. A. Pino, “Applying dempster–shafer theory for developing a flexible, accurate and interpretable classifier,” *Expert Systems with Applications*, vol. 148, p. 113262, 2020.
- [8] S. Peñafiel, N. Baloian, H. Sanson, and J. A. Pino, “Predicting stroke risk with an interpretable classifier,” *IEEE Access*, vol. 9, pp. 1154–1166, 2021.
- [9] N. C. Institute, “What is cancer?,” 2021.
- [10] L. S. A. C. E. Cáncer, “Estado de her2 del cáncer de seno,” 2022.
- [11] N. C. Institute, “Chemotherapy to treat cancer,” 2022.
- [12] N. C. Institute, “Radiation therapy to treat cancer,” 2022.
- [13] N. C. Institute, “Targeted therapy to treat cancer,” 2022.
- [14] N. C. Institute, “Immunotherapy to treat cancer,” 2022.
- [15] N. C. Institute, “Biomarker testing for cancer treatment,” 2022.

- [16] C. Selli and A. H. Sims, “Neoadjuvant therapy for breast cancer as a model for translational research,” *Breast Cancer: Basic and Clinical Research*, vol. 13, p. 1178223419829072, 2019.
- [17] M. C. Winkler, “Perspectivas económicas,” 2022.
- [18] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*. Pearson Education, 2016.
- [19] T. A. A. Abdullah, M. S. M. Zahid, and W. Ali, “A review of interpretable ml in healthcare: Taxonomy, applications, challenges, and future directions,” *Symmetry*, vol. 13, p. 2439, Dec 2021.
- [20] P. Smets, “Decision making in the tbm: the necessity of the pignistic transformation,” *International Journal of Approximate Reasoning*, vol. 38, no. 2, pp. 133–147, 2005.
- [21] B. R. Cobb and P. P. Shenoy, “On the plausibility transformation method for translating belief function models to probability models,” *International journal of approximate reasoning*, vol. 41, no. 3, pp. 314–330, 2006.
- [22] G. Shafer, “Dempster’s rule of combination,” *International Journal of Approximate Reasoning*, vol. 79, pp. 26–40, 2016. 40 years of Research on Dempster-Shafer Theory.
- [23] IBM, “Software development,” 2020.
- [24] E. Gamma, R. Helm, R. Johnson, and J. Vlissides, *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley Professional, 1994.
- [25] J. Nickoloff and S. Kuenzli, *Docker in Action*. Manning Publications, 2019.
- [26] S. M. Swain, J. Baselga, S.-B. Kim, J. Ro, V. Semiglazov, M. Campone, E. Ciruelos, J.-M. Ferrero, A. Schneeweiss, S. Heeson, E. Clark, G. Ross, M. C. Benyunes, and J. Cortés, “Pertuzumab, trastuzumab, and docetaxel in her2-positive metastatic breast cancer,” *New England Journal of Medicine*, vol. 372, no. 8, pp. 724–734, 2015. PMID: 25693012.
- [27] G.-B. M. Díaz-Acedo R, Artacho-Criado S and L. Álvarez P., “Real world study of pertuzumab-trastuzumab-chemotherapy versus trastuzumab-chemotherapy in neoadjuvant treatment of breast cancer,” *Farmacía Hospitalaria*, vol. 44, pp. 96 – 99, 06 2020.
- [28] K. J. Rothman, S. Greenland, and T. L. Lash, *Case–Control Studies*. John Wiley Sons, Ltd, 2008.
- [29] J. McDonald, “Wilcoxon signed-rank test,” 2014.
- [30] R. A. Fisher, “On the interpretation of 2 from contingency tables, and the calculation of p,” *Journal of the Royal Statistical Society*, vol. 85, no. 1, pp. 87–94, 1922.
- [31] P. D. Sasieni, “From genotypes to genes: Doubling the sample size,” *Biometrics*, vol. 53, no. 4, pp. 1253–1261, 1997.

- [32] D. Livingston-Rosanoff, J. Schumacher, K. Vande Walle, T. Stankowski-Drengler, C. C. Greenberg, H. Neuman, and L. G. Wilke, “Does tumor size predict response to neoadjuvant chemotherapy in the modern era of biologically driven treatment? a nationwide study of us breast cancer patients,” *Clin Breast Cancer*, vol. 19, pp. e741–e747, Dec 2019.
- [33] G. Houvenaeghel, A. de Nonneville, M. Cohen, F. Viret, S. Rua, L. Sabiani, M. Buttarelli, E. Charaffe, A. Monneur, A. Jalaguier-Coudray, M. Bannier, R. Sabatier, and A. Gonçalves, “Neoadjuvant chemotherapy for breast cancer: Evolution of clinical practice in a french cancer center over 16 years and pathologic response rates according to tumor subtypes and clinical tumor size: Retrospective cohort study,” *J Surg Res (Houst)*, vol. 5, pp. 511–525, Sep 2022.
- [34] C. University, “Predict breast cancer.” <https://breast.predict.nhs.uk/tool>, 2010. Fecha de acceso: julio de 2023.
- [35] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2017.
- [36] P. Domingos, “A few useful things to know about machine learning,” *Commun. ACM*, vol. 55, p. 78–87, oct 2012.
- [37] A. Géron, *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. O’Reilly Media, 2017.
- [38] L. Arnould, P. Arveux, J. Couturier, M. Gelly-Marty, C. Loustalot, F. Ettore, C. Sagan, M. Antoine, F. Penault-Llorca, B. Vasseur, P. Fumoleau, and B. P. Coudert, “Pathologic Complete Response to Trastuzumab-Based Neoadjuvant Therapy Is Related to the Level of HER-2 Amplification,” *Clinical Cancer Research*, vol. 13, pp. 6404–6409, 11 2007.
- [39] A. Schneeweiss, S. Chia, T. Hickish, V. Harvey, A. Eniu, R. Hegg, C. Tausch, J. Seo, Y.-F. Tsai, J. Ratnayake, V. McNally, G. Ross, and J. Cortés, “Pertuzumab plus trastuzumab in combination with standard neoadjuvant anthracycline-containing and anthracycline-free chemotherapy regimens in patients with her2-positive early breast cancer: a randomized phase ii cardiac safety study (trypaena),” *Annals of Oncology*, vol. 24, no. 9, pp. 2278–2284, 2013.
- [40] E. Luporsi, F. André, F. Spyrtatos, and et al., “Ki-67: level of evidence and methodological considerations for its role in the clinical management of breast cancer: analytical and critical review,” *Breast Cancer Research and Treatment*, vol. 132, no. 3, pp. 895–915, 2012.
- [41] “Java Design Patterns.” <https://java-design-patterns.com/patterns/registry/>. Accedido el día mes año.
- [42] A. C. Society, “Cómo entender su informe de patología: cáncer de seno.”
- [43] N. C. Institute, “Diccionarios del nci.”
- [44] N. C. Institute, “Diccionarios del nci.”
- [45] N. C. Institute, “Hormone therapy for breast cancer,” 2022.

- [46] IBM, “Machine learning,” 2020.
- [47] A. S. Tanenbaum and M. V. Steen, *Distributed Systems: Principles and Paradigms*. Pearson, 2006.
- [48] S. Allamaraju, *RESTful Web Services Cookbook: Solutions for Improving Scalability and Simplicity*. O’Reilly Media, 2010.

## **Anexo A**

### **Datos clasificador interpretable en pacientes con esquema completo**



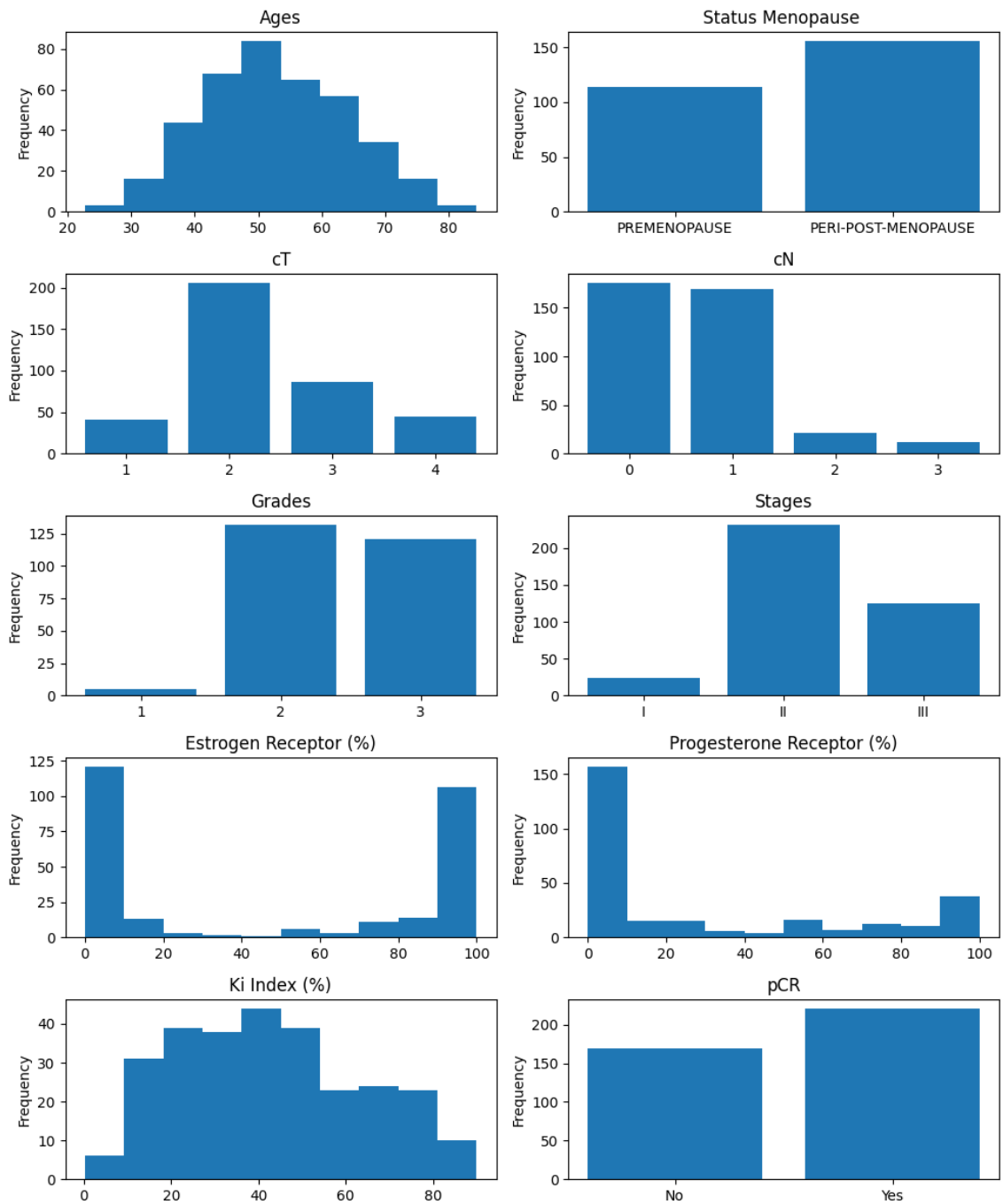


Figura A.1: Distribución de variables usadas para entrenar el modelo.

## Anexo B

# Clasificador interpretable en pacientes tratados con Trastuzumab

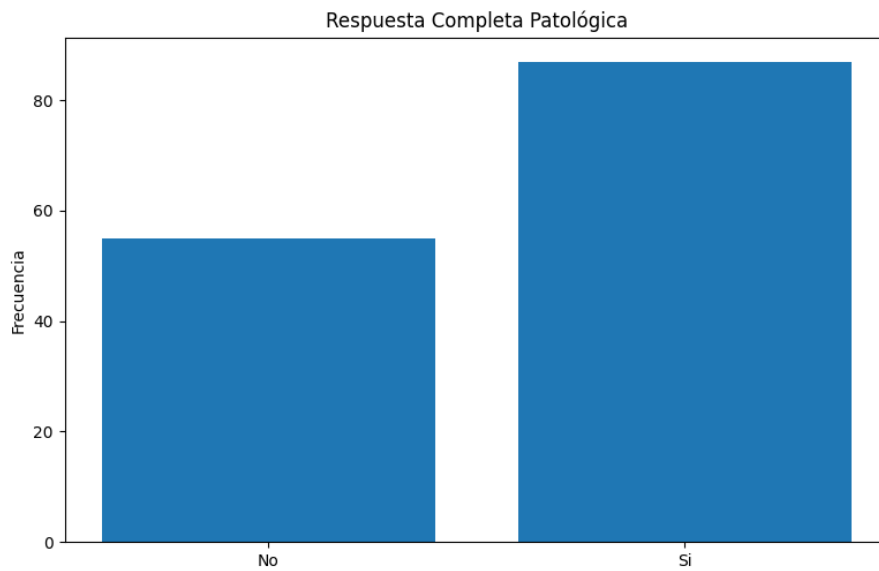


Figura B.1: Distribución de pCR de pacientes.

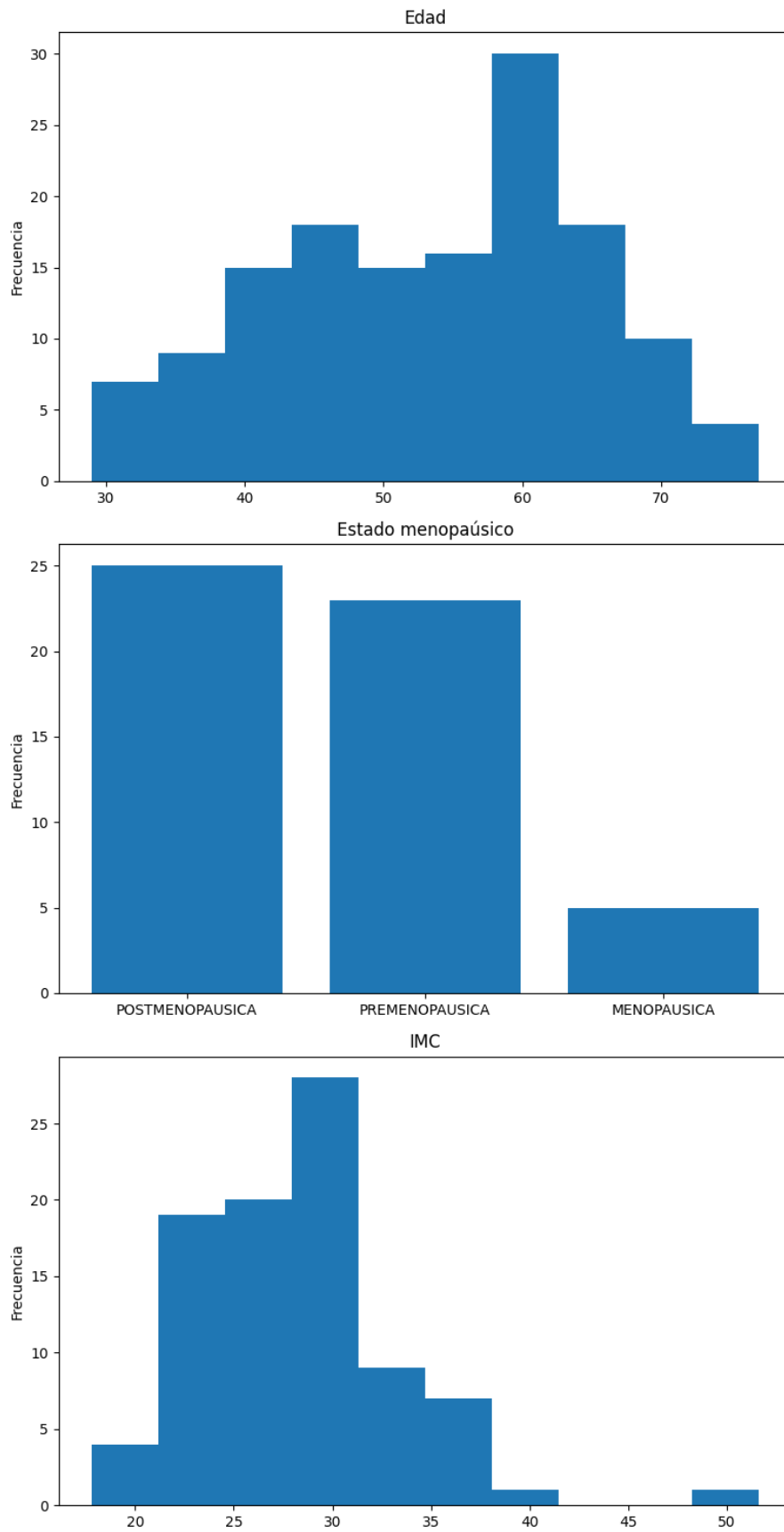


Figura B.2: Distribución de edad, estado menopáusico e IMC en pacientes de la muestra.

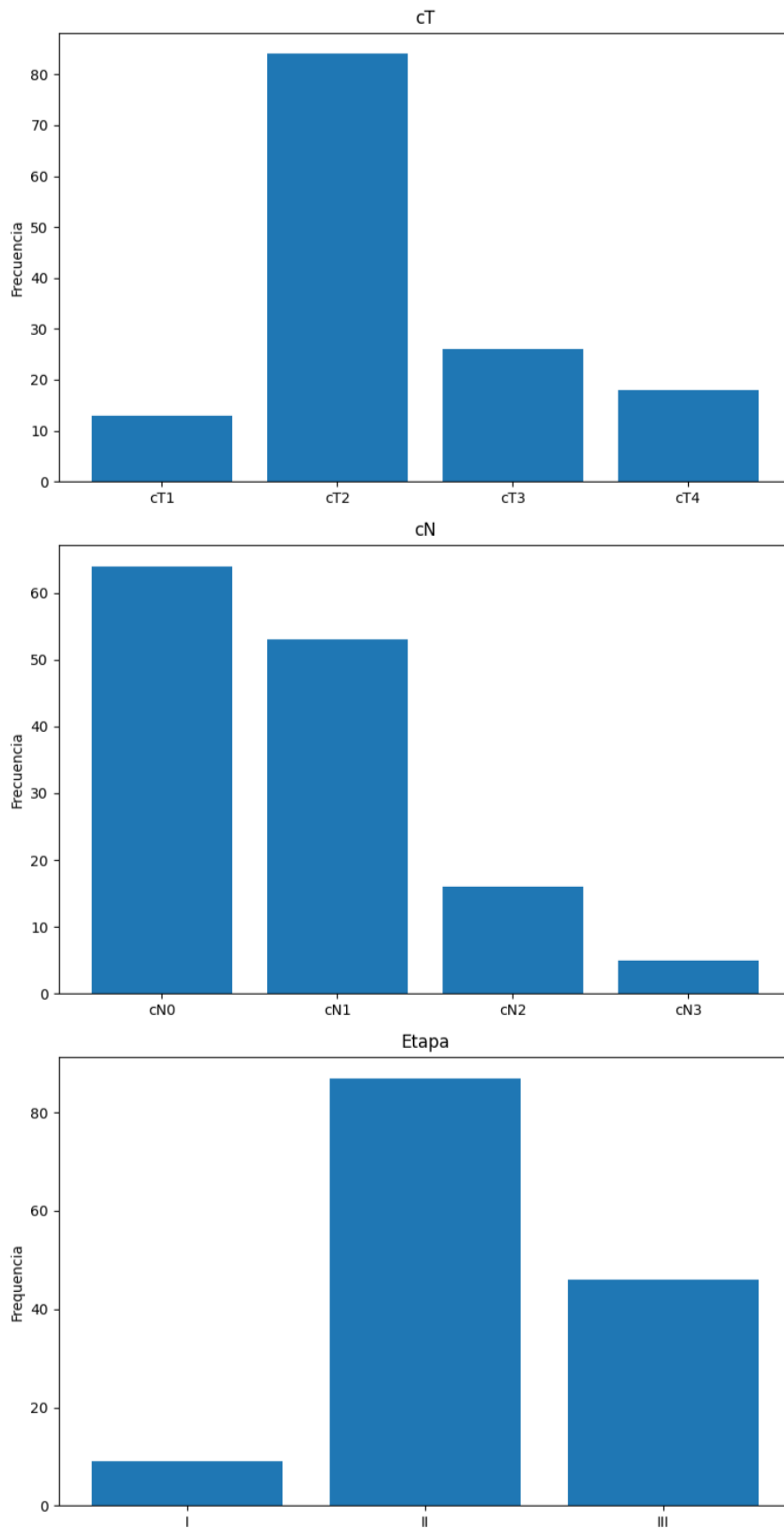


Figura B.3: Distribución de cT (tamaño), cN (linfonodos comprometidos) y etapa del cáncer en pacientes de la muestra.

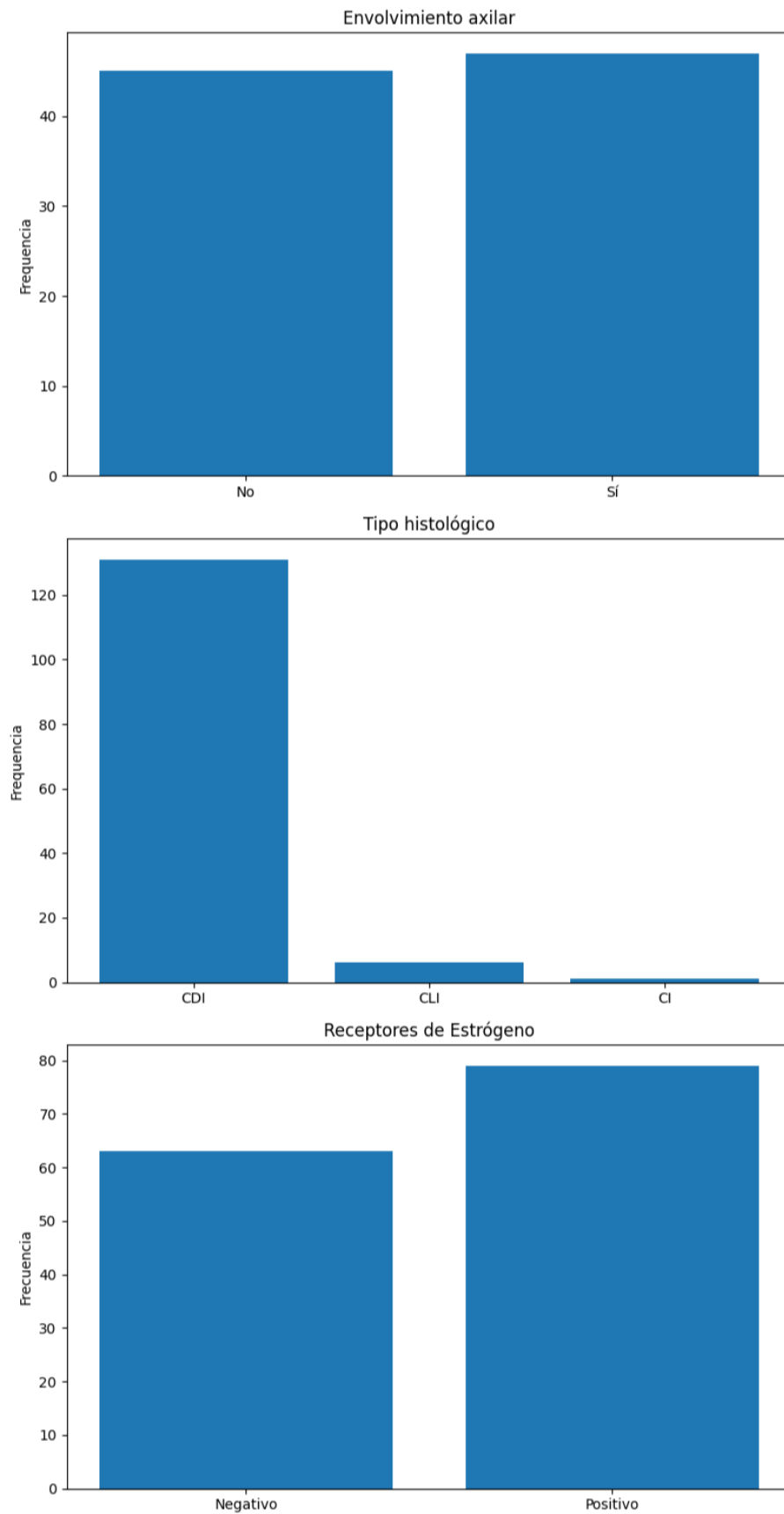


Figura B.4: Distribución de envolvimento axilar, tipo histológico y nivel receptores de estrógeno en pacientes de la muestra.

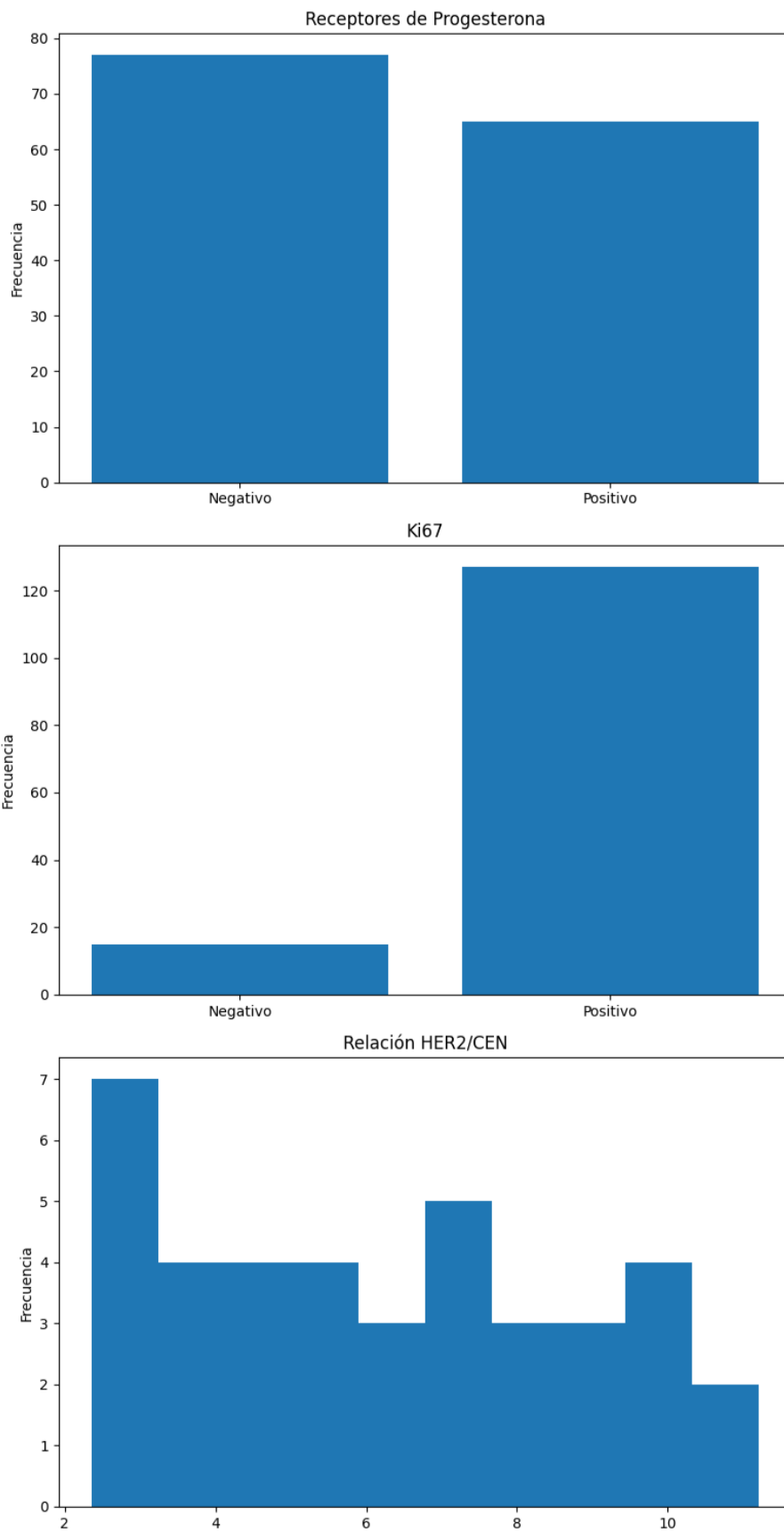


Figura B.5: Distribución de receptores de progesterona, índice Ki67 y relación del gen HER2/CENTROMERO en pacientes de la muestra.

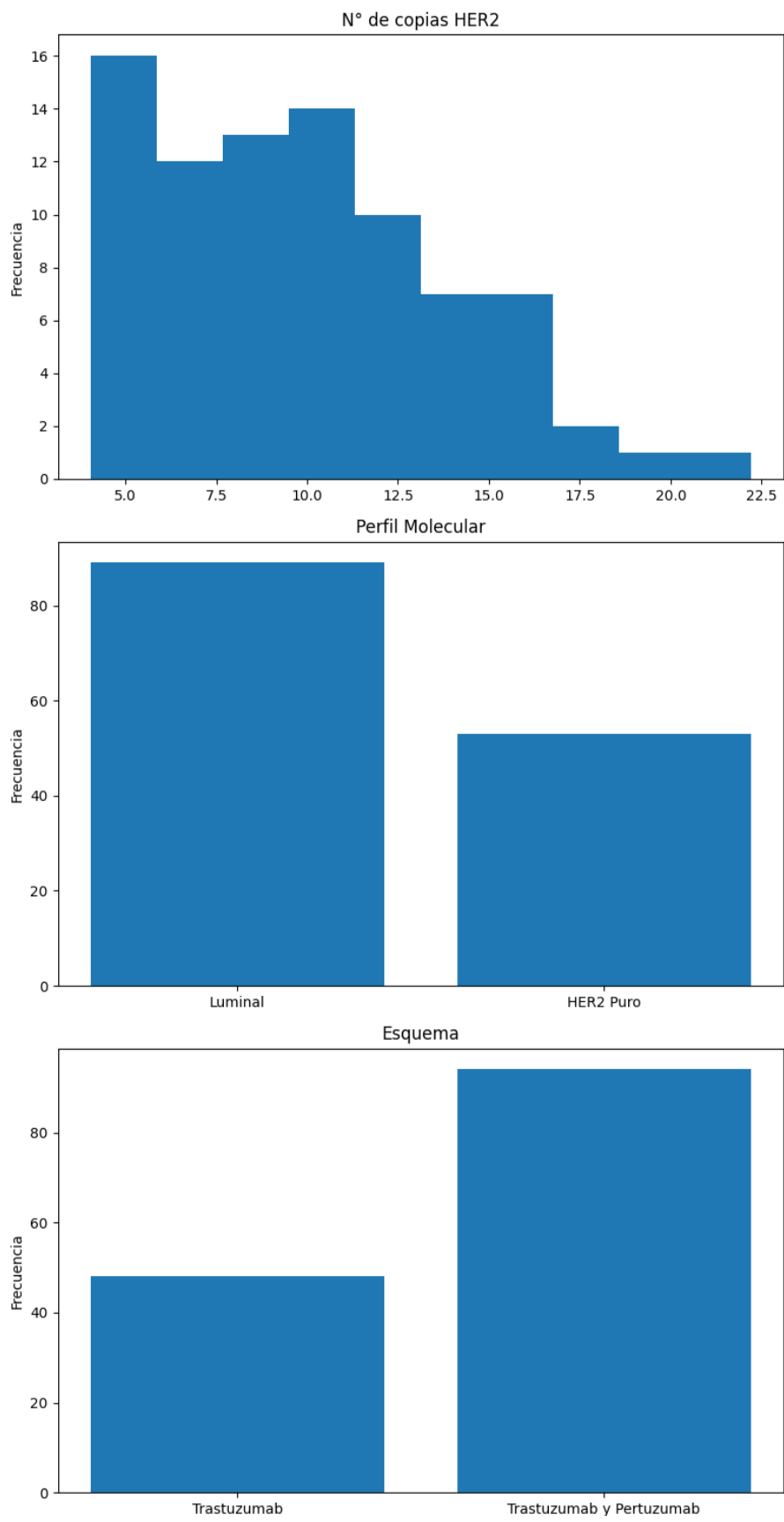


Figura B.6: Distribución de amplificación de HER2, perfil molecular y esquema de tratamiento en pacientes de la muestra.

# Anexo C

## Diseño previo de la aplicación

En esta parte se muestra el diseño propuesto para la aplicación, donde además se presenta el flujo de uso.



Nombre completo	Rut	Edad
Juana Ayala	12.345.678-9	51
IMC	Etapa	Perfil Molecular
22	Etapa I	Luminal
Estado menopausia	Receptor de Progesterona	Receptor de Estrógeno
PRE MENOPAUSICO	80 %	80 %
Indice Ki67	cN	cT
80 %	cN1	cT1
Nº de copias	Relación cen	Tipo histológico
4	5	CDI

**Calcular**

**Efectividad de tratamiento Neoadyuvante**

No has realizado una predicción

**Últimas predicciones**

Fecha	Nombre	T	TP
21-6-2023	Paciente 1	56%	67%
20-6-2023	Paciente 2	23%	88%
19-6-2023	Paciente 3	22%	88%
17-6-2023	Paciente 4	88%	56%
10-6-2023	Paciente 5	82%	44%

Figura C.1: El usuario llena el formulario con datos de un paciente. Envía el formulario.



Nombre completo  
Juana Ayala

IMC  
22

Estado menopausia  
PRE MENOPAUSICO

Indice Ki67  
80 %

Nº de copias  
4

### Resultados

A partir de las características clínicas e histopatológicas entregadas, la paciente Juana Ayala tiene:

- 80% de probabilidad de tener respuesta al tratamiento neoadyuvante siendo tratada solo con **Trastuzumab**
- 48% de probabilidad de tener respuesta al tratamiento neoadyuvante siendo tratada con **Trastuzumab y Pertuzumab**.

**Cerrar**

#### Efectividad de tratamiento neoadyuvante

realizado una predicción

#### Últimas predicciones

Nombre	T	TP
Paciente 1	56%	67%
Paciente 2	23%	88%
19-6-2023 Paciente 3	22%	88%
17-6-2023 Paciente 4	88%	56%
10-6-2023 Paciente 5	82%	44%

**Calcular**

Figura C.2: La solicitud es enviada a la API, realiza el cálculo de predicción y responde con los resultados.

Nombre completo  
Juana Ayala

IMC  
22

Estado menopausia  
PRE MENOPAUSICO

Indice Ki67  
80 %

Nº de copias  
4

Rut  
12.345.678-9

Etapa  
Etapa I

Receptor de Progesterona  
80 %

cN  
cN1

Relación cen  
5

Edad  
51

Perfil Molecular  
Luminal

Receptor de Estrógeno  
80 %

cT  
cT1

Tipo histológico  
CDI

**Calcular**

#### Efectividad de tratamiento Neoadyuvante

Juana Ayala, 51 años

- 80% con esquema **Trastuzumab**
- 48% con esquema **Trastuzumab y Pertuzumab**

#### Últimas predicciones

Fecha	Nombre	T	TP
22-6-2023	Juana Ayala	80%	48%
21-6-2023	Paciente 1	56%	67%
20-6-2023	Paciente 2	23%	88%
19-6-2023	Paciente 3	22%	88%
17-6-2023	Paciente 4	88%	56%

Figura C.3: Su última predicción se puede visualizar en la barra lateral.

# Anexo D

## Aplicación implementada

En esta sección se muestra la aplicación ya desplegada en el servidor de prueba de la FALP, mostrando el flujo de uso. Se oculta el nombre y rut de los pacientes.

The screenshot displays the application interface. On the left, a 'Datos Paciente' modal is open, showing a list of patient data extracted from a database. The data includes: Nombre, Edad (48), Informe del comité (Enterado del caso, hipotiroidismo, no fuma, en con...), cT (cT2), cN (cN0), Examen físico (historia ca mama der tdo 2016 T2N0 RE1%, RP1%, H...), Receptor de estrógeno (NEGATIVO), Receptor de progesterona (NEGATIVO), Ki67 (NEGATIVO), Fecha de diagnóstico (2019-04-05), and Lateralidad (2). A 'Rellenar' button is visible at the bottom of the modal. The main form contains input fields for 'Nombre completo', 'Edad' (48), 'Etapa del cáncer', 'Perfil Molecular', 'Receptor de estrógeno' (Negativo), 'Receptor de progesterona' (Negativo), 'cT' (2), and 'Relación HER2/CEN'. A 'Calcular' button is located below the form. On the right, the 'Efectividad de tratamiento neoadyuvante' section shows 'No has hecho ninguna predicción'. Below this, the 'Últimas predicciones' table is displayed.

Fecha	Nombre	T	TP
17-07-2023	[Redacted]	30 %	36 %
17-07-2023	[Redacted]	35 %	42 %
17-07-2023	[Redacted]	35 %	42 %
17-07-2023	[Redacted]	29 %	35 %
17-07-2023	[Redacted]	21 %	29 %

Figura D.1: Una vez ingresado un rut, la aplicación realiza una consulta directamente a la API de la FALP y obtiene los datos del paciente. Es posible rellenar con un solo clic los valores que están bien estructurados e individualizados.

**Datos**

zona eritematosa 29 julio: retro expansor. 2 agosto: retiro drenaje. Derivada para evaluación de reconstrucción: Muy buena candidata para DIEP, cicatriz cesarea ok Plan: Solicito TAC abdominal. DIEP. [Ver menos](#)

- cT : cT2
- cN : cN0

Examen físico : historia ca mama der tdo 2016 T2N0 RE1%, RP1%, HER2 3+, KI67 >20% Qt neoadyuvante 6 ciclos CD-TP 29 abril 2016: MP+GC CDIS residual de 4 mm GC neg RT herceptin hasta dic 2016 TMX agosto 2016 23 marzo 2018 42 años, trae examen eco ginec ; endometrio 11 mmm (2017 tenía 10 mm ) eco abdominal megativo mx+eco : BIRADS " 2 ex sangre TSH alta 16.96 al ex NED cicatriz periareolar sup der control en 6 meses , a endocrinologo 28 sept 18 hipotiroidismo en tto eutirox eco ginec . endometrio 8 mm ex NED control 6 meses con ex 29 marzo 19 eco (21/3/19) BIRADS 2 Mx+ mag (21/3/19) UCSder microc agrupadas en area de 28 mm BIRADS 4C eco gine endometrio 13 mm eco abd neg bx stx der 17 abril 19 bx stx der : CDIS de alto grado con comedo y calcif y microinvacion del estroma RE neg RP neg , KI67 >15%,

Nombre completo

Edad

Estadío del cáncer

Perfil Molecular

Receptor de estrógeno

Receptor de progesterona

Clasificación HER2/CEN

cT

Tipo Histológico

**Calcular**

### Efectividad de tratamiento neoadyuvante

No has hecho ninguna predicción

---

#### Últimas predicciones

Fecha	Nombre	T	TP
17-07-2023	<div style="width: 100%; height: 10px; background-color: black;"></div>	30 %	36 %
17-07-2023	<div style="width: 100%; height: 10px; background-color: black;"></div>	35 %	42 %
17-07-2023	<div style="width: 100%; height: 10px; background-color: black;"></div>	35 %	42 %
17-07-2023	<div style="width: 100%; height: 10px; background-color: black;"></div>	29 %	35 %
17-07-2023	<div style="width: 100%; height: 10px; background-color: black;"></div>	21 %	29 %

Figura D.2: Además, se pueden ver informes más completos sobre el paciente para extraer mas información. Por ejemplo, el tipo histológico.

### Resultados

A partir de las características clínicas e histopatológicas entregadas, la paciente [REDACTED] tiene:

- 71% de probabilidad de tener respuesta al tratamiento neoadyuvante siendo tratada solo con **Trastuzumab**.
- 77% de probabilidad de tener respuesta al tratamiento neoadyuvante siendo tratada con **Trastuzumab y Pertuzumab**.

**Cerrar**

Figura D.3: Al enviar el formulario, se muestra un modal con los resultados de la predicción.

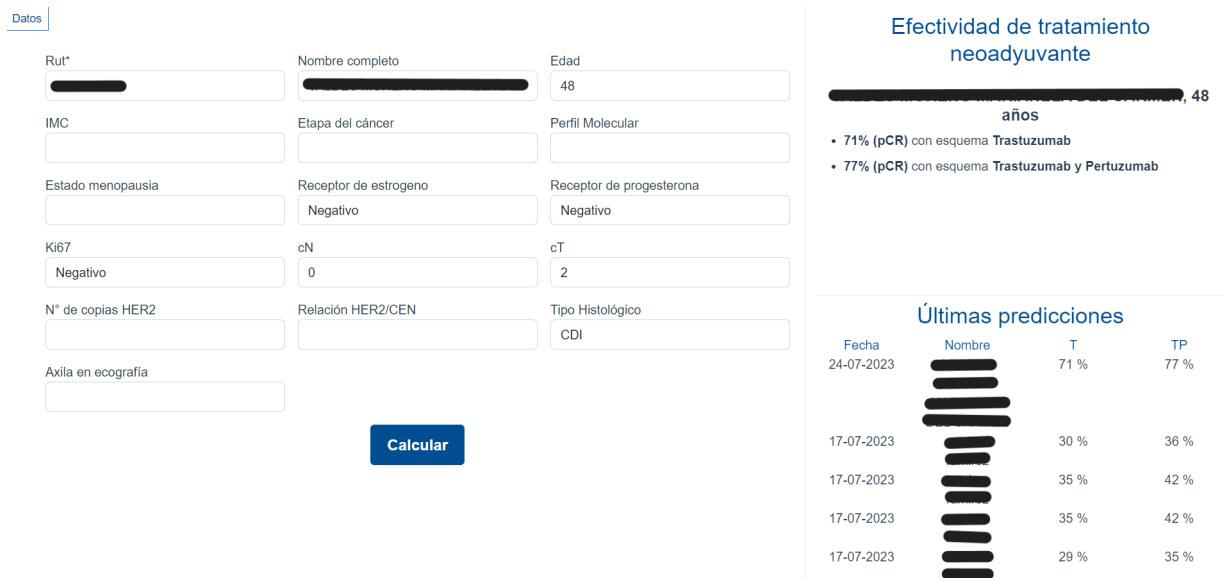


Figura D.4: Los resultados se muestran en la barra lateral derecha.

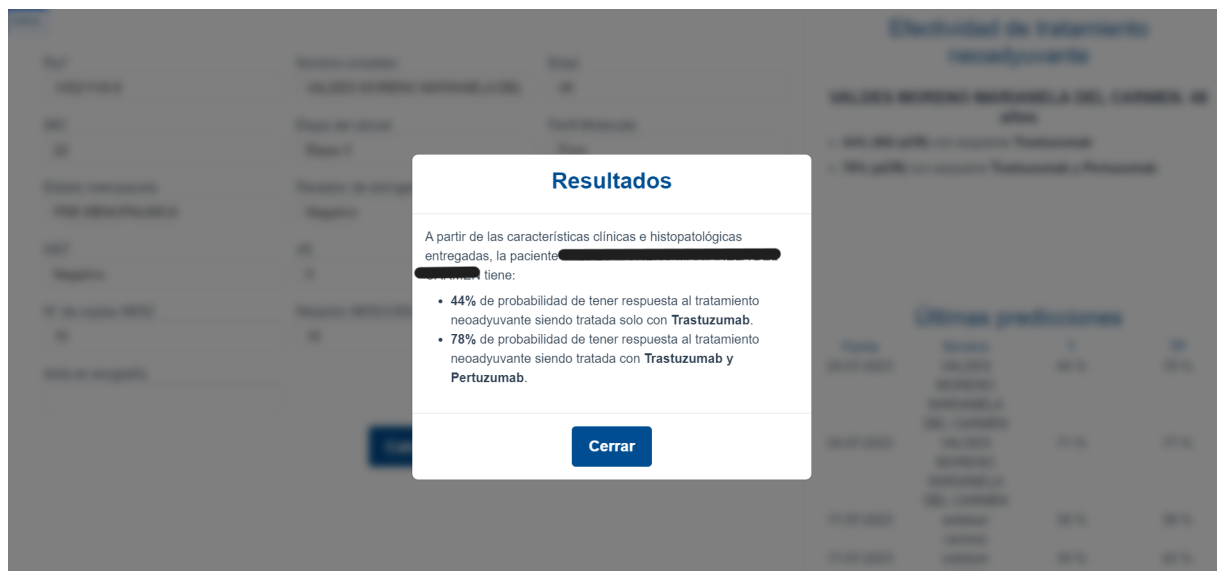


Figura D.5: Es posible realizar otro cálculo para el mismo paciente, agregando algunas de las variables ausentes.

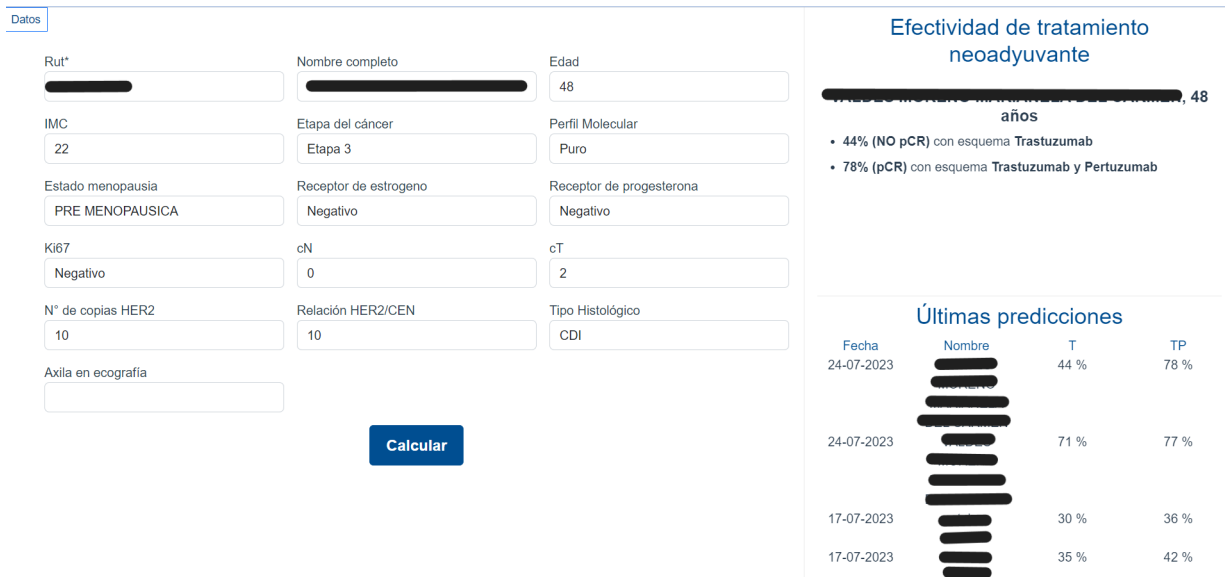


Figura D.6: Es posible comparar el nuevo output con los resultados anteriores.