

Draftable Comparison Export

This document is an exported comparison with limited functionality, generated by Draftable Desktop. To access full functionality, use Draftable's powerful comparison viewer in any of our products.

Left document: TI_Scordamaglia_Ezequiel_V3.pdf

Right document: TI_Scordamaglia_Ezequiel_V4.pdf

What is this document?

This is a comparison of two documents. The two documents are interleaved such that the left document is displayed on even pages and the right document is displayed on odd pages.

Is there a specific way I should view this file?

This document is intended to be viewed in Two Page Continuous mode (or sometimes called 'Two Page Scrolling'). It should open in this mode by default when using Adobe Acrobat and most popular PDF readers.

If the document opens in a different view, you can often change this in the settings. In Adobe Acrobat, go to **View > Page Display > Two Page Scrolling**.

Why are there blank pages?

Blank pages are inserted to keep both documents as aligned as much as possible.

How do I read the changes?

Text deleted from the left document and, hence, not in right document is highlighted red. Text added to the right document and, hence, not in left document is highlighted green.

Tip for printing

When printing this document, we recommend printing double-sided and include this first page. This will result in the matching text being displayed on different pages and easily readable, much like a book.

For more information

Draftable offers powerful document comparison solutions for all use-cases. To view our products, please visit our website: draftable.com.



CARRERA DE ESPECIALIZACIÓN EN INTELIGENCIA ARTIFICIAL

MEMORIA DEL TRABAJO FINAL

Construcción de un modelo para predecir la mortalidad en pacientes en diálisis renal

Autor:

Lic. Ezequiel Scordamaglia

Director:

Esp. Ing. Trinidad Monreal (FIUBA)

Jurados:

Nombre del jurado 1 (pertenencia)

Nombre del jurado 2 (pertenencia)

Nombre del jurado 3 (pertenencia)

*Este trabajo fue realizado en la Ciudad de Lanús, Buenos Aires,
entre diciembre de 2023 y junio de 2024.*



CARRERA DE ESPECIALIZACIÓN EN INTELIGENCIA ARTIFICIAL

MEMORIA DEL TRABAJO FINAL

Construcción de un modelo para predecir la mortalidad en pacientes en diálisis renal

Autor:

Lic. Ezequiel Scordamaglia

Director:

Esp. Ing. Trinidad Monreal (FIUBA)

Jurados:

Nombre del jurado 1 (pertenencia)

Nombre del jurado 2 (pertenencia)

Nombre del jurado 3 (pertenencia)

*Este trabajo fue realizado en la Ciudad de Lanús, Buenos Aires,
entre diciembre de 2023 y junio de 2024.*

Resumen

En esta memoria se describe el diseño y la implementación de un modelo de inteligencia artificial y la arquitectura necesaria para su utilización, desarrollado para una empresa médica que opera centros de atención en todo el país. El algoritmo utiliza datos clínicos de pacientes en tratamiento de diálisis renal con el propósito de predecir el riesgo de mortalidad. Como resultado este trabajo permite al personal médico definir estrategias de tratamiento personalizadas para mejorar la salud de los pacientes en riesgo. Se utilizaron técnicas de estadística y análisis de datos junto con modelos de aprendizaje automático y aprendizaje profundo.

Resumen

En esta memoria se describe el diseño y la implementación de un modelo de inteligencia artificial y la arquitectura necesaria para su utilización, desarrollado para una empresa médica que opera centros de atención en todo el país. El algoritmo utiliza datos clínicos de pacientes en tratamiento de diálisis renal con el propósito de predecir el riesgo de mortalidad. Como resultado este trabajo permite al personal médico definir estrategias de tratamiento personalizadas para mejorar la salud de los pacientes en riesgo. Se utilizaron técnicas de estadística y análisis de datos junto con modelos de aprendizaje automático y aprendizaje profundo.

Agradecimientos

Esta sección es para agradecimientos personales y es totalmente **OPCIONAL**.

Agradecimientos

Esta sección es para agradecimientos personales y es totalmente **OPCIONAL**.

Índice general

Resumen	I
1. Introducción general	1
1.1. Conceptos básicos de la diálisis renal	1
1.1.1. Diálisis renal y tipos de tratamiento	1
1.2. Contexto y motivación	2
1.3. Objetivos, alcance y requerimientos	3
1.3.1. Objetivos	3
1.3.2. Alcance	3
1.3.3. Requerimientos	4
1.4. Estado del arte	4
2. Introducción específica	7
2.1. Tratamiento de los datos	7
2.2. Desbalance de clases	9
2.3. Modelos de inteligencia artificial	10
2.4. Evaluación de modelos de clasificación	12
2.5. Plataforma de gestión de modelos	14
2.6. Servicios web	14
3. Diseño e implementación	17
3.1. Arquitectura propuesta	17
3.2. Adquisición de datos	18
3.3. Preprocesamiento de datos	21
4. Ensayos y resultados	23
4.1. Pruebas funcionales del hardware	23
5. Conclusiones	25
5.1. Conclusiones generales	25
5.2. Próximos pasos	25
Bibliografía	27

Índice general

Resumen	I
1. Introducción general	1
1.1. Conceptos básicos de la diálisis renal	1
1.1.1. Diálisis renal y tipos de tratamiento	1
1.2. Contexto y motivación	2
1.3. Objetivos, alcance y requerimientos	3
1.3.1. Objetivos	3
1.3.2. Alcance	3
1.3.3. Requerimientos	4
1.4. Estado del arte	4
2. Introducción específica	7
2.1. Tratamiento de los datos	7
2.1.1. Valores faltantes	7
2.1.2. Técnicas de imputación	8
2.1.3. Discretización	9
2.1.4. Codificación	9
2.1.5. Valores atípicos	9
2.1.6. Normalización	9
2.2. Desbalance de clases	10
2.3. Modelos de inteligencia artificial	12
2.4. Evaluación de modelos de clasificación	14
2.5. Plataforma de gestión de modelos	15
2.6. Servicios web	16
3. Diseño e implementación	17
3.1. Arquitectura propuesta	17
3.2. Adquisición de datos	18
3.3. Preprocesamiento de datos	21
3.3.1. Preprocesamiento de variables numéricas	21
3.3.2. Preprocesamiento de variables categóricas	23
3.3.3. Normalización de los datos	25
3.4. Estrategias aplicadas para balancear las clases	25
3.5. Diseño y desarrollo de modelos	25
3.5.1. Modelo de <i>Machine Learning</i>	25
3.5.2. Modelo de <i>Deep Learning</i>	26
3.5.3. Selección de funciones de activación, función de pérdida y algoritmo de optimización	28
3.5.4. Estrategias implementadas para evitar el <i>overfitting</i>	28
3.6. Conjunto de herramientas para acceso al modelo	29
4. Ensayos y resultados	31

4.1.	Resultados de los modelos	31
4.2.	Resultados en modelos de <i>Machine Learning</i>	31
4.3.	Resultados en modelos de <i>Deep Learning</i>	31
5.	Conclusiones	33
5.1.	Conclusiones generales	33
5.2.	Próximos pasos	33
	Bibliografía	35

Índice de figuras

1.1. Diferencia entre hemodiálisis y diálisis peritoneal.	2
1.2. Arquitectura de solución.	3
2.1. Submuestreo y sobremuestreo.	11
2.2. Machine Learning vs Deep Learning. ¹	13
2.3. Matriz de confusión.	14
3.1. Arquitectura del sistema.	17
3.2. Outliers en variables numéricas.	22
3.3. Distribuciones de variables numéricas discretizadas.	22
3.4. Distribuciones de variables numéricas.	24
4.1. Matriz de Confusión.	32

Índice de figuras

1.1. Diferencia entre hemodiálisis y diálisis peritoneal.	2
1.2. Arquitectura de solución.	3
2.1. Submuestreo y sobremuestreo.	10
2.2. Machine Learning vs Deep Learning. ¹	12
2.3. Matriz de confusión.	13
3.1. Arquitectura del sistema.	17

Índice de tablas

Índice de tablas

3.1. Variables discretizadas.	23
3.2. Valores faltantes en variables numéricas no discretizadas.	23
3.3. Valores faltantes en variables categóricas.	24
3.4. Desbalance de pacientes por estrategia.	25
3.5. Resultados de diferentes clasificadores	26
3.6. Comparación de modelos con diferentes configuraciones de capas ocultas.	27
3.7. Resultados de diferentes modelos.	27
3.8. Parámetros utilizados durante los entrenamientos.	28
4.1. Matriz de Confusión	32

Dedicado a... [OPCIONAL]

Dedicado a... [OPCIONAL]

Capítulo 1

Introducción general

En este capítulo se presentan los conceptos básicos de la diálisis renal. Además se mencionan las motivaciones que impulsan este trabajo de investigación, se establecen los objetivos, el alcance y los requerimientos, y se revisa el estado del arte en el campo de estudio.

1.1. Conceptos básicos de la diálisis renal

En esta sección se abordan las definiciones de diálisis renal y sus principales tipos de tratamiento.

1.1.1. Diálisis renal y tipos de tratamiento

La diálisis renal es un tratamiento en el que se extraen las toxinas y el exceso de agua de la sangre. Se utiliza como terapia renal sustitutiva cuando los riñones no funcionan correctamente debido a su deterioro. Los riñones desempeñan un papel crucial al eliminar las toxinas y el líquido de la sangre, evitando que los productos de desecho se acumulen en el cuerpo. Cuando los riñones no pueden realizar esta función, la diálisis se convierte en una herramienta vital.

Existen dos tipos principales de tratamientos de diálisis renal [1]:

- Hemodiálisis (HD): en este tratamiento se utiliza una membrana artificial. La purificación de la sangre se lleva a cabo mediante un riñón artificial, que elimina el exceso de agua, residuos y toxinas antes de devolverla al cuerpo. Cada sesión de hemodiálisis puede durar aproximadamente 4 horas y debe realizarse unas 3 veces por semana.
- Diálisis Peritoneal (DP): en este método de tratamiento, la filtración de la sangre se realiza en la cavidad peritoneal del paciente. Se utiliza un catéter permanente que se coloca en el abdomen, a través del cual se introduce una solución especial llamada líquido de diálisis en la cavidad peritoneal. Esta solución absorbe los desechos y el exceso de líquido del cuerpo a través de la membrana peritoneal, que actúa como una barrera semipermeable. Luego, después de un período de tiempo especificado (generalmente varias horas), el líquido de diálisis se drena del abdomen, llevando consigo los desechos y el exceso de líquido. Hay dos variantes de diálisis peritoneal:
 - Diálisis Peritoneal Continua Ambulatoria (DPCA): el paciente realiza los intercambios de líquido de diálisis manualmente varias veces al día, mientras sigue con sus actividades diarias. No se requiere una

Capítulo 1

Introducción general

En este capítulo se presentan los conceptos básicos de la diálisis renal. Además, se mencionan las motivaciones que impulsan este trabajo de investigación, se establecen los objetivos, el alcance y los requerimientos, y se revisa el estado del arte en el campo de estudio.

1.1. Conceptos básicos de la diálisis renal

En esta sección se abordan las definiciones de diálisis renal y sus principales tipos de tratamiento.

1.1.1. Diálisis renal y tipos de tratamiento

La diálisis renal es un tratamiento en el que se extraen las toxinas y el exceso de agua de la sangre. Se utiliza como terapia renal sustitutiva cuando los riñones no funcionan correctamente debido a su deterioro. Los riñones desempeñan un papel crucial al eliminar las toxinas y el líquido de la sangre, evitando que los productos de desecho se acumulen en el cuerpo. Cuando los riñones no pueden realizar esta función, la diálisis se convierte en una herramienta vital.

Existen dos tipos principales de tratamientos de diálisis renal [1]:

- Hemodiálisis (HD): en este tratamiento se utiliza una membrana artificial. La purificación de la sangre se lleva a cabo mediante un riñón artificial, que elimina el exceso de agua, residuos y toxinas antes de devolverla al cuerpo. Cada sesión de hemodiálisis puede durar aproximadamente 4 horas y debe realizarse unas 3 veces por semana.
- Diálisis Peritoneal (DP): en este método de tratamiento, la filtración de la sangre se realiza en la cavidad peritoneal del paciente. Se utiliza un catéter permanente que se coloca en el abdomen, a través del cual se introduce una solución especial llamada líquido de diálisis en la cavidad peritoneal. Esta solución absorbe los desechos y el exceso de líquido del cuerpo a través de la membrana peritoneal, que actúa como una barrera semipermeable. Luego, después de un período de tiempo especificado (generalmente varias horas), el líquido de diálisis se drena del abdomen, llevando consigo los desechos y el exceso de líquido. Hay dos variantes de diálisis peritoneal:
 - Diálisis Peritoneal Continua Ambulatoria (DPCA): el paciente realiza los intercambios de líquido de diálisis manualmente varias veces al día, mientras sigue con sus actividades diarias. No se requiere una

máquina para realizar los intercambios y el proceso es llevado a cabo por el paciente o su cuidador.

- **Diálisis Peritoneal Automática (DPA):** en este método, se utiliza una máquina cicladora para realizar los intercambios de líquido de diálisis durante la noche mientras el paciente duerme. La máquina administra automáticamente el líquido de diálisis, lo retira y lo reemplaza según un programa preestablecido. Esto permite una mayor flexibilidad en el tratamiento y puede ser más conveniente para algunos pacientes.

En la figura 1.1 se muestra la diferencia entre ambos tratamientos.

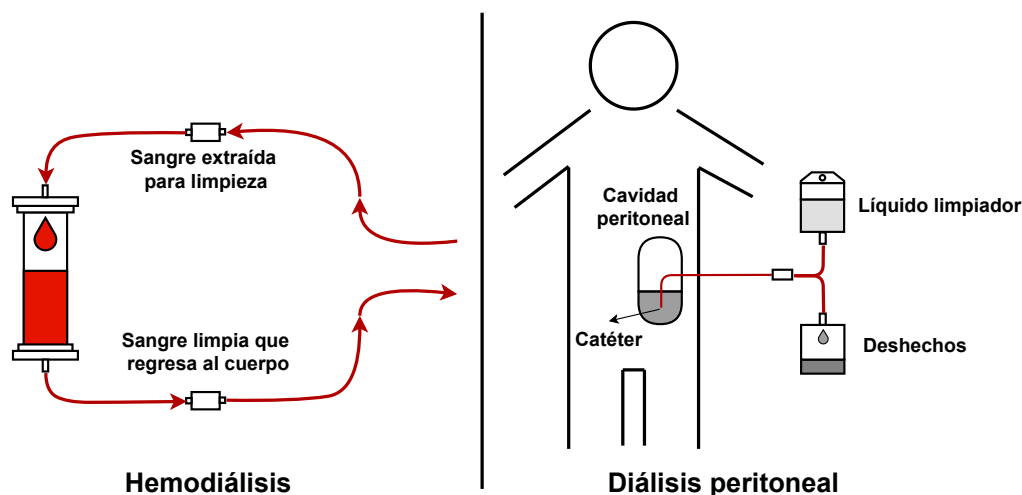


FIGURA 1.1. Diferencia entre hemodiálisis y diálisis peritoneal.

1.2. Contexto y motivación

La motivación principal de este trabajo es incorporar técnicas de inteligencia artificial (IA) en el campo de la medicina, dado que han demostrado una notable capacidad para anticipar eventos futuros basándose en datos históricos [2]. Sin embargo, uno de los desafíos recurrentes a la hora de entrenar modelos de IA es la obtención de datos representativos. En el ámbito de la medicina este desafío también se hace presente ya que se suele contar con pocos datos médicos de una población muy reducida. Para el desarrollo de este trabajo se contó con datos médicos de unos 14 000 pacientes que estuvieron bajo tratamiento de diálisis renal. Tanto el equipo de sistemas como el equipo médico de esta empresa de diálisis renal, si bien no tienen experiencias en herramientas de IA para predicción de eventos, conocen el potencial de estos modelos para identificar patrones, por lo que colaboraron en el desarrollo de este trabajo para lograr el cumplimiento del objetivo.

máquina para realizar los intercambios y el proceso es llevado a cabo por el paciente o su cuidador.

- **Diálisis Peritoneal Automática (DPA):** en este método, se utiliza una máquina cicladora para realizar los intercambios de líquido de diálisis durante la noche mientras el paciente duerme. La máquina administra automáticamente el líquido de diálisis, lo retira y lo reemplaza según un programa preestablecido. Esto permite una mayor flexibilidad en el tratamiento y puede ser más conveniente para algunos pacientes.

En la figura 1.1 se muestra la diferencia entre ambos tratamientos.

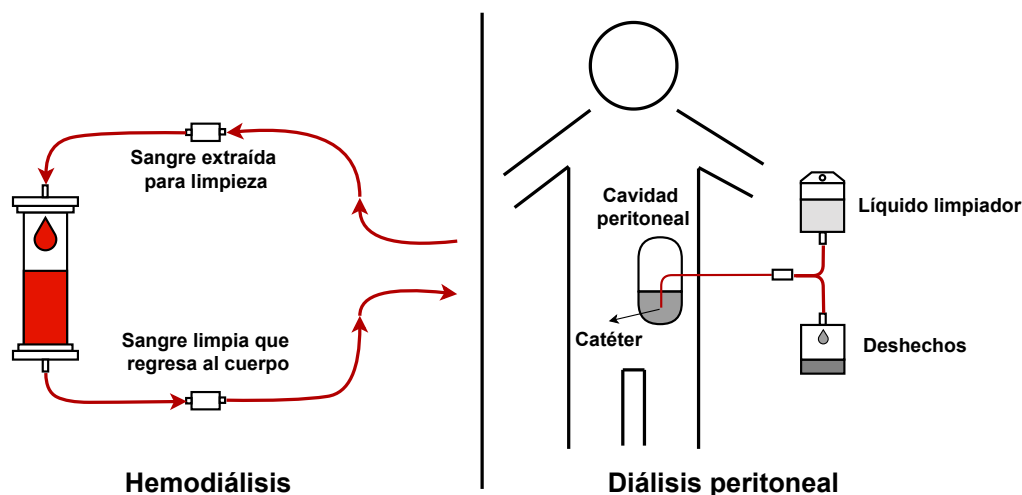


FIGURA 1.1. Diferencia entre hemodiálisis y diálisis peritoneal.

1.2. Contexto y motivación

La motivación principal de este trabajo es incorporar técnicas de inteligencia artificial (IA) en el campo de la medicina, dado que han demostrado una notable capacidad para anticipar eventos futuros basándose en datos históricos [2]. Sin embargo, uno de los desafíos recurrentes a la hora de entrenar modelos de IA es la obtención de datos representativos. En el ámbito de la medicina este desafío también se hace presente ya que se suele contar con pocos datos médicos de una población muy reducida. Para el desarrollo de este trabajo se contó con datos médicos de unos 14 000 pacientes que estuvieron bajo tratamiento de diálisis renal. Tanto el equipo de sistemas como el equipo médico de esta empresa de diálisis renal, si bien no tienen experiencias en herramientas de IA para predicción de eventos, conocen el potencial de estos modelos para identificar patrones, por lo que colaboraron en el desarrollo de este trabajo para lograr el cumplimiento del objetivo.

1.3. Objetivos, alcance y requerimientos

1.3.1. Objetivos

El propósito de este trabajo fue el desarrollo de un modelo de IA que permite predecir el riesgo de mortalidad en pacientes en diálisis renal, junto con la configuración de una plataforma de administración de modelos, una interfaz de comunicación con el modelo y un proceso que solicite las predicciones continuamente. Este conjunto de herramientas provee una predicción actualizada del riesgo de mortalidad de los pacientes, lo que permite al personal médico adaptar el tratamiento y la medicación prescrita para mejorarles su calidad de vida y, en última instancia, salvar vidas. En la figura 1.2 se muestra la arquitectura de la solución propuesta.

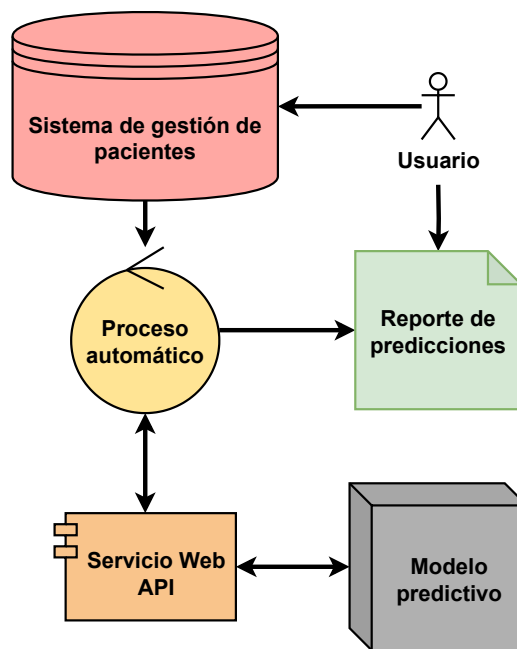


FIGURA 1.2. Arquitectura de solución.

1.3.2. Alcance

Se encuentra dentro del alcance del trabajo el desarrollo de un modelo de predicción de mortalidad, la instalación y configuración de una plataforma que permita desplegarlo en distintos ambientes, y la construcción de una interfaz y un proceso automático que soliciten predicciones y devuelvan la información al usuario. Asimismo, se incluye en el alcance la obtención y el preprocesamiento del conjunto de datos, que deben estar en cumplimiento con la ley 25.326 para garantizar el derecho al honor y a la intimidad de los pacientes, como así también el análisis de las distintas métricas para evaluar el correcto desempeño del modelo. Sin embargo, no se encuentra dentro del alcance del proyecto el desarrollo de una plataforma de gestión (sino que se eligió una existente que cumple con los requerimientos del trabajo) ni el desarrollo de una interfaz web orientada al usuario final. Tampoco se encuentra dentro del alcance la instalación del modelo predictivo en el entorno productivo de la empresa médica. Lo único que se instala será el proceso que recupera datos de los pacientes y llama al servicio web cada cierto período de tiempo para recuperar las predicciones.

1.3. Objetivos, alcance y requerimientos

1.3.1. Objetivos

El propósito de este trabajo fue el desarrollo de un modelo de IA que permite predecir el riesgo de mortalidad en pacientes en diálisis renal, junto con la configuración de una plataforma de administración de modelos, una interfaz de comunicación con el modelo y un proceso que solicite las predicciones continuamente. Este conjunto de herramientas provee una predicción actualizada del riesgo de mortalidad de los pacientes, lo que permite al personal médico adaptar el tratamiento y la medicación prescrita para mejorarles su calidad de vida y, en última instancia, salvar vidas. En la figura 1.2 se muestra la arquitectura de la solución propuesta.

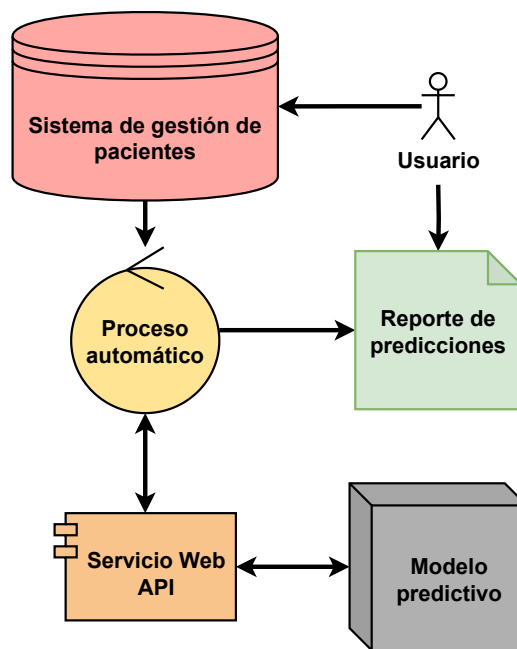


FIGURA 1.2. Arquitectura de solución.

1.3.2. Alcance

Se encuentra dentro del alcance del trabajo el desarrollo de un modelo de predicción de mortalidad, la instalación y configuración de una plataforma que permita desplegarlo en distintos ambientes, y la construcción de una interfaz y un proceso automático que soliciten predicciones y devuelvan la información al usuario. Asimismo, se incluye en el alcance la obtención y el preprocesamiento del conjunto de datos, que deben estar en cumplimiento con la ley 25.326 para garantizar el derecho al honor y a la intimidad de los pacientes, como así también el análisis de las distintas métricas para evaluar el correcto desempeño del modelo. Sin embargo, no se encuentra dentro del alcance del proyecto el desarrollo de una plataforma de gestión (sino que se eligió una existente que cumple con los requerimientos del trabajo) ni el desarrollo de una interfaz web orientada al usuario final. Tampoco se encuentra dentro del alcance la instalación del modelo predictivo en el entorno productivo de la empresa médica. Lo único que se instala será el proceso que recupera datos de los pacientes y llama al servicio web cada cierto período de tiempo para recuperar las predicciones.

1.3.3. Requerimientos

A continuación, se listan los requerimientos principales del trabajo agrupados por afinidad:

1. Requerimientos funcionales

- a) La plataforma de gestión de modelos deberá permitir desplegar modelos en diversos ambientes.
- b) La interfaz por servicio web deberá recibir datos médicos de uno o varios pacientes y devolver las predicciones asociadas a ellos.
- c) El modelo predictivo deberá tener una precisión de al menos un 75 %.
- d) El proceso que solicita predicciones y genera el reporte al usuario deberá poder ejecutarse automáticamente cada cierto período de tiempo.
- e) El reporte de predicciones que le llegue al usuario final deberá tener un formato claro y comprensible.
- f) Se utilizará GIT como repositorio para el control de versión de código.

2. Requerimientos de datos a utilizar

- a) Durante el entrenamiento del modelo se deberá resguardar la confidencialidad de los datos de los pacientes.

3. Requerimientos de documentación

- a) Se redactará una memoria técnica con la información del proyecto.
- b) La documentación de la interfaz por servicio web deberá incluir la lista de métodos disponibles con su detalle.
- c) La documentación del modelo predictivo incluirá información sobre el origen de los datos utilizados para el entrenamiento, las características que se usaron, el detalle del modelo seleccionado y la información que haya sobre la explicabilidad del modelo.

1.4. Estado del arte

Se llevó a cabo una exhaustiva revisión de la literatura relacionada con la predicción de la mortalidad de pacientes en diálisis renal y se ha encontrado principalmente una tesis doctoral [3] muy relevante que plantea un objetivo similar pero cuenta con muchos menos datos para el entrenamiento de los modelos. Si bien las variables médicas seleccionadas para realizar las predicciones en dicha tesis son muy similares a las que se seleccionaron en este trabajo, allí se plantea la discriminación de los casos según el tiempo que los pacientes llevan en diálisis, ya sea 3 meses, 6 meses, 1 año o más. Los modelos utilizados en dicha tesis incluyen *Random Forest* y Regresión Logística, algunos de los cuales también fueron utilizados en este trabajo. En cuanto a las conclusiones, para evaluar el desempeño de los modelos se utilizó la métrica de Área bajo la curva (AUC), que llega a valores entre 70 % y 73 %. En esta tesis también se muestra qué variables tienen más influencia al realizar la predicción de mortalidad del paciente, lo que resulta sumamente importante para el personal médico. También se ha encontrado

1.3.3. Requerimientos

A continuación, se listan los requerimientos principales del trabajo agrupados por afinidad:

1. Requerimientos funcionales

- a) La plataforma de gestión de modelos deberá permitir desplegar modelos en diversos ambientes.
- b) La interfaz por servicio web deberá recibir datos médicos de uno o varios pacientes y devolver las predicciones asociadas a ellos.
- c) El modelo predictivo deberá tener una precisión de al menos un 75 %.
- d) El proceso que solicita predicciones y genera el reporte al usuario deberá poder ejecutarse automáticamente cada cierto período de tiempo.
- e) El reporte de predicciones que le llegue al usuario final deberá tener un formato claro y comprensible.
- f) Se utilizará GIT como repositorio para el control de versión de código.

2. Requerimientos de datos a utilizar

- a) Durante el entrenamiento del modelo se deberá resguardar la confidencialidad de los datos de los pacientes.

3. Requerimientos de documentación

- a) Se redactará una memoria técnica con la información del proyecto.
- b) La documentación de la interfaz por servicio web deberá incluir la lista de métodos disponibles con su detalle.
- c) La documentación del modelo predictivo incluirá información sobre el origen de los datos utilizados para el entrenamiento, las características que se usaron, el detalle del modelo seleccionado y la información que haya sobre la explicabilidad del modelo.

1.4. Estado del arte

Se llevó a cabo una exhaustiva revisión de la literatura relacionada con la predicción de la mortalidad de pacientes en diálisis renal y se ha encontrado principalmente una tesis doctoral [3] muy relevante que plantea un objetivo similar pero cuenta con muchos menos datos para el entrenamiento de los modelos. Si bien las variables médicas seleccionadas para realizar las predicciones en dicha tesis son muy similares a las que se seleccionaron en este trabajo, allí se plantea la discriminación de los casos según el tiempo que los pacientes llevan en diálisis, ya sea 3 meses, 6 meses, 1 año o más. Los modelos utilizados en dicha tesis incluyen *Random Forest* y Regresión Logística, algunos de los cuales también fueron utilizados en este trabajo. En cuanto a las conclusiones, para evaluar el desempeño de los modelos se utilizó la métrica de Área bajo la curva (AUC), que llega a valores entre 70 % y 73 %. En esta tesis también se muestra qué variables tienen más influencia al realizar la predicción de mortalidad del paciente, lo que resulta sumamente importante para el personal médico. También se ha encontrado

una investigación [4] donde se utilizan técnicas de IA para predecir la mortalidad en pacientes con enfermedad renal crónica. Esta también cuenta con muy pocos datos de pacientes en diálisis renal y, aunque no se da mucho detalle sobre el entrenamiento de los modelos, concluye indicando que con un modelo de red neuronal se obtiene una predicción superior al 90 %, mejor que con una Regresión Logística. Por otro lado, se ha encontrado otra tesis [5] orientada a comprobar si el índice neutrófilo/linfocito es un predictor de mortalidad en pacientes que inician hemodiálisis. Si bien no se entrena ningún modelo de IA en la investigación, se concluye en que dicho índice no es un predictor de la mortalidad pero que la edad mayor a 60 años sí representa un factor de riesgo. Existen muchos otros trabajos relacionados con la temática de la predicción de mortalidad en pacientes con enfermedades renales, donde algunos abordan el beneficio que aporta la IA en la detección y predicción de eventos en medicina [6][7][8], otros la predicción de mortalidad teniendo enfermedad renal junto con enfermedad coronaria [9], y otros la predicción de contraer algún cáncer renal [10].

Si bien existe literatura académica que aborda el tema de la predicción de mortalidad de pacientes en diálisis renal, la mayoría de los trabajos se centra en investigaciones de carácter teórico y experimental. Además parten de conjuntos de datos muy chicos, lo que no permite a los modelos generalizar el conocimiento para poder realizar predicciones correctas.

Este trabajo se destaca por su enfoque práctico, ya que se desarrolla una herramienta concreta que pueda ser implementada en una empresa médica dedicada a la diálisis renal. Se espera obtener predicciones en tiempo real para los pacientes en diálisis renal y en base a su riesgo de mortalidad adecuar el tratamiento y la medicación prescrita. Esta iniciativa ofrece una solución práctica y viable en el ámbito clínico, lo que contribuye a mejorar la atención y seguimiento de los pacientes.

una investigación [4] donde se utilizan técnicas de IA para predecir la mortalidad en pacientes con enfermedad renal crónica. Esta también cuenta con muy pocos datos de pacientes en diálisis renal y, aunque no se da mucho detalle sobre el entrenamiento de los modelos, concluye indicando que con un modelo de red neuronal se obtiene una predicción superior al 90 %, mejor que con una Regresión Logística. Por otro lado, se ha encontrado otra tesis [5] orientada a comprobar si el índice neutrófilo/linfocito es un predictor de mortalidad en pacientes que inician hemodiálisis. Si bien no se entrena ningún modelo de IA en la investigación, se concluye en que dicho índice no es un predictor de la mortalidad pero que la edad mayor a 60 años sí representa un factor de riesgo. Existen muchos otros trabajos relacionados con la temática de la predicción de mortalidad en pacientes con enfermedades renales, donde algunos abordan el beneficio que aporta la IA en la detección y predicción de eventos en medicina [6][7][8], otros la predicción de mortalidad teniendo enfermedad renal junto con enfermedad coronaria [9], y otros la predicción de contraer algún cáncer renal [10].

Si bien existe literatura académica que aborda el tema de la predicción de mortalidad de pacientes en diálisis renal, la mayoría de los trabajos se centra en investigaciones de carácter teórico y experimental. Además parten de conjuntos de datos muy chicos, lo que no permite a los modelos generalizar el conocimiento para poder realizar predicciones correctas.

Este trabajo se destaca por su enfoque práctico, ya que se desarrolla una herramienta concreta que pueda ser implementada en una empresa médica dedicada a la diálisis renal. Se espera obtener predicciones en tiempo real para los pacientes en diálisis renal y en base a su riesgo de mortalidad adecuar el tratamiento y la medicación prescrita. Esta iniciativa ofrece una solución práctica y viable en el ámbito clínico, lo que contribuye a mejorar la atención y seguimiento de los pacientes.

Capítulo 2

Introducción específica

El objetivo de este capítulo es proporcionar una base teórica para comprender las herramientas y métodos utilizados en el desarrollo de este trabajo. En particular, se mencionarán las técnicas para el preprocesamiento de los datos, estrategias para el balanceo de clases, un repaso por los distintos modelos de inteligencia artificial y las métricas utilizadas para su evaluación. También se mencionarán conceptos básicos sobre plataformas de gestión de modelos de IA y servicios web.

2.1. Tratamiento de los datos

Una vez que se obtienen los datos para el entrenamiento de modelos de inteligencia artificial, es fundamental realizar un proceso de tratamiento de los datos para asegurar su calidad y adecuación para el análisis. En primer lugar, la división de los datos en conjuntos de entrenamiento y prueba es una práctica común. Esto se hace para evaluar el rendimiento del modelo de manera efectiva, utilizando un conjunto de datos separado para el entrenamiento y otro para la evaluación, lo que ayuda a evitar el sobreajuste y a evaluar la capacidad de generalización del modelo.

Un punto clave en el tratamiento de datos es la corrección de valores inconsistentes o nulos. Esto implica identificar y corregir valores atípicos, faltantes o errores de entrada que podrían distorsionar el análisis y afectar la precisión del modelo. El valor faltante puede darse por diversas razones, como errores durante el ingreso manual de datos, mediciones incorrectas, fallas en el experimento, entre otras [11]. Esta situación suele darse en cualquier conjunto de datos y puede tener un impacto significativo en el análisis y la interpretación de los resultados. Por lo tanto, es crucial entender las causas subyacentes y abordarlas adecuadamente durante el proceso de tratamiento de datos. Usualmente se clasifican en los siguientes grupos:

- Falta completamente al azar (*Missing completely at random* o MCAR): la probabilidad de que un registro tenga un valor faltante para un atributo no depende ni de los datos observados ni de los datos faltantes. Por ejemplo, una muestra de laboratorio que se pierde.
- Falta al azar (*Missing at random* o MAR): indica que la probabilidad de que un registro tenga un valor faltante para un atributo podría depender de los datos observados, pero no del valor del dato faltante en sí mismo. Por ejemplo, las personas con ingresos más altos pueden ser menos propensas a revelarlos en una encuesta, pero si la falta de respuesta es aleatoria dentro de las clases de ingresos, los datos de ingresos son faltantes al azar.

Capítulo 2

Introducción específica

El objetivo de este capítulo es proporcionar una base teórica para comprender las herramientas y métodos utilizados en el desarrollo de este trabajo. En particular, se mencionarán las técnicas para el preprocesamiento de los datos, estrategias para el balanceo de clases, un repaso por los distintos modelos de inteligencia artificial y las métricas utilizadas para su evaluación. También se mencionarán conceptos básicos sobre plataformas de gestión de modelos de IA y servicios web.

2.1. Tratamiento de los datos

Una vez que se obtienen los datos para el entrenamiento de modelos de inteligencia artificial, es fundamental realizar un proceso de tratamiento de los datos para asegurar su calidad y adecuación para el análisis. En primer lugar, la división de los datos en conjuntos de entrenamiento y prueba es una práctica común. Esto se hace para evaluar el rendimiento del modelo de manera efectiva, utilizando un conjunto de datos separado para el entrenamiento y otro para la evaluación, lo que ayuda a evitar el sobreajuste y a evaluar la capacidad de generalización del modelo. Tanto al conjunto de entrenamiento como al conjunto de prueba se le aplican todos los pasos del preprocesamiento que se describen a continuación.

2.1.1. Valores faltantes

Un punto clave en el tratamiento de datos es la corrección de valores inconsistentes o nulos. Esto implica identificar y corregir valores atípicos, faltantes o errores de entrada que podrían distorsionar el análisis y afectar la precisión del modelo. El valor faltante puede darse por diversas razones, como errores durante el ingreso manual de datos, mediciones incorrectas, fallas en el experimento, entre otras [11]. Esta situación suele darse en cualquier conjunto de datos y puede tener un impacto significativo en el análisis y la interpretación de los resultados. Por lo tanto, es crucial entender las causas subyacentes y abordarlas adecuadamente durante el proceso de tratamiento de datos. Usualmente se clasifican en los siguientes grupos:

- Falta completamente al azar (*Missing completely at random* o MCAR): la probabilidad de que un registro tenga un valor faltante para un atributo no depende ni de los datos observados ni de los datos faltantes. Por ejemplo, una muestra de laboratorio que se pierde.
- Falta al azar (*Missing at random* o MAR): indica que la probabilidad de que un registro tenga un valor faltante para un atributo podría depender de los datos observados, pero no del valor del dato faltante en sí mismo. Por

- Falta no al azar (*Missing not at random* o MNAR): implica que la probabilidad de que un registro tenga un valor faltante para un atributo podría depender del valor del atributo mismo. Por ejemplo, un sensor que no detecta temperaturas por debajo de cierto umbral o personas que no completan los ingresos anuales en encuestas si superan cierto valor.

Para resolver el problema de los valores faltantes, existen diversos métodos de imputación. Estos son algunos de los más utilizados:

- Imputación por la media o mediana: se remplazan los valores faltantes con la media o la mediana de la variable correspondiente. Esto es simple y rápido, pero puede introducir sesgos si la distribución de los datos es asimétrica o tiene valores atípicos.
- Imputación por el valor más frecuente: se remplazan los valores faltantes con el valor más común de la variable. Es útil para variables categóricas o variables con una distribución de frecuencia clara.
- Imputación por regresión: se utilizan modelos de regresión para predecir los valores faltantes a partir de las variables restantes. Esto puede ser más preciso que los métodos anteriores, pero puede ser computacionalmente intensivo y requiere asumir una relación lineal entre las variables.
- Imputación por vecinos más cercanos (*K Nearest Neighbor* o KNN): los valores faltantes se estiman a partir de los valores de observaciones similares en el espacio de características. Este método puede ser efectivo en conjuntos de datos con estructuras de vecindario claras, pero puede ser sensible a la elección de la métrica de distancia y número de vecinos.
- Imputación Múltiple por Ecuaciones Encadenadas (MICE): se imputan los valores faltantes mediante la estimación secuencial de modelos predictivos, utilizando las variables restantes como predictores. Este método captura la incertidumbre asociada con la imputación de valores faltantes y puede proporcionar estimaciones más precisas.

Si los datos faltantes se consideran completamente aleatorios (MCAR), los métodos de imputación simples como la imputación por la media o mediana pueden ser apropiados. Cuando los datos faltantes son aleatorios (MAR), se pueden utilizar métodos más sofisticados como la imputación por regresión que utilizan información de otras variables observadas. Para datos no aleatorios (MNAR), la elección del método de imputación es más compleja y puede requerir técnicas específicas que modelen la relación entre los datos faltantes y los observados. Siempre es importante tener en cuenta las limitaciones y el contexto específico del conjunto de datos al tomar decisiones sobre el tratamiento de valores faltantes.

Otra de las técnicas que se puede aplicar a los datos es la discretización, que es la conversión de variables continuas en variables discretas o categóricas. Esta técnica ayuda a que los modelos generen reglas más breves y comprensibles, lo que reduce la complejidad, y también es útil para aumentar la generalización y precisión del conocimiento [12].

ejemplo, las personas con ingresos más altos pueden ser menos propensas a revelarlos en una encuesta, pero si la falta de respuesta es aleatoria dentro de las clases de ingresos, los datos de ingresos son faltantes al azar.

- Falta no al azar (*Missing not at random* o MNAR): implica que la probabilidad de que un registro tenga un valor faltante para un atributo podría depender del valor del atributo mismo. Por ejemplo, un sensor que no detecta temperaturas por debajo de cierto umbral o personas que no completan los ingresos anuales en encuestas si superan cierto valor.

2.1.2. Técnicas de imputación

Para resolver el problema de los valores faltantes, existen diversos métodos de imputación. Estos son algunos de los más utilizados:

- Imputación por la media o mediana: se remplazan los valores faltantes con la media o la mediana de la variable correspondiente. Esto es simple y rápido, pero puede introducir sesgos si la distribución de los datos es asimétrica o tiene valores atípicos.
- Imputación por el valor más frecuente: se remplazan los valores faltantes con el valor más común de la variable. Es útil para variables categóricas o variables con una distribución de frecuencia clara.
- Imputación por regresión: se utilizan modelos de regresión para predecir los valores faltantes a partir de las variables restantes. Esto puede ser más preciso que los métodos anteriores, pero puede ser computacionalmente intensivo y requiere asumir una relación lineal entre las variables.
- Imputación por vecinos más cercanos (*K Nearest Neighbor* o KNN): los valores faltantes se estiman a partir de los valores de observaciones similares en el espacio de características. Este método puede ser efectivo en conjuntos de datos con estructuras de vecindario claras, pero puede ser sensible a la elección de la métrica de distancia y número de vecinos.
- Imputación Múltiple por Ecuaciones Encadenadas (MICE): se imputan los valores faltantes mediante la estimación secuencial de modelos predictivos, utilizando las variables restantes como predictores. Este método captura la incertidumbre asociada con la imputación de valores faltantes y puede proporcionar estimaciones más precisas.

Si los datos faltantes se consideran completamente aleatorios (MCAR), los métodos de imputación simples como la imputación por la media o mediana pueden ser apropiados. Cuando los datos faltantes son aleatorios (MAR), se pueden utilizar métodos más sofisticados como la imputación por regresión que utilizan información de otras variables observadas. Para datos no aleatorios (MNAR), la elección del método de imputación es más compleja y puede requerir técnicas específicas que modelen la relación entre los datos faltantes y los observados. Siempre es importante tener en cuenta las limitaciones y el contexto específico del conjunto de datos al tomar decisiones sobre el tratamiento de valores faltantes.

2.1.3. Discretización

Otra de las técnicas que se puede aplicar a los datos es la discretización, que es la conversión de variables continuas en variables discretas o categóricas. Esta técnica ayuda a que los modelos generen reglas más breves y comprensibles, lo que reduce la complejidad, y también es útil para aumentar la generalización y precisión del conocimiento [12].

2.1.4. Codificación

Para las variables categóricas, se suele realizar una codificación, como *one-hot*, que convierte las variables categóricas en representaciones numéricas, asignando un valor binario a cada categoría. Además, existen otras técnicas de codificación, como la de etiquetas (*label encoding*) y la de frecuencia (*frequency encoding*), que también se utilizan para transformar variables categóricas en datos numéricos compatibles con algoritmos de aprendizaje automático [13].

2.1.5. Valores atípicos

Los valores atípicos, también conocidos como *outliers*, son puntos de datos que se desvían significativamente del resto de la distribución de los datos. Estos pueden ser causados por errores en la recopilación de datos, eventos raros o simplemente representar variaciones genuinas en el fenómeno estudiado. En el contexto del entrenamiento de modelos, los valores atípicos pueden distorsionar los resultados y reducir la precisión del modelo al introducir ruido en los datos. Esto puede llevar a que el modelo aprenda patrones incorrectos o sesgados, lo que afecta negativamente su capacidad para generalizar en datos nuevos.

Para mitigar el impacto de los valores atípicos en el entrenamiento del modelo, existen varias técnicas que se pueden aplicar:

- Eliminación basada en percentiles: consiste en eliminar los valores que se encuentran por encima o por debajo de ciertos percentiles predefinidos en la distribución de los datos. Por ejemplo, podríamos eliminar los puntos por encima del percentil 95 o por debajo del percentil 5.
- *Capping* o truncamiento: esta técnica implica establecer un límite superior e inferior para los valores en el conjunto de datos y luego reemplazar cualquier valor que esté por encima o por debajo de estos límites con el valor del límite correspondiente. Esto ayuda a limitar el impacto de los *outliers* sin eliminar completamente los datos.

Además, algunos algoritmos de modelado son inherentemente robustos a los valores atípicos, como los árboles de decisión y los modelos basados en reglas, lo que los hace preferibles en situaciones donde los *outliers* son comunes o difíciles de eliminar.

2.1.6. Normalización

Como último punto se encuentra la normalización de datos, que implica ajustar los valores de las variables de entrada para que tengan una escala similar. Algunas razones para normalizar los datos incluyen:

Y por último, para obtener un conjunto de datos completamente numérico, se suele utilizar la codificación de variables categóricas. Esta técnica incluye métodos como la codificación *one-hot* que convierte las variables categóricas en representaciones numéricas, asignando un valor binario a cada categoría. Además existen otras técnicas de codificación, como la de etiquetas (*label encoding*) y la de frecuencia (*frequency encoding*), que también se utilizan para transformar variables categóricas en datos numéricos compatibles con algoritmos de aprendizaje automático [13].

Las técnicas anteriormente mencionadas proveen soluciones a los problemas que suelen encontrarse en los conjuntos de datos y ayudan a garantizar la calidad y la adecuación para su posterior análisis y entrenamiento de modelos.

2.2. Desbalance de clases

El desbalance de clases es un problema común en el aprendizaje automático, donde una o más clases están subrepresentadas en comparación con otras en el conjunto de datos [14]. Esto puede ser problemático porque los algoritmos de aprendizaje automático tienden a favorecer las clases mayoritarias y pueden tener dificultades para aprender patrones en las clases minoritarias. El desbalance de clases puede llevar a modelos sesgados y poco precisos, especialmente en problemas de clasificación donde la precisión de las clases minoritarias es de particular interés, como la detección de fraudes, enfermedades o anomalías.

Algunas técnicas comunes para abordar el desbalance de clases incluyen:

- Submuestreo: se reduce el número de muestras de las clases mayoritarias para equilibrar la proporción de clases. Esto puede ayudar a prevenir el sesgo hacia las clases mayoritarias, pero también puede resultar en la pérdida de información valiosa.
- Sobremuestreo: implica aumentar el número de muestras de las clases minoritarias mediante técnicas como la replicación de instancias o la generación de nuevas muestras sintéticas. Esto puede ayudar a mejorar la representación de las clases minoritarias y a evitar el sesgo hacia las clases mayoritarias.
- Ponderación de clases: algunos algoritmos de aprendizaje automático permiten asignar pesos diferentes a las clases para tener en cuenta el desbalance de clases durante el entrenamiento del modelo. Esto puede ayudar a compensar la falta de representación de las clases minoritarias.

En la figura 2.1 se muestra la diferencia entre las técnicas de submuestreo y sobremuestreo.

Para el caso de submuestreo, existen varias técnicas utilizadas para equilibrar el desbalance de clases. Una de ellas es el submuestreo aleatorio (*Random Undersampling*), donde se elimina aleatoriamente un subconjunto de muestras de la clase mayoritaria para igualar el número de muestras de la clase minoritaria. Otra técnica es la de muestras cercanas (*NearMiss*), que elimina muestras de la clase mayoritaria que están próximas a las muestras de la clase minoritaria en el espacio de características, y preserva así la información relevante.

- Evitar problemas de escala: si las características tienen escalas muy diferentes, algunas características pueden dominar el proceso de entrenamiento, lo que puede llevar a un rendimiento deficiente del modelo.
- Facilitar el proceso de entrenamiento: la normalización puede ayudar al optimizador a converger más rápido durante el entrenamiento, lo que puede resultar en tiempos de entrenamiento más cortos.
- Mejorar la generalización: la normalización también puede contribuir a que el modelo sea más capaz de generalizar correctamente a nuevos datos, al prevenir que se sobreajuste a ciertos rangos específicos de valores de las características.

Existen varias técnicas de normalización, como la normalización min-max, la estandarización (*z-score*), la normalización por rangos, entre otras [14]. La elección de la técnica de normalización dependerá del tipo de datos y del problema en cuestión.

Las técnicas anteriormente mencionadas proveen soluciones a los problemas que suelen encontrarse en los conjuntos de datos y ayudan a garantizar la calidad y la adecuación para su posterior análisis y entrenamiento de modelos.

2.2. Desbalance de clases

El desbalance de clases es un problema común en el aprendizaje automático, donde una o más clases están subrepresentadas en comparación con otras en el conjunto de datos [15]. Esto puede ser problemático porque los algoritmos de aprendizaje automático tienden a favorecer las clases mayoritarias y pueden tener dificultades para aprender patrones en las clases minoritarias. El desbalance de clases puede llevar a modelos sesgados y poco precisos, especialmente en problemas de clasificación donde la precisión de las clases minoritarias es de particular interés, como la detección de fraudes, enfermedades o anomalías.

Algunas técnicas comunes para abordar el desbalance de clases incluyen:

- Submuestreo: se reduce el número de muestras de las clases mayoritarias para equilibrar la proporción de clases. Esto puede ayudar a prevenir el sesgo hacia las clases mayoritarias, pero también puede resultar en la pérdida de información valiosa.
- Sobremuestreo: implica aumentar el número de muestras de las clases minoritarias mediante técnicas como la replicación de instancias o la generación de nuevas muestras sintéticas. Esto puede ayudar a mejorar la representación de las clases minoritarias y a evitar el sesgo hacia las clases mayoritarias.
- Ponderación de clases: algunos algoritmos de aprendizaje automático permiten asignar pesos diferentes a las clases para tener en cuenta el desbalance de clases durante el entrenamiento del modelo. Esto puede ayudar a compensar la falta de representación de las clases minoritarias.

En la figura 2.1 se muestra la diferencia entre las técnicas de submuestreo y sobremuestreo.

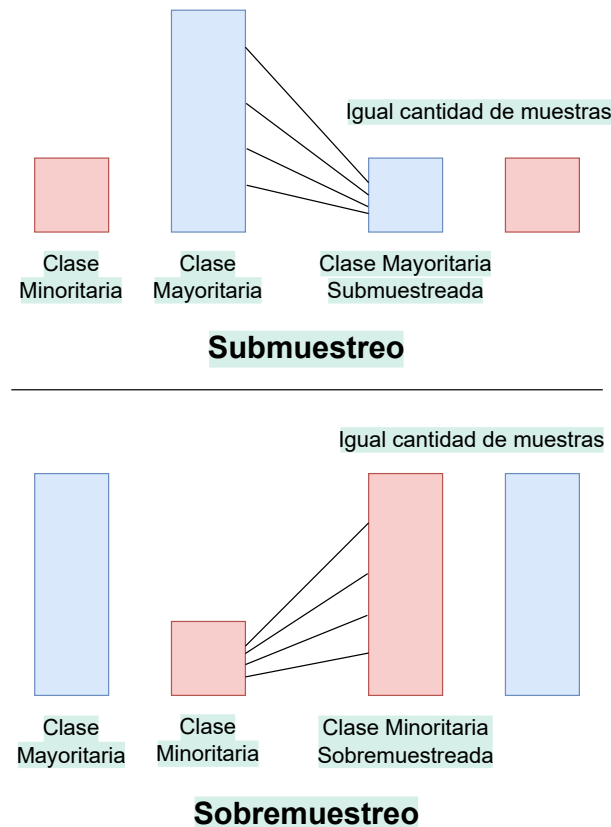


FIGURA 2.1. Submuestreo y sobremuestreo.

En cuanto al sobremuestreo, se puede aplicar la técnica de duplicación de muestras aleatorias (*Random Oversampling*), que consiste en duplicar las muestras de la clase minoritaria para aumentar su representación en el conjunto de datos. Pero también existen técnicas para la generación de muestras sintéticas, como SMOTE (*Synthetic Minority Over-sampling Technique*), que crea nuevas muestras interpoladas entre las muestras existentes de la clase minoritaria, lo que ayuda a mejorar su representación sin duplicar directamente las muestras existentes.

Cada técnica tiene sus ventajas y desventajas y muchas veces se usa una combinación de ellas. La elección adecuada depende del conjunto de datos específico y del problema en cuestión. Es importante experimentar con diferentes enfoques y evaluar su rendimiento para determinar la estrategia más efectiva.

2.3. Modelos de inteligencia artificial

Los modelos de inteligencia artificial son comúnmente utilizados para reconocer patrones en grandes conjuntos de datos y obtener predicciones. Se dice que los modelos aprenden cuando logran mejorar sus resultados en una tarea específica luego de procesar muchos datos y sin obtener instrucciones explícitas de un programador [2]. Los tipos de aprendizaje se dividen en los siguientes:

- Aprendizaje supervisado: para entrenar al modelo se utiliza un conjunto de datos etiquetados. Esto quiere decir que se le provee tanto las características como el valor objetivo esperado. El modelo aprende a hacer predicciones

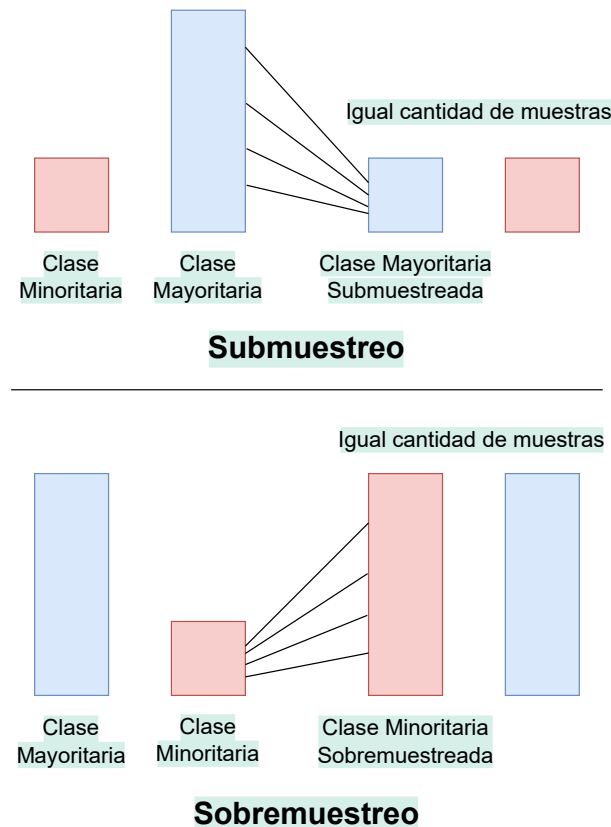


FIGURA 2.1. Submuestreo y sobremuestreo.

Para el caso de submuestreo, existen varias técnicas utilizadas para equilibrar el desbalance de clases. Una de ellas es el submuestreo aleatorio (*Random Undersampling*), donde se elimina aleatoriamente un subconjunto de muestras de la clase mayoritaria para igualar el número de muestras de la clase minoritaria. Otra técnica es la de muestras cercanas (*NearMiss*), que elimina muestras de la clase mayoritaria que están próximas a las muestras de la clase minoritaria en el espacio de características, y preserva así la información relevante.

En cuanto al sobremuestreo, se puede aplicar la técnica de duplicación de muestras aleatorias (*Random Oversampling*), que consiste en duplicar las muestras de la clase minoritaria para aumentar su representación en el conjunto de datos. Pero también existen técnicas para la generación de muestras sintéticas, como SMOTE (*Synthetic Minority Over-sampling Technique*), que crea nuevas muestras interpoladas entre las muestras existentes de la clase minoritaria, lo que ayuda a mejorar su representación sin duplicar directamente las muestras existentes.

Y por último, también se pueden ajustar ciertas configuraciones de los modelos para que las predicciones incorrectas en las clases minoritarias sean penalizadas durante el entrenamiento. Una técnica ampliamente utilizada para esto es la *Focal Loss* (Pérdida Focal) [16], que aborda el problema del desbalance de clases mediante la modulación de la pérdida de los ejemplos difíciles durante el entrenamiento del modelo. Esta se basa en dos conceptos principales: la reducción de la importancia de los ejemplos fáciles (parámetro de modulación) y el enfoque en los ejemplos difíciles (parámetro de focalidad).

basadas en estos ejemplos y se ajusta para minimizar los errores entre las predicciones y las etiquetas conocidas.

- Aprendizaje no supervisado: no se utilizan etiquetas en los datos de entrenamiento. El modelo explora patrones y estructuras en los datos sin una guía explícita. Este enfoque es útil cuando no se conoce las categorías de antemano y se quiere descubrir patrones ocultos.
- Aprendizaje por refuerzo: el modelo aprende a tomar decisiones a través de la interacción con un entorno. Para cada acción recibirá recompensas o castigos según su desempeño. El objetivo es maximizar las recompensas a lo largo del tiempo.

Teniendo los datos y el tipo de aprendizaje que se desea implementar, se debe buscar también una arquitectura de modelo que tenga la capacidad de aprender de los datos y devolver predicciones. Dentro de la inteligencia artificial, existe un universo que se conoce como *Machine Learning* (ML) o aprendizaje automático, que refiere a aquellos algoritmos que utilizan métodos estadísticos para analizar datos, aprender de ellos y elaborar predicciones o sugerencias. Entre los modelos de ML más conocidos se encuentran los siguientes:

- Regresión Lineal: se utiliza para predecir valores continuos basados en variables independientes. Busca la relación lineal entre las variables de entrada y salida.
- Regresión Logística: se utiliza para clasificación binaria. Estima la probabilidad de que una instancia pertenezca a una determinada clase.
- Árboles de Decisión: organizan las características de los datos en una estructura similar a un árbol. Cada nodo en este árbol representa una pregunta sobre una característica específica de los datos. Por ejemplo, podría ser ¿Tiene diabetes? o ¿Es menor de 30 años?. Las ramas del árbol representan las posibles respuestas a estas preguntas, como sí o no. Siguiendo las ramas del árbol, eventualmente se llega a una hoja que representa la decisión o predicción final.
- *Random Forest*: es una técnica de conjunto que combina múltiples árboles de decisión. Cada árbol se entrena con una muestra aleatoria del conjunto de datos y luego las predicciones se promedian para obtener la salida final.
- *Support Vector Machine* (SVM): es un algoritmo de clasificación que encuentra el hiperplano que mejor separa las clases en un espacio de características de alta dimensión. Puede ser usado tanto para clasificación como para regresión.

Dentro del universo ML, hay otro grupo más chico que se denomina *Deep Learning* (DL) o aprendizaje profundo, en donde los algoritmos utilizan una arquitectura de redes neuronales que simulan el comportamiento del cerebro humano, por lo que suelen ser mucho más grandes y complejos. Estos últimos suelen usarse para tareas de visión por computadora o procesamiento de lenguaje natural, y requieren mucha potencia de cómputo y grandes cantidades de datos. En la figura 2.2 se puede visualizar la diferencia entre utilizar modelos de *Machine Learning* y modelos de *Deep Learning* para obtener predicciones.

Cada técnica tiene sus ventajas y desventajas y muchas veces se usa una combinación de ellas. La elección adecuada depende del conjunto de datos específico y del problema en cuestión. Es importante experimentar con diferentes enfoques y evaluar su rendimiento para determinar la estrategia más efectiva.

2.3. Modelos de inteligencia artificial

Los modelos de inteligencia artificial son comúnmente utilizados para reconocer patrones en grandes conjuntos de datos y obtener predicciones. Se dice que los modelos aprenden cuando logran mejorar sus resultados en una tarea específica luego de procesar muchos datos y sin obtener instrucciones explícitas de un programador [2]. Los tipos de aprendizaje se dividen en los siguientes:

- **Aprendizaje supervisado:** para entrenar al modelo se utiliza un conjunto de datos etiquetados. Esto quiere decir que se le provee tanto las características como el valor objetivo esperado. El modelo aprende a hacer predicciones basadas en estos ejemplos y se ajusta para minimizar los errores entre las predicciones y las etiquetas conocidas.
- **Aprendizaje no supervisado:** no se utilizan etiquetas en los datos de entrenamiento. El modelo explora patrones y estructuras en los datos sin una guía explícita. Este enfoque es útil cuando no se conoce las categorías de antemano y se quiere descubrir patrones ocultos.
- **Aprendizaje por refuerzo:** el modelo aprende a tomar decisiones a través de la interacción con un entorno. Para cada acción recibirá recompensas o castigos según su desempeño. El objetivo es maximizar las recompensas a lo largo del tiempo.

Teniendo los datos y el tipo de aprendizaje que se desea implementar, se debe buscar también una arquitectura de modelo que tenga la capacidad de aprender de los datos y devolver predicciones. Dentro de la inteligencia artificial, existe un universo que se conoce como *Machine Learning* (ML) o aprendizaje automático, que refiere a aquellos algoritmos que utilizan métodos estadísticos para analizar datos, aprender de ellos y elaborar predicciones o sugerencias. Entre los modelos de ML más conocidos se encuentran los siguientes:

- **Regresión Lineal:** se utiliza para predecir valores continuos basados en variables independientes. Busca la relación lineal entre las variables de entrada y salida.
- **Regresión Logística:** se utiliza para clasificación binaria. Estima la probabilidad de que una instancia pertenezca a una determinada clase.
- **Árboles de Decisión:** organizan las características de los datos en una estructura similar a un árbol. Cada nodo en este árbol representa una pregunta sobre una característica específica de los datos. Por ejemplo, podría ser ¿Tiene diabetes? o ¿Es menor de 30 años?. Las ramas del árbol representan las posibles respuestas a estas preguntas, como sí o no. Siguiendo las ramas del árbol, eventualmente se llega a una hoja que representa la decisión o predicción final.

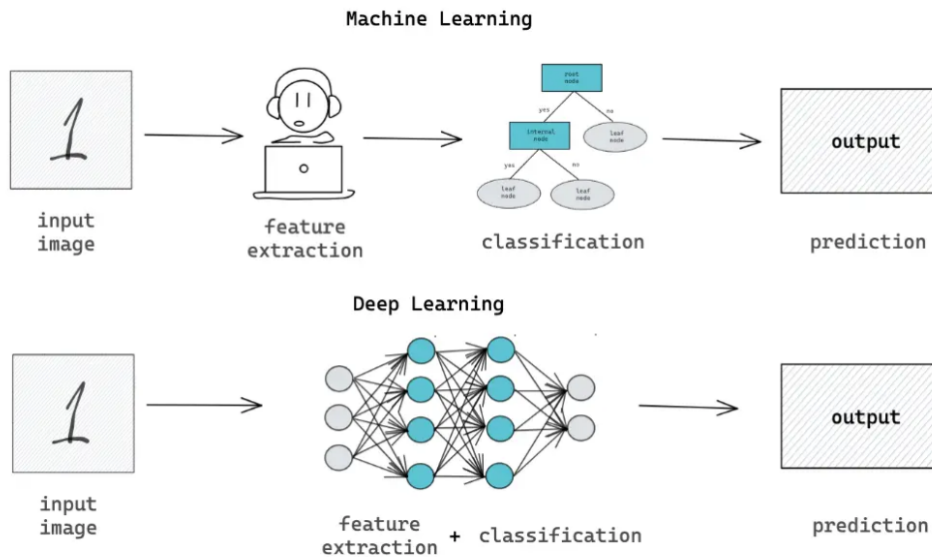


FIGURA 2.2. Machine Learning vs Deep Learning. ¹.

Los modelos también se diferencian por el tipo de predicción que realizan. A un modelo que predice un valor continuo, como el precio de una vivienda, se lo conoce como modelo de regresión. Mientras que a un modelo que predice una etiqueta o clase, ya sea binaria o multi-clase, se lo conoce como modelo de clasificación.

Este último tipo de modelo es muy utilizado en medicina, ya que sirve para predecir la presencia o ausencia de cierta enfermedad o para predecir su tipo específico. En este trabajo en particular se utilizaron modelos de clasificación con aprendizaje supervisado y se entrenaron modelos de *Machine Learning* y *Deep Learning*.

2.4. Evaluación de modelos de clasificación

Para evaluar el comportamiento de un modelo de clasificación binaria comúnmente se recurre a una herramienta llamada matriz de confusión. Esta matriz, como se muestra en la figura 2.3, presenta las clases predichas en las columnas y las clases reales en las filas. A partir de esta tabla, se derivan cuatro métricas clave:

- Verdaderos positivos (*True Positive* o TP): representan las predicciones correctas de una condición positiva.
- Verdaderos negativos (*True Negative* o TN): son las predicciones correctas de una condición negativa.
- Falsos positivos (*False Positive* o FP): se dan cuando el modelo predice incorrectamente una condición positiva que en realidad no lo es.
- Falsos negativos (*False Negative* o FN): se dan cuando el modelo predice incorrectamente la ausencia de una condición positiva que en realidad está presente.

¹Imagen tomada de <https://www.stratascratch.com/blog/data-science-vs-machine-learning-vs-deep-learning-the-difference/>

- *Random Forest*: es una técnica de conjunto que combina múltiples árboles de decisión. Cada árbol se entrena con una muestra aleatoria del conjunto de datos y luego las predicciones se promedian para obtener la salida final.
- *Support Vector Machine (SVM)*: es un algoritmo de clasificación que encuentra el hiperplano que mejor separa las clases en un espacio de características de alta dimensión. Puede ser usado tanto para clasificación como para regresión.

Dentro del universo ML, hay otro grupo más chico que se denomina *Deep Learning* (DL) o aprendizaje profundo, en donde los algoritmos utilizan una arquitectura de redes neuronales que simulan el comportamiento del cerebro humano, por lo que suelen ser mucho más grandes y complejos. Estos últimos suelen usarse para tareas de visión por computadora o procesamiento de lenguaje natural, y requieren mucha potencia de cómputo y grandes cantidades de datos. En la figura 2.2 se puede visualizar la diferencia entre utilizar modelos de *Machine Learning* y modelos de *Deep Learning* para obtener predicciones.

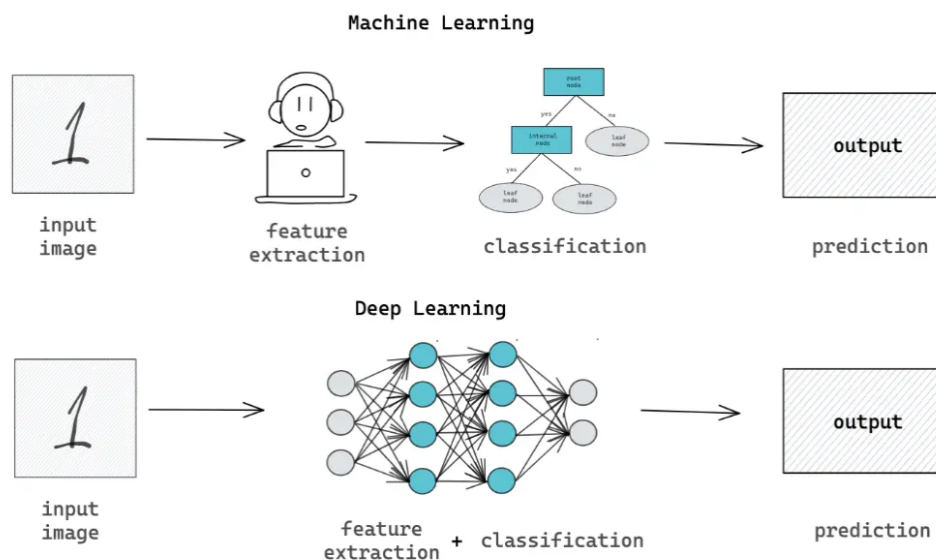


FIGURA 2.2. Machine Learning vs Deep Learning. ¹.

Los modelos también se diferencian por el tipo de predicción que realizan. A un modelo que predice un valor continuo, como el precio de una vivienda, se lo conoce como modelo de regresión. Mientras que a un modelo que predice una etiqueta o clase, ya sea binaria o multi-clase, se lo conoce como modelo de clasificación.

Este último tipo de modelo es muy utilizado en medicina, ya que sirve para predecir la presencia o ausencia de cierta enfermedad o para predecir su tipo específico. En este trabajo en particular se utilizaron modelos de clasificación con aprendizaje supervisado y se entrenaron modelos de *Machine Learning* y *Deep Learning*.

¹Imagen tomada de <https://www.stratascratch.com/blog/data-science-vs-machine-learning-vs-deep-learning-the-difference/>

Observación	Negativo	Verdadero Negativo (<i>True Negative</i> o TN)	Falso Positivo (<i>False Positive</i> o FP)
	Positivo	Falso Negativo (<i>False Negative</i> o FN)	Verdadero Positivo (<i>True Positive</i> o TP)
		Negativo	Positivo
		Predicción	

FIGURA 2.3. Matriz de confusión.

En contextos médicos, los falsos negativos pueden tener consecuencias significativas para la salud del paciente, ya que se está prediciendo que un paciente no tiene cierta condición cuando en realidad la tiene.

Particularmente en este trabajo, los falsos negativos podrían implicar no tomar medidas necesarias para un tratamiento adecuado, lo que aumentaría el riesgo de mortalidad del paciente. Por lo tanto, es fundamental minimizar la tasa de falsos negativos tanto como sea posible, incluso si esto conlleva un aumento en los falsos positivos.

Derivados de estos 4 valores de la matriz de confusión se definen las siguientes métricas para evaluar el rendimiento de un modelo de clasificación:

- *Precision*: es la proporción de ejemplos positivos que fueron correctamente clasificados como positivos respecto al total de ejemplos clasificados como positivos. Es decir, mide la calidad de las predicciones positivas del modelo. Se calcula como $TP / (TP + FP)$.
- *Recall*: es la proporción de ejemplos positivos que fueron correctamente clasificados como positivos respecto al total de ejemplos que son realmente positivos. Es decir, mide la capacidad del modelo para encontrar todos los ejemplos positivos. Se calcula como $TP / (TP + FN)$.
- *F1-Score*: es la media armónica de la *precision* y el *recall*. Proporciona un equilibrio entre ambas métricas y es útil cuando hay un desequilibrio entre las clases. Se calcula como $2 * (precision * recall) / (precision + recall)$.
- *Accuracy*: es la proporción de ejemplos clasificados correctamente (tanto positivos como negativos) respecto al total de ejemplos. Es una métrica general de la calidad del modelo en todas las clases. Se calcula como $(TP + TN) / (TP + TN + FP + FN)$.
- *AUC (Area Under the Curve)*: el AUC es el área bajo la curva ROC (*Receiver Operating Characteristic*). La curva ROC es una representación gráfica de la sensibilidad frente a la tasa de falsos positivos para diferentes umbrales de clasificación. El AUC mide la capacidad del modelo para distinguir entre

2.4. Evaluación de modelos de clasificación

Para evaluar el comportamiento de un modelo de clasificación binaria comúnmente se recurre a una herramienta llamada matriz de confusión. Esta matriz, como se muestra en la figura 2.3, presenta las clases predichas en las columnas y las clases reales en las filas. A partir de esta tabla, se derivan cuatro métricas clave:

Observación	Negativo	Verdadero Negativo (<i>True Negative</i> o TN)	Falso Positivo (<i>False Positive</i> o FP)
	Positivo	Falso Negativo (<i>False Negative</i> o FN)	Verdadero Positivo (<i>True Positive</i> o TP)
		Negativo	Positivo
		Predicción	

FIGURA 2.3. Matriz de confusión.

- Verdaderos positivos (*True Positive* o TP): representan las predicciones correctas de una condición positiva.
- Verdaderos negativos (*True Negative* o TN): son las predicciones correctas de una condición negativa.
- Falsos positivos (*False Positive* o FP): se dan cuando el modelo predice incorrectamente una condición positiva que en realidad no lo es.
- Falsos negativos (*False Negative* o FN): se dan cuando el modelo predice incorrectamente la ausencia de una condición positiva que en realidad está presente.

En contextos médicos, los falsos negativos pueden tener consecuencias significativas para la salud del paciente, ya que se está prediciendo que un paciente no tiene cierta condición cuando en realidad la tiene.

Particularmente en este trabajo, los falsos negativos podrían implicar no tomar medidas necesarias para un tratamiento adecuado, lo que aumentaría el riesgo de mortalidad del paciente. Por lo tanto, es fundamental minimizar la tasa de falsos negativos tanto como sea posible, incluso si esto conlleva un aumento en los falsos positivos.

Derivados de estos 4 valores de la matriz de confusión se definen las siguientes métricas para evaluar el rendimiento de un modelo de clasificación:

- *Precision*: es la proporción de ejemplos positivos que fueron correctamente clasificados como positivos respecto al total de ejemplos clasificados como

clases positivas y negativas. Un valor de AUC cercano a 1 indica un buen rendimiento del modelo, mientras que un valor cercano a 0.5 indica un rendimiento aleatorio.

En este trabajo se utilizaron las métricas antes mencionadas para la evaluación de los modelos entrenados, y se prestó especial atención a las métricas de *F1-Score* y *Recall*, ya que existe un desbalance entre las clases y es importante detectar correctamente a la mayoría de casos positivos.

2.5. Plataforma de gestión de modelos

Una plataforma de gestión de modelos de IA es una herramienta diseñada para ayudar a los equipos de desarrollo y científicos de datos a gestionar, monitorear y desplegar modelos de manera eficiente. Estas plataformas ofrecen las siguientes ventajas:

- Centralización: permiten centralizar todos los modelos en un único lugar, facilitando su gestión y acceso.
- Implementación automatizada: facilitan la implementación de modelos en entornos de producción, proporcionando herramientas para la integración y la implementación continua (Continuous integration and delivery/deployment o CI/CD).
- Versionado: ofrecen capacidades de versionado para los modelos, lo que facilita el seguimiento de cambios y la colaboración entre equipos.

Existen múltiples plataformas para gestión de modelos *Open Source*, tales como MLFlow [15], Apache Airflow [16] o Jenkins [17], las cuales se conectan directamente con repositorios en la nube, como GitHub, y descargan modelos para aplicar en distintos ambientes. También cuentan con la posibilidad de realizar rentrenamientos automáticos de los modelos cuando se cuente con nuevos conjuntos de datos, aunque este punto queda fuera del alcance. En este trabajo se configura Jenkins como plataforma de gestión de modelos.

2.6. Servicios web

Una vez que el modelo de IA fue entrenado y desplegado en un algún ambiente, el siguiente paso es exponerlo a través de un servicio web. Generalmente se desarrolla utilizando una API (*Application Programming Interface*), que es una interfaz que define cómo comunicarse con el servicio web. Comúnmente se utilizan APIs para la interacción entre diferentes sistemas informáticos, lo que facilita el intercambio de datos y la ejecución de funciones remotas. Se suelen utilizar para solicitar datos de alguna fuente de información desde aplicaciones web y móviles. Los servicios web se caracterizan por estar estructurados con métodos que requieren parámetros específicos y proporcionan respuestas predefinidas, lo que simplifica la interacción entre sistemas. Esta estandarización les otorga independencia de plataforma y lenguaje, lo que los convierte en herramientas altamente adaptables y versátiles.

En este caso, la API actúa como un intermediario entre la aplicación cliente y el modelo de IA, lo que facilita la comunicación y la interacción. Como primer

positivos. Es decir, mide la calidad de las predicciones positivas del modelo. Se calcula como $TP / (TP + FP)$.

- *Recall*: es la proporción de ejemplos positivos que fueron correctamente clasificados como positivos respecto al total de ejemplos que son realmente positivos. Es decir, mide la capacidad del modelo para encontrar todos los ejemplos positivos. Se calcula como $TP / (TP + FN)$.
- *F1-Score*: es la media armónica de la *precision* y el *recall*. Proporciona un equilibrio entre ambas métricas y es útil cuando hay un desequilibrio entre las clases. Se calcula como $2 * (precision * recall) / (precision + recall)$.
- *Accuracy*: es la proporción de ejemplos clasificados correctamente (tanto positivos como negativos) respecto al total de ejemplos. Es una métrica general de la calidad del modelo en todas las clases. Se calcula como $(TP + TN) / (TP + TN + FP + FN)$.
- *AUC (Area Under the Curve)*: el AUC es el área bajo la curva ROC (*Receiver Operating Characteristic*). La curva ROC es una representación gráfica de la sensibilidad frente a la tasa de falsos positivos para diferentes umbrales de clasificación. El AUC mide la capacidad del modelo para distinguir entre clases positivas y negativas. Un valor de AUC cercano a 1 indica un buen rendimiento del modelo, mientras que un valor cercano a 0.5 indica un rendimiento aleatorio.

En este trabajo se utilizaron las métricas antes mencionadas para la evaluación de los modelos entrenados, y se prestó especial atención a las métricas de *F1-Score* y *Recall*, ya que existe un desbalance entre las clases y es importante detectar correctamente a la mayoría de casos positivos.

2.5. Plataforma de gestión de modelos

Una plataforma de gestión de modelos de IA es una herramienta diseñada para ayudar a los equipos de desarrollo y científicos de datos a gestionar, monitorear y desplegar modelos de manera eficiente. Estas plataformas ofrecen las siguientes ventajas:

- Centralización: permiten centralizar todos los modelos en un único lugar, facilitando su gestión y acceso.
- Implementación automatizada: facilitan la implementación de modelos en entornos de producción, proporcionando herramientas para la integración y la implementación continua (Continuous integration and delivery/deployment o CI/CD).
- Versionado: ofrecen capacidades de versionado para los modelos, lo que facilita el seguimiento de cambios y la colaboración entre equipos.

Existen múltiples plataformas para gestión de modelos *Open Source*, tales como MLFlow [17], Apache Airflow [18] o Jenkins [19], las cuales se conectan directamente con repositorios en la nube, como GitHub, y descargan modelos para aplicar en distintos ambientes. También cuentan con la posibilidad de realizar rentrenamientos automáticos de los modelos cuando se cuente con nuevos conjuntos de datos, aunque este punto queda fuera del alcance. En este trabajo se configura Jenkins como plataforma de gestión de modelos.

2.6. Servicios web

Una vez que el modelo de IA fue entrenado y desplegado en un algún ambiente, el siguiente paso es exponerlo a través de un servicio web. Generalmente se desarrolla utilizando una API (*Application Programming Interface*), que es una interfaz que define cómo comunicarse con el servicio web. Comúnmente se utilizan APIs para la interacción entre diferentes sistemas informáticos, lo que facilita el intercambio de datos y la ejecución de funciones remotas. Se suelen utilizar para solicitar datos de alguna fuente de información desde aplicaciones web y móviles. Los servicio web se caracterizan por estar estructurados con métodos que requieren parámetros específicos y proporcionan respuestas predefinidas, lo que simplifica la interacción entre sistemas. Esta estandarización les otorga independencia de plataforma y lenguaje, lo que los convierte en herramientas altamente adaptables y versátiles.

En esta caso, la API actúa como un intermediario entre la aplicación cliente y el modelo de IA, lo que facilita la comunicación y la interacción. Como primer paso, recibe los datos del cliente en un formato especificado y realiza su preprocesamiento, tal como se realizó durante el entrenamiento del modelo. Una vez procesados los datos, se realiza la consulta al modelo de IA para obtener una predicción como respuesta. Dicha predicción es la que se devuelve al cliente como respuesta del llamado a la API.

paso, recibe los datos del cliente en un formato especificado y realiza su preprocesamiento, tal como se realizó durante el entrenamiento del modelo. Una vez procesados los datos, se realiza la consulta al modelo de IA para obtener una predicción como respuesta. Dicha predicción es la que se devuelve al cliente como respuesta del llamado a la API.

Capítulo 3

Diseño e implementación

En este capítulo, se presentan las tareas realizadas para el cumplimiento de los objetivos. Se abordan temas como la adquisición de los datos, su preprocesamiento, las técnicas usadas para balancear las clases, los modelos de IA utilizados y las herramientas para el acceso al modelo.

3.1. Arquitectura propuesta

Para cumplir con los objetivos planteados en este trabajo se pensó en una solución integral, donde fue necesario desarrollar y configurar una serie de herramientas que interactúen entre si para lograr que se generen predicciones de mortalidad de forma automática.

En la figura 3.1 se muestra la arquitectura desarrollada.

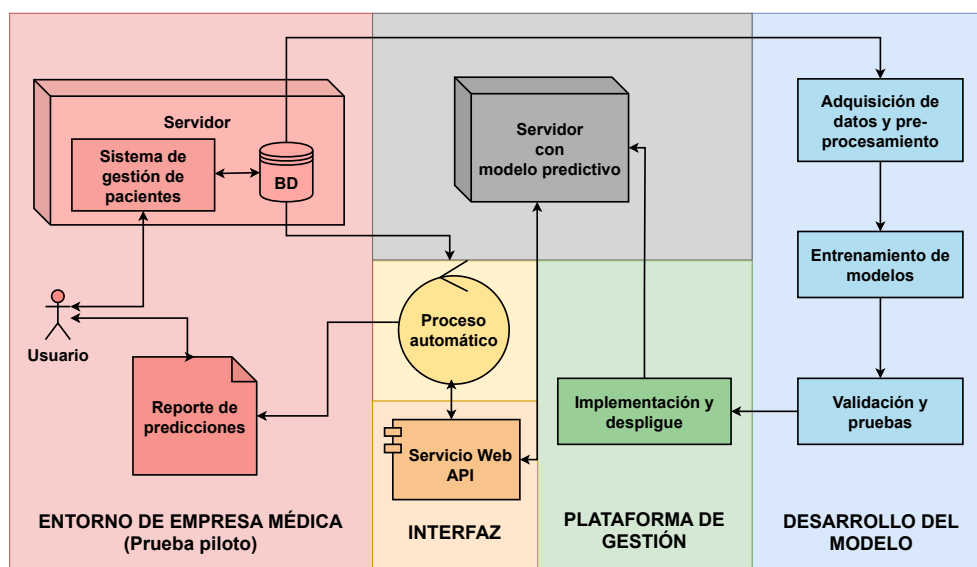


FIGURA 3.1. Arquitectura del sistema.

La solución implementada comienza con la adquisición de datos y entrenamiento de modelos, seguida por la selección e implementación del mejor modelo en un servidor. Luego se desarrolla un servicio web para el acceso al modelo y, por último, una aplicación automática que solicite las predicciones y se las devuelva al usuario.

En las próximas secciones se irán comentando con más detalle cada una de estas etapas.

Capítulo 3

Diseño e implementación

En este capítulo, se presentan las tareas realizadas para el cumplimiento de los objetivos. Se abordan temas como la adquisición de los datos, la selección de características, el preprocesamiento, las técnicas usadas para balancear las clases y los modelos de IA utilizados.

3.1. Arquitectura propuesta

Para cumplir con los objetivos planteados en este trabajo se pensó en una solución integral, donde se fue necesario desarrollar y configurar una serie de herramientas que interactúen entre si para lograr que se generen predicciones de mortalidad de forma automática.

En la figura 3.1 se muestra la arquitectura desarrollada.

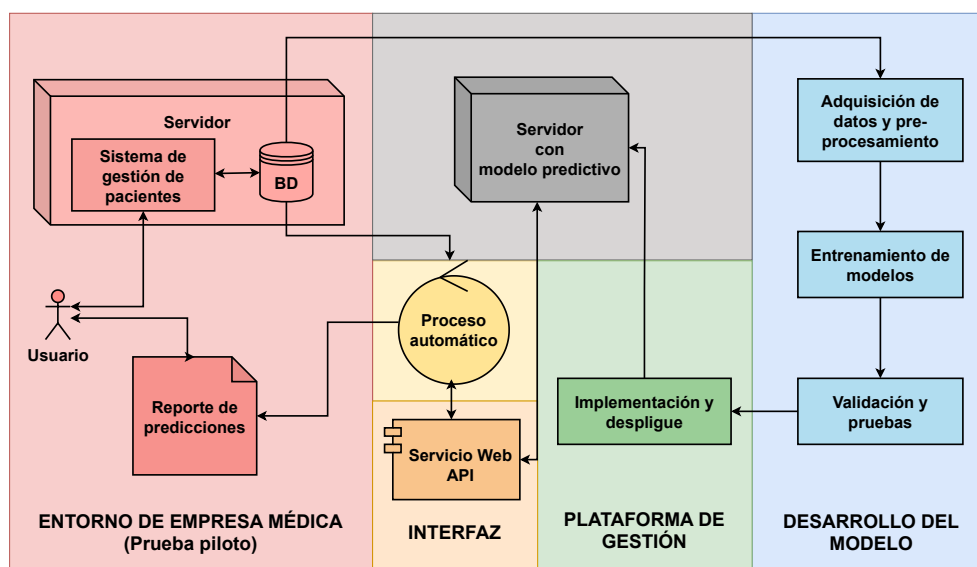


FIGURA 3.1. Arquitectura del sistema.

La solución implementada comienza con la adquisición de datos y entrenamiento de modelos, seguida por la selección e implementación del mejor modelo en un servidor. Luego se desarrolla un servicio web para el acceso al modelo y, por último, una aplicación automática que solicite las predicciones y se las devuelva al usuario.

En las próximas secciones se irán comentando con más detalle cada una de estas etapas.

3.2. Adquisición de datos

En el sistema de gestión de pacientes de la empresa médica se registran datos clínicos de pacientes que se encuentran en tratamiento de diálisis renal. Entre los datos que se registran se encuentran:

- Datos personales de los pacientes, como nombre, apellido, documento, sexo y fecha de nacimiento.
- Datos del tratamiento de diálisis, como turno, grupo de días, motivo de ingreso, centro de atención, primera diálisis en la vida, etiología de la enfermedad renal primaria, diuresis residual.
- Evaluación clínica, donde se registra el peso, la altura, la presión arterial y el tiempo en diálisis.
- Examen de Laboratorio, de los cuales algunas variables se analizan todos los meses, otras cada 3 meses y otras cada 6 meses.
- Accesos Vasculares, donde se registra el tipo, la localización, la fecha de confección, la fecha de primer uso, los eventos y los procedimientos realizados.
- Prescripción de medicamentos, como por ejemplo la administración de Eritropoyetina y Hierro.
- Condiciones médicas, como diabetes, en espera para trasplante, y otras.

Cada centro de atención utiliza un sistema de gestión de pacientes para registrar los datos, los cuales se consolidan a nivel nacional.

Como primer paso en la adquisición de los datos, se seleccionaron las variables médicas que tienen más influencia en la mortalidad de los pacientes. Esto se realizó en conjunto con un médico de la empresa de diálisis renal, quien aportó información muy valiosa sobre el impacto que tiene cada una. Se propusieron más de 80 variables que se registran en el sistema, de las cuales solo se seleccionaron 39. Las variables seleccionadas son las siguientes:

- Motivo de Ingreso: indica si el motivo de ingreso fue programado o no programado.
- Edad: representa la edad en años del paciente.
- Sexo: representa el sexo del paciente (masculino/femenino).
- Altura: representa la altura del paciente en el último mes en centímetros.
- Peso: representa el peso del paciente en el último mes en kg.
- Balance: refiere a la diferencia entre el peso pre diálisis y el peso post diálisis en kg.
- Tiempo en Diálisis: indica el tiempo que el paciente se dializa expresado en minutos.
- Mean Arterial Pressure (MAP): representa la presión arterial media y se calcula con la formula siguiente: $\text{Tensión diastólica} + ((\text{Tensión sistólica} - \text{Tensión diastólica}) / 3)$.
- Hemoglobina: representa el valor de Hemoglobina del último mes en G/dl.

3.2. Adquisición de datos

En el sistema de gestión de pacientes de la empresa médica se registran datos clínicos de pacientes que se encuentran en tratamiento de diálisis renal. Entre los datos que se registran se encuentran:

- Datos personales de los pacientes, como nombre, apellido, documento, sexo y fecha de nacimiento.
- Datos del tratamiento de diálisis, como turno, grupo de días, motivo de ingreso, centro de atención, primera diálisis en la vida, etiología de la enfermedad renal primaria y diuresis residual.
- Evaluación clínica, donde se registra el peso, la altura, la presión arterial y el tiempo en diálisis.
- Examen de Laboratorio, de los cuales algunas variables se analizan todos los meses, otras cada 3 meses y otras cada 6 meses.
- Accesos Vasculares, donde se registra el tipo, la localización, la fecha de confección, la fecha de primer uso, los eventos y los procedimientos realizados.
- Prescripción de medicamentos, como por ejemplo la administración de Eritropoyetina y Hierro.
- Condiciones médicas, como diabetes, en espera para trasplante, y otras.

Cada centro de atención utiliza el sistema de gestión de pacientes para registrar los datos, los cuales se consolidan a nivel nacional.

Como primer paso en la adquisición de los datos, se seleccionaron las variables médicas que tienen más influencia en la mortalidad de los pacientes. Esto se realizó en conjunto con un médico de la empresa de diálisis renal, quien aportó información muy valiosa sobre el impacto que tiene cada una. Se propusieron más de 80 variables que se registran en el sistema, de las cuales solo se seleccionaron 39. Las variables seleccionadas son las siguientes:

- Motivo de Ingreso: indica si el motivo de ingreso fue programado o no programado.
- Edad: representa la edad en años del paciente.
- Sexo: representa el sexo del paciente (masculino/femenino).
- Altura: representa la altura del paciente en el último mes en centímetros.
- Peso: representa el peso del paciente en el último mes en kg.
- Balance: refiere a la diferencia entre el peso pre diálisis y el peso post diálisis en kg.
- Tiempo en Diálisis: indica el tiempo que el paciente se dializa expresado en minutos.
- *Mean Arterial Pressure* (MAP): representa la presión arterial media y se calcula con la formula siguiente: $\text{Tensión diastólica} + ((\text{Tensión sistólica} - \text{Tensión diastólica}) / 3)$.
- Hemoglobina: representa el valor de Hemoglobina del último mes en G/dl.

- Urea pre: representa el valor de Urea pre diálisis del último mes en Mg/dl.
- Urea post: representa el valor de Urea post diálisis del último mes en Mg/dl.
- Potasio: representa el valor de Potasio del último mes en mEq/l.
- Glucemia: representa el valor de Glucemia del último mes en Mg/dl.
- Calcio: representa el valor de Calcio del último mes en Mg/dl.
- Fósforo: representa el valor de Fósforo del último mes en Mg/dl.
- Albúmina: representa el valor de Albúmina del último mes en G/dl.
- Bicarbonato (HCO_3): representa el valor de Bicarbonato del último mes en mEq/l.
- Saturación Transferrina: representa el porcentaje de Saturación Transferrina. Se toma el último valor de los últimos 3 meses.
- Ferritina Plasmática: representa el valor de Ferritina plasmática en Ng/ml. Se toma el último valor de los últimos 3 meses.
- Pth (immulite): representa el valor de Pth en Pg/ml. Se toma el último valor de los últimos 3 meses.
- Anti-HIV: representa el valor de Anticuerpos contra el Virus de la Inmunodeficiencia Humana por el método ELISA. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = No reactivo / R = Reactivo).
- Anti-HCV: representa el valor de Anticuerpos contra el Virus de la Hepatitis C por el método ELISA. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = No reactivo / R = Reactivo).
- HBsAg: representa el valor de Antígeno de Superficie del Virus de la Hepatitis B por el método ELISA. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = No reactivo / R = Reactivo).
- Anti-HBs: representa el valor de Anticuerpos contra el Antígeno de Superficie del Virus de la Hepatitis B por el método ELISA en UI/l. Se toma el último valor de los últimos 6 meses.
- HCV PCR: representa el valor de Anticuerpos contra el Virus de la Inmunodeficiencia Humana por el método PCR. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = Negativo / P = Positivo).
- HBsAg PCR: representa el valor de Antígeno de Superficie del Virus de la Hepatitis B por el método PCR. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = Negativo / P = Positivo).
- Hemoglobina Glicosilada: representa el porcentaje de Hemoglobina Glicosilada del último mes.
- Vitamina D 25oh: representa el valor de Vitamina D 25oh del último mes en ng/ml.
- Kt/V: medida de la eficacia del tratamiento de diálisis. Se calcula en base a la urea pre y post, el balance, el peso post, el tiempo en diálisis, el sexo, la edad y la altura.

- Urea pre: representa el valor de Urea pre diálisis del último mes en Mg/dl.
- Urea post: representa el valor de Urea post diálisis del último mes en Mg/dl.
- Potasio: representa el valor de Potasio del último mes en mEq/l.
- Glucemia: representa el valor de Glucemia del último mes en Mg/dl.
- Calcio: representa el valor de Calcio del último mes en Mg/dl.
- Fósforo: representa el valor de Fósforo del último mes en Mg/dl.
- Albúmina: representa el valor de Albúmina del último mes en G/dl.
- Bicarbonato (HCO_3): representa el valor de Bicarbonato del último mes en mEq/l.
- Saturación Transferrina: representa el porcentaje de Saturación Transferrina. Se toma el último valor de los últimos 3 meses.
- Ferritina Plasmática: representa el valor de Ferritina plasmática en Ng/ml. Se toma el último valor de los últimos 3 meses.
- Pth (immulite): representa el valor de Pth en Pg/ml. Se toma el último valor de los últimos 3 meses.
- Anti-HIV: representa el valor de Anticuerpos contra el Virus de la Inmunodeficiencia Humana por el método ELISA. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = No reactivo / R = Reactivo).
- Anti-HCV: representa el valor de Anticuerpos contra el Virus de la Hepatitis C por el método ELISA. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = No reactivo / R = Reactivo).
- HBsAg: representa el valor de Antígeno de Superficie del Virus de la Hepatitis B por el método ELISA. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = No reactivo / R = Reactivo).
- Anti-HBs: representa el valor de Anticuerpos contra el Antígeno de Superficie del Virus de la Hepatitis B por el método ELISA en UI/l. Se toma el último valor de los últimos 6 meses.
- HCV PCR: representa el valor de Anticuerpos contra el Virus de la Inmunodeficiencia Humana por el método PCR. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = Negativo / P = Positivo).
- HBsAg PCR: representa el valor de Antígeno de Superficie del Virus de la Hepatitis B por el método PCR. Se toma el último valor de los últimos 6 meses. Los valores posibles son (N = Negativo / P = Positivo).
- Hemoglobina Glicosilada: representa el porcentaje de Hemoglobina Glicosilada del último mes.
- Vitamina D 25oh: representa el valor de Vitamina D 25oh del último mes en ng/ml.
- Kt/V: medida de la eficacia del tratamiento de diálisis. Se calcula en base a la urea pre y post, el balance, el peso post, el tiempo en diálisis, el sexo, la edad y la altura.

- Diabético: indica si el paciente tiene diabetes del tipo 1 o 2, o si tiene Nefropatía diabética.
- Etiología de la enfermedad renal primaria: indica la causa subyacente que conduce al desarrollo de la enfermedad renal.
- Acceso Vascular (Tipo): indica el tipo del primer Acceso Vascular (Catéter Liso, Tunelizado, Prótesis o Fístula Arteriovenosa).
- Días al primer uso del Acceso Vascular: representa los días que pasaron entre la primera diálisis en la vida y el primer uso del primer acceso vascular.
- Cantidad de complicaciones en el Acceso Vascular: representa la cantidad de complicaciones que tuvo en el acceso vascular.
- Días en diálisis en la vida: representa la cantidad de días desde la primera diálisis en la vida hasta el día de la fecha.
- Grupo Sanguíneo: indica el tipo de sangre y factor.
- Lista de espera para trasplante: representa la situación en relación a la lista de espera para el trasplante.
- EPO (Dosis Mensual): representa la dosis mensual de Eritropoyetina prescrita en el último mes.
- Hierro (Dosis Mensual): representa la dosis mensual de Hierro prescrita en el último mes.

Se definió también que solo se usarían datos de pacientes en tratamiento de hemodiálisis y que tengan más de 90 días en tratamiento, ya que antes de los 90 días la mortalidad es mucho mayor.

Una vez definidas las variables y las condiciones de los pacientes, se prepararon las consultas a la base de datos para obtener dicha información.

Se definieron dos estrategias para el armado de los conjuntos de datos de entrenamiento y pruebas:

1. Pacientes que lleven más de 90 días en tratamiento HD, fallecidos y no fallecidos, con datos del último examen de laboratorio. Son 15300 registros de los cuales el 61.7 % no falleció.
2. Pacientes que lleven más de 90 días en tratamiento HD, fallecidos entre 2013 y 2023, y no fallecidos que estuvieron en diálisis el último día del año de 2020, 2021 y 2022 (existiendo la posibilidad de que se repitan los pacientes, pero con estado de salud y datos distintos). Son 14100 registros, de los cuales el 58.5 % no falleció.

Ambas estrategias contienen las mismas 39 variables por cada registro. Los conjuntos de datos de ambas estrategias se trabajaron en proyectos separados, aunque compartieron gran parte del preprocesamiento de los datos.

También se armó un conjunto de datos para realizar la validación, el cual incluyó a los pacientes que llevan más de 90 días en tratamiento de hemodiálisis con su estado al final del mes de enero de 2024, incluyendo fallecidos y no fallecidos. Esta es la situación más parecida a la realidad, ya que el modelo estará constantemente evaluando el riesgo de pacientes que estén activos para detectar a aquellos que tengan alto riesgo.

- Diabético: indica si el paciente tiene diabetes del tipo 1 o 2, o si tiene Nefropatía diabética.
- Etiología de la enfermedad renal primaria: indica la causa subyacente que conduce al desarrollo de la enfermedad renal.
- Tipo de Acceso Vascular (AV): indica el tipo del primer Acceso Vascular: Catéter Liso, Tunelizado, Prótesis o Fístula Arteriovenosa.
- Días al primer uso del Acceso Vascular: representa los días que pasaron entre la primera diálisis en la vida y el primer uso del primer acceso vascular.
- Cantidad de complicaciones en el Acceso Vascular: representa la cantidad de complicaciones que tuvo en el acceso vascular.
- Días en diálisis en la vida: representa la cantidad de días desde la primera diálisis en la vida hasta el día de la fecha.
- Grupo Sanguíneo: indica el tipo de sangre y factor.
- Lista de espera para trasplante: representa la situación en relación a la lista de espera para el trasplante.
- EPO (Dosis Mensual): representa la dosis mensual de Eritropoyetina prescrita en el último mes.
- Hierro (Dosis Mensual): representa la dosis mensual de Hierro prescrita en el último mes.

Se definió también que solo se usarían datos de pacientes en tratamiento de hemodiálisis y que tengan más de 90 días en tratamiento, ya que antes de los 90 días la mortalidad es mucho mayor.

Una vez definidas las variables y las condiciones de los pacientes a utilizar, se prepararon las consultas a la base de datos para obtener dicha información. Para esto se utilizó una base de información histórica, donde se guarda al final de cada mes una foto del estado de salud de los pacientes en ese momento, incluyendo a los pacientes activos y a los que fallecieron en dicho mes.

Se definieron tres estrategias para el armado de los conjuntos de datos de entrenamiento y pruebas:

1. Pacientes que lleven más de 90 días en tratamiento HD, fallecidos y no fallecidos, con datos del último examen de laboratorio. Son 15300 registros de los cuales el 61.7 % no falleció.
2. Pacientes que lleven más de 90 días en tratamiento HD, fallecidos entre 2013 y 2023, y no fallecidos que estuvieron en la foto mensual del último mes del año en 2020, 2021 y 2022 (existiendo la posibilidad de que se repitan los pacientes, pero con estado de salud y datos distintos). Son 14100 registros, de los cuales el 58.5 % no falleció.
3. Todos los pacientes que existan en todas las fotos mensuales de todos los meses entre los años 2016 y 2023, que lleven más de 90 días en tratamiento HD al momento de tomar la foto. Son 284000 registros, donde solo el 1.5 % falleció.

3.3. Preprocesamiento de datos

Como primer paso, cada conjunto de datos se dividió en 2 partes: 75 % para realizar el entrenamiento y el 25 % restante para probar el desempeño del modelo. Esto permite evaluar qué tan bien generaliza el modelo con datos nuevos y ayuda medir su eficacia en situaciones reales.

Durante el análisis exploratorio de los datos recibidos, se identificaron ciertas inconsistencias que requerían corrección, como edades superiores a 120 años, alturas de 0 cm o pesos inferiores a 10 kg. Estos datos inconsistentes se marcaron como faltantes para su posterior completitud.

Algunas de las variables seleccionadas, si bien tenían información sumamente importante, contenían gran cantidad de valores faltantes, por lo que se optó por la técnica de discretización. Las variables en esta situación fueron Glucemia, Hemoglobina Glicosilada, Vitamina D 25oh, Bicarbonato, Días al primer uso del Acceso Vascular, EPO y Hierro.

Todos los puntos del preprocesamiento mencionados anteriormente se aplicaron tanto al conjunto de entrenamiento, como al de pruebas y validación.

Las tres estrategias contienen las mismas 39 variables por cada registro. Cada conjunto de datos de las distintas estrategias se trabajaron en proyectos separados, aunque compartieron gran parte del preprocesamiento de los datos.

También se armó un conjunto de datos para realizar la validación, el cual incluyó a los pacientes que llevan más de 90 días en tratamiento HD con su estado al final del mes de enero de 2024, incluyendo fallecidos y no fallecidos. Son unos 2817 registros, donde el 1 % falleció. Esta es la situación más parecida a la realidad, ya que el modelo estará constantemente evaluando a pacientes activos para detectar a ese 1 % con alto riesgo de mortalidad.

3.3. Preprocesamiento de datos

Como primer paso, cada conjunto de datos se dividió en 2 partes: 75 % para realizar el entrenamiento y el 25 % restante para probar el desempeño del modelo. Esto permite evaluar qué tan bien generaliza el modelo con datos nuevos y ayuda medir su eficacia en situaciones reales.

Durante el análisis exploratorio de los datos recibidos, se identificaron ciertas inconsistencias que requerían corrección, como edades superiores a 120 años, alturas de 0 cm o pesos inferiores a 10 kg. Estos datos inconsistentes se marcaron como faltantes para su posterior completitud.

Algunas de las variables seleccionadas, si bien contenían información sumamente importante, contaban con gran cantidad de valores faltantes. Esto se resolvió de forma distinta para las variables numéricas y categóricas.

3.3.1. Preprocesamiento de variables numéricas

Como primer punto, se analizaron los valores atípicos (*outliers*) en las variables numéricas. En la figura 3.2 se observan los diagramas de caja o *box plot* de dichas variables, donde se observa la mediana, los cuartiles de datos y los *outliers*.

Para disminuir el ruido generado por estos valores atípicos se realizaron pruebas con la técnica de *capping* o truncamiento. Se definió como límite inferior y superior el valor de 3 desvíos estándar y se reemplazó todo valor que esté por fuera del intervalo con el valor de cada límite. Los resultados se mantuvieron similares, por lo que la técnica no se aplicó en el conjunto de datos final.

En relación al análisis de datos faltantes en las variables numéricas, se dividió a éstas en dos grupos según su situación:

- Variables con gran cantidad de datos faltantes: en estas variables se aplicó la técnica de discretización. En la Tabla 3.1 se muestran las variables discretizadas con sus 3 grupos. Los valores límite para cada grupo fueron determinados con la ayuda del médico especialista. En la figura 3.3 se muestran las distribuciones de dichas variables.
- Variables con poca cantidad de datos faltantes: para este grupo, se completaron los valores faltantes con la técnica de imputación KNN. En la tabla 3.2 se muestra el porcentaje de valores faltantes para cada una de las variables.

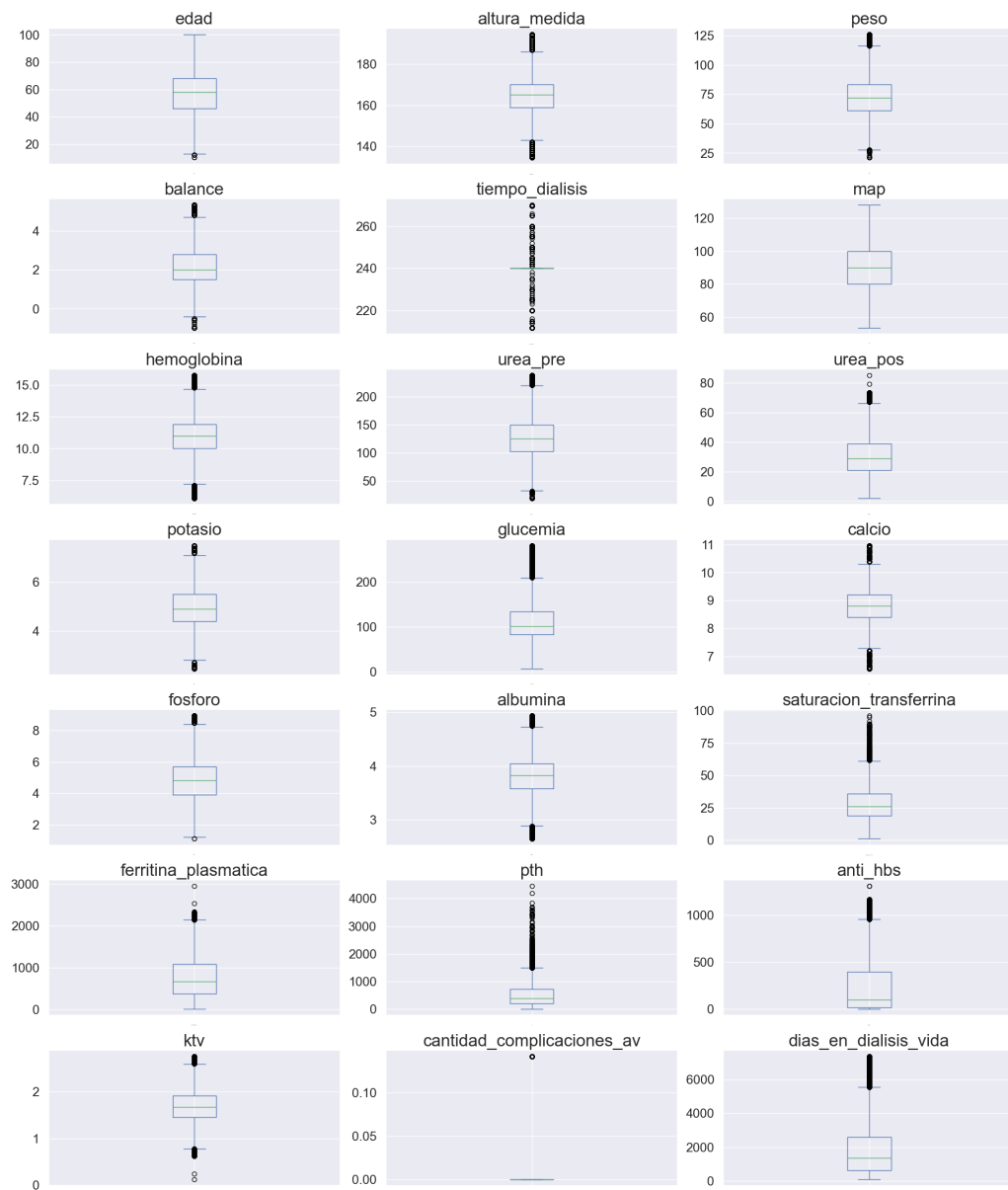


FIGURA 3.2. Outliers en variables numéricas.

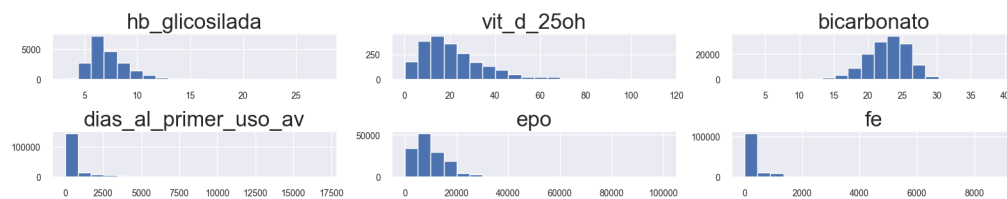


FIGURA 3.3. Distribuciones de variables numéricas discretizadas.

Capítulo 4

Ensayos y resultados

4.1. Pruebas funcionales del hardware

La idea de esta sección es explicar cómo se hicieron los ensayos, qué resultados se obtuvieron y analizarlos.

TABLA 3.1. Variables discretizadas.

Variable	Valor Bajo	Valor Normal	Valor Alto
Hemoglobina Glicosilada	0 a 5.7	5.7 a 6.5	6.5 a 30
Vitamina D 25oh	0 a 20	20 a 30	30 a 120
Bicarbonato	0 a 22	22 a 32	32 a 50
Días al primer uso del AV	0 a 30	30 a 90	90 a 15000
EPO	0 a 3000	3000 a 10000	10000 a 100000
Hierro	0 a 200	200 a 500	500 a 10000

TABLA 3.2. Valores faltantes en variables numéricas no discretizadas.

Variable	Porcentaje
Edad	0.007 %
Peso	0.0113 %
Balance	0.0054 %
Tiempo de Diálisis	0.0161 %
MAP	0.0005 %
Hemoglobina	0.0075 %
Urea Post	0.0354 %
Potasio	0.015 %
Glucemia	0.0097 %
Calcio	0.0005 %
Fósforo	0.0005 %
Albumina	0.0016 %
Saturación Transferrina	0.6617 %
Ferritina Plasmática	0.1073 %
Pth	1.5991 %
Anti-HBs	0.6225 %
Kt/V	0.0531 %

En cuanto a las distribuciones de las variables numéricas no discretizadas, si bien se realizaron pruebas aplicando transformaciones de Yeo Johnson, no se observaron mejoras en los resultados. Por tal motivo no se aplicó ninguna transformación. En la figura 3.4 se muestran las distribuciones.

3.3.2. Preprocesamiento de variables categóricas

Para las variables categóricas se utilizó la técnica de codificación *one-hot*, que crea nuevas variables binarias para las categorías con más cantidad de datos (más del 10 % del total). Esta técnica también resuelve el problema de los datos faltantes en estas variables, ya que en el caso de faltar el dato se colocará el valor 0 en todas las variables binarias del grupo para esa observación. En la tabla 3.3 se muestra la cantidad de valores nulos en cada variable categórica.



FIGURA 3.4. Distribuciones de variables numéricas.

TABLA 3.3. Valores faltantes en variables categóricas.

Variable	Porcentaje
Anti-HIV	0.0639 %
Anti-HCV	0.1728 %
HBSAg	0.0504 %
HCV PCR	98.8162 %
HBSAg PCR	99.4215 %
Acceso vascular	0.0011 %
Grupo sanguíneo	62.2971 %
Lista de espera para trasplante	29.404 %

Capítulo 5

Conclusiones

5.1. Conclusiones generales

La idea de esta sección es resaltar cuáles son los principales aportes del trabajo realizado y cómo se podría continuar. Debe ser especialmente breve y concisa. Es buena idea usar un listado para enumerar los logros obtenidos.

Algunas preguntas que pueden servir para completar este capítulo:

- ¿Cuál es el grado de cumplimiento de los requerimientos?
- ¿Cuán fielmente se pudo seguir la planificación original (cronograma incluido)?
- ¿Se manifestó algunos de los riesgos identificados en la planificación? ¿Fue efectivo el plan de mitigación? ¿Se debió aplicar alguna otra acción no contemplada previamente?
- Si se debieron hacer modificaciones a lo planificado ¿Cuáles fueron las causas y los efectos?
- ¿Qué técnicas resultaron útiles para el desarrollo del proyecto y cuáles no tanto?

5.2. Próximos pasos

Acá se indica cómo se podría continuar el trabajo más adelante.

3.3.3. Normalización de los datos

Para optimizar los tiempos de entrenamiento del modelo y ayudar a la generalización del conocimiento, se aplica una normalización a todas las variables del conjunto de datos. Luego de probar distintas técnicas, se elige la de *z-score*, que usa la media y el desvío estándar. Este proceso transforma todas las variables del conjunto de datos para que tengan un rango similar.

Una vez aplicados todos los puntos anteriores, el conjunto de datos se encuentra completo y posee únicamente valores numéricos, lo que ayuda a los modelos a poder aprender de los datos y generalizar el conocimiento.

Todos los puntos del preprocesamiento mencionados anteriormente se aplicaron tanto al conjunto de entrenamiento, como al conjunto de pruebas y de validación.

3.4. Estrategias aplicadas para balancear las clases

En los conjuntos de datos de las tres estrategias se cuenta con un gran desbalance entre las clases. En la tabla 3.4 se muestra la proporción de cada clase en cada conjunto de datos.

TABLA 3.4. Desbalance de pacientes por estrategia.

Estrategia	Total de Pacientes	No Fallecidos	Fallecidos	Proporción
Estrategia 1	15352	9478	5874	1.61
Estrategia 2	14169	8295	5874	1.41
Estrategia 3	284155	279686	4469	62.58

Los modelos de *Machine Learning* y *Deep Learning* suelen tener problemas al intentar aprender de conjuntos de datos desbalanceados, por lo que resulta obligatorio realizar alguna acción para corregir esta situación. Para este trabajo se realizaron pruebas de entrenamientos utilizando submuestreo, con la técnica de submuestreo aleatorio y *Near Miss*, y sobremuestreo, utilizando la técnica de SMOTE. Pero los resultados empeoraban al aplicar cualquiera de las dos o una combinación de ambas. Por lo que se decidió no aplicar ninguna. Para resolver la situación, se utilizó la técnica de *Focal Loss*, que implica utilizar una función de pérdida durante el entrenamiento del modelo que penalice más las predicciones incorrectas para los ejemplos difíciles. Esto significa que el modelo presta más atención a los casos que son más difíciles de clasificar correctamente, lo que puede incluir aquellos de la clase minoritaria.

3.5. Diseño y desarrollo de modelos

Para este trabajo se probaron arquitecturas de modelos de *Machine Learning* y de *Deep Learning*. A continuación se presentan en detalle las arquitecturas utilizadas y las métricas que evaluaron su desempeño.

3.5.1. Modelo de *Machine Learning*

Si bien se entrenaron varios tipos de modelos de *Machine Learning*, ninguno de ellos llegó a un resultado aceptable o útil para el problema que se busca resolver.

En la tabla 3.5 se pueden visualizar las métricas obtenidas por distintos modelos con el dataset de prueba.

TABLA 3.5. Resultados de diferentes clasificadores

Modelo	Accuracy	AUC	Recall	Precision	F1
CatBoost Classifier	0.9890	0.8623	0.0589	0.6520	0.1078
Ada Boost Classifier	0.9889	0.8476	0.0375	0.6446	0.0707
Gradient Boosting Classifier	0.9889	0.8481	0.0889	0.5636	0.1529
Logistic Regression	0.9888	0.8274	0.0157	0.6546	0.0304
K Neighbors Classifier	0.9887	0.5057	0.0000	0.0000	0.0000
Ridge Classifier	0.9887	0.0000	0.0000	0.0000	0.0000
Dummy Classifier	0.9887	0.5000	0.0000	0.0000	0.0000
Extreme Gradient Boosting	0.9886	0.8311	0.0813	0.4765	0.1383
Random Forest Classifier	0.9883	0.7903	0.0005	0.0125	0.0009
Light Gradient Boosting Machine	0.9883	0.8446	0.0984	0.4373	0.1602
Extra Trees Classifier	0.9881	0.7902	0.0005	0.0091	0.0009
SVM - Linear Kernel	0.9871	0.0000	0.0119	0.1726	0.0168
Linear Discriminant Analysis	0.9864	0.8434	0.1754	0.3162	0.2250
Decision Tree Classifier	0.9777	0.5562	0.1250	0.1026	0.1126
Naive Bayes	0.9163	0.7771	0.3498	0.0491	0.0862
Quadratic Discriminant Analysis	0.9129	0.5006	0.0789	0.0132	0.0195

Se puede observar que ninguno llega a un buen resultado para las métricas de *f1 score* y *recall*.

3.5.2. Modelo de Deep Learning

Como modelo de *Deep Learning* se diseña una red neuronal multicapa que devuelva un único valor binario. Al no contar con conocimientos previos sobre el diseño de una red neuronal óptima para resolver un problema de este tipo, se probaron distintas arquitecturas para determinar cual llega a mejores resultados. En la tabla 3.6 se muestran las diferentes combinaciones de modelos que se probaron en esta exploración.

En la tabla 3.7 se muestran los valores promedio de las diferentes métricas obtenidos en los entrenamientos de los modelos, utilizando el conjunto de datos de validación.

En todas las pruebas se utilizaron los parámetros indicados en la tabla 3.8.

TABLA 3.6. Comparación de modelos con diferentes configuraciones de capas ocultas.

Modelo	Capa 1	Capa 2	Capa 3
1 capa oculta			
1	100		
2	200		
3	500		
4	1000		
2 capas ocultas			
5	200	100	
6	200	200	
7	500	200	
8	500	500	
9	1000	500	
10	1000	1000	
3 capas ocultas			
11	500	200	100
12	500	500	200
13	1000	500	200
14	1000	1000	500

TABLA 3.7. Resultados de diferentes modelos.

Modelo	Accuracy	Precision	Recall	F1 Score	AUC
1	0.7454	0.0339	0.8333	0.0651	0.7890
2	0.7568	0.0328	0.7666	0.0629	0.7620
3	0.7621	0.0348	0.8000	0.0668	0.7810
4	0.7735	0.0365	0.8000	0.0699	0.7870
5	0.7902	0.0409	0.8333	0.0780	0.8120
6	0.8029	0.0402	0.7666	0.0765	0.7850
7	0.7774	0.0386	0.8333	0.0738	0.8050
8	0.8143	0.0409	0.7333	0.0776	0.7940
9	0.8004	0.0397	0.7666	0.0756	0.7840
10	0.7653	0.0353	0.8000	0.0677	0.7820
11	0.8207	0.0406	0.7000	0.0767	0.7610
12	0.7532	0.0336	0.8000	0.0646	0.7760
13	0.7919	0.0397	0.8000	0.0757	0.7960
14	0.8221	0.0409	0.7000	0.0773	0.7620

TABLA 3.8. Parámetros utilizados durante los entrenamientos.

Parámetro	Valor
Número de épocas de entrenamiento	10
Función de activación luego de cada capa	Relu
Capas de <i>dropout</i> luego de cada función de activación	0.5
Función de pérdida (<i>Focal Loss</i>). Parámetro Gamma	9
Función de pérdida (<i>Focal Loss</i>). Parámetro Alpha	0.991
Optimizador	Adam
<i>Learning Rate</i>	0.0001
Regularización L2	0.001
Tipo de algoritmo de entrenamiento	<i>Mini-Batch</i>

En base a los resultados de las comparaciones realizadas, se seleccionó el modelo 5, que cuenta con tres capas densamente conectadas. La primera capa tiene 75 neuronas de entrada y 200 de salida, seguida de una función de activación ReLU (*Rectified Lineal Unit*) y una capa de *dropout* con una probabilidad del 0.5. La segunda capa reduce las características a 100, también utiliza función de activación ReLU y una capa de *dropout*. La capa de salida tiene una sola neurona, ya que el modelo devuelve un resultado binario.

3.5.3. Selección de funciones de activación, función de pérdida y algoritmo de optimización

Se seleccionó la función de activación ReLU para cada una de las capas ocultas de la red neuronal. Esta función es comúnmente utilizada por ser muy simple, ya que devuelve cero para valores negativos y el mismo valor para valores positivos. Esto significa que no requiere cálculos complejos como las funciones sigmoideas o tangenciales hiperbólicas, lo que puede hacer que el entrenamiento de la red neuronal sea más rápido y eficiente computacionalmente. Como función de pérdida, se decide utilizar *focal loss*, ya que fue diseñada específicamente para abordar el problema del desequilibrio de clases en problemas de clasificación. En cuanto al algoritmo de optimización se eligió Adam, que se destaca por su adaptabilidad y eficiencia computacional. Combina conceptos como el *momentum* para acelerar la convergencia y la adaptación de la tasa de aprendizaje para cada parámetro, lo que lo hace especialmente efectivo en problemas con gradientes variables. Además, su regularización implícita y su capacidad para manejar una amplia gama de aplicaciones lo convierten en una opción popular y confiable para el entrenamiento de redes neuronales.

3.5.4. Estrategias implementadas para evitar el *overfitting*

El *overfitting*, o sobreajuste, es un fenómeno en el aprendizaje automático donde un modelo se adapta demasiado bien a los datos de entrenamiento, y captura no solo la relación subyacente entre las características y las etiquetas, sino también el ruido aleatorio o las peculiaridades específicas del conjunto de entrenamiento. Esto resulta en un rendimiento deficiente cuando el modelo se enfrenta a datos nuevos, no vistos durante el entrenamiento, ya que el modelo no puede generalizar correctamente. En otras palabras, el modelo memoriza los datos de entrenamiento en lugar de aprender la relación subyacente entre las características y las etiquetas, lo que conduce a una pérdida de capacidad predictiva en datos no

vistos. Para minimizar el *overfitting* y mejorar la robustez del modelo, se realizó lo siguiente:

- Mantener el modelo lo más simple posible, sin agregar capas o neuronas innecesarias.
- Utilizar las características mínimas que proporcionen información relevante al modelo para poder aprender y generalizar el conocimiento.
- Incorporar capas de *dropout* en la arquitectura de la red neuronal, lo que fuerza a las neuronas a ser más independientes entre sí.
- Agregar regularización *Ridge* (L2) que añade un término de penalización a la función de pérdida del modelo. Con esto se penalizan los valores grandes de los pesos, lo que tiende a reducir su magnitud durante el entrenamiento.
- Utilizar la estrategia de *early stopping* para detener el proceso de entrenamiento al detectar un aumento del error en validación durante 3 épocas. [20].

3.6. Conjunto de herramientas para acceso al modelo

Como parte del objetivo de este trabajo, se configuraron y desarrollaron una serie de herramientas que permiten hacer uso del modelo para obtener predicciones.

En primer lugar, se configuró la plataforma Jenkins para desplegar al modelo en distintos ambientes. Al trabajar con GIT y tener un repositorio en la nube, se puede configurar esta plataforma para que descargue la última versión del modelo entrenado y lo despliegue en algún servidor.

También se desarrolló una API para el acceso al modelo que permite a una aplicación cliente solicitar una o varias predicciones juntas al modelo predictivo. La API fue desarrollada con arquitectura RESTful en lenguaje Python, usando la librería FastAPI [21]. Para solicitar las predicciones se disponibilizaron dos métodos:

- POST `/predict`: este método recibe por parámetro un archivo con extensión CSV (*comma separated values*) con uno o varios registros. El archivo contiene todas las variables necesarias para realizar la predicción mas un identificador del paciente, que sirve para asociar la predicción devuelta. El método devuelve un archivo con las predicciones asociadas a cada paciente, incluyendo la probabilidad de cada clase.
- POST `/predict_single`: este método recibe por parámetro los valores de cada una de las variables y devuelve una predicción individual, que incluye la probabilidad de cada clase.

Ambos métodos cargan un modelo entrenado y realizan la ingeniería de características, esto quiere decir que se aplica todo el preprocesamiento de datos realizado durante el entrenamiento del modelo sobre los nuevos datos recibidos. Una vez que los datos tengan el formato correcto, se computan sobre el modelo para obtener las predicciones como respuesta. La misma plataforma utilizada para desplegar el modelo también se pudo configurar para que despliegue automáticamente la última versión de la API cuando sea necesario.

Por último, se desarrolla una aplicación de consola en lenguaje C#. Esta se encarga de realizar una consulta a la base de datos de la organización médica para

recuperar el listado de pacientes activos a ese momento, que hayan estado en tratamiento de hemodiálisis por más de 90 días. Con el resultado obtenido genera un archivo en formato CSV y realiza una llamada a la API utilizando el método `/predict`. Este método devuelve las predicciones de cada paciente y estas se graban en una nueva tabla de la base de datos. Esta aplicación de consola se configuró para ejecutarse una vez por día para mantener actualizada la información sobre el riesgo de mortalidad de los pacientes activos. Estas predicciones se muestran al usuario final en formato de reporte dentro de la aplicación de gestión de pacientes, donde pueden exportarla en formato de planilla de cálculos.

Capítulo 4

Ensayos y resultados

En este capítulo se exponen los resultados obtenidos de los entrenamientos de los modelos de IA utilizados. Se comentarán los resultados de los modelos de *Machine Learning* pero se hará hincapié sobre los resultados del modelo de *Deep Learning*, que llegó a resultados aceptables.

4.1. Resultados de los modelos

Una vez que se realizó el preprocesamiento tanto a los datos de entrenamiento como a los datos de prueba y validación, se entrenaron múltiples modelos de distintas arquitecturas. A continuación se presentan los resultados obtenidos utilizando el conjunto de datos de la estrategia 3, que contaba con mayor cantidad de datos y fue el que dio mejores resultados.

4.2. Resultados en modelos de *Machine Learning*

Si bien se probó con múltiples modelos de *Machine Learning*, ninguno fue capaz de lograr resultados aceptables, ya que la desproporción de clases era muy grande y las técnicas de submuestreo y sobremuestreo no resolvieron la situación. Algunos modelos como el *Random Forest* tenían ciertos parámetros para darle más peso a la clase minoritaria durante el entrenamiento, pero aún así no llegaban al 0.05 en la métrica de *recall*. Por este motivo no se continuó entrenando modelos de *Machine Learning* y se trabajó en mejorar los resultados en los modelos de *Deep Learning*.

4.3. Resultados en modelos de *Deep Learning*

En cuanto a los modelos de *Deep Learning*, se pudo aplicar la técnica de *Focal Loss* durante el entrenamiento de las redes neuronales, lo que ayudó mucho a compensar el desbalance de las clases y permitió a los modelos dar resultados más aceptables.

A continuación, se enumeran las métricas obtenidas con el el conjunto de validación sobre el modelo seleccionado descrito en el punto 3.5.2, luego de 10 épocas de entrenamiento:

- $F1\ Score = 0.0948$
- $Accuracy = 0.8442$
- $Precision = 0.0505$
- $Recall = 0.7666$
- $AUC = 0.8060$

En la tabla 4.1 se observa la matriz de confusión y en la figura 4.1 se muestra la curva ROC:

TABLA 4.1. Matriz de Confusión

	Predicción Negativa	Predicción Positiva
Clase Negativa	2356	432
Clase Positiva	7	23

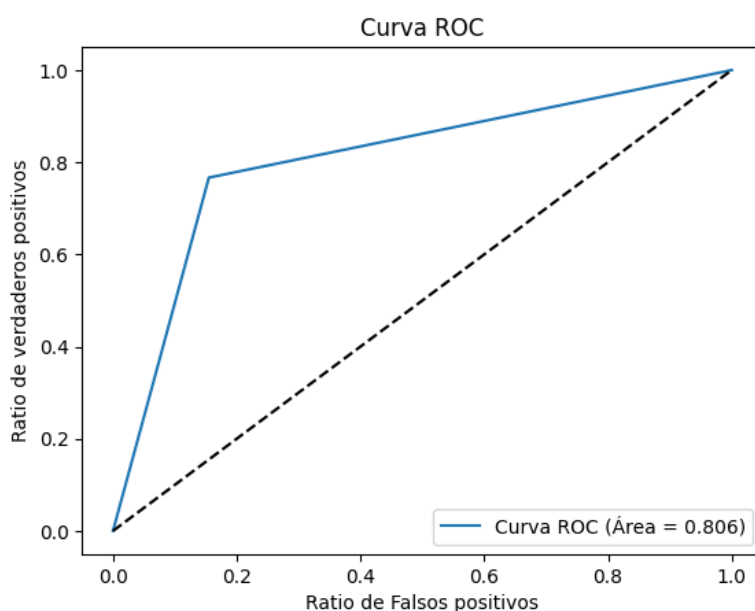


FIGURA 4.1. Matriz de Confusión.

En resumen, solo el modelo de red neuronal supera el requisito de obtener un *accuracy* por encima del 75 %, con un *recall* que también se ubica por encima del 75 %. Aunque se disponga de un conjunto de datos tan grande, el hecho de haber tanta desproporción entre las clases hace que los modelos tengan muchas dificultades para aprender correctamente. De todas formas los resultados obtenidos son muy útiles para el personal médico, ya que permiten identificar un conjunto más reducido de pacientes con alto riesgo para brindarles una mejor atención.

Capítulo 5

Conclusiones

En este capítulo se presentan las conclusiones y los aportes realizados por este trabajo. También, se detallan futuras líneas de investigación para mejorar los resultados.

5.1. Conclusiones generales

A continuación, se resaltan las conclusiones y los objetivos cumplidos con el desarrollo de este trabajo:

- Se logró implementar un modelo de aprendizaje profundo que realiza predicciones de mortalidad en pacientes en tratamiento de diálisis renal.
- El trabajo cumple con todos los requisitos principales definidos en la subsección 1.3.3.
- La desproporción de clases en los conjuntos de datos presentó un problema difícil de resolver que limitó la capacidad predictiva del modelo.
- Se pudo seguir el cronograma establecido sin mayores inconvenientes.

De los riesgos identificados para éste proyecto se tuvieron que ejecutar algunas acciones para mitigar dos de ellos:

- Riesgo 1: no disponer de un conjunto de datos adecuado para el entrenamiento del modelo. En este caso las acciones propuestas en el plan de proyecto no fueron suficientes para mitigar el problema, por lo que se tuvo que realizar una investigación más exhaustiva para encontrar a una solución.
- Riesgo 2: no lograr que el modelo entrenado realice predicciones correctas. Para mitigar el riesgo se invirtió mucho más tiempo del planificado en realizar entrenamientos de modelos y en investigar sobre arquitecturas y estrategias que sean útiles para resolver este tipo de problemas.

5.2. Próximos pasos

Para realizar este trabajo se utilizó la información clínica disponible de los pacientes que realizaron tratamientos de diálisis renal. Esto llevó a la recuperación de algunos datos inconsistentes y gran cantidad de datos faltantes. La calidad de los datos utilizada para entrenar los modelos podría mejorarse desde la etapa de registración, incluso definiendo algunos datos como obligatorios para que sean registrados en todos los pacientes. Esto ayudaría a contar con información valiosa para entrenar nuevos modelos de predicción en un futuro.

Sería beneficioso explorar nuevas arquitecturas de modelos que sirvan para resolver problemas de este tipo, donde hay tanta desproporción en los datos. También se podrían realizar nuevos conjuntos de datos desde una perspectiva distinta, que capturen la evolución de las distintas variables clínicas y de laboratorio para poder predecir como estarán dichas variables en un futuro.

Bibliografía

- [1] Peñaranda-Florez Devi Geesel
Torrado-Navarro Yoryely Pereira-Rodríguez Javier Boada-Morales Lorena.
«Dialisis y hemodialisis. Una revisión actual según la evidencia». En:
Journal (2017). URL: https://www.nefrologiaargentina.org.ar/numeros/2017/volumen15_2/articulo2.pdf (visitado 11-03-2024).
- [2] Mónica Valeria Cortés Badilla Marianella Álvarez Vega Laura María Quirós Mora.
«Inteligencia artificial y aprendizaje automático en medicina». En: *Journal* (2020). URL: <https://www.revistamedicasinergia.com/index.php/rms/article/view/557/923> (visitado 11-03-2024).
- [3] Victoria Eugenia García Montemayor. «Mortalidad en pacientes en diálisis: Importancia y desarrollo de nuevos métodos fiables de predicción». En: *Journal* (2021). URL: <https://helvia.uco.es/bitstream/handle/10396/22708/2022000002402.pdf?sequence=1&isAllowed=y> (visitado 11-03-2024).
- [4] Beatriz Ricardo Paez Sergio Orlando Escalona González Zoraida Caridad González Milán. «Predicción de mortalidad en pacientes con enfermedad renal crónica mediante el uso de la inteligencia artificial». En: *Journal* (2022). URL: <https://convencionsalud.sld.cu/index.php/convencionsalud22/2022/paper/viewFile/44/167> (visitado 11-03-2024).
- [5] Br. Shirley Aracelly Herrera Arce. «Índice neutrófilo/linfocito como predictor de mortalidad en pacientes que inician hemodiálisis». En: *Journal* (2022). URL: https://repositorio.upao.edu.pe/bitstream/handle/20.500.12759/9293/REP_SHIRLEY.HERRERA_INDICE.NEUTROFILO.pdf?sequence=1 (visitado 11-03-2024).
- [6] Mengqin Zhang Xing Chen Jun Zhang Jiyi Huang
Lu Zhang Lijing Yao Hengyuan Zhang. «Application of artificial intelligence in renal disease». En: *Journal* (2022). URL: <https://www.sciencedirect.com/science/article/pii/S2588914121000083> (visitado 11-03-2024).
- [7] Charat Thongprayoon Pattharawin Pattharanitima
Wisit Cheungpasitporn Pajaree Krisanapan Supawit Tangpanithandee.
«Revolutionizing Chronic Kidney Disease Management with Machine Learning and Artificial Intelligence». En: *Journal* (2023). URL: <https://www.mdpi.com/2077-0383/12/8/3018> (visitado 11-03-2024).
- [8] Yingxue Li Tingyu Chen Xiang Li Zhihong Liu Guotong Xie Tiange Chen.
«Artificial Intelligence in Nephrology: How Can Artificial Intelligence Augment Nephrologists' Intelligence?» En: *Journal* (2020). URL: <https://karger.com/kdd/article/6/1/1/186225> (visitado 11-03-2024).
- [9] Yanxiang Gao Enmin Xie Xuecheng Zhao Ziyu Guo Yike Li Nan
Shen-Jingyi Ren Jingang Zheng Zixiang Ye Shuoyan An. «The prediction of in-hospital mortality in chronic kidney disease patients with coronary artery disease using machine learning models». En: *Journal* (2023). URL:

Bibliografía

- [1] Peñaranda-Florez Devi Geesel
Torrado-Navarro Yoryely Pereira-Rodríguez Javier Boada-Morales Lorena.
«Dialisis y hemodialisis. Una revisión actual según la evidencia». En:
Journal (2017). URL: https://www.nefrologiaargentina.org.ar/numeros/2017/volumen15_2/articulo2.pdf (visitado 11-03-2024).
- [2] Mónica Valeria Cortés Badilla Marianella Álvarez Vega Laura María Quirós Mora.
«Inteligencia artificial y aprendizaje automático en medicina». En: *Journal* (2020). URL: <https://www.revistamedicasinergia.com/index.php/rms/article/view/557/923> (visitado 11-03-2024).
- [3] Victoria Eugenia García Montemayor. «Mortalidad en pacientes en diálisis: Importancia y desarrollo de nuevos métodos fiables de predicción». En: *Journal* (2021). URL: <https://helvia.uco.es/bitstream/handle/10396/22708/2022000002402.pdf?sequence=1&isAllowed=y> (visitado 11-03-2024).
- [4] Beatriz Ricardo Paez Sergio Orlando Escalona González Zoraida Caridad González Milán. «Predicción de mortalidad en pacientes con enfermedad renal crónica mediante el uso de la inteligencia artificial». En: *Journal* (2022). URL: <https://convencionsalud.sld.cu/index.php/convencionsalud22/2022/paper/viewFile/44/167> (visitado 11-03-2024).
- [5] Br. Shirley Aracelly Herrera Arce. «Índice neutrófilo/linfocito como predictor de mortalidad en pacientes que inician hemodiálisis». En: *Journal* (2022). URL: https://repositorio.upao.edu.pe/bitstream/handle/20.500.12759/9293/REP_SHIRLEY.HERRERA_INDICE.NEUTROFILO.pdf?sequence=1 (visitado 11-03-2024).
- [6] Mengqin Zhang Xing Chen Jun Zhang Jiyi Huang
Lu Zhang Lijing Yao Hengyuan Zhang. «Application of artificial intelligence in renal disease». En: *Journal* (2022). URL: <https://www.sciencedirect.com/science/article/pii/S2588914121000083> (visitado 11-03-2024).
- [7] Charat Thongprayoon Pattharawin Pattharanitima
Wisit Cheungpasitporn Pajaree Krisanapan Supawit Tangpanithandee.
«Revolutionizing Chronic Kidney Disease Management with Machine Learning and Artificial Intelligence». En: *Journal* (2023). URL: <https://www.mdpi.com/2077-0383/12/8/3018> (visitado 11-03-2024).
- [8] Yingxue Li Tingyu Chen Xiang Li Zhihong Liu Guotong Xie Tiange Chen.
«Artificial Intelligence in Nephrology: How Can Artificial Intelligence Augment Nephrologists' Intelligence?» En: *Journal* (2020). URL: <https://karger.com/kdd/article/6/1/1/186225> (visitado 11-03-2024).
- [9] Yanxiang Gao Enmin Xie Xuecheng Zhao Ziyu Guo Yike Li Nan
Shen-Jingyi Ren Jingang Zheng Zixiang Ye Shuoyan An. «The prediction of in-hospital mortality in chronic kidney disease patients with coronary artery disease using machine learning models». En: *Journal* (2023). URL:

- <https://link.springer.com/article/10.1186/s40001-023-00995-x> (visitado 11-03-2024).
- [10] Florian Jungmann Christina Glasner Philipp Stenzel Stephanie Strobl Aurélie Fernandez Daniel-Christoph Wagner Axel Haferkamp Peter Mildenerger Wilfried Roth Sebastian Foersch Stefan Schulz Ann-Christin Woerl. «Multimodal Deep Learning for Prognosis Prediction in Renal Cancer». En: *Journal* (2021). URL: <https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2021.788740/full> (visitado 11-03-2024).
- [11] Jiří Kaiser. «Dealing with Missing Values in Data». En: *Journal* (2014). URL: https://www.researchgate.net/profile/Kamarul-Ariffin-Mansor-2/post/After_I_ran_EFA_should_I_use_Factor_scores_or_Mean_or_Summed_scales_in_a_multiple_regression_analysis/attachment/59d63ee3c49f478072ea9518/AS%3A273773436047361%401442284081197/download/178-697-1-PB+Handling+Missing+Vaues.pdf (visitado 18-03-2024).
- [12] K.S.Vijaya Lakshmi B. Hemada. «A Study On Discretization Techniques». En: *Journal* (2013). URL: <https://www.ijert.org/a-study-on-discretization-techniques> (visitado 18-03-2024).
- [13] Adrián Rocha Íñigo. «Codificación de variables categóricas en aprendizaje automático». En: *Journal* (2020). URL: <https://idus.us.es/handle/11441/108887> (visitado 18-03-2024).
- [14] Malak Abdullah Roweida Mohammed Jumanah Rawashdeh. «Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results». En: *Journal* (2020). URL: https://www.researchgate.net/profile/Malak-Abdullah/publication/340978368_Machine_Learning_with_Oversampling_and_Undersampling_Techniques_Overview_Study_and_Experimental_Results/links/5ecd39764585152945121352/Machine-Learning-with-Oversampling-and-Undersampling-Techniques-Overview-Study-and-Experimental-Results.pdf (visitado 18-03-2024).
- [15] MFlow. *MFlow WebSite*. <https://mlflow.org/docs/latest/index.html>. (Visitado 24-03-2024).
- [16] Apache. *Apache Airflow WebSite*. <https://airflow.apache.org/>. (Visitado 24-03-2024).
- [17] Jenkins. *Jenkins WebSite*. <https://www.jenkins.io/>. (Visitado 24-03-2024).

- <https://link.springer.com/article/10.1186/s40001-023-00995-x> (visitado 11-03-2024).
- [10] Florian Jungmann Christina Glasner Philipp Stenzel Stephanie Strobl Aurélie Fernandez Daniel-Christoph Wagner Axel Haferkamp Peter Mildenerger Wilfried Roth Sebastian Foersch Stefan Schulz Ann-Christin Woerl. «Multimodal Deep Learning for Prognosis Prediction in Renal Cancer». En: *Journal* (2021). URL: <https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2021.788740/full> (visitado 11-03-2024).
- [11] Jiří Kaiser. «Dealing with Missing Values in Data». En: *Journal* (2014). URL: https://www.researchgate.net/profile/Kamarul-Arifin-Mansor-2/post/After_I_ran_EFA_should_I_use_Factor_scores_or_Mean_or_Summed_scales_in_a_multiple_regression_analysis/attachment/59d63ee3c49f478072ea9518/AS%3A273773436047361%401442284081197/download/178-697-1-PB+Handling+Missing+Vaues.pdf (visitado 18-03-2024).
- [12] K.S.Vijaya Lakshmi B. Hemada. «A Study On Discretization Techniques». En: *Journal* (2013). URL: <https://www.ijert.org/a-study-on-discretization-techniques> (visitado 18-03-2024).
- [13] Adrián Rocha Íñigo. «Codificación de variables categóricas en aprendizaje automático». En: *Journal* (2020). URL: <https://idus.us.es/handle/11441/108887> (visitado 18-03-2024).
- [14] Jonathan G. Kolo Suleiman O. E. Sadiku Abdullahi M. Orire Taliha A. Folorunso Abiodun M. Aibinu. «Effects of data normalization on water quality model in a recirculatory aquaculture system using artificial neural network». En: *Journal* (2019). URL: <http://repository.futminna.edu.ng:8080/jspui/bitstream/123456789/2033/1/IJ9.pdf> (visitado 02-04-2024).
- [15] Malak Abdullah Roweida Mohammed Jumanah Rawashdeh. «Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results». En: *Journal* (2020). URL: https://www.researchgate.net/profile/Malak-Abdullah/publication/340978368_Machine_Learning_with_Oversampling_and_Undersampling_Techniques_Overview_Study_and_Experimental_Results/links/5ecd39764585152945121352/Machine-Learning-with-Oversampling-and-Undersampling-Techniques-Overview-Study-and-Experimental-Results.pdf (visitado 18-03-2024).
- [16] R. Girshick K. He T. Y. Lin P. Goyal y P. Dollár. «Focal Loss for Dense Object Detection». En: *Journal* (2017). URL: <https://ieeexplore.ieee.org/document/8237586> (visitado 02-04-2024).
- [17] MLFlow. *MLFlow WebSite*. <https://mlflow.org/docs/latest/index.html>. (Visitado 24-03-2024).
- [18] Apache. *Apache Airflow WebSite*. <https://airflow.apache.org/>. (Visitado 24-03-2024).
- [19] Jenkins. *Jenkins WebSite*. <https://www.jenkins.io/>. (Visitado 24-03-2024).
- [20] Lee Giles Rich Caruana Steve Lawrence. «Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping». En: *Journal* (2000). URL: <https://proceedings.neurips.cc/paper/2000/file/059fdcd96baeb75112f09fa1dcc740cc-Paper.pdf> (visitado 03-04-2024).
- [21] FastAPI. *FastAPI WebSite*. <https://fastapi.tiangolo.com/>. (Visitado 04-04-2024).