

UNIVERSIDAD DE VALLADOLID  
MÁSTER UNIVERSITARIO  
**Ingeniería Informática**



TRABAJO FIN DE MÁSTER

**Pronóstico del Manejo de la Enfermedad  
de Síncope en el Departamento de  
Urgencias usando Algoritmos de Machine  
Learning**

Realizado por **ROSMERI MARTÍNEZ LICORT**



**Universidad de Valladolid**  
**13 de septiembre de 2022**

Tutores: ISABEL DE LA TORRE DÍEZ  
JESÚS MARÍA VEGAS HERNÁNDEZ



# Universidad de Valladolid



## Máster universitario en Ingeniería Informática

D. ISABEL DE LA TORRE DÍEZ, profesora del departamento de Teoría de la Señal, Comunicaciones e Ingeniería Telemática, área de Telemática. y D. JESÚS MARÍA VEGAS HERNÁNDEZ, profesor del departamento de Informática, área de Ciencias de la Computación e Inteligencia Artificial.

### **Exponen:**

Que el alumno D. ROSMERI MARTÍNEZ LICORT, ha realizado el Trabajo final de Máster en Ingeniería Informática titulado "PRONÓSTICO DEL MANEJO DE LA ENFERMEDAD DE SÍNCOPE EN EL DEPARTAMENTO DE URGENCIAS USANDO ALGORITMOS DE MACHINE LEARNING".

Y que dicho trabajo ha sido realizado por el alumno bajo la dirección de los que suscriben, en virtud de lo cual se autoriza su presentación y defensa.

En Valladolid, 13 de septiembre de 2022

Vº. Bº. del Tutor:

Vº. Bº. del co-tutor:

D. JESÚS MARÍA VEGAS  
HERNÁNDEZ

D. ISABEL DE LA TORRE DÍEZ



## **Resumen**

El síncope es una enfermedad alarmante y una causa frecuente de evaluación en la sala de urgencias, donde no se cuenta con herramientas eficientemente rápidas para su manejo. La predicción temprana del ingreso y el manejo acertado de la enfermedad, permitiría: ahorrar tiempo, recursos y la personalización de la atención. La finalidad de este trabajo es usar machine learning para crear modelos de clasificación binaria que permitan, tras la estratificación de riesgo de la enfermedad, pronosticar: si un paciente necesita hospitalización, si necesita ser admitido en la UCI y su probable fallecimiento durante la atención sanitaria. Es un tema del que poco se concreta en la bibliografía científica accedida, aún así en esta investigación se explotan las bibliotecas de python siguiendo una metodología de trabajo elaborada. A pesar de no contar con una interpretación médica profunda de los errores en los datos, como resultado de esta investigación se obtuvo un modelo de alto rendimiento para el pronóstico de la hospitalización de pacientes, no así para los otros dos pronósticos. Con este punto de partida, se cree que se pueden obtener mejores resultados en trabajos futuros.

## **Descriptores**

Manejo de síncope, Machine learning, Departamento de Emergencia, Estratificación de Riesgo.

## **Abstract**

Syncope is an alarming disease and a frequent cause of evaluation in the emergency department, where there are no efficient rapid tools for its management. The early prediction of admission and the correct management of the disease, would: save time, resources and the personalization of attention. The purpose of this work is to use machine learning to create binary classification models that allow, after disease risk stratification, to predict: if a patient needs hospitalization, if they need to be admitted to the ICU (Intensive Care Unit) and their probable death during health care. It is a topic about which little is specified in the scientific bibliography accessed, even so in this research the python libraries are exploited following an elaborate work methodology. Although not having a deep medical interpretation of those data noise, as a result of this research a high-performance model was obtained for the prognosis of patient hospitalization, not so for the other two prognoses. With this starting point, it is believed that better results can be obtained in future works.

## **Keywords**

Syncope management, Machine learning, Emergency department, Risk stratification

---

# Índice general

---

Índice general	III
Índice de figuras	V
Índice de tablas	VIII
<b>1. Introducción</b>	<b>1</b>
1.1. Objetivos	2
1.2. Marco Contextual	3
1.3. Estructura de la Memoria	4
<b>2. Marco Teórico</b>	<b>5</b>
2.1. Estado del arte en la ciencia	5
2.2. Machine Learning aplicado a la medicina clínica	10
2.3. Acerca del síncope	18
2.4. Conclusión del capítulo	20
<b>3. Desarrollo del Proyecto</b>	<b>22</b>
3.1. Descripción del conjunto de datos	22
3.2. Técnicas y herramientas	28
3.3. Selección y construcción del modelo	30
3.4. Conclusión del capítulo	82
<b>4. Resultados y Discusión</b>	<b>83</b>
4.1. Resultados	83
4.2. Selección de modelo por clasificación	95
4.3. Discusión	98
<b>5. Conclusiones y Líneas de Trabajo Futuras</b>	<b>100</b>
5.1. Conclusiones	100
5.2. Trabajo Futuro	101

<i>Índice general</i>	IV
<b>Apéndices</b>	<b>102</b>
<b>Apéndice A Plan de Proyecto</b>	<b>103</b>
A.1. Introducción . . . . .	103
A.2. Planificación temporal . . . . .	103
A.3. Viabilidad legal . . . . .	103
<b>Bibliografía</b>	<b>105</b>



---

# Índice de figuras

---

2.1. Metodología Prisma utilizada . . . . .	7
2.2. Los principales modelos de aprendizaje supervisado [24] . . . . .	11
2.3. Ecuación que describe relaciones en los modelos de ML lineales . . . . .	13
2.4. Ecuación que describe el modelo de regresión logística . . . . .	13
2.5. División de clases por el modelo de SVM . . . . .	14
2.6. Ejemplo de un árbol DT sencillo [2] . . . . .	14
2.7. Ecuación de la tasa de error para el aprendiz j . . . . .	16
2.8. Ecuación del peso para el aprendiz j . . . . .	16
2.9. Actualización de pesos para el aprendiz j . . . . .	16
2.10. Posibles aplicaciones de los modelos de ML [24] . . . . .	18
2.11. Factores de riesgo en el pronóstico del síncope [70] . . . . .	20
2.12. Puntaje Canadiense de Riesgo del Síncope [70] . . . . .	21
3.13. Tabla de valores normales de las componentes de tensión arterial [16] . . . . .	26
3.14. Visualización de variable “Hospital” . . . . .	32
3.15. Visualización de variable “UVI” . . . . .	33
3.16. Visualización de variable “Coronarias” . . . . .	34
3.17. Visualización de variable “ICTUS” . . . . .	34
3.18. Visualización de variable “Mortalidad Hospitalaria” . . . . .	35
3.19. Distribución de valores de la variable “Sexo” . . . . .	36
3.20. Distribución de valores de la variable “Triage” . . . . .	36
3.21. Descripción estadística de las variables “GCS.Ocular”, “GCS.Verbal”, “GCS.Motora”, respectivamente . . . . .	37
3.22. Gráfica de caja de variable “Edad” . . . . .	40
3.23. Distribución original de variable “Edad” . . . . .	40
3.24. Gráfica de caja de variable “Frecuencia Respiratoria” . . . . .	41
3.25. Distribución original de variable “Frecuencia Respiratoria” . . . . .	42
3.26. Gráfica de caja de variable “Saturación de Oxígeno en Sangre” . . . . .	43
3.27. Distribución original de variable “Saturación de Oxígeno en Sangre” . . . . .	44
3.28. Gráfica de caja de variable “Oxígeno en Aire Inspirado” . . . . .	45
3.29. Distribución original de variable “Oxígeno en Aire Inspirado” . . . . .	46

3.30. Gráfica de caja de variable “Tensión Arterial Sistólica” . . . . .	47
3.31. Distribución original de variable “Tensión Arterial Sistólica” . . . . .	48
3.32. Gráfica de caja de variable “Tensión Arterial Diastólica” . . . . .	50
3.33. Distribución original de variable “Tensión Arterial Diastólica” . . . . .	50
3.34. Gráfica de caja de variable “Frecuencia Cardíaca” . . . . .	51
3.35. Distribución original de variable “Frecuencia Cardíaca” . . . . .	52
3.36. Distribución original de variable “Tiempo de Trombina” . . . . .	54
3.37. Gráfica de caja de variable “Hematocrito” . . . . .	55
3.38. Distribución original de variable “Hematocrito” . . . . .	56
3.39. Gráfica de caja de variable “Leucocito” . . . . .	57
3.40. Distribución original de variable “Leucocito” . . . . .	58
3.41. Distribución original de variable “Plaquetas” . . . . .	59
3.42. Gráfica de caja de variable “Glucosa” . . . . .	60
3.43. Distribución original de variable “Glucosa” . . . . .	61
3.44. Gráfica de caja de variable “Creatinina” . . . . .	62
3.45. Distribución original de variable “Creatinina” . . . . .	63
3.46. Distribución original de variable “Troponina” . . . . .	64
3.47. Los valores de “Edad” con respecto a la salida “Hospital” . . . . .	65
3.48. Grupos de “Edad” con respecto a la salida “Hospital” . . . . .	65
3.49. “Saturación de oxígeno en sangre” con respecto a la salida “Hospital” . . . . .	66
3.50. Código de “Saturación de oxígeno en sangre” con respecto a la salida “Hospital” . . . . .	66
3.51. Valores de “Glucosa” con respecto a la salida “Hospital” . . . . .	67
3.52. Código de valores de “Glucosa” con respecto a la salida “Hospital” . . . . .	67
3.53. Los valores de “Creatinina” con respecto a la salida “Hospital” . . . . .	67
3.54. Los valores de “FiO2” con respecto a la salida “Hospital” . . . . .	68
3.55. Respiración “FiO2” con respecto a la salida “Hospital” . . . . .	68
3.56. Histograma de “Hospital” para “ST” . . . . .	69
3.57. Relación de valores entre “Frecuencia cardíaca” y “Frecuencia respiratoria” . . . . .	69
3.58. Matriz de correlación con la salida “Hospital” . . . . .	70
3.59. Los valores de “Edad” con respecto a la salida “Mortalidad Hospitalaria” . . . . .	70
3.60. Los valores de “GCS.O” con respecto a la salida “Hospital” . . . . .	71
3.61. Histograma de “Mortalidad Hospitalaria” para “ST” . . . . .	71
3.62. Histograma de “Mortalidad Hospitalaria” para “Hospital” . . . . .	72
3.63. Histograma de “UCI” para “Troponina” . . . . .	72
3.64. Correlación de las variables independientes respecto a las variables dependientes . . . . .	74
3.65. Distribución de valores de la variable “UCI” . . . . .	76
3.66. Modelos de clasificación incluidos en la biblioteca de <i>pyCaret</i> . . . . .	81
4.67. Rendimiento de los modelos rápidos en “Hospitalización” en el <i>round 1</i> . . . . .	84
4.68. Rendimiento de los modelos <i>Random Forest Classifier</i> y <i>Light Gradient Boosting Machine</i> en la predicción sobre el conjunto de prueba, para “Hospitalización” . . . . .	85
4.69. Matriz de confusión de “Hospitalización” <b>Logistic Regression</b> . . . . .	85
4.70. Matriz de confusión de “Hospitalización” para <b>Random Forest Classifier</b> . . . . .	86
4.71. Rendimiento de los modelos rápidos en “Mortalidad Hospitalaria” en el <i>round 1</i> . . . . .	87
4.72. Rendimiento de los modelos rápidos en “UCI” en el <i>round 1</i> . . . . .	87

4.73. Rendimiento de los modelos rápidos en “Hospitalización” en el <i>round 2</i> . . . . .	88
4.74. Rendimiento de los modelos optimizados, para “Hospitalización” . . . . .	89
4.75. Rendimiento en la predicción sobre el conjunto de prueba, para “Hospitalización” . . . . .	89
4.76. Matriz de confusión de “Hospitalización” <b>Extra Trees Classifier</b> . . . . .	90
4.77. Matriz de confusión de “Hospitalización” <b>Light Gradient Boosting Machine</b> . . . . .	90
4.78. Matriz de confusión de “Hospitalización” <b>Random Forest</b> . . . . .	91
4.79. Rendimiento de los modelos rápidos en “UCI” en el <i>round 2</i> . . . . .	92
4.80. Matriz de confusión de “Mortalidad Hospitalaria” <b>Discriminant Analysis</b> . . . . .	92
4.81. Rendimiento de los modelos rápidos en “Hospitalización” en el <i>round 3</i> . . . . .	93
4.82. Rendimiento de los modelos del <i>round3</i> en la predicción sobre el conjunto de prueba, para “Hospitalización” . . . . .	93
4.83. Matriz de confusión de “Mortalidad Hospitalaria” <b>Extra Trees Classifier</b> . . . . .	94
4.84. Matriz de confusión de “Mortalidad Hospitalaria” <b>Random Forest Classifier</b> . . . . .	94
4.85. Matriz de confusión de “Mortalidad Hospitalaria” <b>Gradient Boosting Classifier</b> . . . . .	95
4.86. Rendimiento de los modelos rápidos en “UCI” en el <i>round 3</i> . . . . .	96
4.87. Parámetros del mejor modelo para “Hospitalización” . . . . .	97
4.88. Importancia para las variables en el mejor modelo de “Hospitalización” . . . . .	98
4.89. Resultados de estudio anterior sobre los datos, para la variable “Hospitalización” . . . . .	99

---

## Índice de tablas

---

3.1. Frecuencia de registros con cada valor de “Ritmo” . . . . .	38
3.2. Frecuencia de registros con cada valor de “ST” . . . . .	39
3.3. Descripción de variable “Edad” . . . . .	39
3.4. Descripción de variable “Frecuencia Respiratoria” . . . . .	41
3.5. Descripción de variable “Saturación de Oxígeno en Sangre” . . . . .	43
3.6. Descripción de variable “Oxígeno en Aire Inspirado” . . . . .	45
3.7. Descripción de variable “Tensión Arterial Sistólica” . . . . .	47
3.8. Descripción de variable “Tensión Arterial Diastólica” . . . . .	49
3.9. Descripción de variable “Frecuencia Cardíaca” . . . . .	51
3.10. Descripción de variable “Tiempo de Trombina” . . . . .	53
3.11. Descripción de variable “Hematocrito” . . . . .	55
3.12. Descripción de variable “Leucocito” . . . . .	57
3.13. Descripción de variable “Plaquetas” . . . . .	58
3.14. Descripción de variable “Glucosa” . . . . .	60
3.15. Descripción de variable “Creatinina” . . . . .	62
3.16. Descripción de variable “Troponina” . . . . .	63
3.17. Comparación de las salidas del conjunto de datos . . . . .	73
3.18. Comparación de salidas de la futura variable “UCI” . . . . .	75
A.1. Plan de Trabajo . . . . .	104

---

# 1: Introducción

---

En el mundo de la medicina, los clínicos conforman estructuras o comisiones con el fin de analizar y resolver problemas que resultan complicados de solucionar para un solo individuo. Actualmente, la medicina se ha tornado aún más compleja, pues se dispone de más terapias, medicamentos y exámenes; procesar tales datos excede la capacidad de comprensión de la mente humana, por ello se hace necesario buscar herramientas con la capacidad de integrar grandes cantidades de datos, reconocer patrones y crear modelos que permitan: compensar las limitaciones humanas, disminuir carga médica, acelerar la atención y brindar manejos más personalizados [55],[63],[32].

Por ello se ha popularizado la salud digital, que describe el uso de la información digital, los datos y las tecnologías de la comunicación para recopilar, compartir y analizar la información sanitaria con el fin de mejorar la salud del paciente, la prestación de asistencia sanitaria y la educación médica. Este concepto engloba la telesalud, las historias clínicas electrónicas, la monitorización de dispositivos implantables, los datos de los sensores portátiles, la salud conductual, la medicina personalizada, la analítica y la Inteligencia Artificial (IA) [61].

La IA y, específicamente el Aprendizaje Automático (mencionado más en el mundo técnico como Machine Learning o ML, por sus siglas en inglés), han recibido recientemente una atención significativa y se han descrito como el “próximo gran avance tecnológico que afectará la prestación de atención médica”. Una de las aplicaciones más citadas es la estratificación de riesgo de los pacientes para predecir resultados específicos. Se han desarrollado múltiples puntajes de riesgo clínico y reglas de decisión para pacientes del Departamento de Emergencias (ED, por sus siglas en inglés) que utilizan métodos ML [31],[10],[22],[34],[15],[99],[87],[98],[60],[4],[41]. Por otro lado, la estratificación del riesgo había sido un foco de investigación en medicina de emergencia durante muchos años antes de que la IA se hiciera popular y se conoce que existen enfoques estadísticos para el desarrollo y la notificación de herramientas de riesgo a partir de datos recopilados prospectivamente. Muchas de estas herramientas se han convertido en una parte estándar del entrenamiento y la práctica en la especialidad [31],[38],[81],[76],[11],[43],[44],[40],[46],[50][17],[83],[82],[93],[78].

El síncope es una causa frecuente de evaluación en la sala de urgencias y es un evento alarmante que conlleva la posibilidad de resultados graves, incluida la Muerte Súbita Cardíaca (MSC). Por lo tanto, la estratificación de riesgo inmediata debe aplicarse siempre que ocurra un síncope, especialmente en el ED que es donde ocurren las presentaciones más dramáticas y a menudo se diagnostica erróneamente. Durante las últimas dos décadas, se han desarrollado varias herramientas de predicción de síncope para refinar la estratificación de los pacientes y clasificar a los pacientes que necesitan atención hospitalaria experta de aquellos que pueden recibir atención experta no urgente en la comunidad. Sin embargo, a pesar de los resultados prometedores, las herramientas de pronóstico para el síncope siguen siendo un desafío y, a menudo, poco efectivas. Tanto por la falta de pruebas que respalden el diagnóstico como por la ineficiencia del pronóstico de evolución del paciente, la atención a las personas que llegan con esta enfermedad al departamento de urgencias representa un problema desafiante para el médico que trabaja en este departamento [24],[31],[70]. El caso específico de la predicción temprana del ingreso hospitalario podría generar un avance significativo en la estratificación del riesgo individual y la personalización de la atención, así como puede contribuir a optimizar los recursos disponibles y la asignación de camas, además de que se pueden acortar los tiempos de ingreso y se limita el hacinamiento, que es tan común en esta sección primaria de atención hospitalaria [24],[31].

En [24] se realiza una revisión exhaustiva y actualizada en este sentido e identifican los conceptos de ML aplicados a la medicina clínica, además, exploran el caso particular del manejo del síncope en el departamento de urgencias; concluyendo que son pocos los estudios que han abordado el tema de la aplicación de ML al pronóstico de la atención requerida por el paciente con síncope que llega al departamento de urgencias. Aun así, los resultados obtenidos son alentadores y llevan a pensar que los algoritmos de ML tienen todavía potencial para ser explotado y podrían ayudar a hacer más eficiente el manejo del paciente con síncope en el departamento de urgencias.

Este manejo se interpreta, tanto en la bibliografía consultada como en las indicaciones médicas recibidas para esta investigación, como la asignación del paciente al servicio sanitario específico que se cree requiere el paciente luego de haberse identificado y evaluado el riesgo de afrontar algún estado de la enfermedad, otra enfermedad asociada o la posibilidad de morir.

## 1.1. Objetivos

Para la realización de este trabajo, se sitúa el posible uso de los algoritmos de ML en el siguiente entorno: la atención médica del cuerpo de urgencias no cuenta con herramientas eficientemente rápidas para el manejo de los pacientes con la enfermedad de síncope. De ahí que, el **problema de investigación** planteado es la necesidad de manejar de forma más oportuna a los pacientes con la enfermedad de síncope en el servicio de urgencias, pronosticando qué admisiones de emergencia derivan en hospitalización, cuáles a la Unidad de Cuidados Intensivos (UCI) y cuál es la probabilidad de que los pacientes de síncope fallezcan durante la estancia en las instalaciones hospitalarias.

El **objeto principal de investigación** a transformar con el resultado obtenido en este trabajo, son los datos médicos proporcionados. De ahí que el modelado de datos es la actividad esencial de la práctica y el método científico que mejor aborda el desarrollo de la práctica es la ciencia de la computación teórica.

Para la solución del problema en el desarrollo de la investigación, se ha encontrado evidencia suficiente, en la bibliografía consultada, que sugiere el establecimiento de la siguiente **hipótesis**: se pueden usar modelos de ML enfocados en la clasificación, para pronosticar (oportunamente y desde el servicio urgencias) la hospitalización del paciente con enfermedad de síncope, así como el trasladado directamente a la UCI y la posibilidad de fallecer en las horas siguientes durante su atención sanitaria; todo eso se pretende lograra a partir de procesar la información de descripciones personales, las variables fisiológicas y analíticas conformadas con las mediciones realizadas a los pacientes en el servicio de urgencias y del triaje con que se calificó el estado del paciente tras su traslado en ambulancia.

A partir del problema planteado y de la hipótesis señalada, el **objetivo general de investigación** es crear tres modelos de clasificación binaria para, a partir de los datos disponibles, pronosticar si un paciente necesita hospitalización, si necesita ser admitido en la UCI y, además, su probable fallecimiento durante los días de atención hospitalaria.

Trazando la estrategia de planificación con la que concretizar el objetivo general, se desglosan los siguientes **objetivos específicos**:

1. Confeccionar un estado del arte y un marco teórico alrededor del uso de algoritmos de ML en relación con la enfermedad de síncope y, particularmente, en el pronóstico de la atención requerida por los pacientes recibidos en el departamento de emergencias.
2. Preparar los datos y conformar *features* (traducido del inglés como características o rasgos) para cada clasificación.
3. Crear un modelo de clasificación en cada caso, partiendo de la exploración y la optimización de estos.

Estos objetivos se concretan en tareas que se pueden consultar en el apéndice [A](#).

## 1.2. Marco Contextual

Esta investigación surge a petición de un equipo médico asociado a la docencia en la Facultad de Medicina de la UVA (que también se ha encargado de registrar los datos necesarios y elaborar la base de datos) y se materializa como un proyecto emprendido por el grupo de investigación (GIR<sup>1</sup>) de *Telemedicina y eSalud* (GTe), perteneciente a la Escuela Técnica Superior de Ingenieros de Telecomunicación de la UVA. Se desarrolla en

---

<sup>1</sup>GIR: Grupos de Investigación Reconocidos

el marco de Trabajo de Fin de Máster, para el Título de Máster en Ingeniería Informática, como una colaboración con la Escuela de Ingeniería Informática de la UVA; lo que es posible por la similitud de perfiles entre los profesionales de las escuelas relacionadas, específicamente en la rama de la ciencia de datos.

### **1.3. Estructura de la Memoria**

Esta memoria comprende al registro de cada etapa en el cumplimiento de los objetivos trazados y está seccionada en cinco capítulos. De aquí en adelante, el capítulo 2 recoge el estado del arte en la ciencia, en cuanto al uso de las herramientas de análisis inteligente con bases de datos médicas similares a la que se tiene en esta ocasión, además incluye conceptos que permiten comprender el problema planteado. En el capítulo 3 se describen la base de datos y la metodología construida para lograr el objetivo del trabajo. El capítulo 4 presenta los resultados obtenidos de la aplicación de algoritmos de aprendizaje supervisado para clasificar el manejo de los pacientes de síncope en el cuerpo de urgencias. Finalmente, el capítulo 5 recoge las conclusiones de este trabajo y algunas ideas sobre posibles líneas futuras.



---

## 2: Marco Teórico

---

Después de haber definido el problema como paso introductorio al actual proyecto de ML, este capítulo se conforma como el siguiente paso en la guía de desarrollo del trabajo: mirar el panorama general. En el capítulo se introducen conceptos básicos de ML aplicados a la medicina clínica y se exploran sus principales aplicaciones en el entorno del ED, con un enfoque particular en el manejo del síncope. Esta información fue recopilada a través de una búsqueda bibliográfica, con algunas limitaciones de tiempo, la que se describe también en este apartado y con la que se conforma un estado del arte.

Es de destacar que la bibliografía encontrada es mayormente de fecha reciente, lo que demuestra que el manejo del síncope está representando un problema más considerable en el cuerpo de urgencias desde hace pocos años. Se trata de una enfermedad complicada de identificar entre otras, adicionando a esto que tanto España como cualquier país del mundo ha enfrentado en los últimos años un aumento de saturación en el ED por el brote pandémico de Covid-19.

### 2.1. Estado del arte en la ciencia

La metodología usada para organizar la búsqueda de precedentes del tema y poder sacar información relevante que permita encausar el trabajo actual, fue la metodología prismática (conocida como PRISMA), con la que se organiza la documentación que se va revisando a modo de identificar, seleccionar, evaluar y sintetizar la que se ajuste a lo requerido para el cumplimiento del objetivo. Se siguió esta metodología sobre las bases de datos científicas *PubMed* y *PubMedCentral*, principalmente, por ser reconocidas en el mundo de la investigación médica y probadas en el grupo de investigación GTe como las que abarcan más artículos científicos de ésta índole. En algunos de estos artículos son referenciadas otras bases científicas que también se explotaron un poco para la actual investigación: *ProQuest*, *Oxford Academic* y *ScienceDirect*; las que presentan contenido científico relevante de revistas indexadas, repositorios, archivos y otras colecciones de texto.

La búsqueda se direccionó en el campo de la publicación de investigaciones en artículos científicos, en base a esto, se determinaron los items considerados apropiados. Igualmente, el motor de búsqueda en cada base de datos científica se orientó para que estos elementos fueran mapeados en la estructura de los artículos.

## Búsqueda de investigaciones científicas

Además de buscar aplicaciones y dispositivos ya probados o en proceso de evaluación, se investigó la existencia de algún modelo basado en machine learning y que fuera validado, o al menos creado, en el cumplimiento de los requerimientos del presente estudio. Esta etapa se inició con una búsqueda generalista en la que se usaron solo dos items y luego el número se amplió para direccionar la búsqueda hacia los modelos que cumplieran los objetivos a lograr con este estudio.

“(síncope[Title]) AND (machine learning[Abstract])”

“(syncope) AND (machine learning) AND (risk stratification) AND (emergency)”

En ambos casos, las bases de datos resultaron filtrar artículos relacionados solo en parte con todos los requisitos del problema de investigación. Siguiendo la metodología PRISMA, se fue descartando cada grupo de artículos por las siguientes razones: no tener acceso desde la cuenta de la universidad, por el duplicado de resultados al haber coincidencias en las bases científica y por la total irrelevancia con respecto al tema médico. En este último filtrado de artículos, se tuvo en cuenta que la lectura del Título, del Resumen, de los Métodos y de los Resultados obtenidos en los 161 artículos descartados por el mencionado filtro, evidenció que se trataba de algoritmos de machine learning aplicados para otras enfermedades y en el caso de tratarse de la enfermedad de síncope, siempre era sobre el diagnóstico de la misma. La Figura 2.1 muestra la aplicación de la metodología prismática en la búsqueda de estudios relevantes.

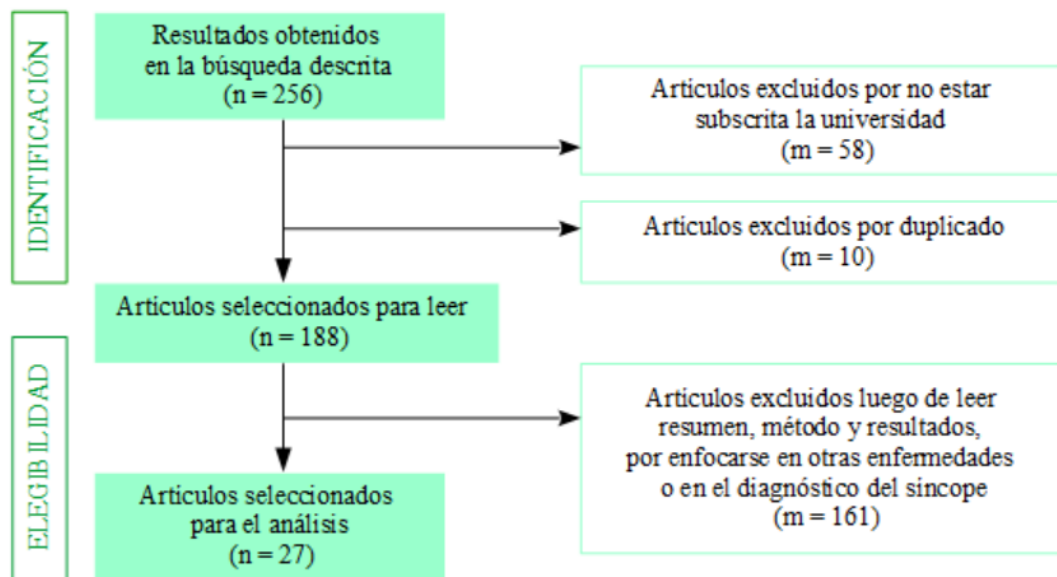


Figura 2.1: Metodología Prisma utilizada

De los artículos seleccionados para analizar, la gran mayoría trata al síncope como una enfermedad asociada a otras más comunes de nombrar y tratar entre la población como son: la diabetes, los problemas neuronales, VIH, epilepsia, fibrilación <sup>2</sup> y, las más mencionadas de todas, las enfermedades cardíacas. Normalmente se mencionan los modelos creados, pero pocas veces se mencionan los algoritmos, los análisis seguidos para elegirlos, así como para seleccionar o conformar rasgos.

En este apartado también se muestra el contenido de algunos artículos, en aras de comparar sus objetivos y las vías de solución que proponen. En todos los casos en que se menciona, la “estratificación de riesgo” se limita a describir eventos adversos de la enfermedad de síncope y a la predicción de estos, pocas veces se llegó a plantear la especificidad de pronosticar el manejo en urgencias. Realmente, los artículos que resultaron ser más cercanos al objetivo de esta investigación, fueron revisados por haber sido citados en los artículos seleccionados para el análisis. Es decir, que no fueron obtenidos directamente en las bases de datos científicas.

En [28] se probó el uso de Redes Neuronales Artificiales, (ANN por sus siglas en inglés) en la estratificación de riesgo de pacientes evaluados en el servicio de urgencias por síncope. Descubrieron que la precisión predictiva de las ANN era comparable, si no superior, a la de las herramientas de pronóstico actualmente disponibles. En [29], las ANN podrían predecir la hospitalización de los pacientes con síncope con una sensibilidad del 100 % y una especificidad del 79 %, lo que podría aumentar la idoneidad del tratamiento médico y, en consecuencia, la eficiencia hospitalaria. Sin embargo, la metodología utilizada se vio limitada por la poca disponibilidad de datos y no se realizó una validación externa del

<sup>2</sup>Fibrilación: Pérdida del control de un músculo por parte de la vía nerviosa [19]

modelo, aunque eventualmente abrió una nueva era para soluciones innovadoras habilitadas por tecnología para personalizar y mejorar la estratificación del riesgo de los pacientes con síncope [24],[70].

En [12] se menciona un algoritmo basado en la combinación de un péptido <sup>3</sup> neuro-endocrino implicado en las causas del síncope cardíaco y el juicio clínico. El mencionado péptido se identifica como MRproANP (del inglés midregional-pro-A-type natriuretic peptide). Este algoritmo no es descrito en su totalidad, pero se menciona que arrojó una sensibilidad del 99 % y un valor predictivo negativo del 99 % para el descarte temprano del síncope cardíaco, que es un tipo específico de síncope [70].

En [47] se utilizaron métodos de ML para mejorar la predicción de eventos cardíacos adversos en pacientes con Miocardiopatía Hipertrófica (MCH). La precisión predictiva del modelo de referencia fue del 73 % (76 % de sensibilidad, 72 % de especificidad), mientras que la del mejor modelo de aprendizaje automático fue del 85 % (p. ej., sensibilidad del 88 %, especificidad del 84 % ). Los 4 modelos de aprendizaje automático superaron significativamente al modelo de referencia; por ejemplo, el área bajo la curva característica operativa del receptor fue de 0,79 con el modelo de referencia frente a 0,93 con la regresión neta elástica ( $p < 0,001$ ). En [79] se construyó un modelo de ML para predecir la mortalidad, utilizando las características de los pacientes, las comorbilidades, las pruebas de laboratorio y el manejo del Departamento de Emergencias (DE) de los pacientes. El modelo creado arrojó un AUC (que se entiende como *área bajo la curva de características operativas del receptor*) de 0,924 con un intervalo de confianza (IC) del 95 %: 0,917-0,930. No se especifica el tipo de modelo, ni se explican los pasos seguidos para obtenerlo.

En [31] compararon modelos predictivos, desarrollados mediante métodos de aprendizaje automático (ML), con el Canadian Syncope Risk Score (CSRS) que es una herramienta de riesgo desarrollada con regresión logística para predecir Eventos Adversos Graves (SAE, por sus siglas en inglés) para el paciente de síncope en el ED. En [31] se le llama SAE a: la muerte, las arritmias, la cardiopatía estructural, la embolia pulmonar, la hemorragia. Se utilizaron 43 variables que incluyen datos demográficos, historia clínica, signos vitales, hallazgos del ECG, análisis de sangre y la impresión diagnóstica del médico de urgencias. El rendimiento de los modelos se evaluó mediante las respectivas AUC en la validación del modelo en los datos de prueba: para CSRS 0,902 (0,877-0,926), para Regresión Regularizada 0,903 (0,877-0,928), para Gradient Boosting 0,914 (0,894-0,934), para Red Neuronal Profunda 0,906 (0,883-0,929), para Gradient Boosting Simplificado 0,904 (0,881-0,927). En general, los modelos ML igualaron el rendimiento del CSRS, aunque algunos modelos utilizaron menos predictores.

En [52] presentan un estudio para la identificación de pacientes con mayor riesgo de ingreso en la UCI desde el servicio de urgencias. Utilizaron variables recogidas de forma rutinaria en el triaje realizado en la ED de hospitales de Portugal y Estados Unidos (signos vitales derivados del estudio de triaje: temperatura, frecuencia cardíaca, frecuencia respiratoria, presión arterial sistólica, presión arterial diastólica, presión arterial media,

---

<sup>3</sup>Péptido: Cualquier componente de las proteínas[19]

oximetría de pulso SpO<sub>2</sub> y escala de dolor), los niveles de glucemia, la Escala de Coma de Glasgow (GCS, por sus siglas en inglés), la prioridad de triaje asignada al paciente, la edad y el sexo del paciente, el modo de llegada a urgencias (ambulancia, por sus medios), discapacidades (camilla, silla de ruedas o ninguna), hora de triaje ( día de la semana, hora y mes), visita a urgencias (primer triaje registrado en el sistema o no), prescripción de medios complementarios de diagnóstico en el triaje (número de exámenes), y tipo de exámenes prescritos (oftalmología, otorrinolaringología, electrocardiograma, radiografía y ortopédico); además de que aplicaron el procesamiento del lenguaje natural a la queja principal del paciente, que fue incluida como predictor del modelo. Así, crearon modelos con las técnicas de ML: regresión logística, bosques aleatorios y un algoritmo de refuerzo de submuestreo aleatorio. Mediante validación cruzada seleccionaron el modelo de regresión logística como el mejor con AUC de 0,91 (intervalo de confianza del 95 %: 0,90 a 0,92) y 0,30 (intervalo de confianza del 95 %: 0,27 a 0,33) para el hospital de los Estados Unidos y de 0,85 (IC 95 % 0,83-0,86) y 0,06 (IC 95 %: 0,05-0,07) para el hospital portugués.

En [98] se utilizaron datos de triaje disponibles de forma rutinaria como predictores (datos demográficos, signos vitales de triaje, quejas principales, comorbilidades). Desarrollaron cuatro modelos de aprendizaje automático: regresión de Lasso, bosque aleatorio, árbol de decisión potenciado por gradiente y Redes Neuronales Profundas (DNN, por sus siglas en inglés). Establecieron un modelo de regresión logística como referencia, usando los cinco niveles del índice de gravedad de emergencia. Los desenlaces clínicos fueron cuidados críticos (ingreso en UCI o muerte intrahospitalaria) y hospitalización (ingreso hospitalario directo o traslado). En la predicción de resultados de cuidados intensivos, los cuatro modelos de aprendizaje automático superaron al modelo de referencia (AUC, 0,86 [IC 95 %: 0,85–0,87] en la DNN frente a 0,74 [IC 95 %: 0,72–0,75] en el modelo de referencia), con menos pacientes sub-triados en los niveles de triaje 3 a 5 (urgente a no urgente). En la predicción de resultados de hospitalización, todos los modelos de aprendizaje automático superaron al modelo de referencia (AUC, 0,82 [IC 95 %: 0,82–0,83] en la DNN frente a 0,69 [IC 95 %: 0,68–0,69] en el modelo de referencia) con menos pacientes con sobre-triajes en los niveles de triaje 1 a 3 (inmediato a urgente).

En [96] se pretendió predecir el ingreso hospitalario utilizando el historial del paciente además de la información recopilada en el triaje realizado en el servicio de urgencias. Con 972 variables por paciente, se entrenaron nueve clasificadores binarios usando regresión logística (LR), aumento de gradiente (XGBoost) y DNN en tres tipos de conjuntos de datos: uno que usa solo información de clasificación, otro que usa solo el historial del paciente y otro que usa todo el conjunto de variables. Los modelos entrenados con información de clasificación arrojaron un AUC de prueba de 0,87 para LR (IC 95 %: 0,86-0,87), 0,87 para XGBoost (IC 95 %: 0,87-0,88) y 0,87 para DNN (IC 95 %: 0,87-0,88). Los modelos entrenados en el historial del paciente arrojaron un AUC de 0,86 para LR (IC 95 %: 0,86-0,87), 0,87 para XGBoost (IC 95 %: 0,87-0,87) y 0,87 para DNN (IC 95 %: 0,87-0,88). Los modelos entrenados en el conjunto completo de variables arrojaron un AUC de 0,91 para LR (IC 95 %: 0,91-0,91), 0,92 para XGBoost (IC 95 %: 0,92-0,93) y 0,92 para DNN (IC 95 %: 0,92-0,92). Todos los algoritmos alcanzaron el rendimiento máximo al 50 % del

conjunto de entrenamiento o menos, obteniendo los mejores resultados con el conjunto completo de variables.

Cabe mencionar la relación entre los términos de redes neuronales y de ML que se han nombrado indistintamente en la bibliografía consultada. Según [26] Una red neuronal (NN, por sus siglas en inglés) es un método de la IA que enseña a las computadoras a procesar datos de una manera que está inspirada en la forma en que lo hace el cerebro humano. Crea un sistema adaptable que las computadoras utilizan para aprender de sus errores y mejorar continuamente. Se trata de un tipo de proceso de ML llamado aprendizaje profundo, que utiliza los nodos o las neuronas interconectados en una estructura de capas que se parece al cerebro humano. De esta forma, las redes neuronales artificiales intentan resolver problemas complicados, con mayor precisión. Una red neuronal profunda (DNN) es una red neuronal artificial (ANN) con varias capas ocultas entre las capas de entrada y salida. Al igual que en las ANN poco profundas, los DNN pueden modelar relaciones no lineales complejas.

En la revisión bibliográfica que presentan los autores de [24], se llega a una conclusión que se mantiene vigente unos meses después, ya que desde la publicación del artículo, las investigaciones enmarcadas en este período de tiempo no han aportado cambios a lo aquí recomendado: se necesita evidencia adicional y más sólida antes de que ML pueda realmente ayudar al médico de la sala de emergencias a llevar a cabo su práctica clínica diaria. En cuanto al síncope, los esfuerzos de investigación deberán apuntar al establecimiento y compartición de grandes bases de datos clínico-administrativas.

## 2.2. Machine Learning aplicado a la medicina clínica

Con el desarrollo tecnológico y desde hace años, en la medicina se usan los algoritmos informáticos como los “sistemas expertos” que funcionan aplicando un conjunto de reglas sobre un tema determinado para resolver preguntas clínicas específicas. A diferencia de estos, los algoritmos de ML aprenden las reglas a partir de los datos y es que el aprendizaje automático (por la traducción de machine learning) es una disciplina del campo de la Inteligencia Artificial cuyo objetivo es la creación, mediante algoritmos, de modelos capaces de aprender identificando patrones en datos masivos. El desarrollo tradicional de herramientas de riesgo se ha basado en una regresión logística aplicada a datos de investigación recogidos prospectivamente, los que normalmente pertenecen a unos pocos miles de pacientes. Estas técnicas de regresión logística se utilizan comúnmente para derivar Reglas de Decisión Clínica (CDR). En comparación con esas metodologías, la ventaja que ofrecen los métodos modernos de ML es la capacidad de utilizar un mayor número de operaciones matemáticas para definir mejor las relaciones complejas entre los factores de riesgo y los resultados [62], [68], [31],[100],[62] [24].

Los métodos de ML permiten a los ordenadores identificar y aprender patrones complejos en grandes conjuntos de datos y utilizar dichos patrones para hacer predicciones en nuevos casos. En principio, los métodos de ML tienen el potencial de superar los modelos de regresión logística al capturar las relaciones no lineales y las interacciones entre los

predictores. Por otro lado, los modelos ML suelen requerir tamaños de muestra más grandes para la estabilidad, la fiabilidad y el ajuste adecuado debido a su mayor flexibilidad [90],[31].

Comúnmente, los modelos de ML se clasifican como supervisados o no supervisados. En el “aprendizaje supervisado”, se le aplican etiquetas a los datos y con ellas se desarrollan modelos capaces de determinar una relación entre los datos de entrada y la etiqueta asociada. De esta forma, el modelo se aproxima a lo que un médico capacitado podría hacer con gran precisión, además de que podría encontrar relaciones novedosas que no son evidentes para los humanos. En el “aprendizaje no supervisado”, los datos no se etiquetan explícitamente sino que se clasifican por patrones o grupos que ocurren naturalmente. Las posibles aplicaciones de estos modelos tienen que ver con la *medicina de precisión*, en la que los esfuerzos se dirigen a redefinir ciertas enfermedades, de acuerdo con los mecanismos por los cuales se producen sus manifestaciones clínicas, para así proporcionar nuevos caminos a la terapia [97],[77], [24].

Los “algoritmos de aprendizaje supervisado” se suelen utilizar para predecir valores numéricos (identificado en la ciencia de datos como *regresión*) y categóricos (reconocido en el análisis de datos como *clasificación*). Estos algoritmos definen modelos que se pueden interpretar como una relación lineal entre las variables predictoras y la variable de resultado. El objetivo del algoritmo de ML es ajustar, a través del proceso de aprendizaje, el valor óptimo de los parámetros que ponderan a cada variable independiente o predictora, con el fin de acertar lo más posible en el valor que se debe obtener en la variable dependiente o resultado [31].

<b>Supervised Machine Learning</b>	
<b>Continuous (numerical) value prediction</b>	<b>Categorical value prediction</b>
Linear regression Polynomial regression Decision trees (regression) Random forests (regression) Artificial neural networks (regression)	Logistic regression Support Vector Machines Decision trees (classification) Random forests (classification) Naive Bayes Nearest neighbor Artificial neural networks (classification)

Figura 2.2: Los principales modelos de aprendizaje supervisado [24]

En la Figura 2.2, se muestran los métodos comúnmente usados para crear modelos de aprendizaje supervisado, según [24]. Se señalan las principales aplicaciones de estos

modelos, pero todos pueden predecir valores continuos o categóricos dependiendo de las variables de interés.

En [24], también se destacan cinco aplicaciones potenciales de los algoritmos de ML para ayudar a los médicos e investigadores:

- **Detección:** identificación retrospectiva de pacientes con la enfermedad a partir de datos históricos (p. ej., series temporales de datos de dispositivos médicos).
- **Diagnóstico:** identificación de la enfermedad a partir de la información disponible (en particular, signos, síntomas y resultados de las pruebas).
- **Predicción:** predicción de la aparición futura de una enfermedad basada en datos actuales e históricos. Para ésta se puede usar cualquier información disponible.
- **Pronóstico:** predicción de la evolución futura de la enfermedad en base a datos actuales e históricos.
- **Terapia:** identificación de la terapia más adecuada para la enfermedad específica y el paciente; esto está estrechamente relacionado con la necesidad de personalización, particularmente en el contexto de la multimorbilidad <sup>4</sup>.

Con el uso cada vez mayor de Registros de Salud Electrónicos (EHR, por sus siglas en inglés), así como imágenes digitales, actualmente hay una gran cantidad de datos que se pueden incorporar a los modelos de ML para mejorar los procesos de diagnóstico y la estratificación del riesgo del paciente [37],[58],[33],[54].

## Modelos de ML para la clasificación

En esta subsección se señalan algunos de los modelos de ML comúnmente relacionados con la tarea de clasificación; así como se describen, a grandes rasgos, sus principios de funcionamiento, basado en [2]. Estos modelos se agrupan en cuatro grupos que distinguen las características generales de cada método de clasificación: **k-NN**, **Modelos Lineales**, **Modelos basados en Árboles de Decisión** y **Ensamblado de modelos**, este último se comenta como parte de los modelos de árboles de decisión.

### k-NN

El algoritmo *k-vecinos más próximos* memoriza los datos de entrenamiento para al realizar la predicción de un nuevo dato, devolver el resultado de considerar los k datos de entrenamiento más próximos. El algoritmo describe un *aprendizaje no generalizable*, pues su objetivo no es desarrollar un modelo que represente los datos de entrenamiento, ni que permita generalizarlo aplicándolo a nuevas muestras.

---

<sup>4</sup>Multimorbilidad: La presencia de dos o más enfermedades crónicas en un mismo individuo [65]



## Modelos lineales

Los modelos lineales asumen la existencia de una relación lineal entre la variable objetivo y las características predictivas.

$$\hat{y} = w_0 + w_1x_1 + w_2x_2 + \dots + w_px_p$$

Figura 2.3: Ecuación que describe relaciones en los modelos de ML lineales

En la Figura 2.3:  $y$  es la variable de clasificación,  $x$  son las variables predictivas y  $w$  son los coeficientes de regresión.

Algunos tipos de modelos lineales son *regresión lineal* (usado en modelos de regresión), *regresión logística* y *máquina de soporte vectorial*.

El modelo de *Regresión Logística* describe la probabilidad de que la variable objetivo pertenezca a una clase o a otra, dado un cierto valor límite; de forma que si el valor de la variable objetivo iguala o excede a dicho umbral, se devuelve como predicción la clase “positiva”. En otro caso se devuelve la clase “negativa”. La función logística que muestra la Figura 2.4, siempre devuelve un valor entre “0” y “1”, lo que describe este tipo de modelo.

$$f(x) = \frac{1}{1 + e^{-x}}$$

Figura 2.4: Ecuación que describe el modelo de regresión logística

La idea de funcionamiento del modelo de *Máquinas de Vectores de Soporte* (SVM, por sus siglas en inglés) parte de separar un conjunto de puntos pertenecientes dos clases distintas en un plano, de forma que esta separación sea realizada por una recta que sea lo más ideal posible: que divida correctamente los puntos, maximizando lo más posible la distancia mínima entre los puntos de ambas clases. Encontrando una recta así, se puede optimizar la capacidad de generalización del algoritmo. La Figura 2.5 muestra lo expuesto hasta este momento. Si los conjuntos de puntos de ambas clases son separables, los puntos más cercanos a la recta ideal (A y B en la Figura 2.5), además de determinar la posición de la recta, se denominan vectores de soporte.

## Modelos basados en Árboles de Decisión

Estos algoritmos básicamente generan una estructura de árbol de decisión conteniendo preguntas del tipo “if-else” que permiten alcanzar el valor buscado. Se construye un árbol a partir del conjunto de datos de entrenamiento y es posible realizar predicciones haciendo pasar a los nuevos puntos por el árbol, respondiendo a cada una de las preguntas. Estos algoritmos no son sensibles a la escala de los datos, lo que simplifica la fase de pre-procesamiento.

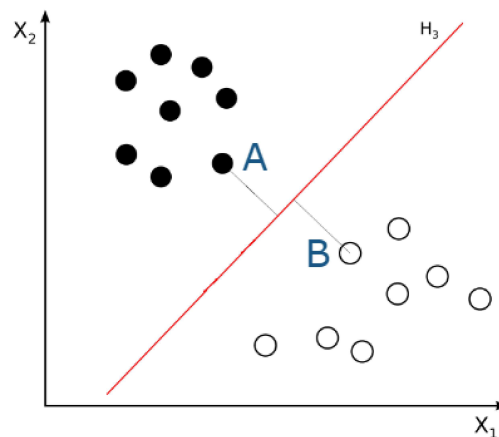


Figura 2.5: División de clases por el modelo de SVM

El modelo de **Árbol de Decisión** (DT por sus siglas en inglés), construye un árbol durante el entrenamiento en el que la raíz constituye un nodo superior del que dependen los consecutivos nodos y de los que se forman las hojas, que serían los nodos que no se pueden seguir dividiendo. En la Figura 2.6 se muestra un ejemplo que parte de un conjunto de datos de dos dimensiones (dos características predictivas) y donde la variable objetivo pertenece a dos clases: “0” y “1”.

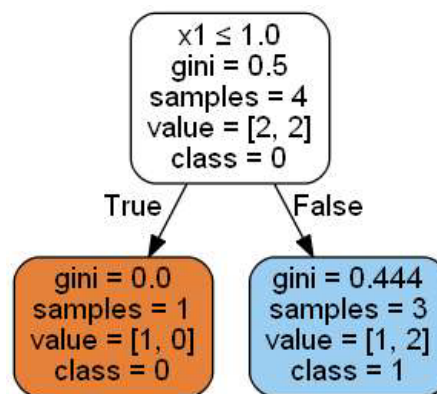


Figura 2.6: Ejemplo de un árbol DT sencillo [2]

Se parte de un espacio con cuatro muestras de las que dos pertenecen a la clase “0” y dos a la clase “1”, lo que significa que el índice de “gini” (criterio utilizado para escoger la característica a la que hacer la pregunta de decisión) es de 0.5, ya que la mitad de los datos está representando a cada clase. El algoritmo busca la característica predictiva al respecto de la que hacer la pregunta y la pregunta en sí, basado en el criterio de minimización del índice de gini ya mencionado, que es una función de impureza<sup>5</sup>. De las dos características

<sup>5</sup>Impureza es una función que evalúa la calidad de la división realizada. Una impureza de 0 significa orden absoluto y una impureza de 1, lo contrario

predictivas, la escogida para hacer la pregunta de decisión es “ $X_1$ ” y la pregunta es: ¿es “ $X_1$ ” menor o igual que 1? La posible respuesta a esta pregunta divide el conjunto de datos en dos bloques más: el que responde positivamente a la pregunta (nodo de color naranja en la Figura 2.6) y el que responde negativamente (nodo de color azul en la Figura 2.6). El primero de los bloques es el subconjunto de datos formado únicamente por la muestra de la clase “0”, lo que implica un índice de gini de 0 y supone que los puntos de este bloque están perfectamente organizados y son de la misma clase. El segundo bloque está formado por tres puntos de ambas clases (uno pertenece a la clase “0” y dos a la clase “1”), lo que supone un índice de gini de 0,444.

El algoritmo de **Random Forest** (traducido como Bosque Aleatorio) se basa en la generación de un cierto número de árboles de decisión siendo entrenado cada uno de ellos con un conjunto aleatorio de muestras. Durante la división de cada nodo, en lugar de considerar todas las características para encontrar el mejor criterio de división, solo se considera un subconjunto de ellas. Cada árbol creado saca el valor medio de las predicciones probabilísticas para votar la clase a asignar. Como consecuencia de la aleatoriedad introducida, el sesgo del modelo aumenta ligeramente pero, gracias a la combinación de los aprendices, decrece la varianza compensando el efecto negativo mencionado, dando como resultado un mejor resultado. También el resultado de este modelo es más difícil de interpretar que el de un único árbol de decisión.

El método de **Extremely Randomized Trees** (traducido como Árboles Extremadamente Aleatorios), lleva la aleatoriedad de Random Forest un paso más adelante. Además de considerar un subconjunto de las características predictivas para cada uno de los árboles a crear (en el paso de escoger una característica y un valor de corte por el que dividir cada nodo) y en lugar de escoger el valor de umbral que mejor divida cada característica (que minimice el criterio de impureza), se genera un valor de corte aleatorio para cada característica planteada, escogiéndose como regla de división el mejor de ellos.

El método de **Gradient Boosting** (conocido también como GRADIENT TREE BOOSTING o GRADIENT BOOSTED REGRESSION TREES) es una familia de algoritmos (conocidos como “meta-algoritmos”) basados en la combinación de modelos predictivos débiles (normalmente se trata de árboles de decisión) para crear un modelo predictivo fuerte. La generación de árboles de decisión débiles se realiza de forma secuencial, creándose cada árbol de forma que corrija los errores del árbol anterior. Los aprendices suelen ser árboles “poco profundos”, de apenas uno, dos o tres niveles de profundidad. Con una tasa de aprendizaje como argumento del modelo, se controla el grado de mejora de un árbol respecto del anterior, de tal forma que una tasa pequeña supone una mejora más lenta pero de mejor adaptación a los datos, con mejoras en los resultados a costa de un mayor consumo de recursos.

El históricamente más conocido algoritmo de boosting es **AdaBoost** y se destacan también los algoritmos **LightBGM** y **XGBoost**, pertenecientes a la familia **Gradient Boosting**. La mayor diferencia entre este último y el primero es que el **Gradient Boosting** entrena cada aprendiz no en los datos de entrenamiento, sino en los errores residuales cometidos por el aprendiz que lo precede.

En el algoritmo **AdaBoost**, todas las muestras reciben un peso inicial que es inverso de la cantidad de muestras con las que cuenta. El primer aprendiz es entrenado y se estima su tasa de error, que para el aprendiz  $j$ -ésimo sería:

$$r_j = \frac{\text{Suma de los pesos de las muestras mal clasificadas}}{\text{Suma de todos los pesos}}$$

Figura 2.7: Ecuación de la tasa de error para el aprendiz  $j$

De forma que si no ha habido muestras mal clasificadas, esta tasa de error sería de valor 0; pero si todas las muestras han sido mal clasificadas, esta tasa sería igual a 1.

$$\alpha_j = \eta \cdot \log \frac{1 - r_j}{r_j}$$

Figura 2.8: Ecuación del peso para el aprendiz  $j$

La Figura 2.8 muestra el peso que recibe el aprendiz y que será tenido en cuenta al final de proceso para estimar una predicción final. Donde  $\eta$  es la llamada *tasa de aprendizaje* que determina la contribución de cada aprendiz y cuya relación con la tasa de error va a depender de la cantidad de muestras. Una vez entrenado el aprendiz y calculada su tasa de error y peso relativo, se actualizan los pesos de las muestras de forma que aquellas para las que la predicción fue incorrecta, vean incrementado su peso tal y como lo muestra la Figura 2.9, o sea que las muestras mal clasificadas multiplican su peso por la exponencial del peso del aprendiz.

$$w_i = \begin{cases} w_i & \text{si la muestra fue bien clasificada} \\ w_i \cdot e^{\alpha_j} & \text{si la muestra no fue bien clasificada} \end{cases}$$

Figura 2.9: Actualización de pesos para el aprendiz  $j$

Una vez actualizados los pesos de las muestras, se normalizan dividiéndolos por la suma total de los pesos. El siguiente aprendiz se ve obligado a concentrarse en aquellas muestras que hayan resultado más difíciles para el aprendiz anterior, que son las que tienen mayor peso. La predicción final es un voto ponderado que depende del peso relativo que haya recibido.

## ML para el manejo del síncope en el departamento de urgencias

El **síncope** es un síntoma común que se encuentra en la práctica clínica y puede manifestarse en un amplio espectro de condiciones que van desde benignas (es decir,

síncope vasovagal) hasta potencialmente mortales (es decir, arritmias sostenidas, infarto agudo de miocardio, embolia pulmonar, disección aórtica). El riesgo de mortalidad a los 7 a 10 días relacionado con el síncope es ligeramente inferior al 1% , mientras que la incidencia de eventos adversos importantes a los 30 días oscila entre el 5% y el 17% [51],[53],[25],[24].

Se estima que el síncope representa del 1 al 3% de todas las visitas al servicio de urgencias y el 6% de todos los ingresos hospitalarios. Una proporción de pacientes, que oscila entre el 12% y el 86% en diferentes países, ingresa en el hospital debido a la incertidumbre diagnóstica en los ED, sin un aumento significativo en el rendimiento diagnóstico a pesar de un alto gasto de recursos económicos [51],[27],[71],[24].

De acuerdo con las pautas internacionales actuales sobre el manejo del síncope, la decisión de hospitalización se basa principalmente en la gravedad de la enfermedad subyacente o la presencia de características de alto riesgo identificadas durante la evaluación inicial en el servicio de urgencias. Por otro lado, las admisiones inapropiadas en pacientes de bajo riesgo pueden aumentar los riesgos relacionados con la hospitalización (incluidas las infecciones adquiridas en el hospital y los errores relacionados con la medicación) y los costos. Para los pacientes considerados de riesgo intermedio, recientemente se propuso el manejo en la *Unidad de Síncope* (US) como alternativa a la hospitalización. Esta unidad puede estar ubicada en el entorno de pacientes hospitalizados o ambulatorios, con derivaciones provenientes del servicio de urgencias o de médicos/cardiólogos de la comunidad. Dos ensayos clínicos aleatorizados evaluaron la US en el servicio de urgencias en comparación con la atención habitual. Demostraron un mayor rendimiento diagnóstico, menores ingresos hospitalarios, costos reducidos y ningún aumento en los resultados adversos de los pacientes asignados al azar a la SU [51],[59],[45],[67],[95],[24].

Por lo tanto, si la causalidad del síncope no puede determinarse durante la evaluación inicial en el servicio de urgencias, una estratificación precisa del riesgo es crucial para garantizar una disposición adecuada en el ED y optimizar la vía diagnóstica-terapéutica de los pacientes[24].

### **Estratificación del riesgo de síncope en el departamento de urgencias**

Recientemente se desarrolló el *Canadian Syncope Risk Score* y se validó externamente en una gran población de pacientes canadienses, mostrando una buena discriminación y calibración para el riesgo de 30 días de eventos adversos graves después de la disposición del servicio de urgencias. Sin embargo, debe validarse en diferentes entornos antes de recomendar su implementación en la práctica clínica [88],[89],[24].

La categorización pronóstica correcta de los pacientes con síncope puede ser un desafío debido a la gran cantidad de posibles determinantes de riesgo y las limitaciones de tiempo del médico de urgencias. Dado que la derivación de pacientes a una US puede no ser factible en todos los entornos de urgencias, podría ser útil tener un sistema basado en reglas que aplique de manera confiable las pautas clínicas a los datos disponibles para evaluar el riesgo del paciente. Además, dado que la mayor parte de la información pronóstica compleja solo

está disponible en forma de texto, un desafío específico es aplicar redes neuronales y ML a EHR para extraer los fenotipos relevantes [24].

En la búsqueda exhaustiva que se presenta en [24], se concluye que es concebible que la implementación de algoritmos de pronóstico pueda conducir a la mejora futura de la clasificación del paciente en el servicio de urgencias en condiciones clínicas como las mostradas, a pesar de que esta clasificación del ED puede ser imprecisa para identificar a los pacientes con síncope de alto riesgo. Sin embargo, en la actualidad, solo unos pocos estudios han analizado la aplicación de ML para la detección y predicción del riesgo de síncope, a pesar de los resultados preliminares alentadores.

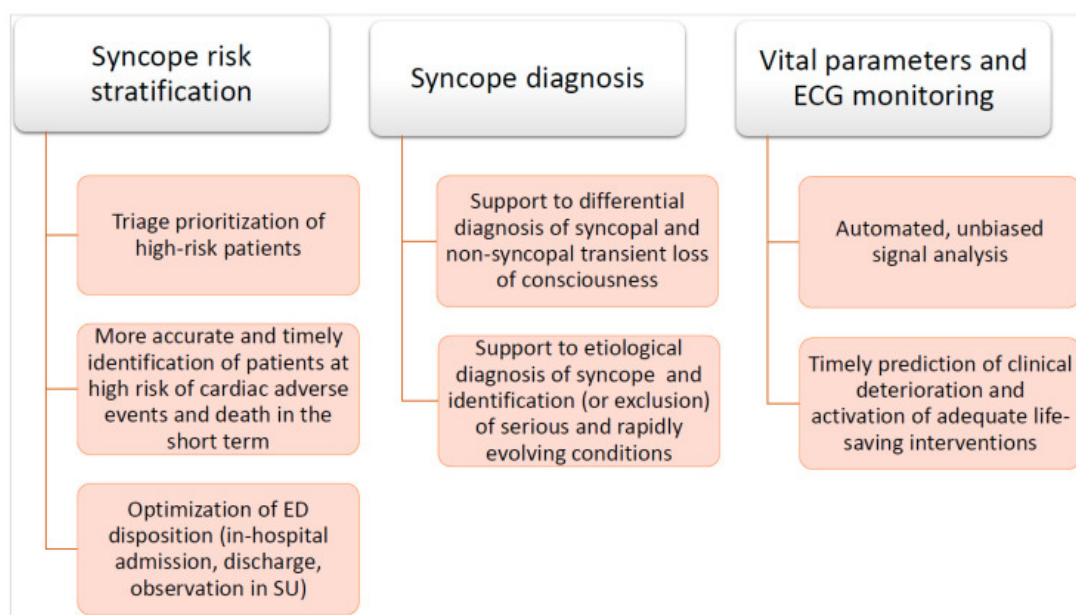


Figura 2.10: Posibles aplicaciones de los modelos de ML [24]

La Figura 2.10 resume las principales perspectivas de futuro para la aplicación de algoritmos de ML en el manejo de pacientes con síncope en urgencias, tomada de [24].

### 2.3. Acerca del síncope

El síncope se define como la Pérdida Transitoria de la Conciencia (TLOC, del inglés Transient Loss of Consciousness) debido a la hipoperfusión<sup>6</sup> cerebral, que se caracteriza por un inicio rápido, una duración corta y una recuperación completamente espontánea. Un diagnóstico correcto de síncope requiere una comprensión de las claves clínicas, la fisiopatología<sup>7</sup> y los patrones epidemiológicos subyacentes, además de un esfuerzo cooperativo multidisciplinario. De hecho, el síncope es solo una de las muchas causas de la TLOC y,

<sup>6</sup>Hipoperfusión: Disminución del flujo de sangre que pasa por un órgano [19].

<sup>7</sup>Fisiopatología: Aspecto del estudio de las enfermedades que se centra en sus posibles alteraciones [19]

antes de que se pueda determinar la causa, es crucial determinar si realmente se produjo el síncope o si el período de inconsciencia se debió a una de las muchas otras condiciones dentro de la espectro más amplio de TLOC como: epilepsia, traumatismo craneoencefálico o ataques psicógenos [91],[14].

La presencia de síncope conlleva la posibilidad de resultados muy graves, incluidas complicaciones cardiovasculares y mortalidad [101],[69],[48] , esta última típicamente como consecuencia de una enfermedad cardíaca. Por esta razón, se requiere un enfoque de estratificación del riesgo cuando y donde sea que se encuentre. El síncope se presenta en toda la profesión médica desde el médico general hasta la medicina interna (cardiología, neurología, geriatría y psiquiatría) y, también, a la ortopedia a través de los departamentos de emergencia.

Los médicos que atienden a pacientes con sospecha de síncope deben poder realizar la investigación inicial y realizar la estratificación del riesgo, si corresponde. Los documentos de las guías de la European Society of Cardiology (ESC) y Heart Rhythm Society (HRS) [14],[74] han abordado el manejo del paciente, y la ESC define la evaluación inicial como antecedentes, examen físico, presión arterial (PA) acostado y de pie como esencial [91]. Es probable que las presentaciones más dramáticas se realicen en el servicio de urgencias, lo que hace que se centre la atención en esta área. No obstante, el enfoque ED tiene pertinencia para todos los demás lugares de presentación. Los primeros intentos de distinguir entre los pacientes que tienen síncope cardiovascular y otros que tienen síncope reflejo es un lugar apropiado para comenzar. En este sentido se ha trabajado desde 1990, tal y como se muestra en [85], donde tanto los esfuerzos europeos como los norteamericanos fueron diligentes, pero se dividieron entre los de los médicos de urgencias (cuyo énfasis está en hacer un diagnóstico rápido y preciso para colocar al paciente en el mejor entorno de atención disponible, ya sea como paciente hospitalizado o ambulatorio), y los cardiólogos, que luego de esta primera mirada en el ED, buscan hacer un diagnóstico completo y tratar al paciente. Estos dos enfoques son potencialmente compatibles, pero en la práctica no son tan fáciles de coordinar. El resultado fue que ninguna de las reglas y puntajes de evaluación del síncope fue satisfactoria, y algunas fallaron cuando se pusieron en práctica.

Se está logrando cierto consenso, sobre todo en Canadá con un sistema de puntuación de riesgo (CRSS) [73],[89] y un breve ejercicio de finalización de formularios de Basilea, Suiza [21]. Ambos enfoques han demostrado eficacia en grandes estudios basados en ED [84]. Cuando se aplica, cualquiera de estos métodos debería acelerar el tiempo de evaluación y hacerlo mucho más preciso. En particular, en el caso de Basilea, se incluyen análisis de sangre (troponina o péptido natriurético cerebral), que hoy en día están bastante disponibles y aumentan la precisión [89],[21],[84].

La herramientas de predicción de síncope usan varias combinaciones de diferentes parámetros clínicos derivados de ECG, historial del paciente, examen físico y biomarcadores<sup>8</sup> [70]

---

<sup>8</sup>Biomarcador: cualquier sustancia, estructura o proceso que pueda ser medido en el cuerpo o que sus productos e influencias puedan predecir el resultado de una enfermedad [86]

La Figura 2.11 muestra los predictores clínicos establecidos por el American College of Cardiology (ACC), American Heart Association (AHA) y la guía correspondiente creada por HRS para el manejo de pacientes con síncope, teniendo en cuenta los factores de riesgo en menos de 30 días y en más de 30 días. Los predictores de más alto riesgo, se basan en la historia del paciente y se relacionan con la presencia de enfermedades cardíacas [70].

Risk factors	History: Outpatient clinic or ED evaluation	Physical examination, ECG, laboratory investigations
Short-term risk factors $\leq 30$ days	<ul style="list-style-type: none"> <li>• Male sex</li> <li>• Older age</li> <li>• No prodrome</li> <li>• Palpitations preceding loss of consciousness</li> <li>• Exertional syncope</li> <li>• Structural heart disease</li> <li>• Heart failure</li> <li>• Cerebrovascular disease</li> <li>• Family history of SCD</li> <li>• Trauma</li> </ul>	<ul style="list-style-type: none"> <li>• Evidence of bleeding</li> <li>• Abnormal ECG</li> <li>• Persistent abnormal vital signs</li> <li>• Positive troponin/BNP</li> <li>• Electrolyte derangement</li> </ul>
Long-term risk factors $>30$ days	<ul style="list-style-type: none"> <li>• Male sex</li> <li>• Older age</li> <li>• Absence of nausea/vomiting preceding syncopal event</li> <li>• Ventricular arrhythmias</li> <li>• Cancer</li> <li>• Structural heart disease</li> <li>• Heart failure</li> <li>• Cerebrovascular disease</li> <li>• Diabetes mellitus</li> <li>• High CHADS-2 score</li> </ul>	<ul style="list-style-type: none"> <li>• Abnormal ECG</li> <li>• Lower GFR</li> </ul>

Figura 2.11: Factores de riesgo en el pronóstico del síncope [70]

Otra de las herramientas de predicción utilizadas en este ambiente médico es el Canadian Syncope Risk Score (CSRS), que es un sistema obtenido desde el servicio de urgencias y del que se obtienen valores positivos cuando un dato clínico indica una condición seria y valores negativos en caso contrario, de esta forma se predice el riesgo de presentar síncope. En la Figura 2.12, se muestran los límites de puntuación de este sistema del que se suele obtener una predicción con alto grado de de discriminación, calibración y exactitud [70].

## 2.4. Conclusión del capítulo

Como resultado de la búsqueda bibliográfica, se encontraron artículos con explicaciones de funcionamiento de dispositivos valorados con posibilidad de brindar servicios médicos y ninguno era para el manejo del síncope en el cuerpo de urgencias. Entre los artículos relacionados con el uso de ML, se encontraron muy pocos estudios que mencionan, sin



Category	Risk factors	Points	Total score (-3 to 11)	Risk
Clinical assessment	Predisposition to VVS <sup>a</sup>	-1	-3	Very low
		1	-2	
	Prevalent CVD SBP < 90 or >180 mm Hg	2	-1 0	Low
Investigations (troponin, ECG)	Elevated troponin	2	1	Medium
	QRS axis $\leq 30^\circ$ or $> 100^\circ$	1	2	
	QRS duration > 130 ms	1	3	
	QTc interval > 480 ms	2		
ED diagnosis	Vasovagal syncope	-2	4	High
	Cardiac syncope	2	5 $\geq 6$	Very high

Figura 2.12: Puntaje Canadiense de Riesgo del Síncope [70]

detalles de procesos de creación, el manejo del paciente con síncope en el ED. En estos últimos, por lo general se usaron más tipos de datos de los que contamos en este estudio (como las comorbilidades) o menos (solo algunas variables de las tantas que tenemos). La mayoría de las investigaciones sobre el uso de ML en relación con la enfermedad del síncope, esta no es más que una enfermedad asociada, cuyos indicadores conforman un predictor más.

Se señala que en todas las regiones del mundo no es igual el resultado a obtener para un mismo modelo, ya que las costumbres alimenticias y el clima pueden influir en hábitos y adaptaciones corporales del ser humano y, consecuentemente, en las fluctuaciones de variables médicas y en los resultados de los análisis.

Los estudios anteriores no muestran claramente la metodología seguida desde la selección de los datos, hasta la selección del modelo, pero sí coinciden en usar modelos supervisados de clasificación para resultados afines. Es tendencia en estos estudios vinculados a la medicina adoptar la métrica de AUC para evaluar los modelos creados, en base a la probabilidad de que este distinga entre una clase y otra.

---

## 3: Desarrollo del Proyecto

---

A partir de las conclusiones sacadas del capítulo anterior, este apartado contiene la experiencia que justifica los caminos tomados en la resolución del problema planteado. En las secciones de este capítulo se presenta, con un nivel de detalle, la base de datos proporcionada, luego se muestran las herramientas usadas y los métodos específicos para desarrollar el trabajo, después de lo cual se describe la metodología experimental seguida.

### 3.1. Descripción del conjunto de datos

En esta sección se da una explicación detallada sobre el origen del conjunto de datos, su descripción general y, a partir de las transformaciones realizadas por indicación del médico responsable, se presentan los campos de la base de datos original con los que se realiza el estudio.

#### Origen de los datos

El equipo médico que colabora en esta ocasión con el grupo de investigación GTe, lo hace desde 2 años atrás planteando posibles enfoques de investigación para la presente base de datos. Uno de los médicos ha sido el encargado de registrar la información disponible en el cuerpo de urgencias y de elaborar tanto la base de datos con la que actualmente se trabaja, como la idea de realización de esta investigación. Los datos han sido presentados en formato excel, y no contienen información confidencial, habiendo sido totalmente anonimizados. Toda la información médica registrada corresponde a tres cohortes <sup>9</sup> de pacientes atendidos en el departamento de urgencias de los centros de salud: **Hospital Clínico Universitario de Valladolid**, **Hospital Universitario Río Hortega de Valladolid**, **Hospital General de Segovia**, **Complejo Asistencial Universitario de Salamanca** y **Complejo Asistencial Universitario de Burgos**; en fechas comprendidas entre febrero de 2018 y julio de 2021. Asimismo, los *Comités Éticos de Investigación Clínica*

---

<sup>9</sup>Cohorte: Término común en ambientes médicos para aludir a conjuntos de pacientes.

de las respectivas instituciones sanitarias, han autorizado el acceso a estos datos para el estudio actual.

## Descripción general

La base de datos está conformada en su totalidad por 573 registros de pacientes, de los que la descripción de su paso por el departamento de urgencias se presenta en 85 variables. Estas variables se desglosan en comorbilidades, mediciones fisiológicas y analíticas realizadas a los pacientes, descripciones médicas de cada caso particular, el manejo que se resuelve dar a cada paciente para desplazarlo del ED, la cantidad de días que convive el paciente en los espacios sanitarios y el fallecimiento de algunas personas en las instalaciones, además de datos personales como la edad y el género de cada persona.

## Presentación de las variables

En la primera reunión con el médico responsable del estudio, se desestimaron algunas variables calificadas como no portadoras de información para el estudio actual, según la valoración médica. Estas 35 variables indicaban: fechas, identificador de cohorte, los días de hospitalización, las descripciones realizadas por el médico que atiende en el ED, las comorbilidades y algunas mediciones médicas. En la segunda reunión, se desecharon 24 variables más por no presentar al menos el 50 % de los valores y, en estos casos, el médico no disponía de los registros correspondientes como para completar la base de datos ni las explicaciones coherentes para justificar estos espacios. Finalmente, la base de datos que se presenta para la descripción de los campos, lo hace con 26 variables introducidas a continuación.

**Edad** Variable numérica con valores enteros que muestra la edad de los pacientes registrados.

**Sexo** Variable categórica de naturaleza nominal que indica el género de las personas registradas:

- 1 = Masculino
- 2 = Femenino

**Crea** Variable numérica con valores decimales que contiene el nivel de creatinina en sangre dado en mg/dL (miligramos de creatinina por cada decilitro de sangre [57]. Cuando el paciente es fememino, se puede considerar un valor de creatinina normal entre 0.6 mg/dL y 1.1 mg/dL; mientras que cuando es masculino, se mueve el rango a valores entre 0.7 mg/dL y 1.3 mg/dL.

**Glu** Variable numérica con valores enteros que presenta los niveles de glucosa en sangre, estos niveles se expresan en mg/dL [56]. En general, menos de 100 mg/dL (5,6 mmol/L) se considera normal.

**Leuco** Variable numérica con valores decimales que presenta el recuento de glóbulos blancos en un microlitro de sangre [56]. Los valores se pueden relacionar con el sexo y de la edad, pero por lo general, deben haber entre 4,5 y 11 leucocitos por mL de sangre (4.500 y 11.000 leucocitos por microlitro ) en adultos.

**Hct** Variable numérica con valores enteros y decimales, que indica qué porcentaje de la sangre son glóbulos rojos (Hematocritos)[57]. Para esta variable se suelen presentar los valores normales con relatividad al género del paciente: para la mujer este valor normal debe estar entre 35 % y 45 % y en el caso de los hombres, normalmente es entre 40 % y 50 %.

**ST** Variable categórica nominal que presenta la forma que tiene la porción de ciclo cardíaco intervalo ST en el ECG. Los posibles valores de la variable son:

- 1 = Normal
- 2 = Elevación
- 3 = Descenso
- 4 = T picudas
- 5 = T negativas
- 6 = Onda Q
- 7 = Otros

**Ritmo** Variable categórica nominal que presenta los tipos de trastornos en el latido del corazón. Los cuales son:

- 1 = Sinusal
- 2 = Arritmia sinusal
- 3 = Fibrilación auricular
- 4 = Flutter auricular
- 5 = Taquicardia auricular
- 6 = Taquicardia supraventricular
- 7 = Taquicardia ventricular
- 8 = Bradicardia sinusal
- 9 = Bloqueo AV de primer grado
- 10 = Bloqueo AV de segundo grado tipo I
- 11 = Bloqueo AV de segundo grado tipo II
- 12 = Bloqueo AV completo
- 13 = Ritmo de marcapasos
- 14 = Ritmo de la unión

- 15 = Ritmo idio-ventricular
- 16 = Bloqueo de rama derecha
- 17 = Bloqueo de rama izquierda
- 18 = Extrasístoles ventriculares
- 19 = Asistolia
- 20 = Fibrilación ventricular
- 21 = Otros

**GCS.M-H** Variable categórica ordinal que muestra la puntuación de la respuesta motora del paciente, según la Escala de coma de Glasgow [23].

- 1 = Ninguna respuesta
- 2 = Extensión
- 3 = Flexión anormal
- 4 = Retirada y flexión
- 5 = Localiza el dolor
- 6 = Orden verbal obedece

**GCS.V-H** Variable que categórica ordinal que muestra la puntuación de la respuesta verbal del paciente, según la Escala de coma de Glasgow [23].

- 1 = Ninguna respuesta
- 2 = Sonidos incomprensibles
- 3 = Palabras inapropiadas
- 4 = Desorientado y hablando
- 5 = Orientado y conversando

**GCS.O-H** Variable que categórica ordinal que muestra la puntuación de la respuesta ocular del paciente, según la Escala de coma de Glasgow [23]

- 1 = No responde
- 2 = Dolor
- 3 = Orden Verbal
- 4 = Espontánea

**FC-H** Variable numérica con valores enteros, con los que establece los latidos por minuto (lpm), en estado de reposo, o sea, la frecuencia cardíaca del paciente [56]. El valor considerado normal en este parámetro debe estar entre 60 lpm y 100 lpm.

**TAD-H** Variable numérica con valores enteros que muestran la tensión arterial que se registra cuando el corazón está en relajación (Tensión arterial diastólica). La unidad de medida es milímetros de mercurio (mmHg) [49]. La figura 3.13 muestra los valores comprendidos como normales para esta variable, los que también pueden limitarse por edad y por género de las personas.

**TAS-H** Variable numérica con valores enteros que muestran la tensión arterial que se registra cuando el corazón está en contracción (Tensión arterial sistólica). La unidad de medida es milímetros de mercurio (mmHg) [49]. La figura 3.13 muestra los valores comprendidos como normales para esta variable, los que también pueden limitarse por edad y por género de las personas.

Categoría	Sistólica (mmHg)		Diastólica (mmHg)
Hipotensión	menor de 80	o	menor de 60
Normal	80-120	y	60-80
Prehipertensión	120-139	o	80-89
Hipertensión grado 1 (HTA 1)	140-159	o	90-99
Hipertensión grado 2 (HTA 2)	160 o superior	o	100 o superior
Crisis hipertensiva (emergencia médica)	superior a 180	o	superior a 110

Fuente: American Heart Association

Figura 3.13: Tabla de valores normales de las componentes de tensión arterial [16]

**Triaje** Variable categórica ordinal cuyos valores presentan la valoración que se hace del grado de gravedad del paciente a su llegada al departamento de urgencias. Los valores establecidos están comprendidos en los niveles siguientes [80]:

- 1 = Prioridad absoluta con atención inmediata y sin demora.
- 2 = Situaciones muy urgentes de riesgo vital, inestabilidad o dolor muy intenso. Demora de asistencia médica hasta 15 minutos.
- 3 = Urgente pero estable hemodinámicamente con potencial riesgo vital que probablemente exige pruebas diagnósticas y/o terapéuticas. Demora máxima de 60 minutos.
- 4 = Urgencia menor, potencialmente sin riesgo vital para el paciente. Demora máxima de 120 minutos.

- 5 = No urgencia. Poca complejidad en la patología o cuestiones administrativas, citaciones, etc. Demora de hasta 240 minutos.

**FR-H** Variable numérica que representa la frecuencia de respiración de cada paciente por minuto. Su valor normal suele ser de entre 12 y 20 respiraciones por minuto.

**SpO2-H** Variable numérica con porcentajes enteros que indica la saturación de oxígeno y mide cuánta hemoglobina hay en la sangre [57]. Los valores normales suelen estar entre el 95 % y el 100 %.

**FiO2-H** Variable numérica de valores decimales que representa la concentración de oxígeno en la mezcla de aire inspirado. La notación se da en mL de oxígeno por cada mL de volumen de aire corriente de un paciente [57]. El aire natural incluye un 21 % de oxígeno, lo que equivale a una fracción inspirada de oxígeno de 0,21; mientras el aire enriquecido con oxígeno puede contener valores superiores hasta 1,00.

**TT-H** Variable numérica que indica el tiempo de trombina, o sea, el tiempo que tarda la porción líquida de la sangre (plasma) en coagularse [57], [35]. Debe presentar valores de 14 a 19 segundos para considerarse leve o ligeramente elevado.

**Plaquetas** Variable numérica con valores enteros cuya medida es la cantidad media de plaquetas que tiene una persona por microlitro [92]. Se consideran valores normales los comprendidos entre 150 uL y 450 uL.

**Trop.** Variable numérica que muestra el resultado de evaluar la cantidad de las proteínas troponinas en sangre. Al corroborar los valores con los términos normalmente manejados, se puede concluir que se da este valor en microgramo por centilitro (ug/cL).

Las variables que siguen son las seleccionadas como variables dependientes para establecer las tres pronósticos que solicita el médico responsable de los datos. Así como este transmitió su interés con el estudio, propuso las variables y las relaciones ellas. El pronóstico temprano de que el paciente ingrese, lo representa la variable *Hospital*, la variable *Mort. Hospi* se hace corresponder con el pronóstico del fallecimiento del paciente en las instalaciones hospitalarias durante el período de atención y las variables *UVI*, *Coronarias* e *ICTUS*, tienen relación directa con la posible predicción del ingreso en Cuidados Intensivos, por lo que conformarán una variable más representativa de este propósito.

**Hospital** Variable categórica binaria que indica si el paciente fue hospitalizado (“1”) o si no lo fue (“0”).

**Mort. Hospi** Variable categórica binaria que indica si el paciente falleció en los días posteriores a su llegada al ED (“1”) o si no (“0”).

**UVI** Variable categórica binaria que indica si el paciente fue llevado a atender a la Unidad de Vigilancia Intensiva (“1”) o si no lo fue (“0”).

**Coronarias** Variable categórica binaria que indica si el paciente fue llevado a atender a la Unidad de Coronarias (“1”) o si no lo fue (“0”).

**ICTUS** Variable categórica binaria que indica si el paciente fue llevado a atender a la Unidad de daños cerebrovascular ICTUS (“1”) o si no lo fue (“0”).

## Observaciones sobre los datos

Todos los pacientes que registra esta base de datos tuvieron algún episodio de síncope y llegaron en ambulancia al servicio de urgencias, precisamente la recopilación de datos comenzó en el traslado hacia el ED. Esto se puede justificar con los valores de la variable *Triaje*, la que siempre muestra tres de los cinco posibles clasificaciones inherentes a sí: las indicadoras de mayor estado de gravedad; pues los pacientes que pudieran clasificarse con 4 y 5, habrían sido capaces de llegar al departamento de urgencias andando o por sus medios.

La base de datos actual es la primera realizada por el médico encargado, él mismo reconoció la posibilidad de arrastrar errores en las mediciones y errores en el traspaso de los datos desde el registro analógico de pacientes. Al analizar las relaciones entre ciertos datos, se detectaron incompatibilidades como es el caso del registro 492, que presenta ritmo sinusoidal y una FC (frecuencia cardíaca) de 44 lpm (latidos por minuto), mientras en varias bibliografías consultadas sobre el tema (una de ellas es [18]), aparece que el ritmo sinusoidal equivale a una respiración normal y sus valores de FC se dan entre 60 y 100 lpm. Estos ruidos en los datos predisponen a un procesamiento no satisfactorio y a la creación de modelos no tan buenos como se podría esperar.

## 3.2. Técnicas y herramientas

Las fases primeras del trabajo, que se recogen en los capítulos anteriores, han sido el resultado de aplicar [metodología Prisma](#) a la revisión bibliográfica.

Tanto el trabajo de exploración y preparación de los datos como el de creación y evaluación de modelos, está registrado en el *Jupyter Notebook* de la plataforma *Anaconda*. Algunas pruebas se realizaron con *Google Colab*. Se usó el lenguaje de programación *Python* en su versión 3.8.13, la que ha sido instalada en un [entorno virtual](#) que se creó para usar la [biblioteca \*pyCaret\*](#) y todas sus dependencias. La elección de otras bibliotecas y funciones se ajustó a este entorno.

Se usaron bibliotecas estándar y populares en el análisis exploratorio de los datos y la búsqueda y creación de modelos; así como otras desconocidas hasta el proceso de elaboración de este trabajo. Se presentan a continuación.

***Pandas*** Es una biblioteca de código abierto de Python que proporciona herramientas de análisis y manipulación de datos de alto rendimiento utilizando sus potentes estructuras de datos. El nombre de *Pandas* se deriva del término “Panel Data” y es



la biblioteca de análisis de datos de Python. Usando esta biblioteca se pueden lograr cinco pasos típicos en el procesamiento y análisis de datos, independientemente del origen de los datos: cargar, preparar, manipular, modelar y analizar [6].

***Numpy*** y ***Tabulate*** Poseen funciones matemáticas de alto nivel. La primera constituye la biblioteca de matriz n-dimensional base y, significando “Numerical Python”, proporciona potentes estructuras de datos, implementando matrices multidimensionales [7]. Estas estructuras de datos garantizan cálculos eficientes con matrices.

***Matplotlib*** y ***Seaborn*** La primera es una biblioteca de trazado conocida por la capacidad de generación de gráficos interactivos y el trazado completo 2D, mientras que la segunda es una biblioteca gráfica para la visualización de datos estadísticos con la fácil generación de elegantes gráficos [8][72].

***Sklearn*** Esta biblioteca es comúnmente usada para la transformación y normalización de datos, así como para la división de estos en los conjuntos que permitirán la creación y validación del modelo. Es de código abierto, está construida sobre SciPy (Scientific Python) y proporciona una gama de algoritmos de aprendizaje supervisados y no supervisados en Python [9]. Para el presente estudio, fue utilizada en una etapa anterior.

***Imblearn*** Imbalanced-learn (importado como imblearn) es una biblioteca de código abierto con licencia del MIT que se basa en scikit-learn (importado como sklearn) y proporciona herramientas cuando se trata de la clasificación con clases desequilibradas [39]. Para el presente estudio, fue utilizada en una etapa anterior.

***kNN*** La biblioteca kNN obtiene sus siglas de k-Nearest Neighbors (traducido como k-vecinos más cercanos). Aunque kNN es un algoritmo de aprendizaje automático supervisado, cuando se trata de detección de anomalías, adopta un enfoque no supervisado. Esto se debe a que no hay un “aprendizaje” real involucrado en el proceso y no hay un etiquetado predeterminado de “valor atípico” o “no valor atípico” en el conjunto de datos, sino que se basa completamente en valores de umbral. La suposición fundamental en la familia del vecino más cercano es que las observaciones similares están próximas entre sí y los valores atípicos suelen ser observaciones solitarias, que se mantienen más lejos del grupo de observaciones similares [5].

***Autoviz*** es una biblioteca de código abierto mantenida por *AutoviML*, que toma datos como entrada y encuentra la estructura básica e identifica características impactantes para trazar hermosas visualizaciones en una sola línea de código [42].

***PyCaret*** Es una biblioteca de código abierto y código bajo en Python y, además, es una herramienta integral de gestión de modelos y aprendizaje automático que acelera, de forma productiva, el ciclo de experimentación [66]. Se auxilia de casi todas las bibliotecas anteriores para seguir el proceso el desarrollo. Es la biblioteca utilizada en este estudio para en la selección y creación de modelos, así como en parte de la exploración de las variables.

A raíz de las conclusiones sacadas del capítulo de búsqueda bibliográfica (capítulo anterior) y a partir de la tenencia de variables de salida binaria, en todos los casos de decisión médica a pronosticar, se establece la exploración de modelos supervisados de clasificación binaria.

## Marco Conceptual

En este espacio se concentran algunos conceptos inherentes al entendimiento de las herramientas y tecnologías utilizadas para el desarrollo del proyecto.

*Metodología PRISMA:* Es una metodología de trabajo que permite documentar de manera transparente las revisiones sistemáticas de los artículos con bases como: el porqué de la búsqueda bibliográfica, qué se hizo y qué se encontró [64].

*Entorno Virtual:* Es un árbol de directorios que contiene archivos ejecutables de Python y otros archivos que indican que es un entorno particular respecto a otros. El intérprete Python, las bibliotecas y los scripts instalados en un entorno virtual están aislados de los instalados en otros entornos virtuales[1].

*Biblioteca:* Es un conjunto de módulos incorporados que brindan acceso a funcionalidades establecidas en directorios y que ayuda a crear una interfaz independiente.

*Función de Python:* Una serie de sentencias que retornan un valor al que las llama. También se le puede pasar cero o más argumentos los cuales pueden ser usados en la ejecución de la misma [1].

*Módulo de Python* Un objeto que sirve como unidad de organización del código Python. Los módulos tienen espacios de nombres conteniendo objetos Python arbitrarios. Los módulos son cargados en Python por el proceso de importing [1]. Es un fichero conteniendo definiciones y declaraciones de Python.

*Paquete de Python* Un módulo Python que puede contener submódulos o recursivamente, subpaquetes. Técnicamente, un paquete es un módulo Python con un atributo “\_path\_” [1]. Colección de módulos que se pueden importar, inclusive individualmente, para el manejo uniforme de archivos y datos.

### 3.3. Selección y construcción del modelo

En esta sección se recorre todo el proceso de creación de modelos finales para caso de clasificación, desde la exploración de los datos, hasta la creación del modelo seleccionado, pasando por la limpieza y transformación de la base de datos, según se ha planificado la metodología a seguir.

## Análisis exploratorio de datos

En este apartado se pasa a explorar el dataset para sacar posibles consideraciones que mejoren los resultados de las etapas posteriores. Esta tarea permitirá, además, encontrar outliers y ruido en los diversos campos. Anteriormente se mencionaba la existencia de valores ausentes que afectaba demasiado a la base de datos, correspondientes a algunas de las mediciones fisiológicas en el departamento de urgencias, y que estos fueron detectados desde una primera mirada a la base de datos en las entrevistas con el médico responsable de estos datos, el que sugirió desechar estas variables del estudio actual por no poder completarlas en el tiempo inminente. Precisamente, con las variables originales de conjunto de datos, el número total de registros con valores perdidos sería 572 (de 573 registros comprendidos) representando el 99,8 % de los del dataset completo. Luego de seleccionar las variables que participarían en el estudio, solo se registró el 26 % de la base de datos con al menos un valor ausente a lo largo de cada registro.

Ya se ha mencionado que la base de datos está contaminada con ruido. Después de varias exploraciones, se ha detectado la incompatibilidad de valores entre diferentes campos, lo que indica para un mismo registro, que algún campo está fuera del rango considerado como posible a pesar del valor de otro campo. Este tipo de ruido es complicado de detectar pero aún si se detectara, es complejo de eliminar. La detección de este ruido, que podríamos llamar *incompatibilidad*, se basaría en conocimientos médicos, que se han rozado en esta investigación gracias al acercamiento a bibliografía básica como páginas de aclaración de dudas a usuario, diccionarios médicos, lecciones públicas de médicos y artículos científicos, algunos ejemplos de esto son [19],[57], [56] y otros que son referenciados, en este mismo capítulo y en todo el documento. El otro problema con este tipo de ruido es que, de ser identificado, implicaría la relación de al menos dos variables, lo que implica, a pesar del tiempo invertido, que está fuera de las capacidades técnicas y tecnológicas establecer una base de decisión y determinar qué variable es correcta y con qué valor se podría sustituir la incorrecta. Concluyendo, este tipo de ruido supondría, para un analista de datos, mucho más tiempo de investigación y de revisión de cada registro para llegar a construir una solución aceptable.

En el desarrollo de este apartado, y habiendo visualizado la base de datos de forma general, se hace un análisis descriptivo y más a lo específico con cada variable y registro, para comparar comportamientos y describir relaciones entre variables, las que puedan tender o no a favorecer a alguna de las clases de las variables de salida. Extrayendo la información que registra el dataframe de la base de datos original, se encuentra que la base de datos tiene 8 920 entradas, de las que las 8 347 que representan la diferencia con respecto a los registros identificados desde un inicio (573), solo tienen valor en una variable y es el 0.

Tras acotar la base de datos con las consideraciones derivadas de las entrevistas al médico, se procede a realizar el análisis correspondiente, para el que se cuentan 573 registros con 26 columnas de las que 5 son identificadas como atributos objetivo.

## Análisis univariable

Para la exploración estadística, se estudian primero las variables dependientes y su distribución de clases, luego las variables categóricas y por último, las numéricas.

### VARIABLES DE SALIDA

Nombre Hospital

**Tipo de dato** Categórico Binario

**Valores faltantes** No tiene

**Ruido y Visualización** No tiene valores atípicos

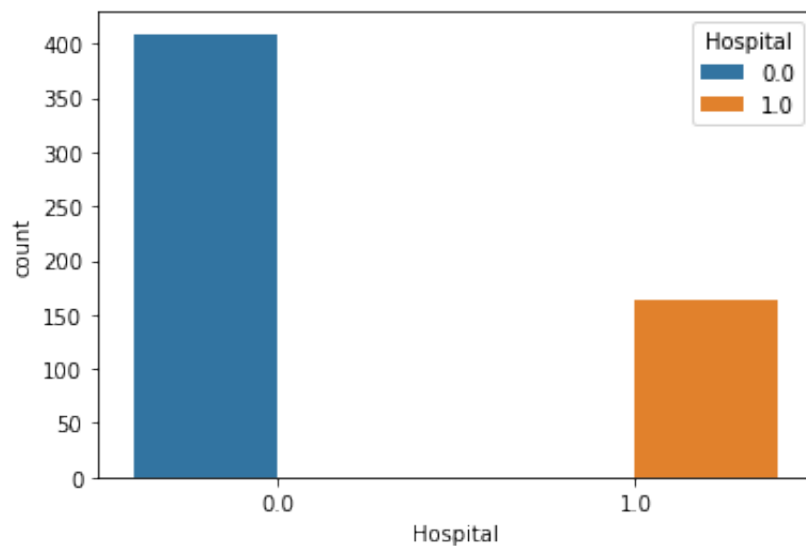


Figura 3.14: Visualización de variable “Hospital”

**Comentario** La etiqueta de clase “no hospitalización” es el resultado más frecuente en el departamento de urgencias y para las condiciones que se han expuesto desde el principio. Todos los registros están representados y de ellos dos séptimas partes (164 valores) identifican a pacientes que sí se hospitalizaron.

Nombre UVI

**Tipo de dato** Categórico Binario

**Valores faltantes** No tiene

**Ruido y Visualización** No tiene valores atípicos

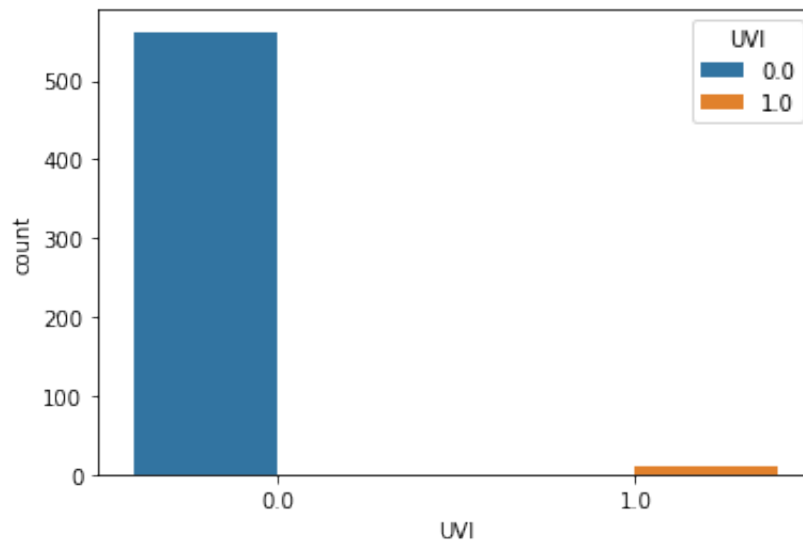


Figura 3.15: Visualización de variable “UVI”

**Comentario** Esta es una variable que presenta valores positivos (o sea, de sí traslado a la UVI) muy desbalanceados en cantidad total, con respecto al número de la clase cero, que es 562. De esta forma, la clase positiva representa solo un 2% de la cantidad de pacientes registrados.

**Nombre** Coronarias

**Tipo de dato** Categórico Binario

**Valores faltantes** No tiene

**Ruido y Visualización** No tiene valores atípicos

**Comentario** Presenta 16 valores con la clase “1” y 557 con la clase “0”, o sea, muy pocos de los pacientes registrados fueron dirigidos al espacio de atención especial de la enfermedad coronaria.

**Nombre** ICTUS

**Tipo de dato** Categórico Binario

**Valores faltantes** No tiene

**Ruido y Visualización** No tiene valores atípicos

**Comentario** Presenta todos los registros en la clase negativa y esto hace que sea una salida difícil de pronosticar tal cual, porque no contiene ejemplos de clasificación positiva para que un modelo hipotético aprenda a reconocerlo.

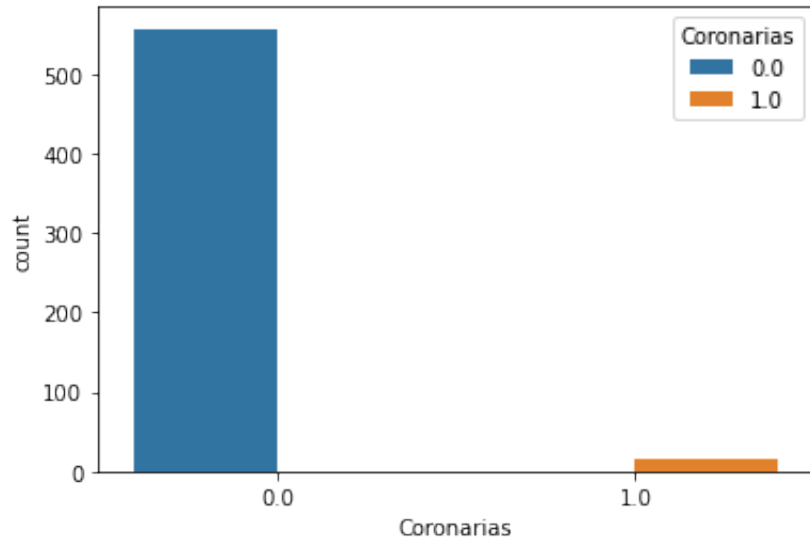


Figura 3.16: Visualización de variable “Coronarias”

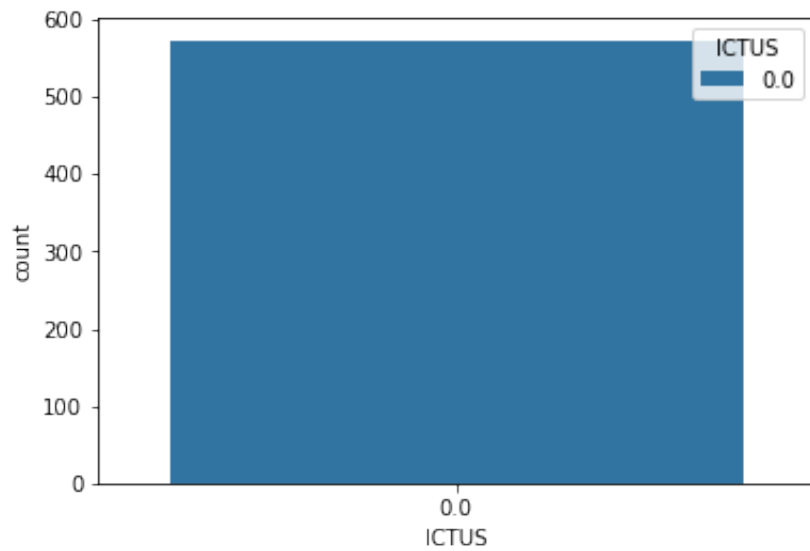


Figura 3.17: Visualización de variable “ICTUS”

**Nombre** Mortalidad Hospitalaria

**Tipo de dato** Categórico Binario

**Valores faltantes** Tiene perdido el valor del registro 568

**Ruido y Visualización** No tiene valores atípicos

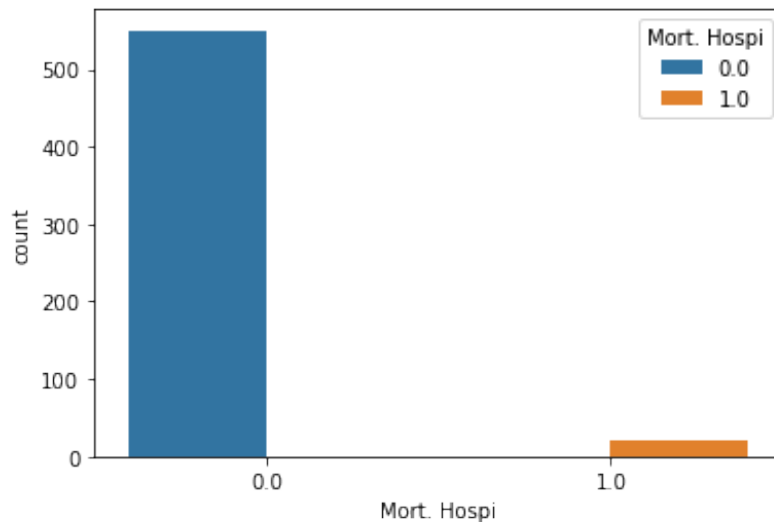


Figura 3.18: Visualización de variable “Mortalidad Hospitalaria”

**Comentario** Esta variable comprende a las personas que han fallecido en el tiempo que permanecieron en las instalaciones hospitalarias (desde el día de la llegada a urgencias), algunos habiendo estado ingresados hasta 10 días. En estas condiciones, 550 de los 573 pacientes registrados, fallecieron mientras eran atendidos en el hospital.

## VARIABLES CATEGÓRICAS

**Nombre** Sexo

**Tipo de dato** Categórico con dos posibles valores

**Valores faltantes** No tiene

**Ruido y Visualización** No presenta valores atípicos

**Comentario** Se evidencia que en el registro de pacientes hay más hombres que mujeres, por un margen de 18%. [Figura 3.19]

**Nombre** Triage

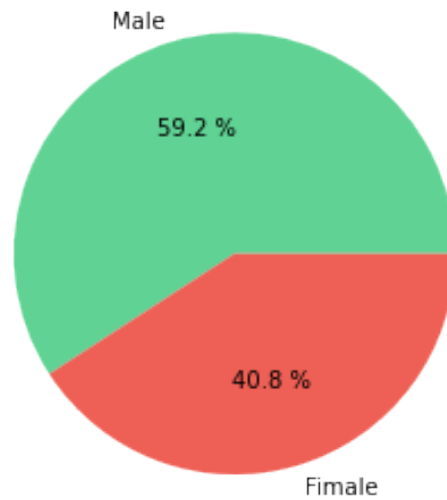


Figura 3.19: Distribución de valores de la variable “Sexo”

**Tipo de dato** Categórico ordinal con 3 posibles valores

**Valores faltantes** No tiene

**Ruido y Visualización** No presenta valores atípicos

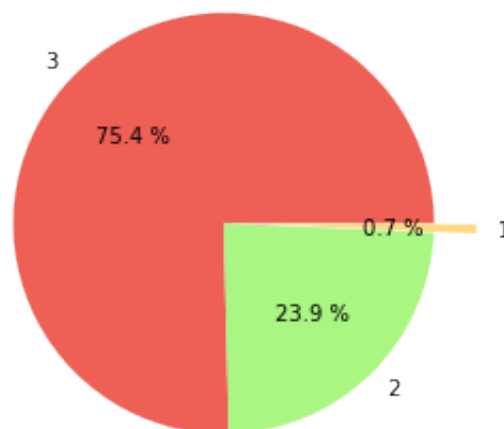


Figura 3.20: Distribución de valores de la variable “Triage”

**Comentario** Ya se aclaró porqué esta variable presenta solo 3 variantes de valores de las 5 con que normalmente se maneja en el ED. La mayor parte de los pacientes (específicamente 432) llegaron a urgencias con un nivel menos grave dentro de estas 3 variantes más graves de valoración. Sin embargo, solo 4 pacientes se reportan con condiciones de tal gravedad que necesitaron resucitación (o sea, con valor “1”). [Figura 3.20]



**Nombre** Estados de coma de Glasgow (GCS)

**Tipo de dato** Categórico ordinal en los tres casos

**Valores faltantes** No tienen

**Ruido y Visualización** No presentan valores atípicos

	GCS.O-H	GCS.V-H	GCS.M-H
<b>count</b>	573.0	573.0	573.0
<b>unique</b>	4.0	4.0	6.0
<b>top</b>	4.0	5.0	6.0
<b>freq</b>	555.0	552.0	561.0

Figura 3.21: Descripción estadística de las variables “GCS.Ocular”, “GCS.Verbal”, “GCS.Motora”, respectivamente

**Comentario** Estas tres variables se asemejan en su distribución de valores y en los tipos de clasificaciones que comprenden. Según la tabla de la Figura 3.21, una considerable mayoría de los pacientes registrados, en los tres casos, presentan el nivel más alto las respectivas clasificaciones, lo que indica que la mayoría de los pacientes (cantidad evidenciada en la última fila) tiene actividad ocular espontánea (nivel “4” de Ocular), que mantienen una conversación (nivel “5” de Verbal) y que su cuerpo obedece a la orden que se le de (nivel “6” de Motora).

**Nombre** Ritmo

**Tipo de dato** Categórico nominal

**Valores faltantes** No tiene

**Ruido y Visualización** No presenta valores atípicos pero sí que 16 registros muestran el nombre de la categoría que identifica el tipo de ritmo interpretado a partir del ECG del paciente, mientras que en los 557 casos restantes, esta variable muestra el número que se le hace corresponder a cada ritmo (Tabla 3.1).

**Comentario** Del total de registros contabilizados, 346 son señalados por tener un ritmo sinusoidal (332 tienen el valor 1 y 14 muestran el nombre correspondiente), que es el equivalente a decir que tanto el ritmo como la respiración son normales.

**Nombre** intervalo ST

Valores de Ritmo	Conteo de valores
1	332
8	58
3	58
5	49
13	18
Ritmo sinusoidal	16
10	14
12	5
4	4
11	4
17	2
14	2
16	1
15	1
6	1
2	1
7	1
18	1
Flutter auricular	1
Fibrilación auricular	1

Tabla 3.1: Frecuencia de registros con cada valor de “Ritmo”

**Tipo de dato** Categórico nominal

**Valores faltantes** No tiene

**Ruido y Visualización** No presenta valores atípicos pero muestra otro tipo de ruido que podría llamarse “ruido de valores” y que también aparece en la variable **Ritmo** y consiste en que no se corresponde la notación que identifica a las categorías en todos los registros: palabras o números. (Tabla 3.2)

**Comentario** La variable presenta 573 registros con valor normal, que es el más frecuente con 484 casos clasificados con la categoría 1 y 13 casos clasificados con el nombre de esa categoría 1.

## VARIABLES NUMÉRICAS

**Nombre** Edad

**Tipo de dato** Numérico entero

Valores de ST	Conteo de valores
1	484
5	30
6	17
Normal	13
3	12
4	7
2	4
7	2
Descenso	2
T picudas	1
T negativas	1

Tabla 3.2: Frecuencia de registros con cada valor de “ST”

Indicadores estadísticos	valores estadísticos
count	573.000000
mean	70.719023
std	16.301238
min	18.000000
25 %	62.000000
50 %	75.000000
75 %	83.000000
max	96.000000

Tabla 3.3: Descripción de variable “Edad”

**Valores faltantes** No tiene

**Ruido y Visualización** En la descripción estadística de esta variable (Tabla 3.3) y la gráfica de caja correspondiente (Figura 3.22), se reconocen como atípicos las edades entre 18 y 30 años, según el cálculo de los correspondientes cuartiles y de las relaciones entre ellos.

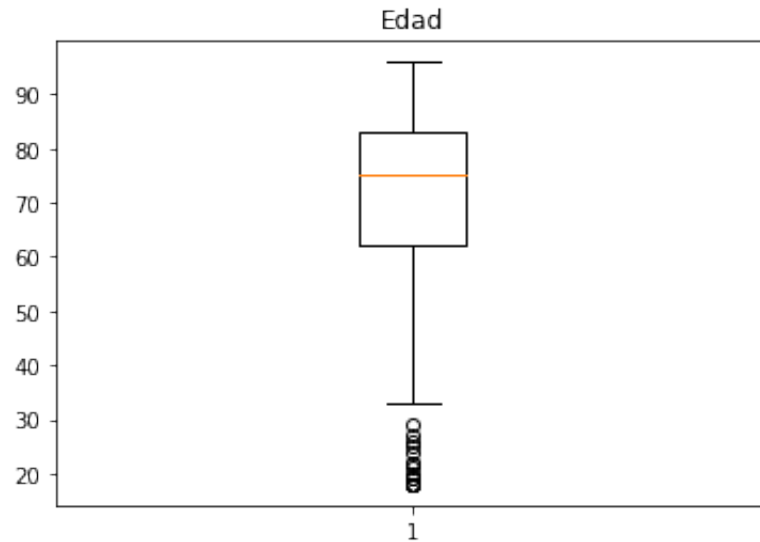


Figura 3.22: Gráfica de caja de variable “Edad”

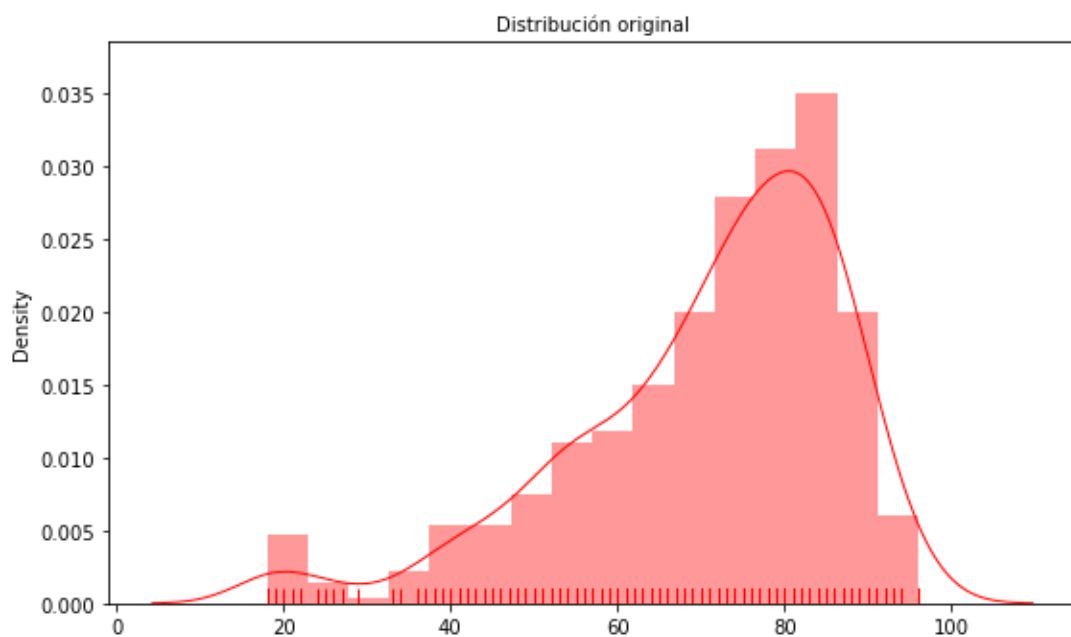


Figura 3.23: Distribución original de variable “Edad”

**Comentario** En las figuras anteriores se muestra que la variable presenta 573 valores distribuidos entre 18 y 96 años. Las edades menos comunes son aquellas que se encuentran entre 18 y 30. Más del 50% de las personas registradas tiene edades entre 75 y 96 años, relación de mayoría que también se evidencia en el hecho de que el promedio de estas edades sea 70.7% del total de edades de los pacientes.

Indicadores estadísticos	valores estadísticos
count	573.000000
mean	14.891798
std	3.977492
min	9.000000
25 %	12.000000
50 %	14.000000
75 %	16.000000
max	48.000000

Tabla 3.4: Descripción de variable “Frecuencia Respiratoria”

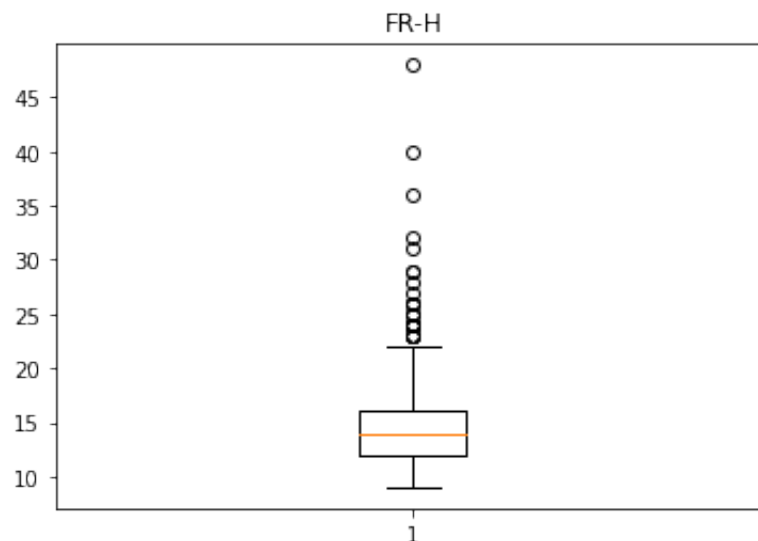


Figura 3.24: Gráfica de caja de variable “Frecuencia Respiratoria”

**Nombre** Frecuencia Respiratoria [FR]

**Tipo de dato** Numérico entero

**Valores faltantes** No tiene

**Ruido y Visualización** Presenta valores atípicos que se señalan en la Figura 3.24 y se esperan encontrar a partir de la alta diferencia que muestra la Tabla 3.4 entre el valor del tercer cuartil y el máximo valor contenido.

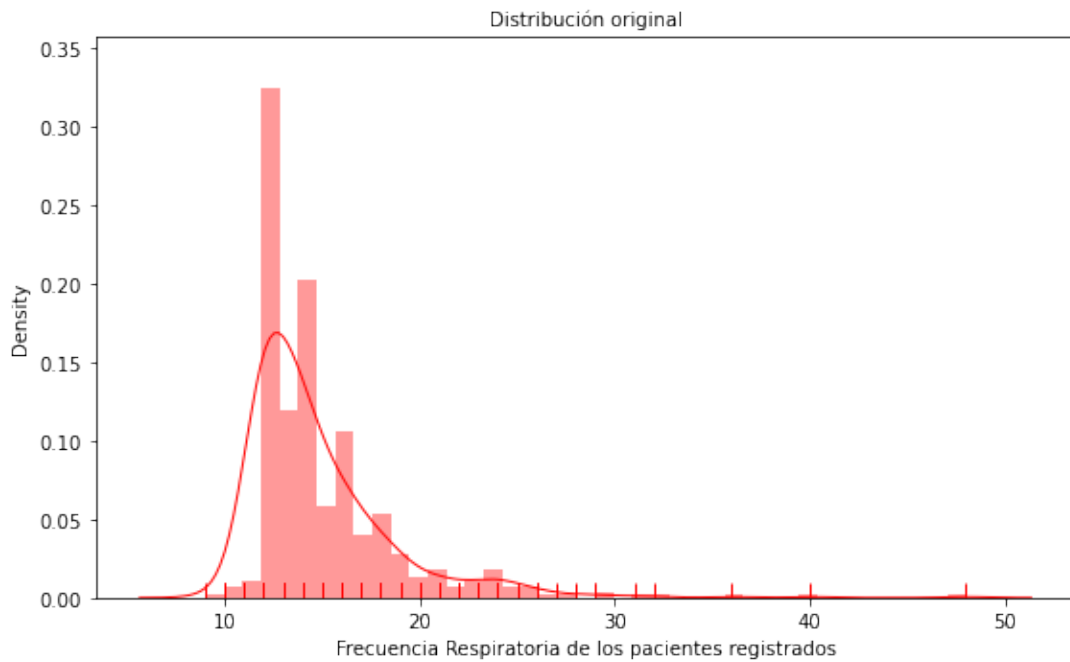


Figura 3.25: Distribución original de variable “Frecuencia Respiratoria”

**Comentario** La variación normal de la variable suele estar entre 12 y 20, es comprensible que la biblioteca “matplotlib” destaque los valores fuera de este rango, en este caso por encima, como atípicos.

**Nombre** Saturación de Oxígeno en Sangre [SpO2]

**Tipo de dato** Numérico entero

**Valores faltantes** No tiene

**Ruido y Visualización** Presenta valores atípicos que se señalan en la Figura 3.26 y en la Figura 3.27.

Indicadores Estadísticos	Valores Estadísticos
count	573.000000
mean	96.497382
std	2.714648
min	85.000000
25 %	95.000000
50 %	97.000000
75 %	99.000000
max	100.000000

Tabla 3.5: Descripción de variable “Saturación de Oxígeno en Sangre”

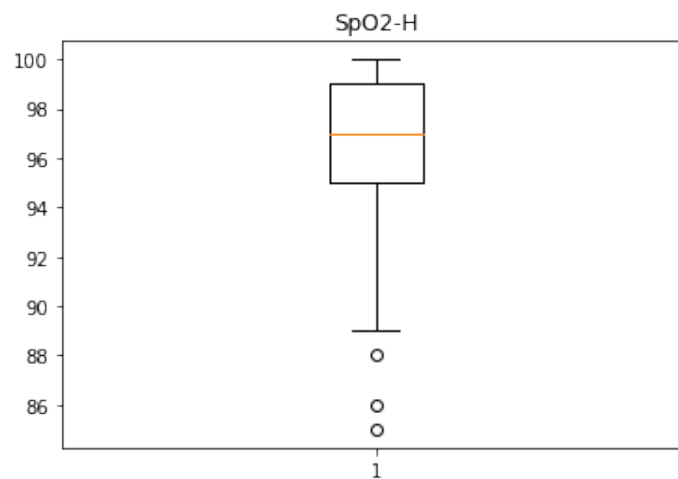


Figura 3.26: Gráfica de caja de variable “Saturación de Oxígeno en Sangre”

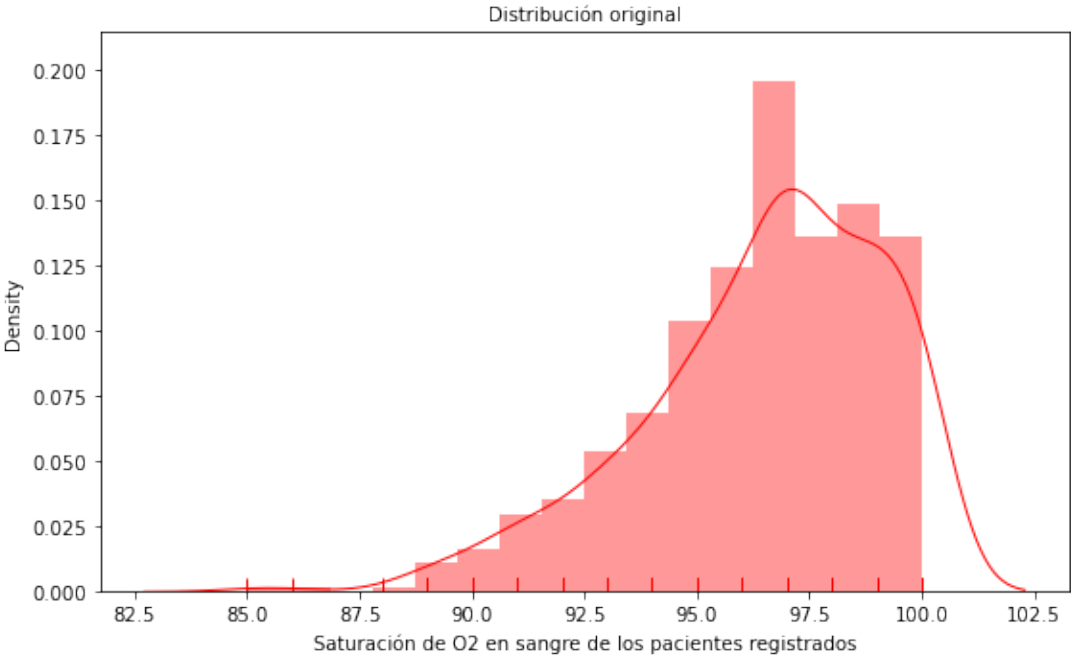


Figura 3.27: Distribución original de variable “Saturación de Oxígeno en Sangre”



**Comentario** Según la distribución estadística de la variable que se muestra en la Tabla 3.5, la relación entre cuartiles da a entender que los valores por debajo de 89 son atípicos para esta variable.

**Nombre** Oxígeno en Aire Inspirado [FiO2]

**Tipo de dato** Numérico entero

**Valores faltantes** No tiene

**Ruido y Visualización** Presenta valores atípicos que se señalan en la Figura 3.28.

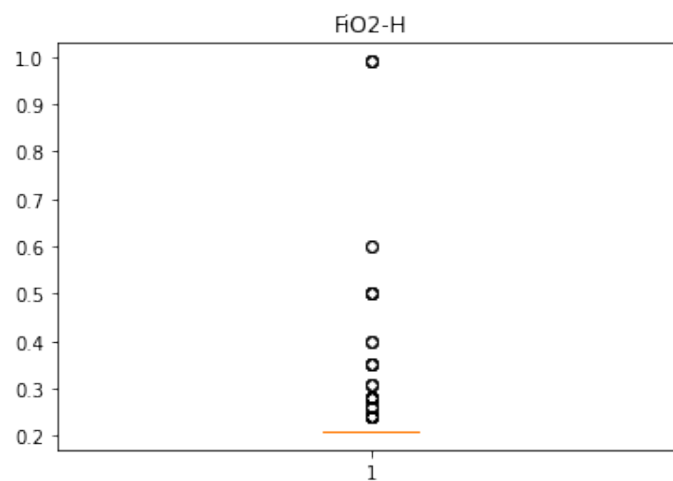


Figura 3.28: Gráfica de caja de variable “Oxígeno en Aire Inspirado”

Indicadores Estadísticos	Valores Estadísticos
count	573.000000
mean	0.225864
std	0.076858
min	0.210000
25 %	0.210000
50 %	0.210000
75 %	0.210000
max	0.990000

Tabla 3.6: Descripción de variable “Oxígeno en Aire Inspirado”

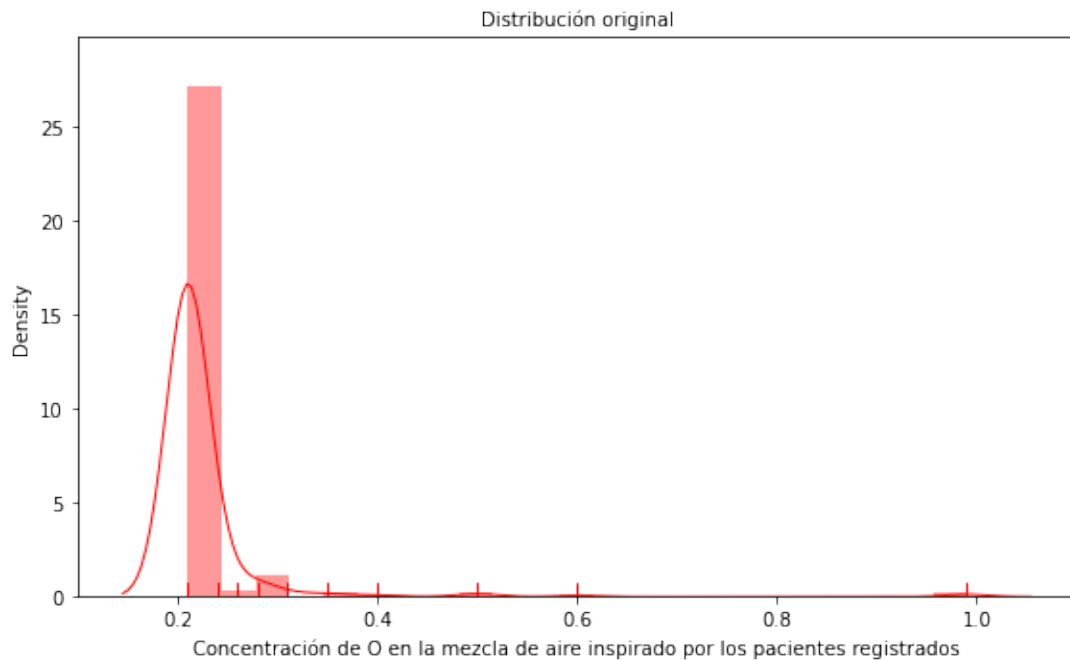


Figura 3.29: Distribución original de variable “Oxígeno en Aire Inspirado”

**Comentario** El valor 0.21 es tan común entre los datos, que la biblioteca de análisis estadístico reconoce a todos los demás como atípicos (la variable tiene, específicamente, 609 registros con valor 0.21) En el boxplot, los tres cuartiles son 0.21 y aplicando las relaciones correspondientes, también el mínimo y máximo de los valores corresponde a 0.21, por ello los valores por debajo de 0.21 y por encima de 0.21 (estos últimos comprendidos entre 0.21 y 0.99, según la descripción estadística de la variable en la Tabla 3.6), son los considerados atípicos.

**Nombre** Tensión Arterial Sistólica [TAS]

**Tipo de dato** Numérico entero

**Valores faltantes** No tiene

**Ruido y Visualización** Presenta valores atípicos que se señalan en la Figura 3.30 y su presencia es reconocida en la diferencia entre el valor mínimo y el cuartil primero de los valores de la variable (Tabla 3.7)

Indicadores Estadísticos	Valores Estadísticos
count	573.000000
mean	124.386056
std	24.863739
min	0.210000
25 %	108.000000
50 %	123.000000
75 %	139.000000
max	232.000000

Tabla 3.7: Descripción de variable “Tensión Arterial Sistólica”

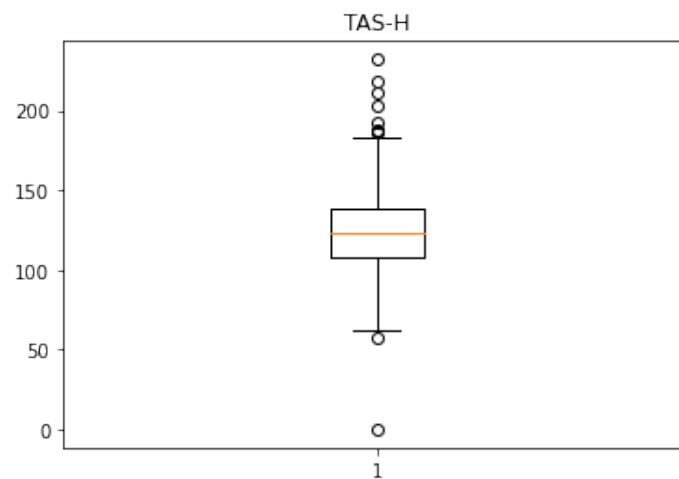


Figura 3.30: Gráfica de caja de variable “Tensión Arterial Sistólica”

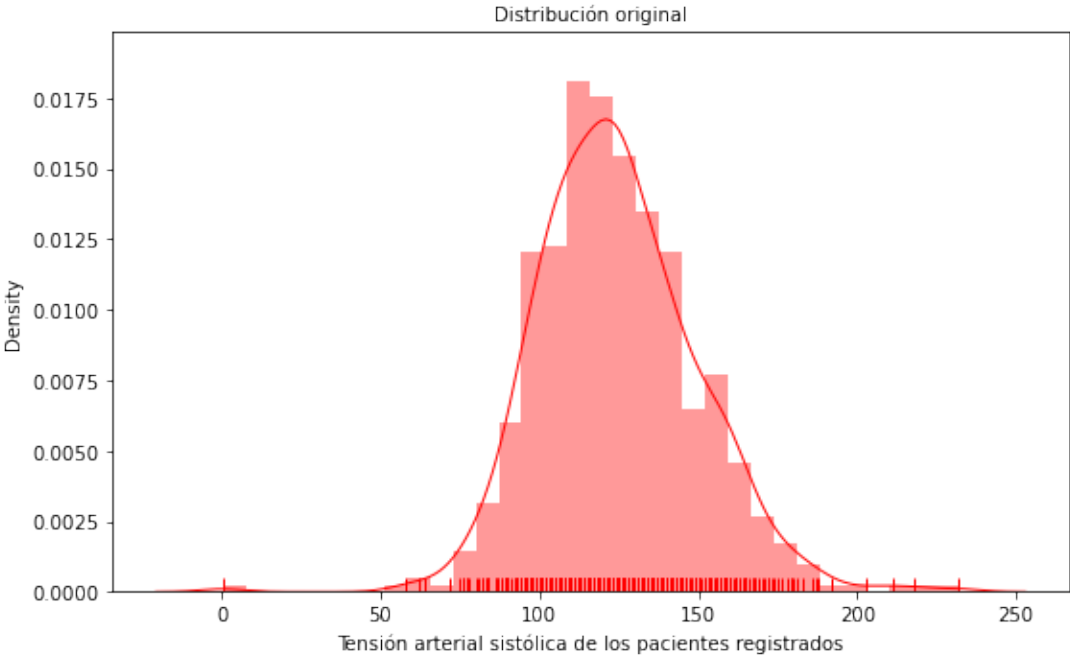


Figura 3.31: Distribución original de variable “Tensión Arterial Sistólica”

**Comentario** Es notable que el valor mínimo presentado por la variable es un valor totalmente sin sentido cuando se habla de tensión sistólica y totalmente acorde con la variable de concentración de oxígeno en aire inspirado: **Oxígeno en Aire Inspirado [FiO2]**. Este valor es reconocido, claro está, como valor atípico dentro de la variable a estudiar y pudiera ser producto de un error en la transferencia de los registros escritos, de los hospitales a la base de datos digital (Ver Figura 3.31).

**Nombre** Tensión Arterial Diastólica [TAD]

**Tipo de dato** Numérico entero

**Valores faltantes** No tiene

**Ruido y Visualización** Presenta valores atípicos que se señalan en la Figura 3.32, en la Tabla 3.8 y en la distribución mostrada en la Figura 3.33.

Indicadores Estadísticos	Valores Estadísticos
count	573.000000
mean	69.120419
std	14.464249
min	29.000000
25 %	29.000000
50 %	68.000000
75 %	78.000000
max	162.000000

Tabla 3.8: Descripción de variable “Tensión Arterial Diastólica”

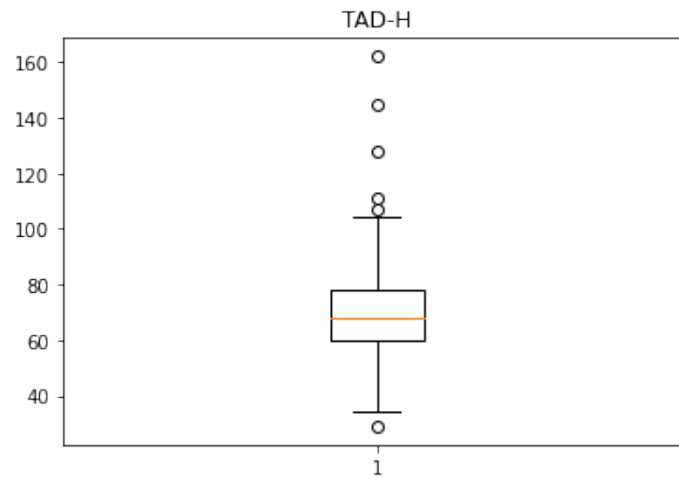


Figura 3.32: Gráfica de caja de variable “Tensión Arterial Diastólica”

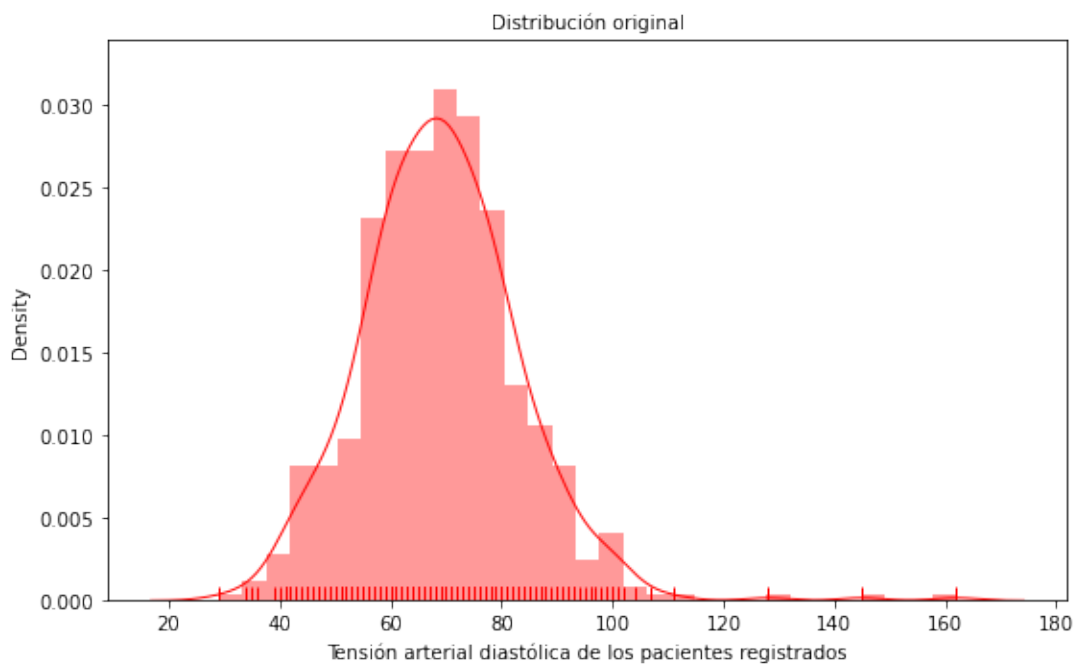


Figura 3.33: Distribución original de variable “Tensión Arterial Diastólica”

**Nombre** Frecuencia Cardíaca [FC]

**Tipo de dato** Numérico entero

**Valores faltantes** No tiene

**Ruido y Visualización** Presenta valores atípicos que se señalan en la Figura 3.34, en la Tabla 3.9 y la Figura 3.35.

Indicadores Estadísticos	Valores Estadísticos
count	573.000000
mean	73.593368
std	16.863617
min	30.000000
25 %	62.000000
50 %	70.000000
75 %	83.000000
max	147.000000

Tabla 3.9: Descripción de variable “Frecuencia Cardíaca”

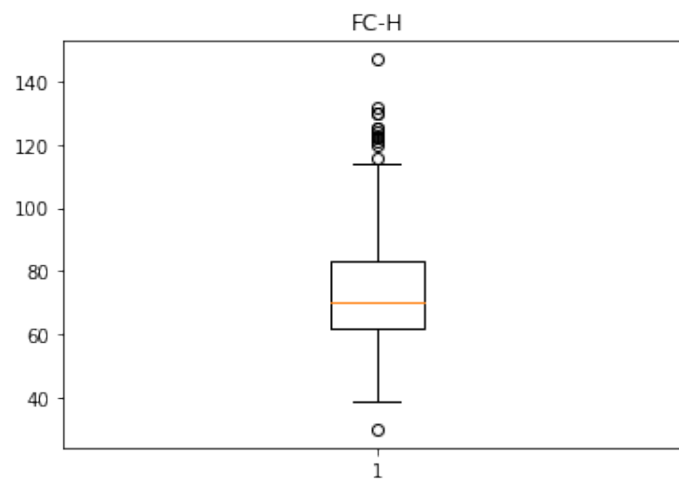


Figura 3.34: Gráfica de caja de variable “Frecuencia Cardíaca”

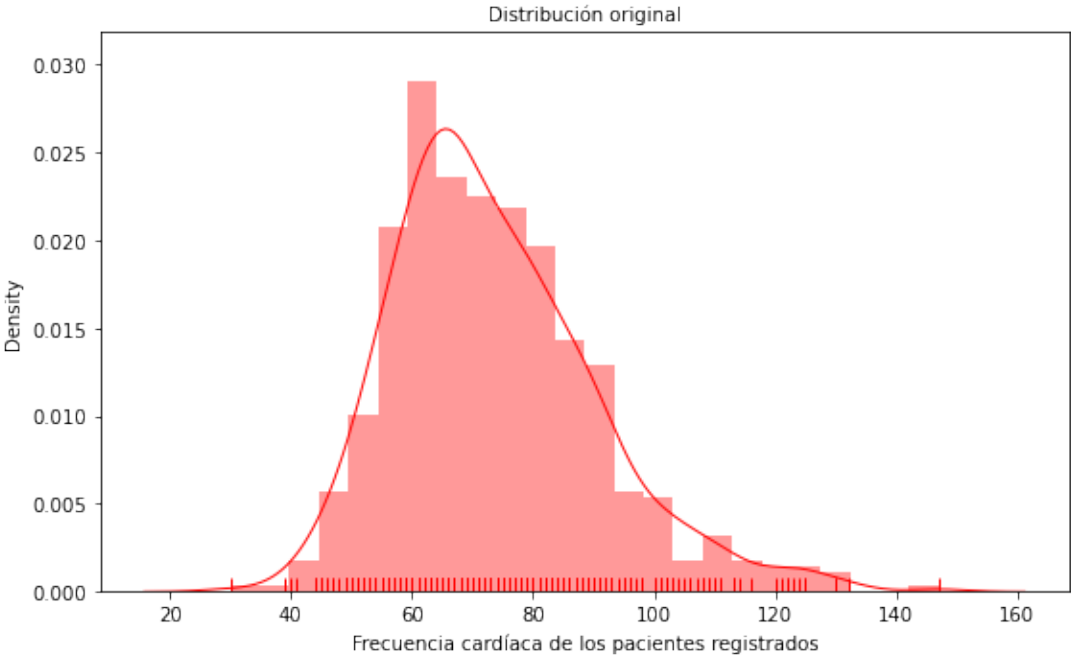


Figura 3.35: Distribución original de variable “Frecuencia Cardíaca”



**Nombre** Tiempo de Trombina [TT]

**Tipo de dato** Numérico decimal

**Valores faltantes** Tiene un valor ausente

**Ruido y Visualización** Presenta un valor atípico que se señala en la Figura 3.36 y en la Tabla 3.10.

Indicadores Estadísticos	Valores Estadísticos
count	572.000000
mean	36.105769
std	3.079607
min	33.500000
25 %	35.500000
50 %	36.000000
75 %	36.400000
max	108.000000

Tabla 3.10: Descripción de variable “Tiempo de Trombina”

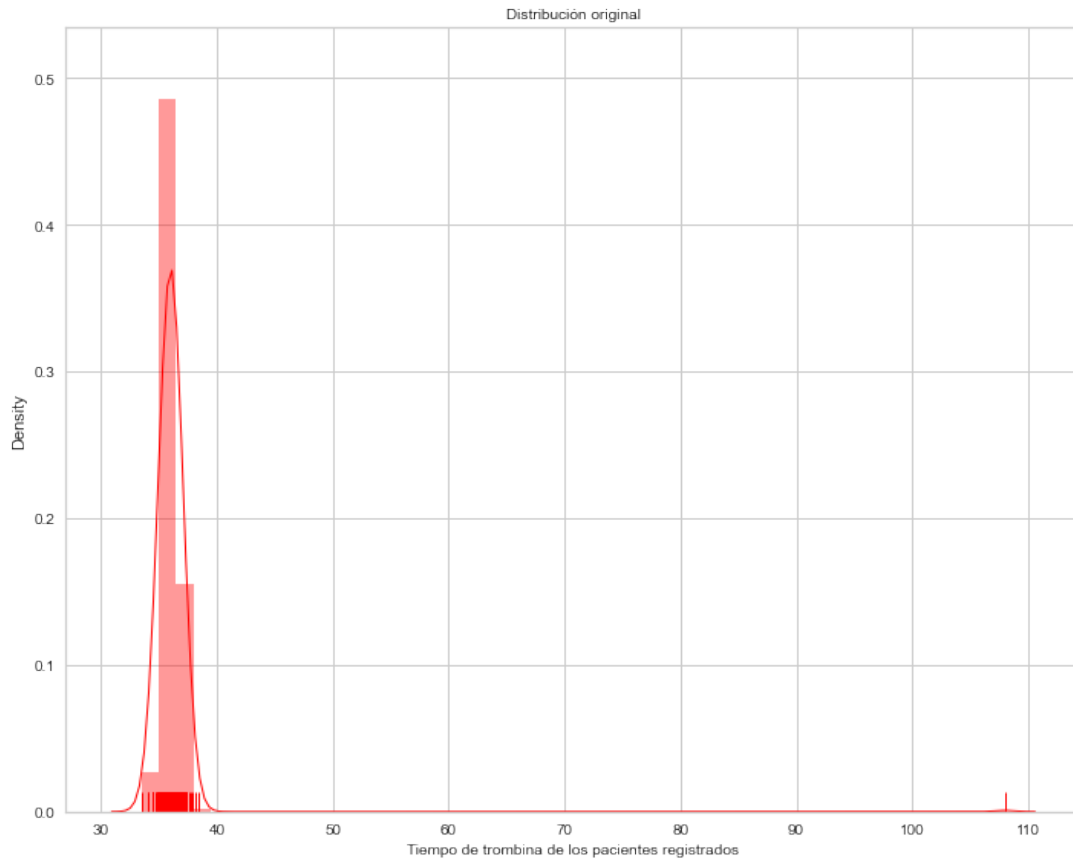


Figura 3.36: Distribución original de variable “Tiempo de Trombina”

**Comentario** Todos los valores están concentrados entre 30 y 40, excepto el 108 que, tras los cálculos correspondientes para sacar los valores límites respecto a las relaciones entre cuartiles (para este caso mín = 33.15 y máx = 36.85), se evidencia como un valor atípico.

**Nombre** Hematocrito [Hct]

**Tipo de dato** Numérico decimal

**Valores faltantes** No tiene

**Ruido y Visualización** Presenta valores atípicos que se señalan en la Figura 3.37, en la Tabla 3.11 (mayormente en la diferencia entre el valor mínimo de la variable y el primer cuartil de la misma) y en la Figura 3.38.

Indicadores Estadísticos	Valores Estadísticos
count	573.000000
mean	39.531588
std	6.129437
min	11.400000
25 %	36.400000
50 %	40.300000
75 %	43.500000
max	59.000000

Tabla 3.11: Descripción de variable “Hematocrito”

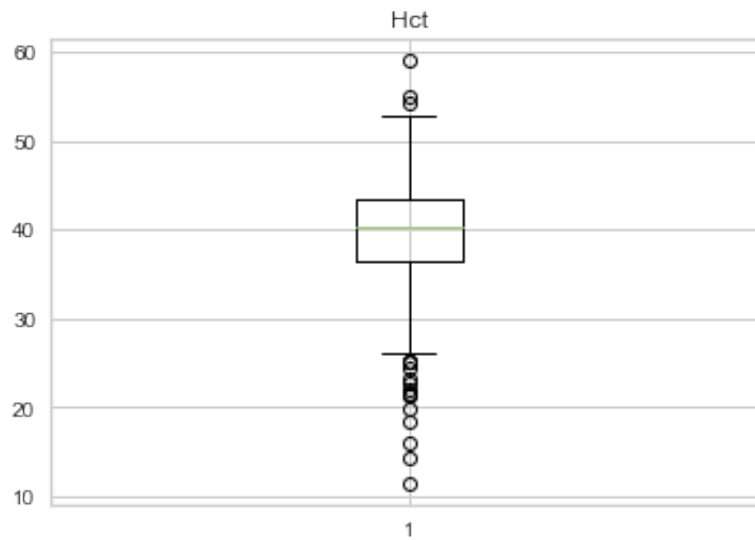


Figura 3.37: Gráfica de caja de variable “Hematocrito”

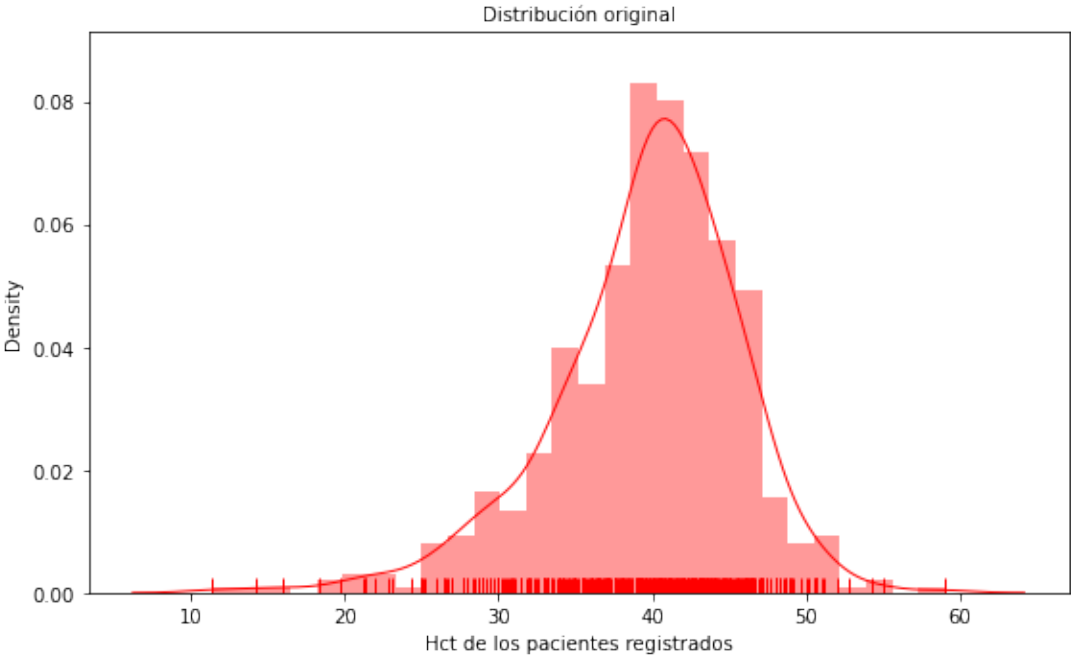


Figura 3.38: Distribución original de variable "Hematocrito"

**Nombre** Leucocito [Leuco]

**Tipo de dato** Numérico decimal

**Valores faltantes** no tiene

**Ruido y Visualización** Presenta valores atípicos que se señalan notablemente en la Figura 3.39, en la diferencia entre el valor del tercer cuartil y el valor máximo de la variable como lo muestra la Tabla 3.12 y en la distribución que muestra la Figura 3.40.

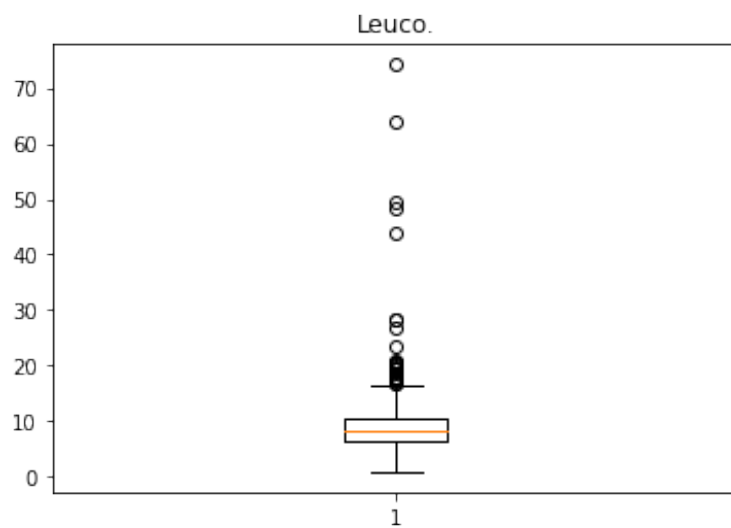


Figura 3.39: Gráfica de caja de variable “Leucocito”

Indicadores Estadísticos	Valores Estadísticos
count	573.000000
mean	9.304101
std	5.809216
min	0.900000
25 %	6.500000
50 %	8.100000
75 %	10.600000
max	74.200000

Tabla 3.12: Descripción de variable “Leucocito”

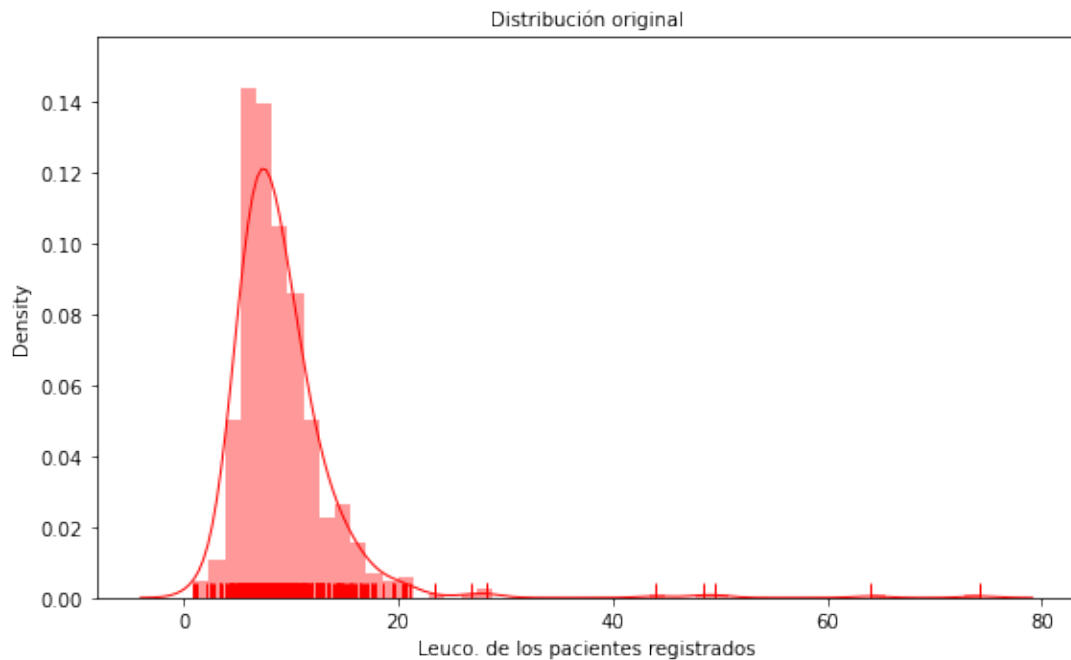


Figura 3.40: Distribución original de variable “Leucocito”

### Nombre Plaquetas

**Tipo de dato** Numérico decimal

**Valores faltantes** Tiene 2 valores ausentes

**Ruido y Visualización** Presenta valores atípicos que se muestran en la Figura 3.41 y cuya presencia se hace resaltar en la diferencia entre las dos últimas filas de la Tabla 3.13.

Indicadores Estadísticos	Valores Estadísticos
count	571.000000
mean	218.887916
std	90.391457
min	33.000000
25 %	171.500000
50 %	206.000000
75 %	250.000000
max	1078.000000

Tabla 3.13: Descripción de variable “Plaquetas”

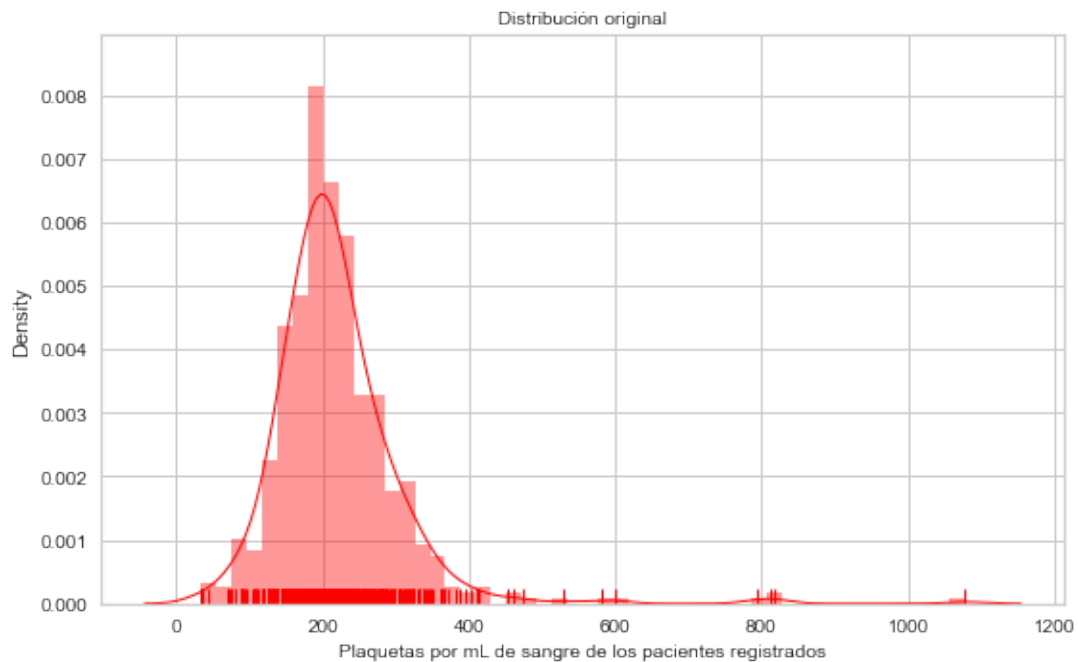


Figura 3.41: Distribución original de variable “Plaquetas”

**Nombre** Glucosa [Glu]

**Tipo de dato** Numérico entero

**Valores faltantes** No tiene valores ausentes

**Ruido y Visualización** Presenta valores atípicos que se muestran en la Figura 3.42, la Figura 3.43 y se evidencian en la Tabla 3.14 tanto en la diferencia entre los límites de valores de la variable (máximo y mínimo) y los cuartiles correspondientemente más cercanos (tercero y primero, respectivamente).

Indicadores Estadísticos	Valores Estadísticos
count	573.000000
mean	141.762653
std	44.284003
min	23.000000
25 %	113.500000
50 %	133.000000
75 %	162.000000
max	355.000000

Tabla 3.14: Descripción de variable “Glucosa”

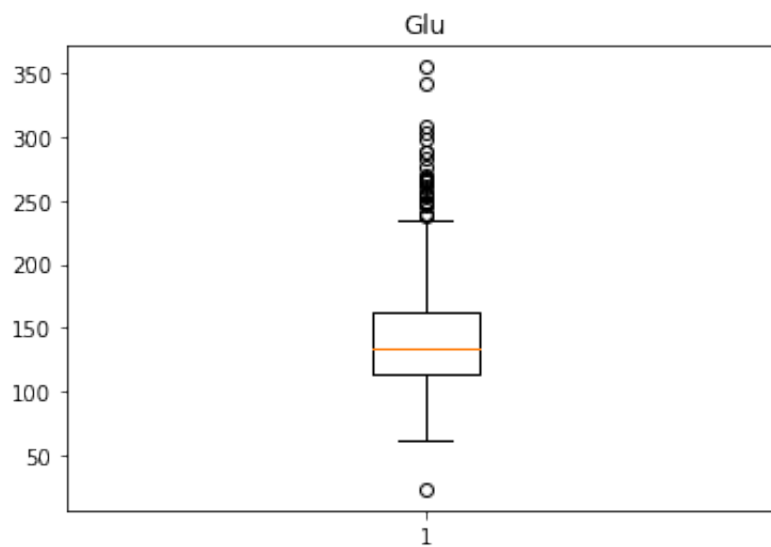


Figura 3.42: Gráfica de caja de variable “Glucosa”



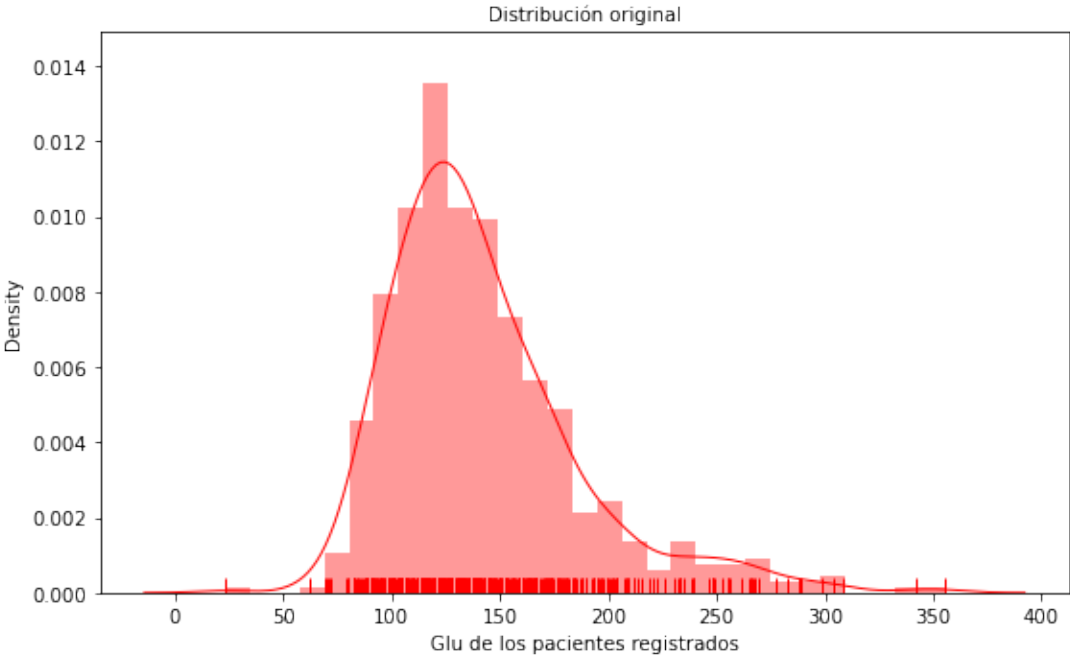


Figura 3.43: Distribución original de variable "Glucosa"

**Nombre** Creatinina [Crea]

**Tipo de dato** Numérico decimal

**Valores faltantes** No tiene valores ausentes

**Ruido y Visualización** Presenta muchos valores atípicos, los que se muestran en la gráfica de cajas de la Figura 3.44, en la Tabla 3.15 y en la Figura 3.45; donde todos los valores fuera de lugar están por encima de 2, parece este mismo ser un valor alto para la variable.

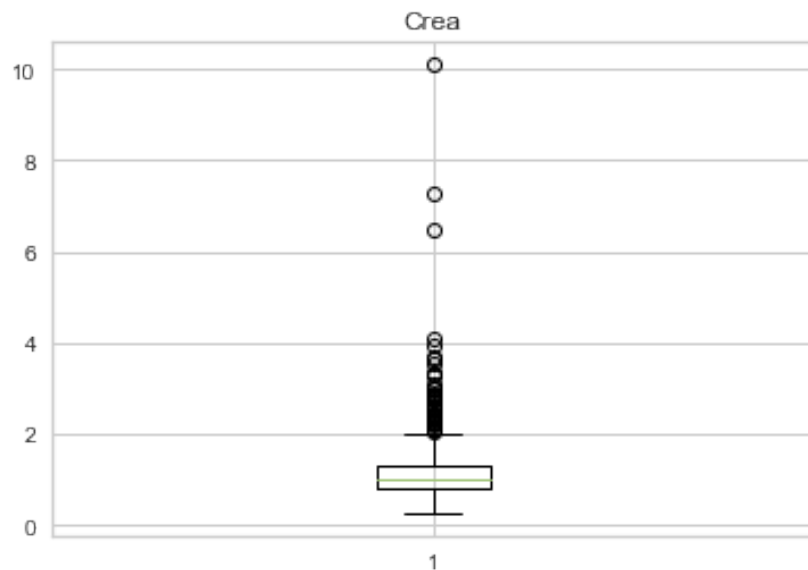


Figura 3.44: Gráfica de caja de variable “Creatinina”

Indicadores Estadísticos	Valores Estadísticos
count	573.000000
mean	1.189738
std	0.743829
min	0.270000
25 %	0.820000
50 %	1.020000
75 %	1.300000
max	10.100000

Tabla 3.15: Descripción de variable “Creatinina”

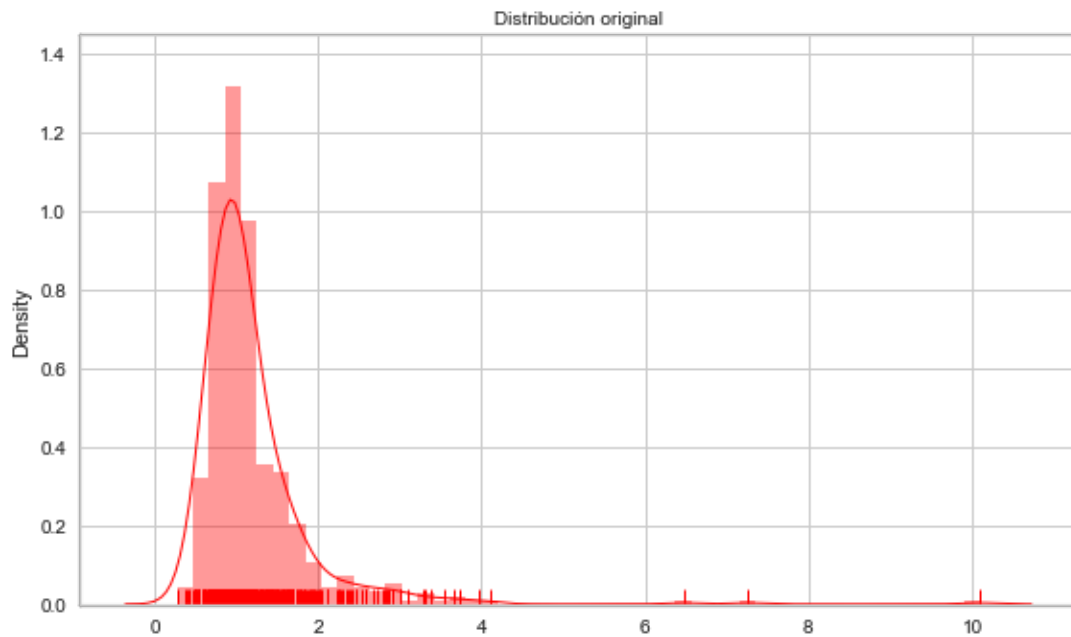


Figura 3.45: Distribución original de variable “Creatinina”

**Nombre** Troponina [Trop]

**Tipo de dato** Numérico decimal

**Valores faltantes** Tiene 147 valores ausentes

**Ruido y Visualización** Presenta valores atípicos que se muestran en la Figura 3.46 y en el conteo de valores mostrado en el la Tabla 3.16. Además de lo anterior, tanto en las últimas dos filas de las estadísticas (Tabla 3.16) como en la gráfica evidencian que los valores alrededor de mil son demasiado altos para esta variable.

Indicadores Estadísticos	Valores Estadísticos
count	426.000000
mean	51.939061
std	326.485653
min	0.760000
25 %	4.600000
50 %	10.495000
75 %	23.537500
max	6293.000000

Tabla 3.16: Descripción de variable “Troponina”

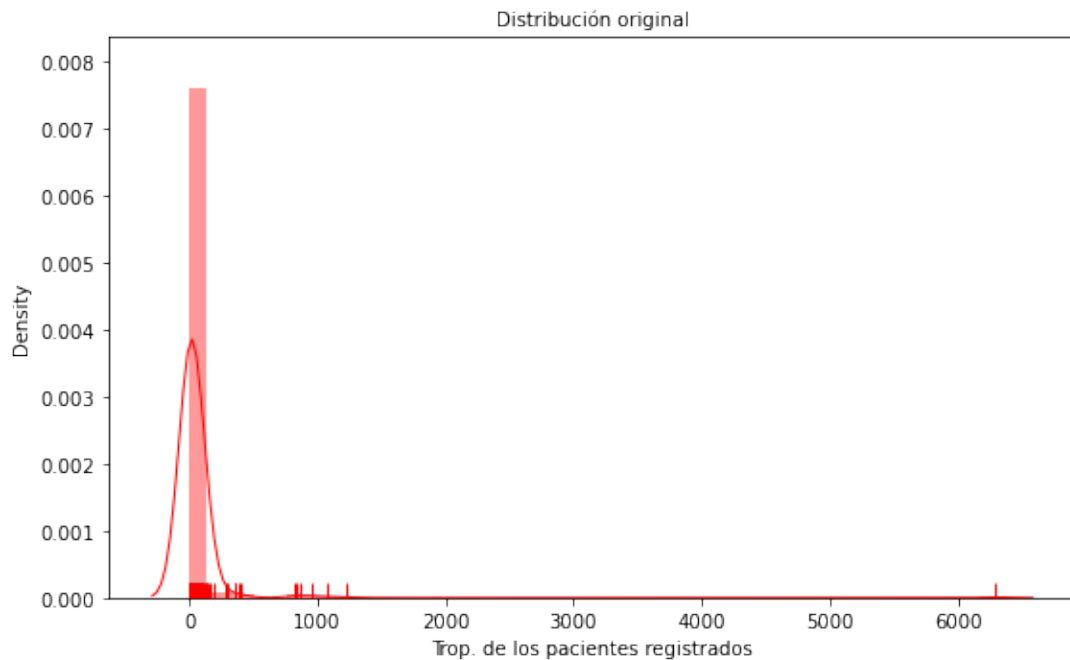


Figura 3.46: Distribución original de variable “Troponina”

Después de visualizar cada atributo, se concluye que es necesario revisar el comportamiento multivariable para poder decidir qué tan útiles son para el estudio cada uno de ellos, pues se encontraron muchos valores ausentes y un elevado número de valores fuera de lugar en las respectivas variables que los comprenden. También se pudo reconocer ruido causado por la inserción de valores de una variable en otra.

### Análisis multivariable

En este apartado se analizan las variables mostrando: el comportamiento individual frente a las variables de salida y la relación entre par de predictores y la variable de salida, así como muestra la matriz de correlación donde intervienen todas. En el caso de las variables a predecir, se relacionan en dos formas estratégicas: la descripción de comportamientos entre ellas y la comparación de las tres que constituyen el servicio UCI (según los requerimientos del médico responsable) y que conformarán una nueva variable.

### “Hospitalización”

La Figura 3.48 muestra la edad de los pacientes separada por cuatro grupos de edad, limitados por regulaciones médicas. En la Figura 3.47 la mayoría de los hospitalizados son personas de avanzada *Edad*, pero esto no resalta ante el hecho de que son mayoría las personas de avanzada edad en ambas clases. También se visualiza que hay personas hospitalizadas menores de 30 años. Teniendo en cuenta los valores atípicos en cada variable

que los presenta, no se visualiza tendencia alguna de comportamiento a favor de una variable u otra en base a la edad de los pacientes registrados, como en el caso de la *Creatinina* (ver Figura 3.53) y la *Saturación de oxígeno en la sangre* (ver Figura 3.54 y Figura 3.49) cuyos valores considerados normales se graficaron dispersos en ambas clases. También se obtuvieron (como resultado de las graficaciones análogas a las figuras anteriores) valores altos, normales o de pre-diabetes de *Glucosa* tanto en “hospitalizados” como en “no hospitalizados” (ver Figura 3.51). Los valores establecidos para los grupos de glucosa, se muestran en el segmento de código de la Figura 3.52.

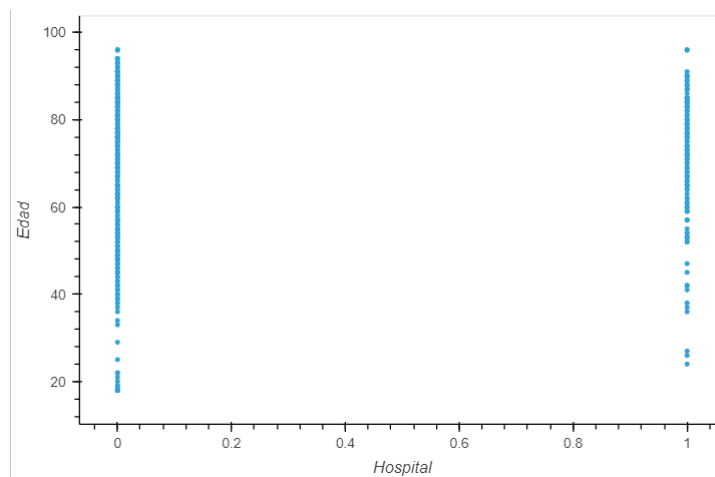


Figura 3.47: Los valores de “Edad” con respecto a la salida “Hospital”

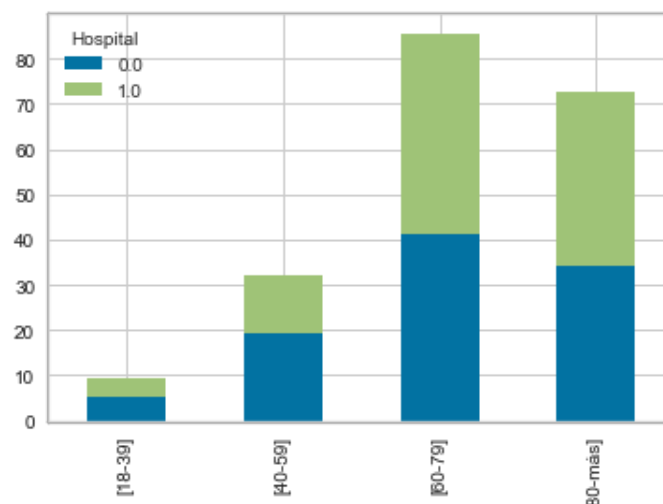


Figura 3.48: Grupos de “Edad” con respecto a la salida “Hospital”

Los valores seguidos para establecer límites de saturación de oxígeno en sangre, se muestran en la Figura 3.50, los que fueron tomados de [57] y las páginas de consulta médica para pacientes consultadas.

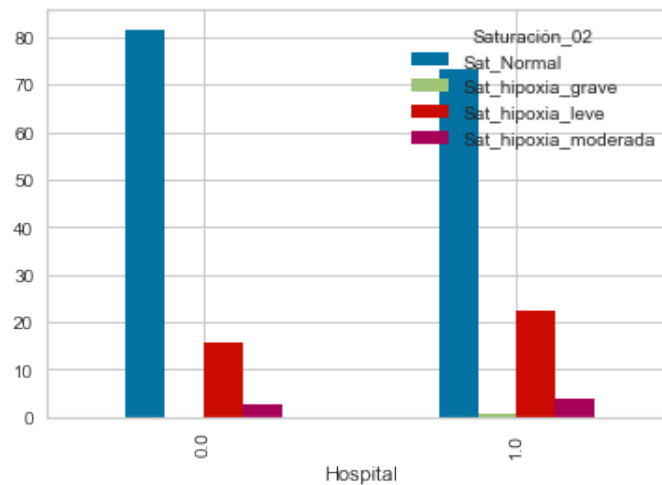


Figura 3.49: “Saturación de oxígeno en sangre” con respecto a la salida “Hospital”

```
df_data.loc[df_data['SpO2-H'].between(95, 100, inclusive = True), 'Saturación_02'] = 'Sat_Normal'
df_data.loc[df_data['SpO2-H'].between(91, 94, inclusive = True), 'Saturación_02'] = 'Sat_hipoxia_leve'
df_data.loc[df_data['SpO2-H'].between(86, 90, inclusive = True), 'Saturación_02'] = 'Sat_hipoxia_moderada'
df_data.loc[df_data['SpO2-H'] <= 85, 'Saturación_02'] = 'Sat_hipoxia_grave'
```

Figura 3.50: Código de “Saturación de oxígeno en sangre” con respecto a la salida “Hospital”

Respecto a la *Frecuencia respiratoria*, hay valores que ninguna de las personas hospitalizadas presentaron, pero esto se puede deber a algún error de llenado de la base de datos o que simplemente no coincidió ese valor con la medición de algún paciente, pues se trata de valores intercalados entre otros que sí están presentes tanto en personas hospitalizadas como en no hospitalizadas.

Los *estados de coma de Glasgow*, se obtienen graficados respecto a la variable de salida *Hospitalización* de forma que son los mismos estados para ambas clasificaciones en el caso del parámetro verbal, para el caso de los otros dos parámetros de la coma: todos los estados tienen representación en la clase positiva (pacientes hospitalizados). Tal como el estado de coma de de Glasgow, los niveles de *Triaje* también se representan por igual en ambas clases de la variable de salida.

La variable que indica la *Concentración de oxígeno en el aire inspirado (FiO2)*, muestra valores de 0,21 (los que se consideran normales y son mayoría entre los datos) para ambas clases, pero casi todos los valores por encima de 0,21 (aunque no todos ni en mayor cantidad) son representativos de las personas que fueron hospitalizadas (ver Figura 3.54 y Figura 3.55). Se debe resaltar que estos valores superiores podrían indicar que estas personas necesitaron algún tipo de ventilación artificial, más no necesariamente que fueron hospitalizadas, ya que la totalidad de los casos de hospitalización no tienen un valor elevado en esta variable, ni todos los valores altos corresponden a personas que resultaron hospitalizadas.

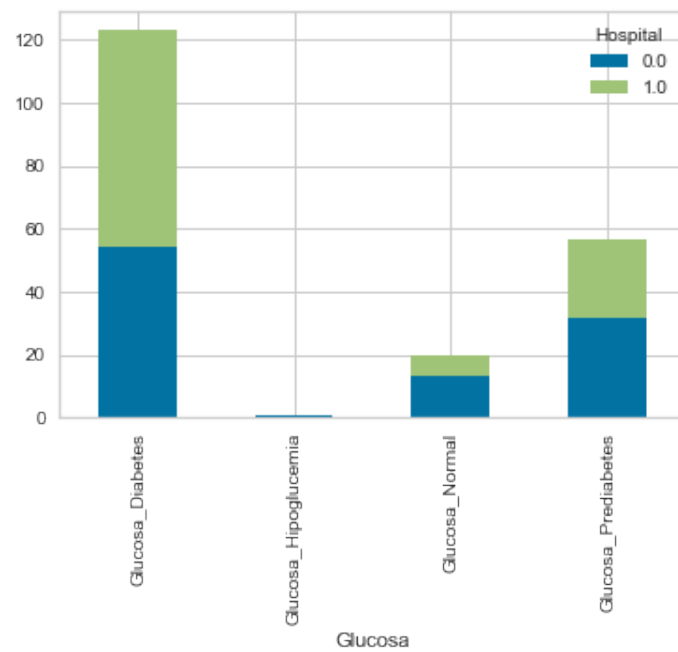


Figura 3.51: Valores de “Glucosa” con respecto a la salida “Hospital”

```
df_data.loc[df_data['Glu'].between(70, 100, inclusive = False), 'Glucosa'] = 'Glucosa_Normal'
df_data.loc[df_data['Glu'].between(100, 125, inclusive = True), 'Glucosa'] = 'Glucosa_Prediabetes'
df_data.loc[df_data['Glu'] < 70, 'Glucosa'] = 'Glucosa_Hipoglucemia'
df_data.loc[df_data['Glu'] >= 126, 'Glucosa'] = 'Glucosa_Diabetes'
```

Figura 3.52: Código de valores de “Glucosa” con respecto a la salida “Hospital”

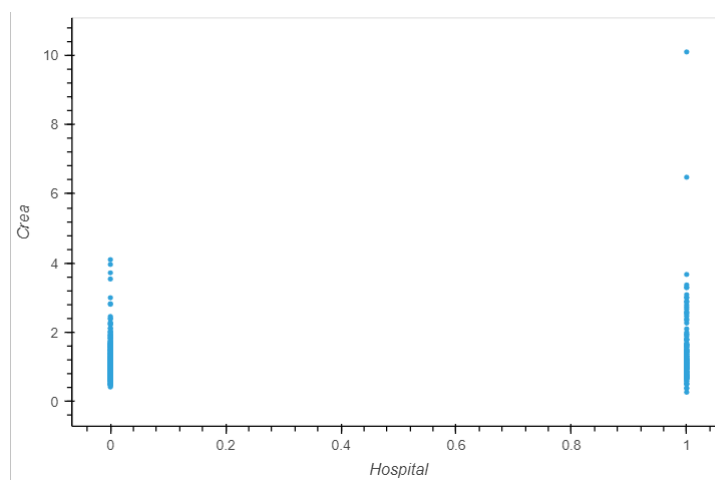


Figura 3.53: Los valores de “Creatinina” con respecto a la salida “Hospital”

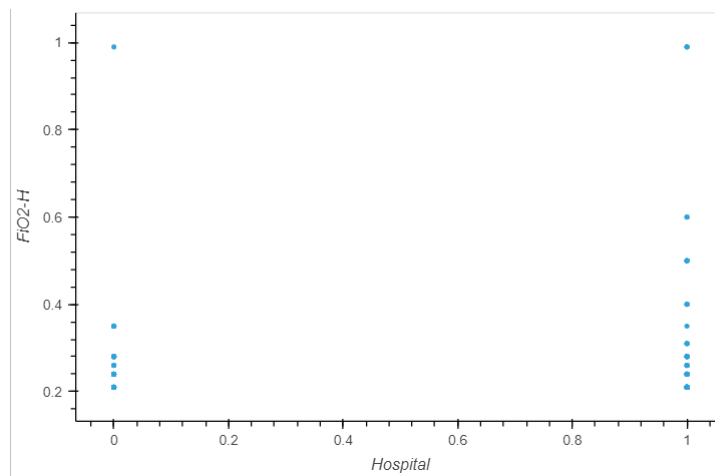


Figura 3.54: Los valores de “FiO2” con respecto a la salida “Hospital”

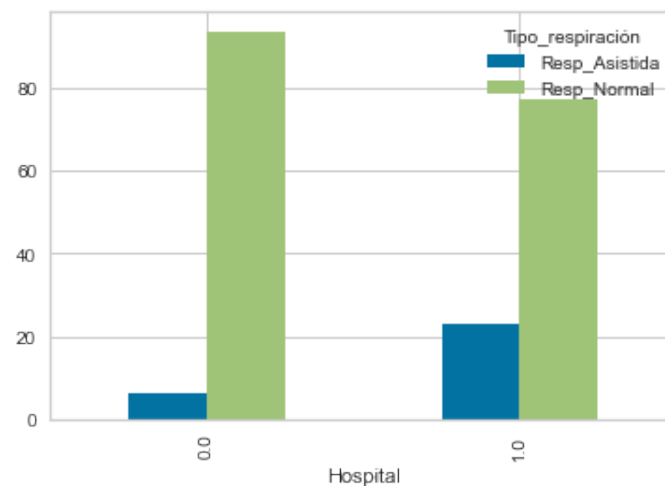


Figura 3.55: Respiración “FiO2” con respecto a la salida “Hospital”

Según la Figura 3.56 no hay un nivel de  $ST$ , que sea característico de las personas hospitalizadas, como se comprobó sucede con *Ritmo*, el resto de variables categóricas. Las clases está distribuidas de forma que en todos los niveles de la variable categórica que se analice, la mayor parte de las personas fueron hospitalizadas.

De la Figura 3.57 se concluye que con *Frecuencia respiratoria* alta y también *Frecuencia cardíaca* alta, aunque no hay establecida una tendencia de los valores a favorecer la hospitalización, pareciera que en esos límites la hospitalización es segura.

En las relaciones de distribución, donde está la mayor probabilidad de ocurrencia de una variable de salida, ahí estará la otra, no se establecen diferencias. Asimismo, en la matriz de correlación mostrada en la Figura 3.58, se evidencia la alta analogía entre las variables de estado de coma de Glasgow, mientras son negativas y débiles las relaciones



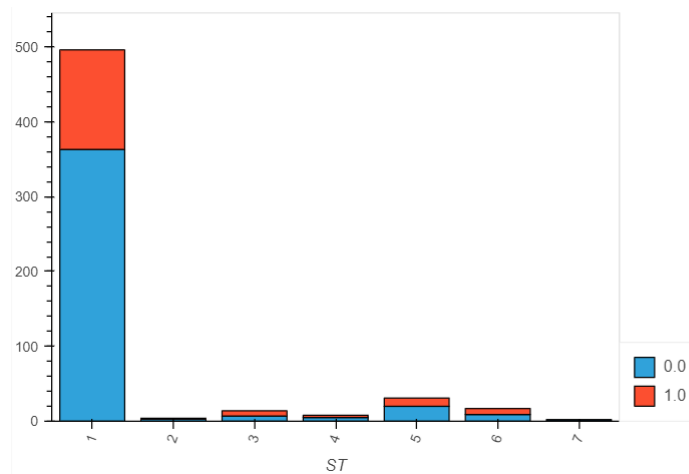


Figura 3.56: Histograma de "Hospital" para "ST"

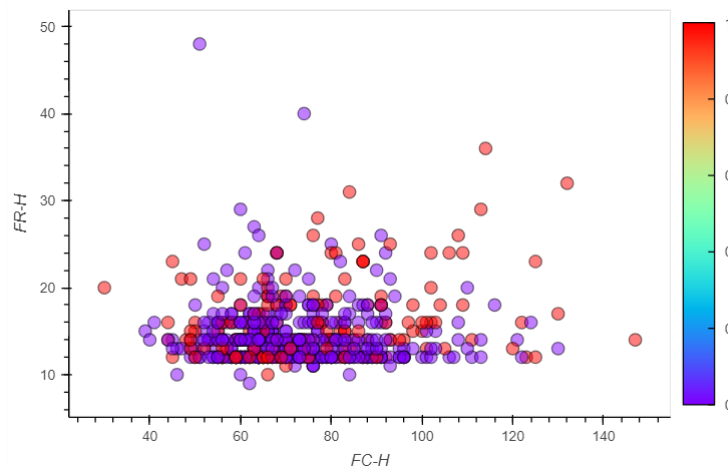


Figura 3.57: Relación de valores entre "Frecuencia cardíaca" y "Frecuencia respiratoria"

con la variable de salida. El resto de la matriz de correlación presenta una analogía muy débil de las variables independientes con respecto a la variable de salida.

### "Mortalidad Hospitalaria"

Revisando el comportamiento de los atributos con respecto a la variable que indica la *Mortalidad hospitalaria*, se encuentra que las personas fallecidas tenían edades avanzadas, por encima de 60 años (ver Figura 3.59). Y las *Frecuencias cardíacas* de estas personas solía estar por encima de 100 lpm y por debajo de 90 lpm.

En cuanto al *estado de coma de Glasgow*, esta variable se comportó de forma similar a como lo hizo respecto a la variable **Hospitalización**, con el agregado de que el nivel "2" (que es el segundo de los niveles más graves) del parámetro ocular (ver Figura 3.60), solo

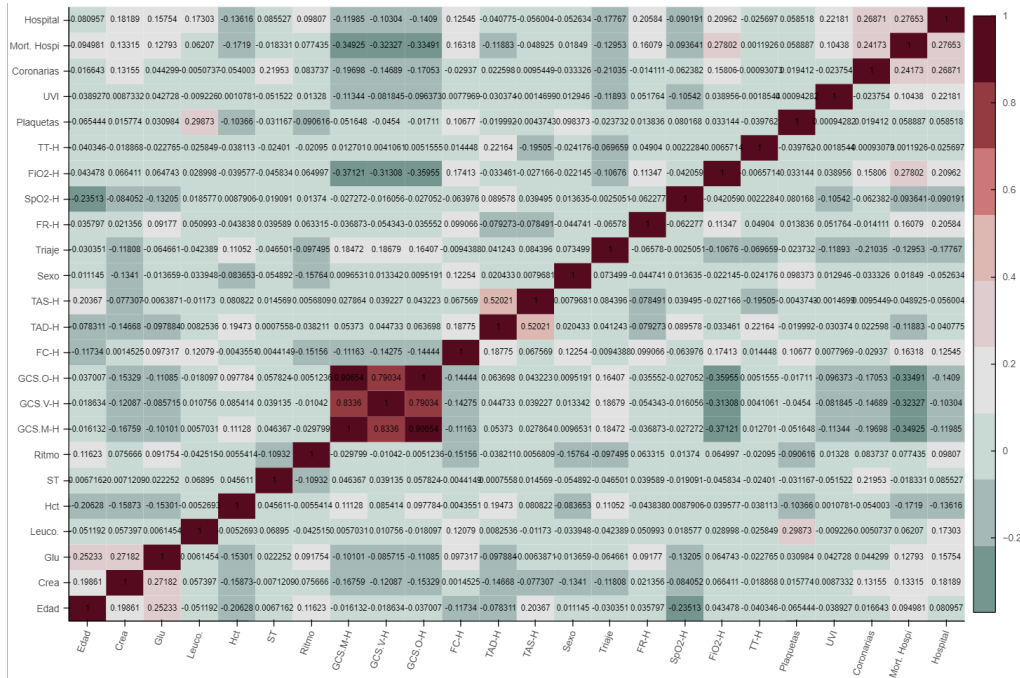


Figura 3.58: Matriz de correlación con la salida “Hospital”

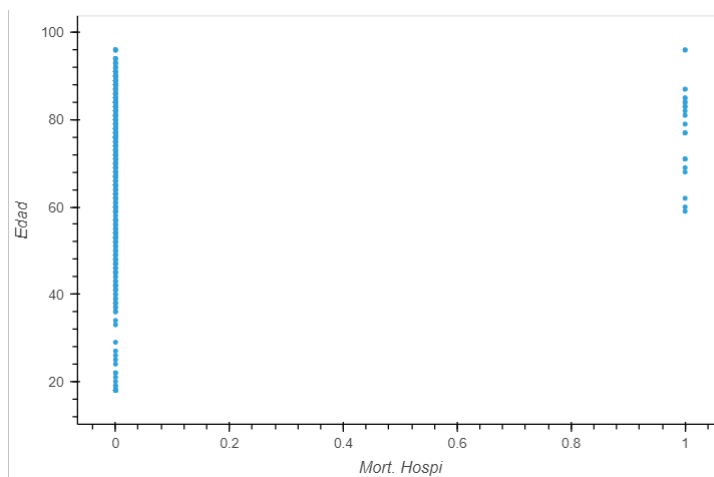


Figura 3.59: Los valores de “Edad” con respecto a la salida “Mortalidad Hospitalaria”

está representado entre las personas fallecidas, lo que no es precisamente un encauce hacia la hospitalización para todos los que sean catalogados con este nivel de estado de coma.

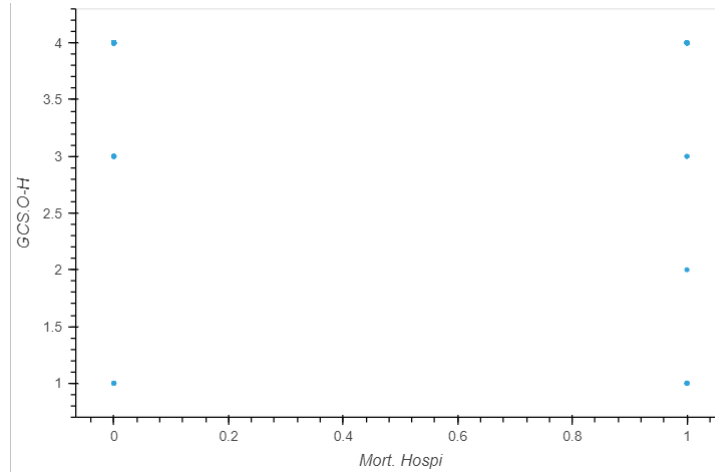


Figura 3.60: Los valores de “GCS.O” con respecto a la salida “Hospital”

En base a lo graficado en la Figura 3.61, todas las personas que fallecieron presentaron una forma normal de la porción de ciclo cardíaco *ST* en el ECG realizado en el ED; así como la mayor parte de los fallecidos fueron hospitalizados también, esta última información se visualiza en la Figura 3.62.

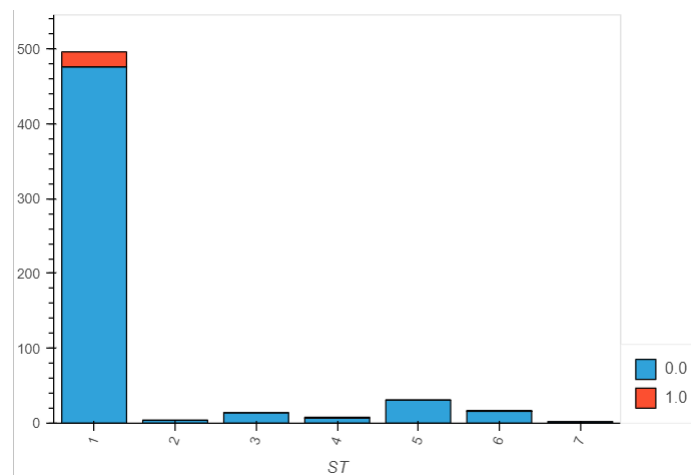


Figura 3.61: Histograma de “Mortalidad Hospitalaria” para “ST”

### “UCI”

Con la clase **UCI**, el comportamiento de las variables independientes es similar al comportamiento que éstas han tenido con el resto de las variables de salida: las personas

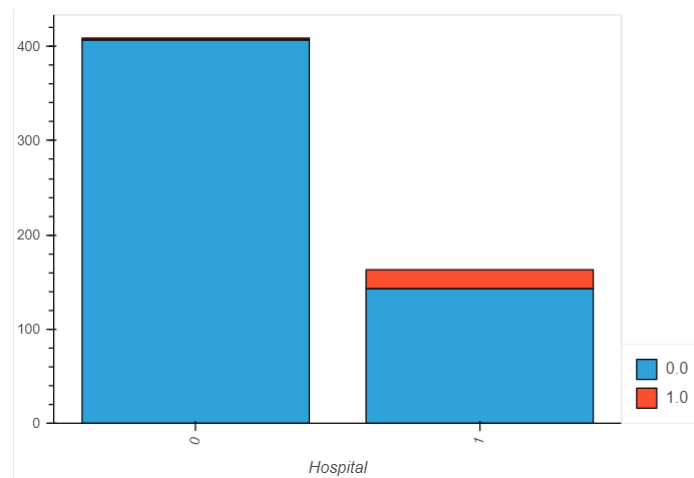


Figura 3.62: Histograma de “Mortalidad Hospitalaria” para “Hospital”

trasladadas a la UCI, suelen ser mayores de edad (de alrededor de 60 años como mínimo) y confluyen en ambas clases de la variable de salida tanto los valores que indican normalidad en el estado del paciente (en el caso de cualquier variable dependiente) como los que no. Similar a lo que ocurre con la clase de *Mortalidad hospitalaria*, el nivel “2” del parámetro Ocular del *estado de coma de Glasgow*, solo se ve representado en las personas ingresadas a a la UCI, aunque los restantes niveles se relacionan con esta clase (la positiva) y con la negativa.

En la Figura 3.63 se evidencia que de las personas trasladadas a la UCI, la mayor parte presenta un nivel de *Troponina* considerado médicamente como irregular.

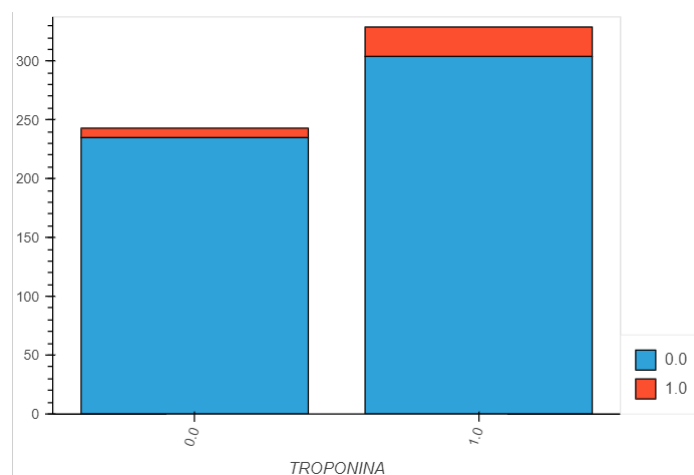


Figura 3.63: Histograma de “UCI” para “Troponina”

### “Variables de salida”

Hospitalización	Mortalidad Hospitalaria	UCI	Frecuencia
0.0	0.0	0.0 0.0 0.0	407
1.0	0.0	0.0 0.0 0.0	123
1.0	1.0	0.0 0.0 0.0	13
1.0	0.0	0.0 1.0 0.0	11
1.0	0.0	1.0 0.0 0.0	9
1.0	1.0	0.0 1.0 0.0	5
1.0	1.0	1.0 0.0 0.0	2
0.0	1.0	0.0 0.0 0.0	2

Tabla 3.17: Comparación de las salidas del conjunto de datos

La Tabla 3.17 está conformada por cuatro columnas de las que las tres primeras reflejan la combinación de valores de las cinco variables de salida identificadas en la base de datos original: **Hospitalización**, **Mortalidad Hospitalaria** y las tres variables que conformarán la variable **UCI** (**ICTUS**, **Coronarias** y **UVI**); de forma que la tercera columna se subdivide en tres valores que se analizan ya como parte de la misma variable **UCI**, la que se ha de reconocer con clase positiva si presenta valor “1” en cualquiera de las tres variables que la componen, y con clase negativa si es unánime el valor “0” entre los atributos componentes.

Cuando se relacionan los posibles valores de **Hospitalización**, **Mortalidad Hospitalaria** y de las variables que conformarán la variable **UCI**, como lo muestra la Tabla 3.17, se comprueba que existe la posibilidad de que una persona haya sido hospitalizada y necesite atención en la UCI (101), así como de que haya sido hospitalizada y fallezca próximamente en las instalaciones hospitalarias (110). Vale señalar, a partir de estas observaciones, que las personas (en el registro de pacientes de la base de datos) que se hayan reconocido como pacientes destinados a los cuidados intensivos (en todos los casos), puntúan como pacientes hospitalizados. También se visualizan casos en los que se dan las tres posibilidades: hospitalizado, con necesidad de atención de cuidados intensivos y con posibilidad de fallecimiento en la estancia hospitalaria (111), lo que tiene total sentido.

La Figura 3.64 muestra las relaciones entre todas las variables independientes y las tres variables de dependientes que se obtendrán finalmente. Por la gama de colores, se evidencia que suelen coincidir los valores de analogía de las variables predictoras con respecto a las dos últimas clasificaciones mencionadas anteriormente (**Mortalidad Hospitalaria** y **UCI**); las que se relacionaban altamente entre sí, mostrando que por lo general las personas que han fallecido han sido ingresadas a la UCI con anterioridad. Los *estados de coma de Glasgow* tienen una correlación mayor, aunque inversa, con estas dos variables de salida, valor que está al rededor de 0,33 de probabilidad. La variable **FiO2** es la segunda que más se relaciona con las tres salidas, con una probabilidad de 0,24.

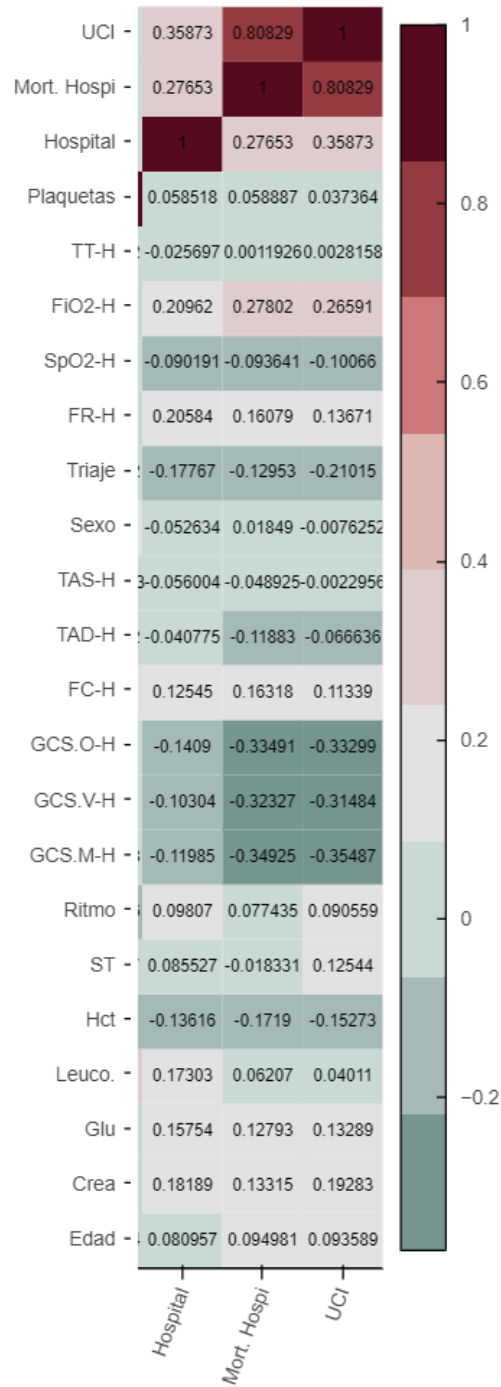


Figura 3.64: Correlación de las variables independientes respecto a las variables dependientes

UVI	ICTUS	Coronarias	Frecuencia
0.0	0.0	0.0	539
0.0	0.0	1.0	17
1.0	0.0	0.0	11
1.0	0.0	1.0	5

Tabla 3.18: Comparación de salidas de la futura variable “UCI”

### Conclusión de la Exploración

Luego de explorar las variables, se concluye que no se visualizan tendencias marcadas a favor de una clase en específico, y sí que puede haber ruido en la base de datos capaz de sesgar los datos.

### Limpieza de los datos

Esta fase tiene una pre-etapa incluida en el análisis exploratorio de los datos y es transversal al resto de los procesos por los que pasan los datos, hasta la creación de un modelo. En la etapa exploratoria se filtraron los 573 registros válidos (en la primera mirada a la base de datos) utilizando como base el valor de la primera variable de la base de datos original (**Cohort**), ya que que no presenta valores ausentes en la búsqueda general de no nulos. También se buscaron las posibles filas duplicadas para su eliminación. Aunque ya se haya realizado una pre-limpieza para obtener resultados más claros de análisis de los datos, en este apartado se explican los pasos seguidos para limpiar la base de datos lo mejor posible y disponerla para la selección de modelos: conformar la variable UCI (tal y como lo propuso el médico responsable de los datos), corregir los valores incoherentes en la estructura de las variables y rectificar los registros con espacios vacíos y valores fuera de lugar en la base de datos.

### Conformación de la variable UCI

El proceso de limpieza implica transformaciones leves de la base de datos y se inicia con la conformación de la variable **UCI**. En la Figura 3.18 se registran los valores que presentan las variables **UVI**, **ICTUS** y **Coronarias**, en ese orden.

Para crear la variable **UCI**, se encontraron cuatro posibles combinaciones de las cinco esperadas para esta variable, y es que en la base de datos no hay pacientes que estén relacionados con el **ICTUS**, además de que el que ingresa en una área no ingresa en la otras, o sea, estos campos que conformarán **UCI** son excluyentes entre sí. Era de esperar que no coincidiera ninguno en la calificación positiva, porque se trata de opciones de tratamiento específicas dentro de la **UCI**. En el apartado de transformación se agrupan las variables en la nueva **UCI**, en forma de variable binaria, que tendrá valor positivo para cualquiera de las opciones de ingreso en las tres unidades que la componen (o sea,

si alguna de las variables componentes es positiva) y tenderá valor negativo en caso que todas sean negativas.

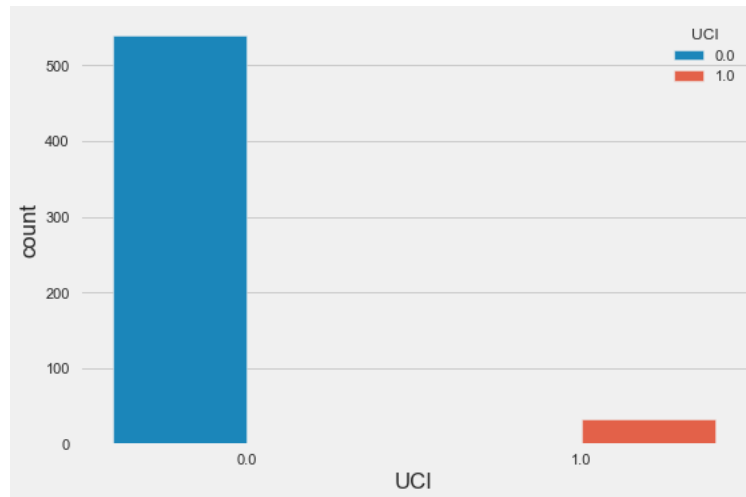


Figura 3.65: Distribución de valores de la variable “UCI”

### Valores incoherentes de las variables

En la etapa de exploración de los datos, se encontraron dos atributos con algunos de sus valores categóricos y originalmente establecidos como números enteros que identifican las categorías correspondientes, representados con el nombre de la categoría. Esta situación crea incoherencia en los tipos de datos de las respectivas variables. Ya que se habían proporcionado los nombres por nivel correspondientes a cada variable categórica numérica (junto con la base de datos), se corrigieron estos valores en las variables afectadas: **Ritmo** y **ST**.

### Valores ausentes

De las variables detectadas con valores ausentes, se mantuvieron en el estudio **TT**, **Plaquetas**, **Mortalidad Hospitalaria** y **Trop**. Las razones de mantener estas variables son las siguientes:

- “*Mortalidad Hospitalaria*”, porque es una variable de salida y la mejor medida ante la ausencia de un valor podría ser eliminar ese registro, aprovechando que es solo uno, para esa predicción en específico.
- “*TT*” y “*Plaquetas*”, porque entre ambas tienen 3 valores ausentes y son muy pocos comparados con la cantidad de registros en general. Estos espacios vacíos se pueden sustituir por alguna medida estadística que, en este caso es el promedio.



- “*Trop.*”, porque el médico le dio un significado a los límites de valores de esta variable numérica, así como comentó las razones por las que quedaron registros vacíos en el momento de rellenar la base de datos.

En el caso del atributo **Trop**, para completar esta variable se parte de que el médico planteó que los valores ausentes se pueden interpretar como son valores por encima de 4, ya que 4 es el límite de que todo está bien con el paciente. De esta forma, los valores por encima de 4 son los interesantes para el médico que evalúa al paciente (o sea, indican que no están bien los niveles de troponina). Comparando esta observación con el resultado de la búsqueda de información en internet [3], se resuelve que los 147 valores ausentes se sustituyen por “5.0”. Tanto la explicación del médico como la investigación realizada, permiten sustituir los valores ausentes por una categoría de normalidad dentro de la descripción de la variable, pero no permite rellenar la variable con valores numéricos, ya que no se tienen. Con base en ello y los rangos de valores considerados de salud normal y de salud afectada, se crea una nueva variable categórica que indica las normalidad o no en los niveles de troponina en sangre de los pacientes. Esta variable se nombró **Troponina**.

### Valores atípicos

Con el auxilio de la función **KNN**, se seleccionaron los registros con valores atípicos y se eliminaron del conjunto de datos. En primer estudio realizando a los datos, se comprobó que estos valores, aunque sin ser técnicamente comprendidos como para establecer límites razonables, entorpecían a la solución de la clasificación. Por ello, mientras no se tenga un respaldo profesional que acompañe en la búsqueda de valores posibles e imposibles para cada atributo (inclusive si se acotan los valores para la enfermedad en específico), se toma la decisión de prescindir de todos los reconocidos por las bibliotecas que realizaron este proceso. Finalmente, se eliminaron 122 registros quedando la base de datos con 451 para el resto de los pasos a seguir.

### Exploración y selección de modelos

Para seleccionar los modelos más competentes, se establece una estrategia de trabajo en la que se parte de introducir los datos obtenidos como resultado del pre-procesamiento de la etapa anterior, tal cual, y se procede a aplicar, paulatinamente, transformaciones a este conjunto de datos, transformaciones que se derivan de lo aprendido en estudios anteriores y del análisis realizado en la etapa de exploración. Cada transformación sobre los datos ha de hacer que los datos transformados se interpreten como hiper-parámetros en la exploración de modelos. El desempeño de los modelos probados se mide mediante validación cruzada y, en base a los errores que obtengan, se buscarán los resultados más prometedores.

Luego de la etapa de limpieza y preparación de los datos, estos se dividen en dos subconjuntos: entrenamiento y prueba. Al principio del proceso de investigación, se estableció que lo ideal sería dividir el conjunto de datos en tres partes: el subconjunto con

el que aprenda el modelo, el de evaluar el modelo, y el de validar el modelo a partir de casos de uso o hipótesis que con la ayuda del médico se pudieran establecer. Al tratarse de un conjunto de datos con tan pocos registros (originalmente y luego de la limpieza realizada) y al no contar con otra base de datos para validación, el tercer subconjunto quedó desestimado en esta ocasión.

Se ha diseñado una “Metodología experimental” a seguir para esta etapa, que se basa en: seleccionar las métricas de evaluación, analizar los modelos para cada hiper-parámetro de selección de datos (según las métricas) y seleccionar le mejor modelo para optimizarlo. Los pasos correspondientes se desglosan en los suba apartados siguientes.

### Selección de medidas de evaluación

Las métricas de evaluación están incorporadas en la biblioteca que se usa para explorar y crear los modelos de clasificación. Son todas métricas que tradicionalmente se relacionan con los problemas de clasificación en el uso de ML. Se parte de la tabla que describe el rendimiento de un modelo supervisado de ML en los datos de prueba, o sea, la “matriz de confusión”, ya que casi todas las métricas de desempeño se basan en ella. Esta se conforma por las cuatro relaciones entre los valores reales y las perdiciones de las salidas, que son los parámetros que se relacionan a continuación.

- Verdadero Positivo (TP, por sus siglas en inglés): cuando la clase real del punto de datos es 1 (Verdadero) y la predicha es también 1 (Verdadero).
- Verdaderos Negativos (TN, por sus siglas en inglés): cuando la clase real del punto de datos es 0 (Falso) y la pronosticada también es 0 (Falso).
- Falsos Positivos (FP, por sus siglas en inglés): cuando la clase real del punto de datos es 0 (False) y la pronosticada es 1 (True).
- Falsos Negativos (FN, por sus siglas en inglés): Cuando la clase real del punto de datos es 1 (Verdadero) y el valor predicho es 0 (Falso).

Según el escenario que impone el significado de los datos, se está buscando minimizar al máximo FP y FN, donde el escenario ideal sería con estos parámetros en 0. En las tres clasificaciones que se están analizando, se quieren capturar todos los casos *Verdaderos* para la salida de cada modelo y se podría hacer una clasificación positiva aún cuando la persona realmente no necesita ingresar a una de las opciones hospitalarias (ingreso de forma general incluyendo a la UCI) o realmente no va a fallecer en el tiempo de atención médica; lo que implicaría un uso innecesario de recursos hospitalarios y esfuerzos médicos, además de la sobrecarga en el servicio. Pero sería un gran error que ocurriera lo contrario y se pierda el proceso de exámenes y de atención que necesitaría un paciente que debería estar hospitalizado o que podría fallecer pronto. Por lo tanto, es importante buscar que se minimicen los FN.

A partir de la relación de estos cuatro parámetros de la matriz de confusión, se construyen las métricas de clasificación que se enumeran a continuación, resaltando el significado de cada una.

1. *Accuracy*: Es el porcentaje total de elementos clasificados correctamente y es la medida más directa de la calidad de los clasificadores con el mejor valor en 1 [36].
2. *AUC*: El área bajo la curva ROC da una idea de qué tan bien funciona el modelo. Es perfectamente capaz de distinguir entre clase positiva y clase negativa. Cuando AUC es 0.7, significa que hay 70 % de probabilidad de que el modelo pueda distinguir entre clase positiva y clase negativa [30].
3. *Recall*: Es el número de elementos identificados correctamente como positivos del total de positivos verdaderos. Nos da el rendimiento de un clasificador con respecto a falsos negativos [36].
4. *Precision*: Es el número de elementos identificados correctamente como positivo de un total de elementos identificados como positivos. Nos da el rendimiento de un clasificador con respecto a los falsos positivos [13].
5. *F1*: Se utiliza para combinar las medidas de precisión y recall en un sólo valor, asumiendo que nos importa de igual forma la precisión y la exhaustividad [36]. Por tanto, es una media aritmética entre ambas métricas, que da idea de rendimiento combinado.
6. *Kappa de Cohen*: El kappa de Cohen es una métrica que se usa a menudo para evaluar el acuerdo entre dos evaluadores. También se puede utilizar para evaluar el rendimiento de un modelo de clasificación. Tiene en cuenta el desequilibrio en la distribución de clases y, por lo tanto, puede ser más complejo de interpretar [94].
7. *MCC*: El Coeficiente de Correlación de Matthews, dicta que cuanto mayor sea la correlación entre los valores reales y predichos, mejor será la predicción. MCC tiene en cuenta los cuatro valores en la matriz de confusión de forma que un valor de MCC alto, significa que ambas clases se predicen bien, el 0 compara el modelo con una moneda al aire y el -1 es el valor mínimo que puede tener [75].

Para comparar el desempeño de los modelos mediante validación cruzada y enfocando la aplicación de este hacia minimizar los FN, *Recall* debe ser lo más cercano posible al 100 %. Se tiene en cuenta, además, AUC es la métrica que más se menciona en la bibliografía médica en cuanto a evaluación de modelos de clasificación y que el Accuracy es la métrica preferencial ante modelos de clasificación, ya que mide las dos clasificaciones con un solo número. Se llega al consenso de elegir la métrica *Accuracy* como la principal para seleccionar los mejores modelos y *AUC* para optimizar el modelo que se llegue a crear. Igualmente, se verificará la elección del modelo con la revisión del valor de *Recall* y de las demás métricas.

## Hiper-parámetros para los casos de selección

Los parámetros de un modelo son las características que va aprendiendo durante el entrenamiento (como los pesos en una NN), mientras que los hiper-parámetros pueden verse como los ajustes del modelo para que este pueda resolver de manera óptima el problema de aprendizaje automático por el que se le creó. En el proceso de selección del modelo, la elección de transformaciones sobre el dataset puede reconocerse como otro hiper-parámetro, especialmente cuando no hay seguridad acerca del efecto de estas modificaciones sobre el comportamiento de los datos. En este acápite, se proponen transformaciones que van desde modificaciones de registros hasta la selección de atributos (proceso conocido como *feature selection*) y la ingeniería de características (proceso conocido como *feature engineering*). Estas se tienen en cuenta por ser muy comunes en el marco de creación de modelos de clasificación y, por tanto, en este estudio se considera que establecen hipótesis sobre el buen funcionamiento de sea cual sea el modelo creado. Se van escalando desde los menores niveles de complejidad en cuanto a la transformación del conjunto de datos y de cada variable de forma específica.

Al final del proceso de exploración de modelos, se evalúan los posibles modelos y se construye un modelo para cada posible combinación de transformaciones de los datos. Luego, se selecciona la arquitectura de datos que produce los mejores resultados. De forma interna al proceso de exploración, se realiza la optimización de cada modelo mediante la búsqueda de la configuración óptima que permita mejorar la métrica seleccionada, en la aplicación del modelo a los datos. Este proceso de mostrará en el siguiente capítulo.

Luego de la limpieza realizada al conjunto de datos, hay una transformación que es inherente a este tipo de proceso de aprendizaje y que se mantiene en todos los casos: la aplicación de **One Hot Encoder**<sup>10</sup> a los variables categóricas. Los hiper-parámetros del proceso de selección de modelos, serán las transformaciones presentadas a continuación por nivel de consideración y uso. Estas transformaciones se organizaron en forma ascendente en cuanto al nivel de modificación de los datos y se agrupan en tres etapas con grupos de transformaciones. La selección de cada etapa se basó en el primer acercamiento a los datos (el ejercicio de *Prácticas Laborales del Máster*) donde se realizaron pruebas de modelado con todas estas transformaciones de forma independiente y los resultados fueron muy cercanos para algunas de ellas.

**Primera** Corrección de valores atípicos y removimiento de colinearidad. Descartar los atributos que tienen multicolinealidad, que son los que presentan alta correlación entre sí.

**Segunda** Estandarización, normalización, data split stratify (que consiste en la separación del los datos del conjunto de entrenamiento y del de prueba, en función de la clase que presenta cada registro) y ponderación de la salida en los casos de las variables de

---

<sup>10</sup>OHE: Método de codificación con el que se crean columnas para cada valor distinto de una característica específica, que para cada registro marca con un 1 la columna a la que pertenece dicho registro mientras las demás quedan en 0

salida **Mortalidad Hospitalaria** y **UCI**, para corregir el desbalance de las clases (poniendo al 50 % ambas clases).

**Tercera** Selección de atributos, agrupación de atributos y creación de nuevas variables.

### Pre-selección de modelos

A partir de la entrada de los datos del conjunto de entrenamiento, se seleccionan los mejores modelos utilizando validación cruzada y se comparan los resultados en base a las medidas seleccionadas (*Accuracy*, *AUC* y *Recall*).

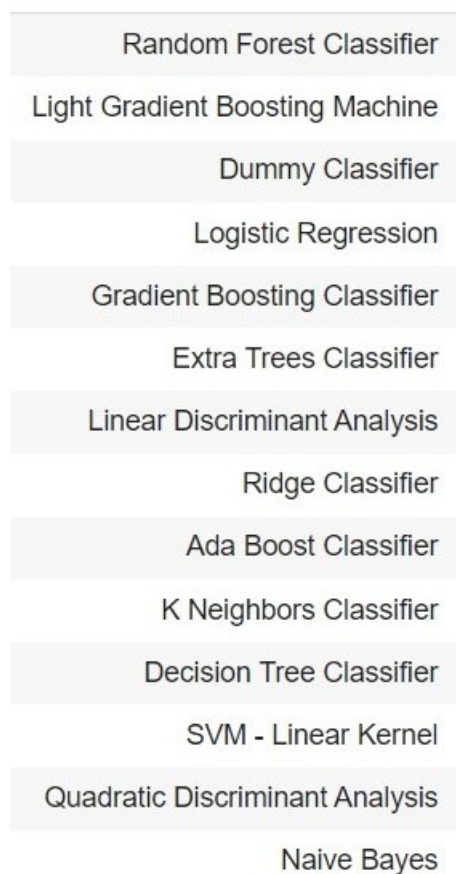


Figura 3.66: Modelos de clasificación incluidos en la biblioteca de *pyCaret*

- Entrenar muchos modelos rápidos, utilizando parámetros estándar de diferentes categorías (Ver Figura 3.66).
- Medir y comparar el desempeño de estos modelos de modo que, para cada uno, se utiliza la validación cruzada de 10 subconjuntos y se calcula la media y la desviación estándar de las medidas de rendimiento en las 10 evaluaciones.
- Seleccionar los mejores modelos para cada prueba y cada clasificación.

## Afinamiento de modelos

El proceso de optimización más bien consiste en encontrar una combinación de hiper-parámetros que devuelva un modelo capaz de reducir las medidas de pérdidas predefinidas y, así, aumentar la precisión en datos independientes. La propia biblioteca *pyCaret* tiene incluidas las medidas de evaluación que ya se han mencionado (que aunque se prioricen las seleccionadas no se minimiza la importancia del resto), predefine un conjunto de parámetros para cada modelo de clasificación binaria, recorre en iteración varias combinaciones y mide la precisión y el resto de métricas hasta que encuentra un modelo “mejor”. De forma que la métrica seleccionada debe ser inclusiva en cuanto a buenos resultados en las clasificaciones a realizar, además, se prefiere que luego del proceso de optimización las mejoras realizadas en función de esta métrica permitan también mejoras en el resto de las métricas de interés.

## Selección de modelos

Se toman los mejores modelos (1, 2 o 3 dependiendo de su rendimiento) y se ajustan los hiper-parámetros a partir de métricas seleccionadas en esta etapa (con *Accuracy* y *AUC*), tras establecer hipótesis fundamentadas en la experiencia técnica y la consulta bibliográfica realizada. Con este modelo final, se mide su rendimiento en el conjunto de prueba para estimar el error de generalización. En esta etapa sí que se mide, además de las métricas anteriores, *Recall*.

## 3.4. Conclusión del capítulo

En este capítulo se evidenciaron los pasos seguidos en cuanto a la exploración, análisis y estudio de los datos, así como se comentó el resto de la metodología seguida para llegar a seleccionar un modelo de clasificación para el caso médico que se está estudiando.

Los datos evidenciaron estar bastante contaminados por razones que fueron comentadas con el médico responsable en uno de los encuentros tenidos: la toma de decisiones equivocadas en la atención del ED (esta situación fue mencionada como típica de este espacio médico por uno de los artículos referenciados en este estudio), la equivocación en el paso de la información de los registros médicos de los hospitales a una base de datos digital (llevando a mover valores entre columnas o entre filas), así como la interpretación personal que haga el médico de la información que está manipulando (que pudo ser decisión del médico que tomó el registro del paciente en el ED o del médico que completó la base de datos sin tener en cuenta que la puede usar otra persona ajena al espacio médico).

Cada información ganada a partir del primer y extenso acercamiento a los datos, fue causa de cada decisión toma para establecer los procesos de limpieza y transformación de los mismos. Este último proceso fue desarrollado por etapas en las que se aplicaron, gradualmente, los grupos de transformaciones para, de igual forma, aplicar los modelos de clasificación y guiar así la exploración de modelos, en donde, desde el primer momento, se valoraron las medidas de rendimiento específicas con las que seleccionar y optimizar el modelo final.

---

## 4: Resultados y Discusión

---

Luego de fundamentar el *cómo se hizo* de la etapa de exploración y selección de modelos, en el capítulo anterior, en el capítulo actual toca presentar el *qué se hizo*, mostrando los resultados obtenidos para cada situación de prueba sobre los datos a lo largo del proceso de búsqueda. Esta búsqueda es guiada por grupos de transformaciones aplicadas a los datos, identificadas como tres casos de pruebas dentro del proceso de exploración y selección de modelos. Al final de este capítulo, se podrá comentar la decisión final alrededor de los modelos.

### 4.1. Resultados

Habiendo establecido los casos de prueba en el capítulo anterior (también llamados hiper-parámetros de la exploración y selección de modelos), se presentan los resultados obtenidos y se fundamentan las decisiones tomadas para cada uno. Cada caso de prueba se nombra **Iteración** y se identifica con el número del orden en que se presenta en el capítulo anterior, además de que en cada uno se dividen los resultados por las 3 clasificaciones pedidas por el médico responsable de los datos. Internamente en cada una de estas divisiones: se toman los modelos de mejor rendimiento en la prueba rápida y se optimizan (todo según las métricas tomadas como fundamentales: *Accuracy*, *AUC* y *Recall*); se evalúan estos modelos valorando las matrices de confusión correspondientes; se selecciona el mejor modelo del grupo y se evalúa la predicción que éste realiza sobre el conjunto de prueba. Cuando se obtiene un modelo para cada hiper-parámetro de la exploración de modelos, se comparan los mejores modelos obtenidos para cada tipo de clasificación y este es el que se guarda como resultado final.

#### Iteración 1

La primera prueba se basó en usar los datos ya limpios (o lo más limpios posible), agregándoles la corrección de valores atípicos y descartando los atributos que no proporcionan información útil para el proyecto y los que tienen multicolinealidad (que son los que presentan alta correlación entre sí).

## “Hospitalización”

La Figura 4.67 muestra la comparación de rendimiento de los modelos de clasificación binaria que comprende la biblioteca *pyCaret*, usada para este proceso, aplicados en forma rápida sobre los datos. Esta comparación es el resultado de realizar validación cruzada con 10 pruebas por cada modelo.

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>rf</b>	Random Forest Classifier	0.8133	0.7099	0.1167	0.5500	0.1858	0.1453	0.2058	0.481
<b>lightgbm</b>	Light Gradient Boosting Machine	0.8074	0.7149	0.3000	0.5750	0.3707	0.2745	0.3031	0.125
<b>dummy</b>	Dummy Classifier	0.7982	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.019
<b>lr</b>	Logistic Regression	0.7981	0.7251	0.1952	0.5317	0.2740	0.1893	0.2293	0.528
<b>gbc</b>	Gradient Boosting Classifier	0.7981	0.6659	0.2881	0.6470	0.3518	0.2543	0.3040	0.125
<b>et</b>	Extra Trees Classifier	0.7921	0.7167	0.1143	0.3500	0.1644	0.0980	0.1198	0.461
<b>lda</b>	Linear Discriminant Analysis	0.7893	0.7111	0.2810	0.5005	0.3401	0.2287	0.2514	0.017
<b>ridge</b>	Ridge Classifier	0.7890	0.0000	0.1310	0.3900	0.1864	0.1110	0.1397	0.014
<b>ada</b>	Ada Boost Classifier	0.7804	0.6178	0.3476	0.4650	0.3870	0.2589	0.2683	0.105
<b>knn</b>	K Neighbors Classifier	0.7621	0.5489	0.0286	0.1500	0.0472	-0.0272	-0.0290	0.117
<b>dt</b>	Decision Tree Classifier	0.7561	0.6067	0.3571	0.3907	0.3596	0.2150	0.2207	0.016
<b>svm</b>	SVM - Linear Kernel	0.5889	0.0000	0.3571	0.1761	0.1979	0.0210	0.0197	0.017
<b>qda</b>	Quadratic Discriminant Analysis	0.5696	0.5110	0.4071	0.2552	0.2931	0.0532	0.0348	0.019
<b>nb</b>	Naive Bayes	0.4670	0.6599	0.6857	0.2437	0.3426	0.0717	0.0562	0.014

Figura 4.67: Rendimiento de los modelos rápidos en “Hospitalización” en el *round 1*

De estos modelos, se seleccionaron para su posterior análisis y una comparación más exhaustiva, el *Random Forest Classifier*, el *Light Gradient Boosting Machine* y el *Logistic Regression*. Estos fueron de los cuatro mejores resultados en cuanto a las métricas que comprenden el reconocimiento de las dos clases de salida. Fueron elegidos, porque el *Dummy Classifier*, que es el tercero en mejor resultado en cuanto a *Acuraccy*, presenta un valor de *AUC* bajo y el *Recall* está en 0.

Se optimizan los modelos en base a las métricas *Acuraccy* y *AUC*, indistintamente, y son ajustados con los datos de entrenamiento. Luego, verificando el reporte de clasificación para el conjunto de prueba, ver Figura 4.68, se evidencia que el mejor resultado corresponde siempre al modelo ajustado con la métrica *AUC*; así como que el mejor resultado es el obtenido con el modelo basado en el *Random Forest Classifier*.

Como era de esperar al visualizar la Figura 4.68, la Figura 4.70 muestra la matriz de confusión de un modelo capaz de reconocer, en al menos el 52% de los casos, la clase



```
predict_model(tuned_rfc_1_auc);
```

Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0 Random Forest Classifier	0.8182	0.8036	0.5806	0.5806	0.5806	0.4646	0.4646

```
predict_model(tuned_rfc_1_acc);
```

Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0 Random Forest Classifier	0.7902	0.8059	0.0645	0.6667	0.1176	0.0825	0.1598

```
predict_model(tuned_lr_1_acc);
```

Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0 Light Gradient Boosting Machine	0.7483	0.6391	0.2581	0.3810	0.3077	0.1607	0.1653

```
predict_model(tuned_lr_1_auc);
```

Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0 Light Gradient Boosting Machine	0.7832	0.7742	0.0000	0.0000	0.0000	0.0000	0.0000

Figura 4.68: Rendimiento de los modelos *Random Forest Classifier* y *Light Gradient Boosting Machine* en la predicción sobre el conjunto de prueba, para “Hospitalización”

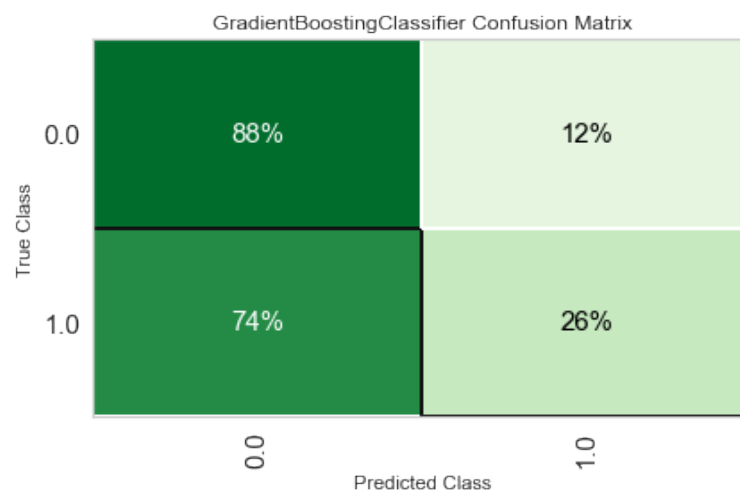


Figura 4.69: Matriz de confusión de “Hospitalización” **Logistic Regression**

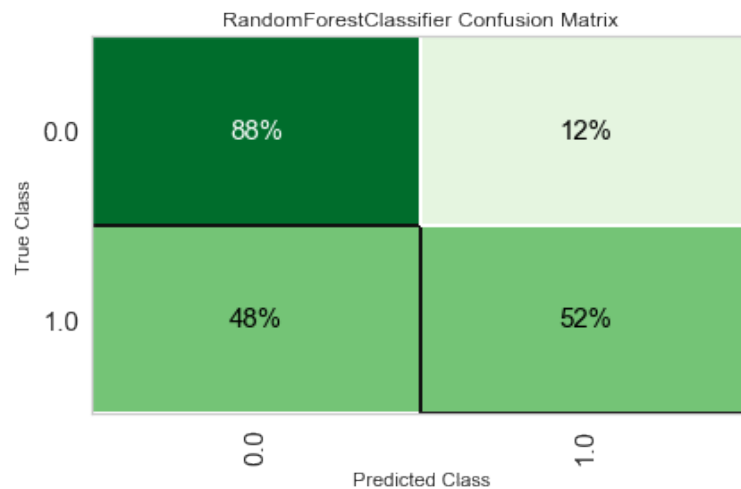


Figura 4.70: Matriz de confusión de “Hospitalización” para **Random Forest Classifier**

positiva como sí misma; mientras tiene un valor alto del 88 % de los casos de clasificación acertada en la clase negativa. Esta relación es superior a la que muestra la Figura 4.69, donde la clase positiva se clasifica acertadamente el 26 % de las veces que aparece en los registros.

#### “Mortalidad Hospitalaria”

En la Figura 4.71 se muestra la tabla de comparación de los rendimientos de las pruebas rápidas con todos los modelos. Aunque casi todos los modelos tengan un valor de *Accuracy* casi perfecto, los valores de *AUC* son muy bajos y los del resto de las métricas son completamente nulos, dando a entender que la clasificación favorecería completamente a la clase negativa. Específicamente, la métrica *MCC* (que tiene en cuenta los cuatro valores en la matriz de confusión) indica con el valor 0 que los modelos comparan a la posible clasificación a realizar, con el lanzamiento de una moneda al aire. El resultado de estas medidas implica que ningún modelo cumple con las características mínimas necesarias para ser creado en el objetivo planteado.

#### “UCI”

Al igual que en el caso de la clasificación anterior, para esta clasificación no es viable crear modelo alguno, pues las medidas de rendimiento en la Figura 4.72, indican un total desbalance de la clasificación futura.

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>knn</b>	K Neighbors Classifier	0.9820	0.2737	0.0000	0.0000	0.0000	nan	0.0000	0.0140
<b>svm</b>	SVM - Linear Kernel	0.9820	0.0000	0.0000	0.0000	0.0000	nan	0.0000	0.0060
<b>rf</b>	Random Forest Classifier	0.9820	0.4118	0.0000	0.0000	0.0000	nan	0.0000	0.0940
<b>et</b>	Extra Trees Classifier	0.9820	0.2140	0.0000	0.0000	0.0000	nan	0.0000	0.0960
<b>xgboost</b>	Extreme Gradient Boosting	0.9820	0.2832	0.0000	0.0000	0.0000	nan	0.0000	0.0650
<b>lightgbm</b>	Light Gradient Boosting Machine	0.9820	0.2595	0.0000	0.0000	0.0000	nan	0.0000	0.0220
<b>dummy</b>	Dummy Classifier	0.9820	0.3000	0.0000	0.0000	0.0000	nan	0.0000	0.0050
<b>lr</b>	Logistic Regression	0.9790	0.2540	0.0000	0.0000	0.0000	nan	0.0000	0.6410
<b>ridge</b>	Ridge Classifier	0.9790	0.0000	0.0000	0.0000	0.0000	nan	-0.0031	0.0060
<b>ada</b>	Ada Boost Classifier	0.9790	0.2723	0.0000	0.0000	0.0000	nan	0.0000	0.0480
<b>lda</b>	Linear Discriminant Analysis	0.9729	0.2451	0.0000	0.0000	0.0000	nan	-0.0062	0.0070
<b>gbc</b>	Gradient Boosting Classifier	0.9701	0.2804	0.0000	0.0000	0.0000	nan	-0.0092	0.0450
<b>dt</b>	Decision Tree Classifier	0.9610	0.2907	0.0000	0.0000	0.0000	nan	-0.0148	0.0060
<b>nb</b>	Naive Bayes	0.4336	0.2346	0.2000	0.0103	0.0195	-0.0155	-0.0448	0.0060
<b>qda</b>	Quadratic Discriminant Analysis	0.0964	0.2920	0.5000	0.0175	0.0338	-0.0011	-0.0043	0.0080

Figura 4.71: Rendimiento de los modelos rápidos en “Mortalidad Hospitalaria” en el *round 1*

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>knn</b>	K Neighbors Classifier	0.9669	0.5018	0.0000	0.0000	0.0000	0.0000	0.0000	0.0110
<b>svm</b>	SVM - Linear Kernel	0.9669	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0050
<b>rf</b>	Random Forest Classifier	0.9669	0.5873	0.0000	0.0000	0.0000	0.0000	0.0000	0.0870
<b>et</b>	Extra Trees Classifier	0.9669	0.5799	0.0000	0.0000	0.0000	0.0000	0.0000	0.0970
<b>lightgbm</b>	Light Gradient Boosting Machine	0.9669	0.5352	0.0000	0.0000	0.0000	0.0000	0.0000	0.0200
<b>dummy</b>	Dummy Classifier	0.9669	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0040
<b>lr</b>	Logistic Regression	0.9640	0.4248	0.0000	0.0000	0.0000	-0.0041	-0.0044	0.5970
<b>ridge</b>	Ridge Classifier	0.9639	0.0000	0.0000	0.0000	0.0000	-0.0031	-0.0031	0.0050
<b>xgboost</b>	Extreme Gradient Boosting	0.9639	0.6528	0.0000	0.0000	0.0000	-0.0031	-0.0031	0.0740
<b>ada</b>	Ada Boost Classifier	0.9578	0.3696	0.0000	0.0000	0.0000	-0.0094	-0.0094	0.0460
<b>gbc</b>	Gradient Boosting Classifier	0.9578	0.5050	0.0000	0.0000	0.0000	-0.0094	-0.0094	0.0430
<b>lda</b>	Linear Discriminant Analysis	0.9367	0.5696	0.0000	0.0000	0.0000	-0.0265	-0.0274	0.0070
<b>dt</b>	Decision Tree Classifier	0.9126	0.4719	0.0000	0.0000	0.0000	-0.0358	-0.0394	0.0060
<b>qda</b>	Quadratic Discriminant Analysis	0.2386	0.5577	0.9000	0.0396	0.0757	0.0132	0.0503	0.0090
<b>nb</b>	Naive Bayes	0.1803	0.3872	0.6500	0.0263	0.0503	-0.0140	-0.1127	0.0060

Figura 4.72: Rendimiento de los modelos rápidos en “UCI” en el *round 1*

## Iteración 2

Para la segunda prueba, se aplicaron las transformaciones señaladas en el capítulo anterior (data split stratify, normalize y transformar fix imbalance), agregadas a las primeras.

### “Hospitalización”

En la Figura 4.73 se muestra la Tabla que compara los resultados de la validación cruzada de los modelos de clasificación aplicados a los datos de entrenamiento. Por los resultados obtenidos, se seleccionan los modelos *Extra Trees Classifier*, *Light Gradient Boosting Machine* y *Random Forest Classifier* por tener un consenso de mejores resultados en las métricas *Accuracy*, *AUC*, *Recall*.

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>et</b>	Extra Trees Classifier	0.7799	0.7232	0.2786	0.5488	0.3545	0.2417	0.2683	0.463
<b>dummy</b>	Dummy Classifier	0.7742	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.019
<b>lightgbm</b>	Light Gradient Boosting Machine	0.7617	0.7187	0.3714	0.4721	0.4059	0.2642	0.2703	0.070
<b>rf</b>	Random Forest Classifier	0.7557	0.7102	0.2518	0.4619	0.3166	0.1828	0.2002	0.514
<b>gbc</b>	Gradient Boosting Classifier	0.7525	0.7349	0.3857	0.4508	0.4014	0.2521	0.2591	0.189
<b>ada</b>	Ada Boost Classifier	0.7226	0.7152	0.3929	0.4067	0.3904	0.2142	0.2186	0.127
<b>dt</b>	Decision Tree Classifier	0.6865	0.5934	0.4250	0.3363	0.3687	0.1666	0.1716	0.025
<b>ridge</b>	Ridge Classifier	0.6591	0.0000	0.4714	0.3126	0.3737	0.1513	0.1597	0.022
<b>lda</b>	Linear Discriminant Analysis	0.6590	0.6381	0.4839	0.3172	0.3806	0.1576	0.1678	0.026
<b>lr</b>	Logistic Regression	0.6530	0.6891	0.5125	0.3243	0.3927	0.1686	0.1792	0.041
<b>svm</b>	SVM - Linear Kernel	0.6296	0.0000	0.5339	0.3283	0.3941	0.1586	0.1725	0.023
<b>knn</b>	K Neighbors Classifier	0.5803	0.6056	0.5482	0.2921	0.3798	0.1122	0.1178	0.122
<b>qda</b>	Quadratic Discriminant Analysis	0.4439	0.4638	0.5446	0.1284	0.2078	-0.0220	-0.0708	0.025
<b>nb</b>	Naive Bayes	0.2196	0.6445	0.9196	0.2137	0.3465	-0.0305	-0.1087	0.024

Figura 4.73: Rendimiento de los modelos rápidos en “Hospitalización” en el *round 2*

Se crearon los modelos de los algoritmos seleccionados. El rendimiento de su optimización (en base a la métrica *AUC* de mejor resultado en el round 1) se presenta en la Figura 4.74.

El rendimiento de los modelos seleccionados prediciendo la clasificación sobre el conjunto de prueba, es mostrado en la Figura 4.75. En este caso, los mejores resultados los presenta el modelo que se basa en el algoritmo *Extra Trees Classifier*.

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	Extra Trees Classifier	0.7105	0.7317	0.3786	0.3508	0.3592	0.1811	0.1792

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	Light Gradient Boosting Machine	0.7193	0.7253	0.4768	0.4113	0.4313	0.2515	0.2568

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	Random Forest Classifier	0.7526	0.7231	0.3857	0.4534	0.4084	0.2558	0.2609

Figura 4.74: Rendimiento de los modelos optimizados, para “Hospitalización”

```
predict_model(tuned_et);
```

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	Extra Trees Classifier	0.7133	0.6926	0.5152	0.4048	0.4533	0.2628	0.2663

```
predict_model(tuned_li);
```

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	Light Gradient Boosting Machine	0.6923	0.689	0.3939	0.3514	0.3714	0.1686	0.1691

```
predict_model(tuned_rf);
```

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	Random Forest Classifier	0.7273	0.6857	0.303	0.3846	0.339	0.1702	0.1721

Figura 4.75: Rendimiento en la predicción sobre el conjunto de prueba, para “Hospitalización”

Las matrices de confusión respectivas (ver Figura 4.76, Figura 4.77 y Figura 4.78) a los modelos seleccionados, refutan lo ya visto en la Figura 4.75: el mejor modelo en esta prueba ha resultado ser el *Extra Trees Classifier*.

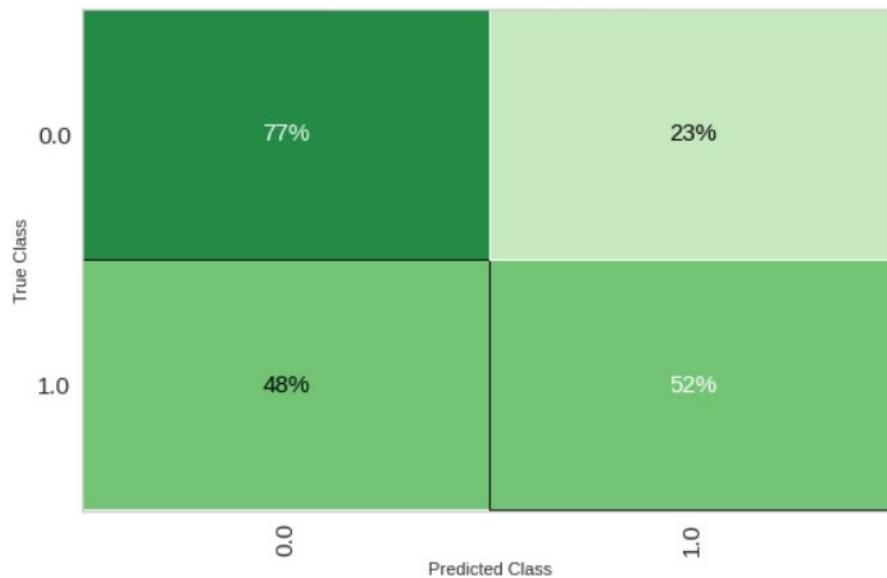


Figura 4.76: Matriz de confusión de “Hospitalización” **Extra Trees Classifier**

En 4.76, la predicción a favor de la clase positiva en el espacio de esta misma clase, está por encima del 50% lo que hace que este resultado sea mejor que el resto.

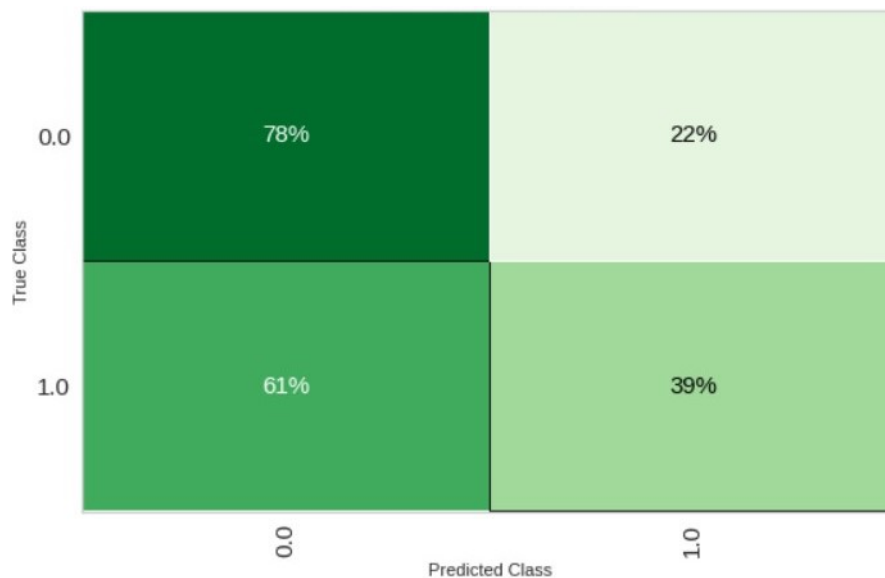


Figura 4.77: Matriz de confusión de “Hospitalización” **Light Gradient Boosting Machine**

“UCI”

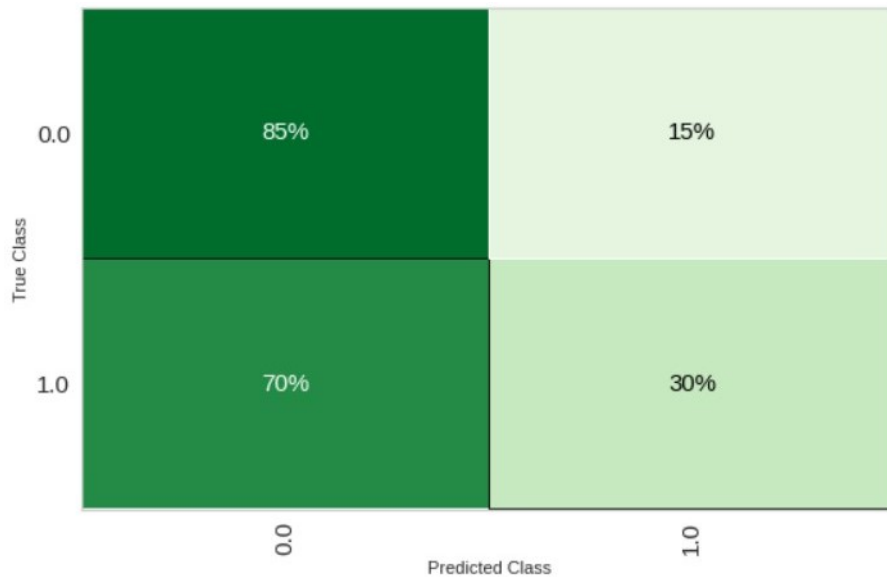


Figura 4.78: Matriz de confusión de “Hospitalización” **Random Forest**

Respecto a las variable **UCI**, la Figura 4.79 muestra valores positivamente altos de *Accuracy* pero bajos de *AUC* y muy bajos de *Recall*. Por lo que, de nuevo, no se llegó a crear ningún modelo, avizorando resultados deficientes.

#### “Mortalidad Hospitalaria”

Al obtener resultados similares a los de **UCI** en la prueba rápida de los modelos sobre el conjunto de entrenamiento, se llegó a crear el modelo correspondiente al algoritmo de mejor resultado en cuanto a *Accuracy*. La Figura 4.80 presenta la matriz de confusión del algoritmo **Discriminant Analysis**, que es también el mejor modelo de la prueba rápida en **UCI**.

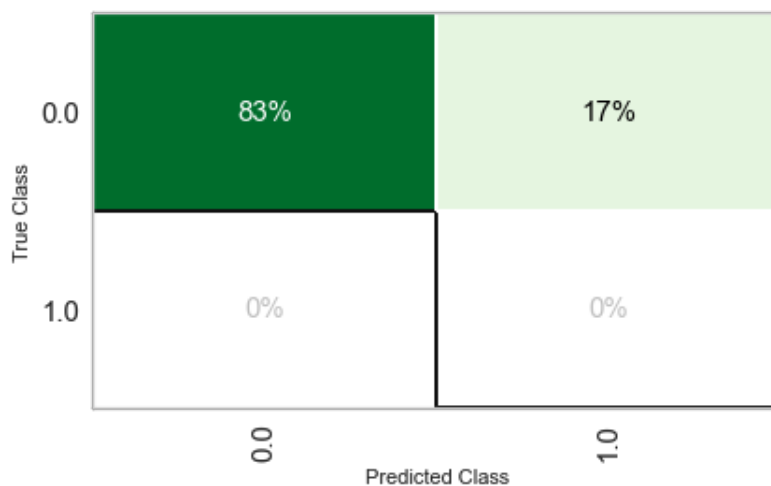
Se puede comprobar que el rendimiento de la prueba rápida indica que todos los modelos tendrían malas predicciones. En la Figura 4.80 no se llega a predecir la clase positiva.

### Iteración 3

En el tercer grupo de transformaciones están comprendidas la selección y agrupación de atributos y la creación de nuevas variables (mediante las funciones de la biblioteca *pyCaret*: feature selection, polynomial feature, create clusters). Estas transformaciones se unen a las de la etapa anterior y se aplican sobre el conjunto de entrenamiento.

#### “Hospitalización”

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>qda</b>	Quadratic Discriminant Analysis	0.9760	0.4485	0.1	0.1000	0.1000	NaN	0.0970	0.023
<b>rf</b>	Random Forest Classifier	0.9759	0.6092	0.0	0.0000	0.0000	NaN	0.0000	0.507
<b>ada</b>	Ada Boost Classifier	0.9759	0.4825	0.1	0.1000	0.1000	NaN	0.1000	0.131
<b>dummy</b>	Dummy Classifier	0.9759	0.4000	0.0	0.0000	0.0000	NaN	0.0000	0.026
<b>lightgbm</b>	Light Gradient Boosting Machine	0.9729	0.4497	0.0	0.0000	0.0000	NaN	-0.0031	0.168
<b>et</b>	Extra Trees Classifier	0.9699	0.5730	0.0	0.0000	0.0000	NaN	0.0000	0.461
<b>gbc</b>	Gradient Boosting Classifier	0.9638	0.4885	0.0	0.0000	0.0000	NaN	-0.0062	0.237
<b>dt</b>	Decision Tree Classifier	0.9428	0.4392	0.1	0.0500	0.0667	NaN	0.0546	0.025
<b>svm</b>	SVM - Linear Kernel	0.9065	0.0000	0.3	0.0783	0.1233	0.0946	0.1236	0.020
<b>lr</b>	Logistic Regression	0.9037	0.4831	0.4	0.1010	0.1583	0.1308	0.1716	0.291
<b>ridge</b>	Ridge Classifier	0.8883	0.0000	0.4	0.1067	0.1667	0.1397	0.1759	0.020
<b>lda</b>	Linear Discriminant Analysis	0.8854	0.4555	0.3	0.0750	0.1186	0.0900	0.1170	0.024
<b>knn</b>	K Neighbors Classifier	0.8101	0.3835	0.1	0.0111	0.0200	-0.0206	-0.0233	0.120
<b>nb</b>	Naive Bayes	0.1508	0.2554	0.5	0.0176	0.0340	-0.0132	-0.0749	0.021

Figura 4.79: Rendimiento de los modelos rápidos en “UCI” en el *round 2*Figura 4.80: Matriz de confusión de “Mortalidad Hospitalaria” **Discriminant Analysis**



Como en las pruebas anteriores, se realiza la comparación de modelos mediante validación cruzada en forma de una prueba rápida, cuyo resultado se muestra en la Figura 4.81.

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>et</b>	Extra Trees Classifier	0.7800	0.6889	0.2143	0.5733	0.3044	0.2040	0.2436	0.468
<b>dummy</b>	Dummy Classifier	0.7742	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.018
<b>rf</b>	Random Forest Classifier	0.7531	0.6917	0.2518	0.4405	0.3114	0.1754	0.1896	0.509
<b>gbc</b>	Gradient Boosting Classifier	0.7528	0.6815	0.3571	0.4753	0.4008	0.2490	0.2572	0.192
<b>lightgbm</b>	Light Gradient Boosting Machine	0.7470	0.6775	0.3321	0.4232	0.3649	0.2128	0.2178	0.068
<b>ada</b>	Ada Boost Classifier	0.7440	0.6644	0.4232	0.4576	0.4336	0.2698	0.2731	0.131
<b>dt</b>	Decision Tree Classifier	0.7143	0.6208	0.4518	0.3876	0.4137	0.2273	0.2304	0.027
<b>svm</b>	SVM - Linear Kernel	0.6718	0.0000	0.5161	0.3473	0.4001	0.1902	0.2109	0.026
<b>lr</b>	Logistic Regression	0.6596	0.6965	0.5589	0.3501	0.4278	0.2060	0.2178	0.044
<b>ridge</b>	Ridge Classifier	0.6537	0.0000	0.5589	0.3406	0.4223	0.1962	0.2083	0.023
<b>lda</b>	Linear Discriminant Analysis	0.6537	0.6711	0.5625	0.3454	0.4261	0.2002	0.2119	0.028
<b>knn</b>	K Neighbors Classifier	0.5990	0.5879	0.5661	0.2964	0.3873	0.1304	0.1472	0.123
<b>qda</b>	Quadratic Discriminant Analysis	0.2684	0.5015	0.8714	0.2188	0.3490	-0.0181	-0.0521	0.024
<b>nb</b>	Naive Bayes	0.2441	0.5387	0.8411	0.2082	0.3335	-0.0443	-0.1194	0.023

Figura 4.81: Rendimiento de los modelos rápidos en “Hospitalización” en el *round 3*

Para la medida de rendimiento *Accuracy*, se muestran ordenados los mejores modelos. De ellos se seleccionaron **Extra Trees Classifier**, **Random Forest Classifier** y **Gradient Boosting Classifier**, teniendo en cuenta los resultados en las tres métricas fundamentales.

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
<b>0</b>	Extra Trees Classifier	0.6783	0.7292	0.4242	0.3415	0.3784	0.1648	0.1666
	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
<b>0</b>	Random Forest Classifier	0.6993	0.6771	0.3333	0.3438	0.3385	0.144	0.144
	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
<b>0</b>	Gradient Boosting Classifier	0.7413	0.6501	0.2121	0.3889	0.2745	0.1333	0.1424

Figura 4.82: Rendimiento de los modelos del *round3* en la predicción sobre el conjunto de prueba, para “Hospitalización”

Luego de optimizar los modelos seleccionados, se aplicaron sobre el conjunto de prueba y se obtuvieron los resultados que muestra la Figura 4.82.

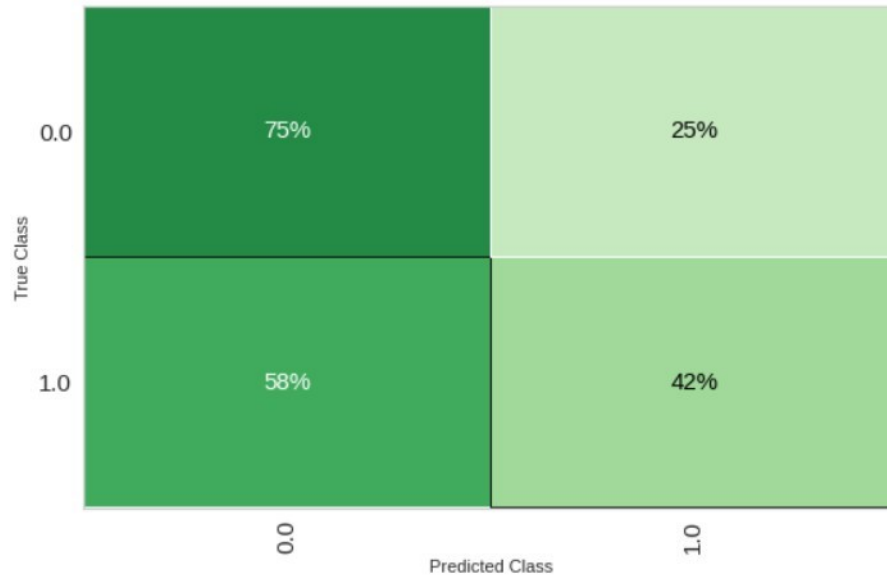


Figura 4.83: Matriz de confusión de “Mortalidad Hospitalaria” **Extra Trees Classifier**

Las matrices de confusión correspondientes a cada modelo, se presentan en Figura 4.83, Figura 4.84 y Figura 4.85.

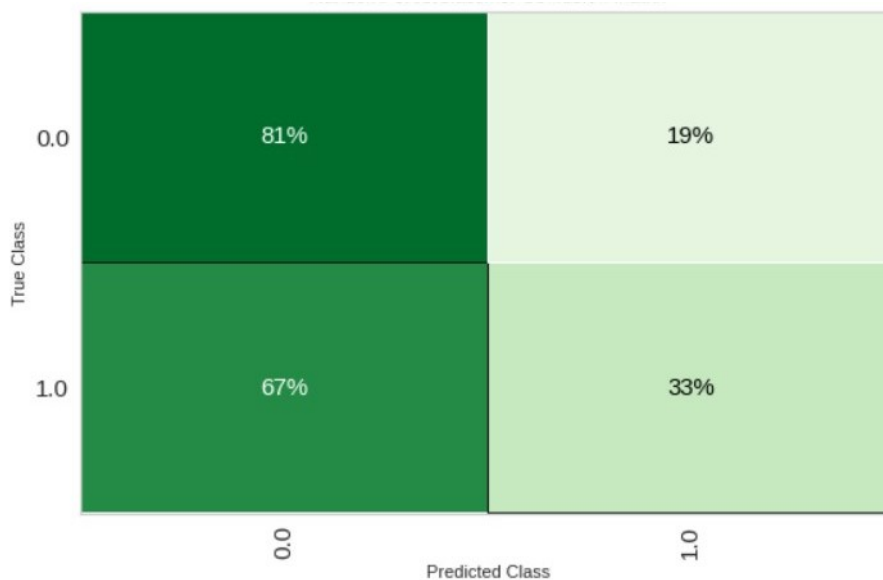


Figura 4.84: Matriz de confusión de “Mortalidad Hospitalaria” **Random Forest Classifier**

El modelo de la Figura 4.83, muestra un resultado regular a favor de la clase negativa y un resultado del 42% de clasificación a favor de la clase positiva. El segundo modelo

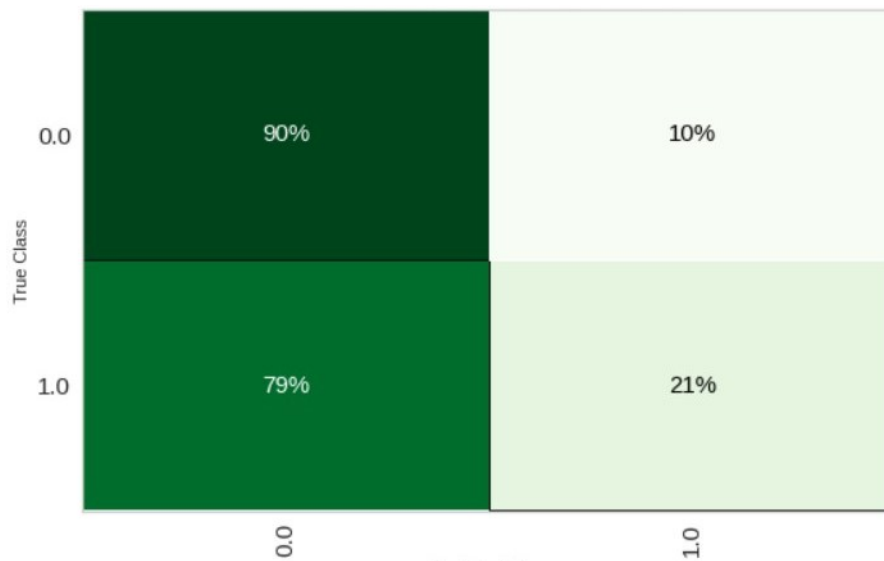


Figura 4.85: Matriz de confusión de “Mortalidad Hospitalaria” **Gradient Boosting Classifier**

(ver Figura 4.84) muestra un resultado un poco más alto en la clase negativa y un poco más bajo en la positiva. El tercer modelo (ver Figura 4.85) muestra un resultado alto de clasificación a favor de la clase negativa y muy bajo a favor de la clase positiva. Tras evaluar las métricas seleccionadas sobre el conjunto de prueba (ver Figura 4.82), el mejor modelo en esta ocasión es el **Extra Trees Classifier**.

#### “Mortalidad Hospitalaria” y “UCI”

En el caso de las variables de salida “**Mortalidad Hospitalaria**” y “**UCI**”, se obtienen resultados similares con respecto al grupo de transformaciones anterior, por lo que no se creó modelo alguno para estas salidas.

En la Figura 4.86, se muestra la tabla de comparación de resultados rápidos de rendimiento de los modelos sobre el conjunto de entrenamiento, para la salida “**UCI**”.

## 4.2. Selección de modelo por clasificación

Habiendo revisado los resultados obtenidos en las secciones anteriores, sólo es posible seleccionar un modelo concluyente, para la variable de salida “**Hospitalización**” y este es el mejor modelo de la primera prueba, el cuál fue analizado y guardado como archivo *pkl*. La implementación guardada como el mejor modelo corresponde a la de un **RandomForestClassifier**, del cual se muestran los parámetros en la Figura 4.87.

De estos parámetros se definen algunos como propios del **RandomForestClassifier** [20]:

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>qda</b>	Quadratic Discriminant Analysis	0.9790	0.4500	0.1	0.1000	0.1000	NaN	0.1000	0.025
<b>rf</b>	Random Forest Classifier	0.9759	0.3576	0.0	0.0000	0.0000	NaN	0.0000	0.518
<b>dummy</b>	Dummy Classifier	0.9759	0.4000	0.0	0.0000	0.0000	NaN	0.0000	0.020
<b>lightgbm</b>	Light Gradient Boosting Machine	0.9730	0.3869	0.0	0.0000	0.0000	NaN	-0.0030	0.138
<b>et</b>	Extra Trees Classifier	0.9699	0.4444	0.0	0.0000	0.0000	NaN	-0.0062	0.483
<b>gbc</b>	Gradient Boosting Classifier	0.9669	0.2887	0.0	0.0000	0.0000	NaN	-0.0093	0.244
<b>ada</b>	Ada Boost Classifier	0.9639	0.2173	0.0	0.0000	0.0000	NaN	-0.0062	0.137
<b>dt</b>	Decision Tree Classifier	0.9246	0.4282	0.1	0.0500	0.0667	0.0434	0.0444	0.026
<b>svm</b>	SVM - Linear Kernel	0.8701	0.0000	0.1	0.0200	0.0333	-0.0029	0.0008	0.024
<b>lr</b>	Logistic Regression	0.8520	0.4129	0.0	0.0000	0.0000	-0.0385	-0.0526	0.044
<b>lda</b>	Linear Discriminant Analysis	0.8220	0.4033	0.2	0.0291	0.0500	0.0114	0.0247	0.024
<b>ridge</b>	Ridge Classifier	0.8190	0.0000	0.2	0.0291	0.0500	0.0110	0.0230	0.022
<b>knn</b>	K Neighbors Classifier	0.6501	0.3138	0.0	0.0000	0.0000	-0.0461	-0.0938	0.124
<b>nb</b>	Naive Bayes	0.2496	0.3412	0.5	0.0186	0.0359	-0.0110	-0.0389	0.022

Figura 4.86: Rendimiento de los modelos rápidos en “UCI” en el *round 3*

**n\_estimators** Número de árboles que va a tener el bosque aleatorio. Normalmente cuantos más mejor, pero a partir de cierto punto deja de mejorar y sólo hace que vaya más lento. Un buen valor por defecto puede ser el uso de 100 árboles.

**n\_jobs** Número de cores que se pueden usar para entrenar los árboles. Cada árbol es independiente del resto, así que entrenar un bosque aleatorio es una tarea muy paralelizable. Por defecto sólo utiliza 1 core de la CPU. Para mejorar el rendimiento se pueden usar tantos cores como se estimen necesarios. Si se usan `n_jobs = -1`, se está indicando que se quieren usar tantos cores como tenga la máquina.

**max\_features** Una forma de garantizar que los árboles son diferentes, es que todos se entrenan con una muestra aleatoria de los datos. Si queremos que todavía sean más diferentes, podemos hacer que distintos árboles usen distintos atributos. Esto puede ser útil especialmente cuando algunos atributos están relacionados entre sí.

Algunos parámetros de regularización, también disponibles para este clasificador, se definen a continuación [20]:

- **max\_depth** Profundidad máxima del árbol.
- **min\_samples\_split** Número mínimo de muestras necesarias antes de dividir un nodo. Se puede expresar en porcentaje.

Parameters	
<b>bootstrap</b>	False
<b>ccp_alpha</b>	0.0
<b>class_weight</b>	balanced
<b>criterion</b>	gini
<b>max_depth</b>	9
<b>max_features</b>	log2
<b>max_leaf_nodes</b>	None
<b>max_samples</b>	None
<b>min_impurity_decrease</b>	0
<b>min_impurity_split</b>	None
<b>min_samples_leaf</b>	5
<b>min_samples_split</b>	9
<b>min_weight_fraction_leaf</b>	0.0
<b>n_estimators</b>	180
<b>n_jobs</b>	-1
<b>oob_score</b>	False
<b>random_state</b>	7508
<b>verbose</b>	0
<b>warm_start</b>	False

Figura 4.87: Parámetros del mejor modelo para “Hospitalización”

- **min\_samples\_leaf** Número mínimo de muestras que debe haber en un nodo final u hoja. También se puede expresar en porcentaje.
- **max\_leaf\_nodes** Número máximo de nodos finales u hojas.

La importancia de las variables que influyen en este modelo, se muestra en al Figura 4.88. Se comprueba que los resultados del modelo fueron influenciados por muy pocas de las variables introducidas y con bajos valores de influencia en los resultados.

La matriz de confusión de este modelo (ver Figura 4.70) muestra un resultado superior a los demás modelos, de forma que más del 50 % de las predicciones de la clase positiva, fueron acertadas. Este resultado es superior al obtenido hasta este momento con estos

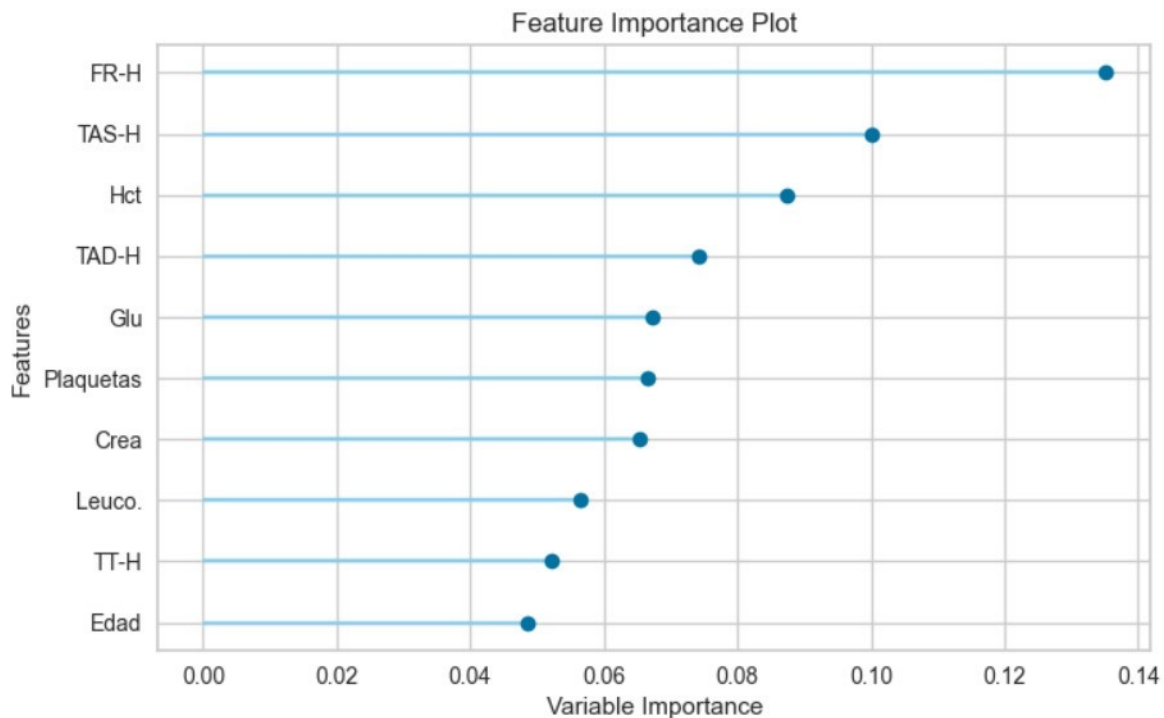


Figura 4.88: Importancia para las variables en el mejor modelo de “Hospitalización”

datos. Para el caso de la clase negativa, se mantiene una predicción muy alta, que ha sido tendencia general en el entrenamiento con estos mismos datos (ver Figura 4.70).

### 4.3. Discusión

Respecto al modelo seleccionado en la clasificación “**Hospitalización**”, la característica más importante resultó ser la *Frecuencia Respiratoria*, influyendo en la decisión del modelo seleccionado con una importancia de 0,14. En el análisis de correlación de variables (ver Figura 3.58), resulta que esta variable tiene el valor más alto de relación con la variable de salida en cuestión, que es un sencillo 0,20584. El resto de variables con correlación positiva con respecto a la salida (siempre por debajo del valor anterior) están en esta gráfica (Figura 4.88) que solo muestra 10 variables de más de 22 (teniendo en cuenta que se crearon nuevas variables tras aplicar OHE a las variables categóricas, específicamente una variable por categoría), contribuyendo poco a la predicción y realzando los errores en la predicción de la clasificación del modelo. Estos errores podrían deberse a la ya mencionada y descrita contaminación de los datos.

Respecto a la clasificación para “**Mortalidad Hospitalaria**” y “**UCI**” no se obtuvo modelo alguno, ni medianamente posible. En la primera prueba realizada, las clases están totalmente desbalanceadas y la clase positiva tiene una representación ínfima en ambos casos de clasificación (casi inexistente) lo que influyó en que los modelos aprendieran

una sola posibilidad de clasificación. En la segunda prueba, se eligen los registros que conformarán los conjuntos de entrenamiento y prueba, seleccionando muestras de ambos tipos de clases y, a pesar de haber sido ponderadas las clases agregándole registros a la de menos representación hasta alcanzar un 50% de los registros generales, no se logró un mejor resultado. Esto puede deberse a que los escasos registros con la clase positiva se caracterizan por ser casos en los que esa solución que representa la clase positiva se manifiesta de forma inusual o inesperada en el servicio de urgencias, o sea, que los predictores indican situaciones similares para todos los pacientes, pero que en algunos casos se puede dar un cambio drástico (no registrado en algún predictor) que propicie la decisión de enviar a un paciente a la UCI (recordemos que es una posibilidad incluida en la “**Hospitalización**”) o que propicie que el paciente fallezca en el tiempo de hospitalización. Teniendo en cuenta lo planteado, la tercera prueba no supone cambio alguno con respecto a la segunda, ya que se crean nuevas variables, pero esto no garantiza, como se pudo comprobar, la creación de tendencias hacia las clases positivas.

Volviendo a “**Hospitalización**”, el resultado de rendimiento del modelo es muy alentador con respecto al obtenido en un estudio anterior sobre los datos, este se muestra en la Figura 4.89, donde solo se lograba obtener una clasificación favorecedora en gran medida de la clase negativa. A pesar de todos los análisis sobre la interpretación médica de cada uno de los tipos de valores de las variables presentadas, en esa ocasión no se eliminaron los valores atípicos que plagaban los datos. En el estudio actual, se eliminaron algunos registros con valores que se consideraron notablemente fuera de lugar y, agregado a esto, con las transformaciones más sencillas sobre el conjunto de datos se obtuvieron los mejores resultados.

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>lr</b>	Logistic Regression	0.7409	0.6890	0.3356	0.6120	0.4222	0.2790	0.3032	0.0370
<b>ridge</b>	Ridge Classifier	0.7409	0.0000	0.3182	0.6095	0.4137	0.2721	0.2959	0.0150
<b>lda</b>	Linear Discriminant Analysis	0.7284	0.6786	0.3356	0.5695	0.4164	0.2578	0.2746	0.0150
<b>rf</b>	Random Forest Classifier	0.7282	0.6933	0.2848	0.6241	0.3760	0.2337	0.2699	0.4820
<b>knn</b>	K Neighbors Classifier	0.7133	0.6328	0.1939	0.5883	0.2797	0.1543	0.1950	0.0240
<b>gbc</b>	Gradient Boosting Classifier	0.7109	0.6828	0.3447	0.5348	0.4121	0.2330	0.2466	0.1930
<b>et</b>	Extra Trees Classifier	0.7108	0.6975	0.2508	0.5210	0.3330	0.1801	0.2006	0.4060
<b>lightgbm</b>	Light Gradient Boosting Machine	0.7034	0.6545	0.3189	0.4915	0.3803	0.2018	0.2111	0.0590
<b>dummy</b>	Dummy Classifier	0.7032	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0120
<b>svm</b>	SVM - Linear Kernel	0.6983	0.0000	0.3841	0.4876	0.3971	0.2193	0.2340	0.0140
<b>dt</b>	Decision Tree Classifier	0.6884	0.6233	0.4629	0.4751	0.4626	0.2463	0.2491	0.0320
<b>nb</b>	Naive Bayes	0.6857	0.6951	0.4939	0.4985	0.4667	0.2558	0.2706	0.0250
<b>ada</b>	Ada Boost Classifier	0.6759	0.6154	0.3705	0.4430	0.3938	0.1803	0.1850	0.1750
<b>qda</b>	Quadratic Discriminant Analysis	0.5507	0.5455	0.5030	0.3651	0.3565	0.0800	0.0720	0.0250

Figura 4.89: Resultados de estudio anterior sobre los datos, para la variable “Hospitalización”

---

## 5: Conclusiones y Líneas de Trabajo Futuras

---

En este capítulo, se realiza una valoración general del trabajo realizado y de los resultados obtenidos en el desarrollo de este. A partir de ahí, se plantean mejoras que permitan continuar trabajando en la línea del proyecto realizado.

### 5.1. Conclusiones

Este trabajo se realizó con la motivación de probar una herramienta de auxilio para que los médicos del departamento de urgencias puedan manejar con más eficiencia a los pacientes con síncope que se reciben en este servicio hospitalario. Todo el tiempo se trató de probar cuán eficiente pueden ser los algoritmos de ML que, se interpretó desde un inicio que identificarían el manejo del paciente con síncope como una clasificación en el aprendizaje supervisado de los modelos. La base de datos con la que se contó para ello es pequeña en cuanto a cantidad de registros y con gran cantidad de atributos, además de estar contaminada por cambio de valores entre variables, errores en el paso de la información del formato analógico al formato digital y errores médicos incluidos.

Llegar a obtener un mejor modelo, ha significado un reto dentro de esta investigación, pues desde el análisis exploratorio de los datos se detectó más ruido del esperado (sobre todo ruido complejo de corregir para alguien fuera del mundo médico), se estudiaron los posibles valores y los límites entre ellos que pudieran ayudar a comprender los atributos e identificar fuentes de ruidos entre ellos y, a partir de ahí, se trazaron estrategias que permitieran comparar rendimientos y llegar a un mejor modelo de ML. Para explorar los modelos, e incluso a los propios datos, la biblioteca de *pyCaret* demostró ser una herramienta eficiente y capaz de ser explotada.

No se obtuvo resultados para dos de las clasificaciones que se pretendía hacer (“Mortalidad Hospitalaria” y “UCI”), las razones observadas a lo largo del proyecto, son expuestas en el capítulo anterior y se basan en la contaminación de los datos y el manejo que se hace de estos. Sí se obtuvo un resultado, y favorable por demás, para la clasificación



de “Hospitalización”, que incluye a la clasificación “UCI”. Este resultado es similar, e incluso mejor, que los obtenidos en los trabajos que se pudieron encontrar en la búsqueda bibliográfica realizada en el capítulo 2. Aunque las condiciones que exponen y los tipos de datos del paciente que incluyen como variables, no son los mismos.

Se puede concluir que la aplicación de estos algoritmos y los resultados obtenidos se encuentran muy relacionados con la calidad y cantidad de los datos, la primera en mayor medida, ya que los otros factores importantes asociados son los parámetros del modelo y la biblioteca *pyCaret* facilitó la búsqueda e identificación de estos.

## 5.2. Trabajo Futuro

Considerando los resultados que se alcanzan y las conclusiones sacadas en cada etapa, se propone partir de mejorar la bases de datos para trabajos futuros. La recomendación consiste en intentar recuperar las variables con valores ausentes y corregir los cambios de valores entre variables. Para lograr este propósito, se necesitaría un trabajo más exhaustivo en colaboración con el médico responsable de los datos.

Se señala que el hecho de haber experimentado con algoritmos de aprendizaje supervisado, no exime de la posibilidad de usar redes neuronales y algoritmos de aprendizaje no supervisado para experimentaciones futuras.

Habiendo obtenido resultados viables solo para un tipo de clasificación de las tres establecidas por el médico, se propone explorar la creación de modelos de clasificación multi-clase, en la que se tenga en cuenta la inclusión-exclusión entre las tres variables de salida actuales para la conformación de categorías, así como la creación de nuevas categorías a partir de la relación de los valores que ya están.

La base de datos contaba con pocos registros y solo se dividieron los datos en conjunto de entrenamiento y de prueba, cuando se deseaba separar también un conjunto de validación. Por ello se espera contar con otra base de datos similar a la anterior, obtenida de la misma fuente, que pueda servir como conjunto de validación de los modelos.

Con base en los resultados obtenidos tanto en la búsqueda bibliográfica realizada como en el desarrollo de ML para la creación de un modelo de clasificación, se traza como siguiente paso, la publicación de sendos artículos en revistas científicas de impacto.

# Apéndices

## Apéndice A

---

# Plan de Proyecto

---

### A.1. Introducción

En este apéndice se presenta el plan de proyecto elaborado para la realización del trabajo actual, así como la legalidad ética y de protección de datos en cuanto a información médica confidencial.

### A.2. Planificación temporal

La Tabla A.1 muestra la planificación realizada al inicio del trabajo, aunque comprende algunos desvíos de tiempo (otros no) que se deben, fundamentalmente, a la búsqueda exhaustiva de información al rededor de los datos y de su uso en el entorno de urgencias.

### A.3. Viabilidad legal

Como se ha mencionado anteriormente, este estudio fue realizado por la petición de un médico del ED, quien se encargó de anonimizar los datos usados y de obtener las autorizaciones pertinentes por parte de los *Comités Éticos de Investigación Clínica* de las respectivas instituciones sanitarias, para el estudio actual.

Toda la información médica registrada corresponde a pacientes atendidos en el ED de los centros de salud: **Hospital Clínico Universitario de Valladolid, Hospital Universitario Río Hortega de Valladolid, Hospital General de Segovia, Complejo Asistencial Universitario de Salamanca y Complejo Asistencial Universitario de Burgos**; en fechas comprendidas entre febrero de 2018 y julio de 2021.

---

**ACTIVIDADES Y TAREAS POR ETAPAS**


---

**PT0 Gestión del proyecto - 180 h**

Definición de la motivación, objetivos y el problema de la investigación

Supervisar la ejecución del proyecto en términos de avance de las actividades en los plazos previstos

Generar notas con los resultados de las actividades de gestión

**PT1 Revisión de la literatura - 20 h**

Definición de las bases de datos científicas, los criterios de inclusión-exclusión a utilizar en la revisión

Identificación y selección de artículos relacionados con el ML para el manejo del síncope en el ED

**PT2 Análisis y diseño - 40 h**

Analizar las técnicas de exploración que se encuentren en la revisión

Definir las herramientas de programación que se pueden usar

Exploración y análisis de la base de datos proporcionada

Limpieza del conjunto de datos para obtener la mejor predicción posible

Analizar las técnicas y algoritmos de ML encontrados en la revisión

Investigar técnicas y algoritmos de ML en la documentación de la herramienta de programación

Diseñar una metodología de trabajo en función de la clasificación propuesta a realizar

**PT3 Desarrollo - 40 h**

Desarrollar el diseño anterior

**PT06 Documentación - 80 h**

Escribir la memoria de la tesis doctoral

Tabla A.1: Plan de Trabajo

---

## Bibliografía

---

- [1] Documentación de python. [Internet; accedido 15-julio-2022].
- [2] Tutorial de machine learning. [Internet; accedido 27-julio-2022].
- [3] Troponina: valores normales y resultados de la prueba, 2021. [Internet; accedido 15-julio-2022].
- [4] ADEDINSEWO D, CARTER RE, A. Z., AND ET AL. Artificial intelligence-enabled ecg algorithm to identify patients with left ventricular systolic dysfunction presenting to the emergency department with dyspnea. *Circ Arrhythm Electrophysiol* 13, 8 (2020).
- [5] ALAM, M. k-nearest neighbors (knn) for anomaly detection.
- [6] APRENDEIA. Introducción a la librería pandas de python parte 1. [Internet; accedido 15-julio-2022].
- [7] APRENDEIA. Introducción a numpy python 1. [Internet; accedido 15-julio-2022].
- [8] APRENDEIA. Libreria pandas de matplotlib tutorial. [Internet; accedido 15-julio-2022].
- [9] APRENDEIA. Libreria scikit-learn de python. [Internet; accedido 15-julio-2022].
- [10] B. GRAHAM, R. BOND, M. Q., AND MULVENNA, M. Using data mining to predict hospital admissions from the emergency department. *IEEE Access* 6 (2018), 10458–10469.
- [11] BACKUS BE, SIX AJ, K. J. E. A. Chest pain in the emergency room: a multicenter validation of the heart score. *Crit Pathw Cardiol* 9, 3 (2010), 164–169.
- [12] BADERTSCHER, P., NESTELBERGER, T., DE LAVALLAZ ANS M. THAN, J., MORAWIEC, B., KAWECKI, D., AND ET AL. Prohormones in the early diagnosis of cardiac syncope. *J. Am. Heart Assoc* 6, 12 (2017).

- [13] BIGDATA, S. Métricas de clasificación de machine learning. parte 3, 2019. [Internet; accedido 27-julio-2022].
- [14] BRIGNOLE, M., MOYA, A., DE LANGE, F., DEHARO, J., ELLIOTT, P., FANCIULLI, A., AND ET AL. 2018 esc guidelines for the diagnosis and management of syncope. *Eur. Heart J* 39, 21 (2018), 1883–1948.
- [15] C, E., U, B., M, K., AND O, E. Artificial neural network, genetic algorithm, and logistic regression applications for predicting renal colic in emergency settings. *Int J Emerg Med* 2, 2 (2009), 99–105.
- [16] CURIOSOANDO.COM. ¿cuáles son los valores normales de presión arterial? [Internet; accedido 27-julio-2022].
- [17] D, S., S, B., L, D. D., MN, H., JA, W., AND LE, N. Pediatric emergency care applied research network head injury clinical prediction rules are reliable in practice. *Arch Dis Child* 99, 5 (2014), 427–431.
- [18] DE IMAGEN CARDÍACA, S. E. Biblioteca de preguntas básicas de imagen cardíaca. [Internet; accedido 25-junio-2022].
- [19] DE NAVARRA, C. U. Diccionario médico. [Internet; accedido 25-junio-2022].
- [20] DIOS LUNA, J. B. *Aplicación de Algoritmos de Machine Learning para predecir la probabilidad de abandono de estudiantes en MOOCs basándose en el modelo Context-aware Feature Interaction Network*. 2021.
- [21] DU FAY DE LAVALLAZ, J., BADERTSCHER, P., ZIMMERMANN, T., NESTELBERGER, T., WALTER, J., STREBEL, I., AND ET AL. Early standardized clinical judgement for syncope diagnosis in the emergency department. *J Intern Med* 290, 3 (2021), 728–739.
- [22] EK, L., F, Y., DA, H., MD, M., AND HK, S. A clinical decision tool for predicting patient care characteristics: patients returning within 72 hours in the emergency department. *AMIA Annu Symp Proc 2012* (2012), 495–504.
- [23] ELSEVIER, G. Escala de coma de glasgow: tipos de respuesta motora y su puntuación. [Internet; accedido 25-junio-2022].
- [24] F, D., D, S., M, G., R, M., M, S., AND R, F. Machine learning and syncope management in the ed: The future is coming. *Medicina (Kaunas)* 57, 4 (2021), 351.
- [25] F, N., G, M., AND G, L. Analysis of temporal and causal relationship between syncope and 30-day events in a cohort of emergency department patients to identify the true rate of short-term outcomes. *J Emerg Med* 55, 5 (2018), 612–619.
- [26] FUENTE, C. Redes neuronales profundas – tipos y características. [Internet; accedido 30-junio-2022].

- [27] G, C., F, D., M, S., M, B., F, B., AND R, F. Is hospital admission valuable in managing syncope? results from the steps study. *Cardiol J* 21, 6 (2014), 606–610.
- [28] G., C., G., F., M., S., AND CASAGRANDA I. AND, S. B. E. A. Neural networks as a tool to predict syncope risk in the emergency department. *Europace* 19, 11 (2016), 2891–1895.
- [29] G, F., G, C., R, F., JV, Q., A, U., AND R, I. Artificial neural networks and risk stratification in emergency departments. *Intern Emerg Med* 14, 2 (2019), 219–299.
- [30] GONZALEZ, L. Curvas roc y Área bajo la curva (auc), 2019. [Internet; accedido 27-julio-2022].
- [31] GRANT L, JOO P, N. M., AND V, T. Machine learning versus traditional methods for the development of risk stratification scores: a case study using original canadian syncope risk score data. *Intern Emerg Med* 17, 4 (2022), 1145–1153.
- [32] GUI C, C. V. Machine learning in medicine. *University of Western Ontario Medical Journal* 86, 2 (2017), 76–78.
- [33] H, L., BY, T., H, N., CCS, V., AND ET AL. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. *Nat Med* 25, 3 (2019), 433–438.
- [34] HARRISON RF, K. R. Artificial neural network models for prediction of acute coronary syndromes using clinical data from the time of presentation. *Ann Emerg Med* 46, 5 (2005), 431–439.
- [35] HEALTH, U. S. D. Tiempo de trombina. [Internet; accedido 25-junio-2022].
- [36] HERAS, J. M. Precision, recall, f1, accuracy en clasificación, 2020. [Internet; accedido 27-julio-2022].
- [37] I, S., AND G, N. Clinical text data in machine learning: Systematic review. *JMIR Med Inform* 8, 3 (2020).
- [38] IG, S., GH, G., RD, M., NAIR RC, M. I., AND JR, W. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Ann Emerg Med* 21, 4 (1992), 384–390.
- [39] IMBALANCED LEARN. Imbalanced-learn documentation, 2022. [Internet; accedido 15-julio-2022].
- [40] JA, K., AM, M., C, K., PB, R., AND DM, C. Clinical criteria to prevent unnecessary diagnostic testing in emergency department patients with suspected pulmonary embolism. *J Thromb Haemost* 2, 8 (2004), 1247–1255.
- [41] JANG D-H, KIM J, J. Y., AND ET AL. Developing neural network models for early detection of cardiac arrest in emergency department. *Am J Emerg Med* 38, 1 (2020), 43–49.

- [42] JOURNALDEV. Autoviz en python: visualización de datos simplificada, 2022. [Internet; accedido 15-julio-2022].
- [43] JR, H., AB, W., K, T., AND WR, M. Selective cervical spine radiography in blunt trauma: methodology of the national emergency x-radiography utilization study (nexus). *Ann Emerg Med* 32, 4 (1998), 461–469.
- [44] JR, H., WR, M., AB, W., KH, T., AND MI, Z. Validity of a set of clinical criteria to rule out injury to the cervical spine in patients with blunt trauma. national emergency x-radiography utilization study group. *N Engl J Med* 343, 2 (2000), 94–99.
- [45] JV, C., E, A., H, H., C, K., AND DJ, M. Unnecessary hospitalization and related harm for patients with low-risk syncope. *JAMA Intern Med* 175, 6 (2015), 1065–1067.
- [46] KLINE JA, COURTNEY DM, K. C. E. A. Prospective multicenter evaluation of the pulmonary embolism rule-out criteria. *J Thromb Haemost* 6, 5 (2008), 772–780.
- [47] KOCHAV, S. M., RAITA, Y., FIFER, M. A., AND TAKAYAMA, H. Predicting the development of adverse cardiac events in patients with hypertrophic cardiomyopathy using machine learning. *International Journal of Cardiology* 327 (2021), 117–124.
- [48] KOENE, R., ADKISSON, W., AND BENDITT, D. Syncope and the risk of sudden cardiac death: evaluation, management, and prevention. *J. Arrhythm* 33, 6 (2017), 533–544.
- [49] KUAIK, D. Tensión arterial elevada en casa. ¿cómo has de interpretarla?, 2019. [Internet; accedido 25-junio-2022].
- [50] KUPPERMANN N, HOLMES JF, D. P. E. A. Identification of children at very low risk of clinically-important brain injuries after head trauma: a prospective cohort study. *Lancet* 374, 9696 (2009), 1160–1170.
- [51] M, B., A, M., DE LANGE FJ, JC, D., AND ET AL. 2018 esc guidelines for the diagnosis and management of syncope. *Eur Heart J* 39, 21 (2018), 1883–1948.
- [52] M, F., R, M., SM, V., F, L., AND ET AL. Predicting intensive care unit admission among patients presenting to the emergency department using machine learning and natural language processing. *PLoS One* 15, 3 (2020).
- [53] M, S., G, C., F, D., AM, R., AND ET AL. Syncope recurrence and mortality: a systematic review. *Europace* 17, 2 (2015), 300–308.
- [54] M, T., WW, C., BA, S., AND ET AL. Comparison of 2 natural language processing methods for identification of bleeding among critically ill patients. *JAMA Netw Open* 1, 6 (2018).



- [55] MARIANELLA ÁLVAREZ VEGA, LAURA MARÍA QUIRÓS MORA, M. V. C. B. Inteligencia artificial y aprendizaje automático en medicina. *Revista Médica Sinergia* 5, 8 (2020).
- [56] MAYOCLINIC. [Internet; accedido 25-junio-2022].
- [57] MEDLINEPLUS. [Internet; accedido 25-junio-2022].
- [58] MEHTA N, D. M. Machine learning, natural language programming, and electronic health records: The next step in the artificial intelligence journey? *J Allergy Clin Immunol* 141, 6 (2018), 2019–2021.
- [59] MEMBERS, W. C., WK, S., AND SHELDON RS, E. A. 2017 acc/aha/hrs guideline for the evaluation and management of patients with syncope: A report of the american college of cardiology/american heart association task force on clinical practice guidelines and the heart rhythm society. *Heart Rhythm* 14, 8 (2017).
- [60] N, L., Z, Z., AF, W. H., AND ME, H. O. Artificial intelligence in emergency medicine. *J Emerg Crit Care Med* (2018).
- [61] N, V., I, C., M, T., H, H., AND ET AL. 2021 ishne/hrs/ehra/aphrs collaborative statement on mhealth in arrhythmia management: Digital medical tools for heart rhythm professionals. *Journal of Arrhythmia* 37, 2 (2021), 271–319.
- [62] OBERMEYER Z, E. E. Predicting the future - big data, machine learning, and clinical medicine. *N Engl J Med* 375, 13 (2016).
- [63] OBERMEYER Z, L. T. The limits of the human mind and the future of medicine. *New England Journal of Medicine* 377, 13 (2017), 1209–1211.
- [64] P., M. J., AND ET AL. Declaración prisma 2020: una guía actualizada para la publicación de revisiones sistemáticas the prisma 2020 statement: an updated guideline for reporting systematic reviews,. *Revista Española de Cardiología* 74 (2021), 790–799.
- [65] PRADOS-TORRESA, A., DEL CURA-GONZÁLEZD, I., PRADOS-TORRES, J. D., AND ET AL. Multimorbilidad en medicina de familia y los principios ariadne. un enfoque centrado en la persona. *Atención Primaria* 49, 5 (2017), 300–307.
- [66] PYCARET. pycaret official. [Internet; accedido 15-julio-2022].
- [67] RA, K., M, B., GA, D., AND ET AL. Syncope unit: rationale and requirement—the european heart rhythm association position statement endorsed by the heart rhythm society. *Europace* 17, 9 (2015), 1325–1340.
- [68] RC, D. Machine learning in medicine. *Circulation* 132, 20 (2015), 1920–1930.

- [69] RICCI, F., SUTTON, R., PALERMI, S., TANA, C., RENDA, G., GALLINA, S., AND ET AL. Prognostic significance of noncardiac syncope in the general population: a systematic review and meta-analysis. *J. Cardiovasc. Electrophysiol* 29, 12 (2018), 1641–1647.
- [70] RICHARD SUTTON, FABRIZIO RICCI, A. F. Risk stratification of syncope: Current syncope guidelines and beyond. *Autonomic Neuroscience* 238 (2022).
- [71] RK, S., AND RS, S. Syncope in the emergency department. *Front Cardiovasc Med* 6 (2019).
- [72] RODRÍGUEZ, D. Visualización de datos en python con seaborn, 2018. [Internet; accedido 15-julio-2022].
- [73] SHELDON, R., MORILLO, C., KRAHN, A., O’NEILL, B., THIRUGANASAMBANDAMOORTHY, V., PARKASH, R., AND ET AL. Standardized approaches to the investigation of syncope: Canadian cardiovascular society position paper. *Can. J Cardiol* 27, 2 (2011), 246–253.
- [74] SHEN, W., SHELDON, R., AAND M.I. COHEN, D. B., FORMAN, D., GOLDBERGER, Z., AND ET AL. 2017 acc/aha/hrs guideline for the evaluation and management of patients with syncope: a report of the american college of cardiology/american heart association task force on clinical practice guidelines and the heart rhythm society. *Circulation* 136, 5 (2017).
- [75] SHMUELI, B. Matthews correlation coefficient is the best classification metric you’ve never heard of, 2019. [Internet; accedido 27-julio-2022].
- [76] SIX AJ, BACKUS BE, K. J. Chest pain in the emergency room: value of the heart score. *Neth Heart J* 16, 6 (2008), 191–196.
- [77] SJ, S., DH, K., S, S., MA, B., CW, Y., AND ET AL. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 131, 3 (2015), 269–279.
- [78] SJ, W., TR, M., KM, F., JP, F., AND DM, A. Prospective validation of wells criteria in the evaluation of patients with suspected pulmonary embolism. *Ann Emerg Med* 44, 5 (2004), 503–510.
- [79] SOFFER, S., KLANG, E., BARASH, Y., GROSSMAN, E., AND ZIMLICHMAN, E. Predicting in-hospital mortality at admission to the medical ward: A big-data machine learning model. *The American Journal of Medicine* 134 (2021).
- [80] SOLER, W., MUÑOZ, M. G., BRAGULAT, E., AND ÁLVAREZ, A. El triaje: herramienta fundamental en urgencias y emergencias. *Anales del Sistema Sanitario de Navarra* 33, 1 (2010).

- [81] STIELL IG, GREENBERG GH, M. R. E. A. Decision rules for the use of radiography in acute ankle injuries. refinement and prospective validation. *JAMA* 269, 9 (1993), 1127–1132.
- [82] STIELL IG, CLEMENT CM, R. B. E. A. Comparison of the canadian ct head rule and the new orleans criteria in patients with minor head injury. *JAMA* 294, 12 (2005), 1511–1518.
- [83] STIELL IG, WELLS GA, V. K. E. A. The canadian ct head rule for patients with minor head injury. *Lancet* 357, 9266 (2001), 1391–1396.
- [84] SUTTON, R. Syncope presenting to the emergency department. *J. Intern. Med* 290, 3 (2021), 755–756.
- [85] SUTTON, R., BRIGNOLE, M., AND BENDITT, D. Key challenges in the current management of syncope. *Nat. Rev. Cardiol.* 9, 10 (2012), 590–598.
- [86] SÁNCHEZ, M. S., ALMONTE, C., CHAVES, G. C., AND ET AL. Li consenso centroamericano y del caribe de insuficiencia cardíaca. biomarcadores séricos. *Costarricense de Cardiología* 20, 4 (2018).
- [87] TAYLOR RA, PARE JR, V. A., AND ET AL. Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data-driven, machine learning approach. *Acad Emerg Med* 23, 3 (2016), 269–278.
- [88] V, T., K, K., GA, W., AND ET AL. Development of the canadian syncope risk score to predict serious adverse events after emergency department assessment of syncope. *CMAJ* 188, 12 (2016), 167–175.
- [89] V, T., MLA, S., N, L. S., AND ET AL. Multicenter emergency department validation of the canadian syncope risk score. *JAMA Intern Med* 180, 5 (2020), 737–744.
- [90] VAN DER PLOEG T, PC, A., AND EW, S. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. *BMC Med Res Methodol* 14, 1 (2014), 137.
- [91] VAN DIJK, J., THIJS, R., BENDITT, D., AND WIELING, W. A guide to disorders causing transient loss of consciousness: focus on syncope. *Nat Rev Neurol* 5, 8 (2009), 438–448.
- [92] VILLINES, Z. ¿qué significan los niveles altos o bajos en el conteo de plaquetas?
- [93] WELLS PS, ANDERSON DR, R. M. E. A. Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and d-dimer. *Ann Intern Med* 135, 2 (2001), 98–107.

- [94] WIDMANN, M. Kappa de cohen: qué es, cuándo usarlo y cómo evitar sus trampas, 2020. [Internet; accedido 27-julio-2022].
- [95] WK, S., WW, D., PA, S., AND ET AL. Syncope evaluation in the emergency department study (seeds): a multidisciplinary approach to syncope management. *Circulation* 110, 24 (2004), 3636–3645.
- [96] WS, H., AD, H., AND RA, T. Predicting hospital admission at emergency department triage using machine learning. *PLoS One* 13, 7 (2018).
- [97] WY, C., TH, O. Y., AND D, A. Biomolecular events in cancer revealed by attractor metagenes. *PLoS Comput Biol* 9, 2 (2013).
- [98] Y, R., T, G., MK, F., DFM, B., CA, C., AND K, H. Emergency department triage prediction of clinical outcomes using machine learning models. *Crit Care* 23, 1 (2019), 64.
- [99] Y, S., BH, H., YT, S., AND E, S. Forecasting daily attendances at an emergency department to aid resource planning. *BMC Emerg Med* 9, 1 (2009).
- [100] Y-H, P., H, K., J, T., AND ET AL. Machine learning methods are comparable to logistic regression techniques in predicting severe walking limitation following total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 28, 10 (2020), 3207–3216.
- [101] YASA, E., RICCI, F., MAGNUSSON, M., SUTTON, R., GALLINA, S., CATERINA, R., AND ET AL. Cardiovascular risk after hospitalisation for unexplained syncope and orthostatic hypotension. *Heart* 104, 6 (2018), 487–493.