

Universidad Internacional de La Rioja (UNIR)

ESIT

**Máster en Análisis y Visualización de Datos
Masivos**

Desarrollo de un entorno de visualización inteligente de análisis de datos médicos

Trabajo Fin de Máster

presentado por: Fernández Pérez, Paula

Director/a: Serrano Fernández, Emilio

Ciudad: Madrid

Fecha: 26 de junio 2019

Resumen

En este trabajo se desarrolla una herramienta que permita al personal sanitario predecir de una manera rápida y sencilla tanto la probabilidad de muerte de un individuo como diagnósticos asociados que puede desarrollar durante su episodio asistencial.

La contribución de este TFM es la inclusión de experiencias pasadas mediante el uso de tecnología MachineLearning junto con constantes de monitorización básicas en la predicción. A pesar de las numerosas contribuciones de la inteligencia artificial en los entornos médicos, en la actualidad los sistemas de predicción de mortalidad más utilizados en el entorno hospitalario son las escalas como el Apache II, pero éstas necesitan de muchas variables históricas del paciente y resultados de laboratorio que no siempre son accesibles.

La herramienta desarrollada cuenta con un LogLoss de 0.438 para la predicción de muerte y de 4.0375 para la predicción de posibles diagnósticos asociados. Por lo tanto, tendrá una alta fiabilidad para la primera y servirá de orientación para la segunda.

Palabras Clave: AutoMachineLearning, predicción, probabilidad de muerte, diagnósticos

Abstract

The aim of this work is to develop a tool for clinicians, which predicts in an easy and fast way both, death and associated diagnostics probability.

The contribution of this work is the inclusion in the prediction of experiences using MachineLearning together with basic vital signs. Even though the fact that many AI applications have been developed for medical issues, nowadays, the most used predicting systems for mortality are still scores such as Apache II. The problem with them is that scores need many historical patient info and laboratory results, which are not always accessible.

The developed tool has a LogLoss score of 0.438 for death prediction and 4.0375 for secondary diagnostics prediction, which means that it has a high reliability for the first one and would be a help for the second one.

Palabras Clave: AutoMachineLearning, prediction, death probability, diagnostics.

Índice

Resumen.....	1
Abstract.....	1
1. Introducción	7
2. Contexto y estado del arte.....	8
Sistemas actuales de predicción.....	9
Aplicaciones actuales de predicción	9
Intellispac Critical Care & Anesthesia de Philips.....	10
Centricity de General Electric	10
MetaVision ICU de iMDSOft.....	11
ehCos de Everis Health.....	12
3. Objetivos concretos y metodología de trabajo	13
Objetivo general	13
Objetivos específicos.....	13
Metodología de trabajo	13
4. Identificación de requisitos	14
5. Descripción de la herramienta software desarrollada	16
6. Conjunto de datos a analizar	17
Motivación para la elección del conjunto de datos	18
Descripción de la base de datos	19
7. Herramientas y tecnologías para el análisis del conjunto de datos.....	21
Docker	21
PostgreSQL	22
DBEaver	22
Spark.....	23
DataBricks	23
8. Herramientas y tecnologías para la creación del modelo	24
H2O	24

9. Herramientas de visualización para la presentación de los datos.....	24
GoogleCharts	24
10. Implementación	25
Acceso al conjunto de datos	25
Importación de los datos.....	25
Limpieza de los datos	27
Creación de tablas.....	31
Implementación del modelo	37
Despliegue del modelo en el servidor	46
Diseño de la interfaz	48
11. Evaluación de la herramienta	49
12. Conclusiones	54
13. Líneas futuras.....	55
14. Referencias bibliográficas.....	56
15. Bibliografía complementaria	58
Anexo I. Extracción variables predictivas	59
Anexo II. Script generación datos para el entrenamiento del modelo.	60
Anexo III: Despliegue del modelo predictivo en el servidor.....	61
Anexo IV: Desarrollo Interfaz.....	63
Anexo V: Respuestas de las encuestas de usabilidad.....	75

Índice de figuras

Figura 1. Dashboard ICCA-Qlik View [7]	10
Figura 2: Pantalla principal de Centricity para la ayuda a la decisión [8].....	10
Figura 3: Dashboard de Centricity-Microsoft [6].....	11
Figura 4: Pantalla de alarmas de Metavision [9]	11
Figura 5: Mapa de camas de ehCos [10].....	12
Figura 6: Requisitos funcionales.....	15
Figura 7: Requisitos no funcionales.....	16
Figura 8: Arquitectura de la herramienta software desarrollada.....	17
Figura 9: Arquitectura Docker [17].....	22
Figura 10. Jupiter Notebook	23
Figura 11: Contenido curso CITI “Data or Specimens Only Research”	25
Figura 12: Importación de la base de datos MIMIC III a PostgreSQL	26
Figura 13: Salida comando docker images.....	26
Figura 14: Salida del comando docker logs –f mimickdocker, antes de terminar la creación de la base de datos.....	27
Figura 15: Salida del comando docker logs –f mimickdocker, confirmando la creación de la base de datos	27
Figura 16: Proceso de limpieza de variables predictivas demográficas	28
Figura 17: Proceso de limpieza de datos de variables predictivas clínicas	29
Figura 18: Tabla variables predictoras clínicas.....	30
Figura 19: Generación datos de entrenamiento.....	34
Figura 20. Exportación tablas	35
Figura 21: Importación de las tablas en Databricks	35
Figura 22. Límites variables predictivas clínicas.....	36
Figura 23: Fases de la creación del modelo	37
Figura 24: Entrada de datos H2O	37
Figura 25: Parámetros evaluación H2O.....	38
Figura 26: Validación cruzada 5-iteraciones	39
Figura 27: Evaluación columna a predecir.....	39
Figura 28: Importancia de las variables en la predicción de probabilidad de muerte	40
Figura 29: Número de modelos para determinar la importancia de parámetros.....	40
Figura 30: Valores de LogLoss para la predicción de probabilidad de muerte	41
Figura 31: Matriz de confusión del algoritmo de probabilidad de muerte	41
Figura 32: Importancia de las variables en la predicción de diagnósticos secundarios.....	45
Figura 33: Valores de LogLoss para la predicción de diagnósticos secundarios	45

Figura 34: Pasos para el despliegue del modelo en el servidor	47
Figura 35: Interfaz	48
Figura 36: Requisito F1	50
Figura 37: Requisitos F2, F3,F5	51
Figura 38: Requisitos F6, F7	51
Figura 39: requisito NF4.....	51

Índice de código fuente

Código fuente 1: Construcción imagen Docker.....	26
Código fuente 2: Despliegue contenedor Docker	26
Código fuente 3: Reducción columnas de la tabla de chartevents.....	30
Código fuente 4: Creación tabla patientsInfo	32
Código fuente 5: Creación tabla cohortDeath	32
Código fuente 6: Creación tabla cohortDiagnostico	33
Código fuente 7: Creación tabla patientsDiag.....	33
Código fuente 8: Creación tabla patientsDiagnostics.....	34
Código fuente 9: Variables predictivas clínicas.....	36

1. Introducción

En los entornos médicos, se almacena una gran cantidad de información que no suele ser usada de manera predictiva sino descriptiva. En los últimos años, con la implementación de sistemas de información clínica, ha aumentado considerablemente la cantidad de datos que se recopilan de forma automática y fiable de dispositivos de cabecera, que, junto con los datos registrados por el personal sanitario, pueden ser usados para predecir diagnósticos y anticiparse a posibles complicaciones durante el ingreso hospitalario.

Sin embargo son varios los profesionales que aseguran que la gestión del conocimiento debería estar dirigida a la capacidad de acción y que se debe organizar el conocimiento dentro del ámbito sanitario [1].

En este trabajo de fin de máster, se pretende crear un dashboard que ayude a los facultativos en la toma de decisiones. Para ello se usará una base de datos con información de pacientes que, tras un análisis exhaustivo, permitirá determinar qué constantes o parámetros vitales son de mayor relevancia para predecir un diagnóstico en concreto y la probabilidad de muerte del individuo. Para evaluar la predicción de un diagnóstico concreto, el estudio se centrará en aquellos pacientes cuyo motivo de ingreso esté relacionado con una patología asociada al sistema cardíaco, mientras que para la predicción de la probabilidad de muerte del individuo no se hará dicho filtrado por motivo de ingreso.

Este método de ayuda a la decisión presenta varias mejoras con respecto a los sistemas utilizados en la actualidad, destacando entre ellos la rapidez de uso, la posibilidad de retroalimentar el modelo de forma fácil y efectiva y la disponibilidad de esta herramienta en entornos de cuidados críticos fuera de las propias unidades de cuidados críticos, por ejemplo en las salas de urgencia donde únicamente contamos con el criterio médico y las constantes tomadas por los dispositivos de cabecera pero no con datos de laboratorio o antecedentes familiares que son necesarios para hacer uso de algunos de los métodos de predicción que existen actualmente en el mercado.

En los primeros capítulos del documento se expondrá el problema identificado y los objetivos que se pretenden alcanzar en este trabajo.

Posteriormente se hará una identificación de requisitos para seguidamente describir la herramienta software desarrollada y analizar de manera precisa el conjunto de datos que se va a utilizar, explicando la motivación para dicho uso y la estructura que presentan los datos, elemento fundamental.

A continuación, se expondrán las herramientas tanto para la creación del modelo como para la implementación de las visualizaciones finales, detallando las características más relevantes para cada una de ellas.

Más adelante, en el capítulo de “Implementación” se explicarán los pasos seguidos desde la importación de los datos a la ejecución de las visualizaciones.

Finalmente se hará una evaluación del trabajo realizado y se elaborarán conclusiones, definiendo también las posibles líneas futuras.

2. Contexto y estado del arte

Durante los últimos años se ha evidenciado un gran auge de los sistemas de información en todos los entornos, desde entornos industriales a médicos, pasando por administrativos y otros sectores.

En el caso concreto de los sistemas de información en el entorno médico, cabe destacar el conjunto de unidades que trata a pacientes críticos entre los que se incluyen unidades de cuidados intensivos, unidades de reanimación, unidades coronarias y cardíacas y quirófanos.

En estas unidades, cada paciente cuenta con varios dispositivos de cabecera como por ejemplo monitores, respiradores, máquinas de hemofiltración, bombas de infusión de medicamentos y nutriciones o monitores de gasto cardíaco que envían continuamente datos tanto al sistema de información con el que interactúa el personal sanitario, como a una base de datos centralizada.

Manor-Shulman et al [2] citado por V. Gomez Tello et al [3]:

“Los servicios de medicina intensiva (SMI) constituyen un área sanitaria muy especializada y con numerosos dispositivos que generan multitud de datos y registros de los pacientes ingresados, llegando hasta más de 1.300 datos por enfermo y día de estancia.”

Estos datos pueden ayudar a los facultativos a la toma de decisiones y a evitar efectos adversos [4] .

“La Agencia de Calidad del Sistema Nacional de Salud (AC-SNS) y la Sociedad Española de Medicina Intensiva, Crítica y Unidades Coronarias (SEMICYUC) han realizado un estudio en 79 UCI y 1.017 pacientes; de ellos, el 58,1% presentaron algún incidente. Se notificaron 1.424 incidentes en 7 591 pacientes, de los cuales 943 fueron incidentes sin daño y 481 eventos adversos. Se produjeron 1,22 incidentes por cada

paciente ingresado. La tasa de aparición de incidentes fue de 5,89 incidentes por cada 100 pacientes y hora. El 74% de los incidentes comunicados estuvieron en relación con la medicación, aparatos, cuidados, accesos vasculares y sondas, vía aérea y ventilación mecánica. El 90% de todos los incidentes y el 60% de los efectos adversos (EA) fueron clasificados como evitables o posiblemente evitables.”

De igual modo en [4] se determina la recomendación de implantar TIC en el entorno sanitario:

“En el estudio de la SEMICYUC sobre la seguridad en la UCI, el 24,5% de los incidentes se relacionaron con la medicación, siendo los relacionados con la prescripción y administración los más frecuentes. Para prevenir estos efectos adversos, se recomienda: [...] Fomentar desde la UCI la implantación de nuevas tecnologías de la información y comunicación (TIC), especialmente la disponibilidad de historia clínica informatizada y de sistemas de prescripción farmacológica asistida electrónica.”

Sistemas actuales de predicción

En la actualidad existen numerosas aplicaciones con el objetivo de informatizar la historia clínica de unidades de cuidados críticos. Sin embargo, no cuentan con un sistema visual y de fácil uso de predicción de diagnósticos o eventos adversos.

En las unidades mencionadas con anterioridad, se utilizan dos métodos de predicción diferentes:

1. **Escalas:** El uso de escalas está estandarizado en la comunidad médica. Se trata de algoritmos definidos tras estudios médicos que permiten pronosticar ciertas variables, siendo la más frecuente la probabilidad de muerte. Para ello, se definió el Apache II [5], cuyo resultado es un número asociado a una probabilidad de muerte. Esta escala, al igual que el resto precisan de un alto conocimiento médico para poder ser rellenadas, ya que la puntuación otorgada a cada variable es determinante en el resultado final y es necesario conocer con detalle a qué hace referencia exactamente cada una de ellas.
2. **Alarmas:** En base a estudios anteriores, se pueden configurar ciertas alarmas que avisen al personal sanitario de posibles adversidades en la salud del paciente. Sin embargo, estas alarmas son reactivas y en la actualidad no existe una manera de que sea el clínico quien de manera proactiva haga un análisis de las posibilidades de futuras incidencias.

Aplicaciones actuales de predicción

En este apartado se presentarán algunas de las aplicaciones que usan los métodos de predicción anteriormente descritos.

Intellispace Critical Care & Anesthesia de Philips

Esta aplicación de Philips [6] cuenta con un módulo Data Analysis and Reporting (DAR) que incluye plantillas y herramientas clínicas inteligentes que permiten realizar análisis e informes de datos individualizados sin necesidad de tener que recurrir a especialistas con capacidades técnicas avanzadas. Además presenta un módulo de query analyzer y un módulo de Qlik View Sense, presentado en la figura 1, que permiten realizar consultas dinámicas y ágiles.

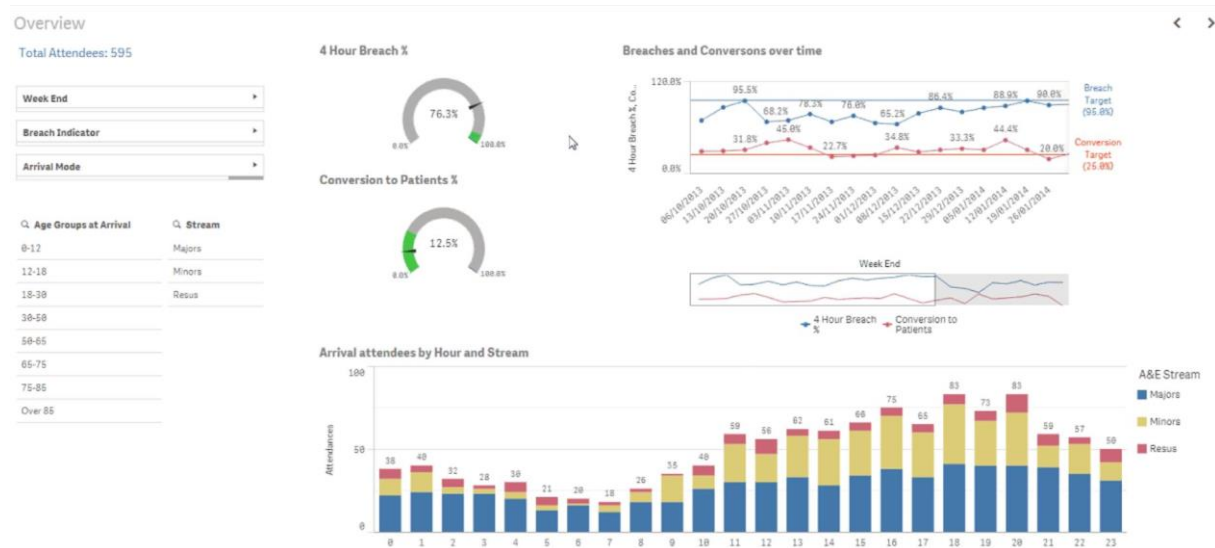


Figura 1. Dashboard ICCA-Qlik View [7]

Centricity de General Electric

La aplicación de escritorio de Centricity, presenta la información clínica en diferentes pantallas, como puede observarse en la figura:

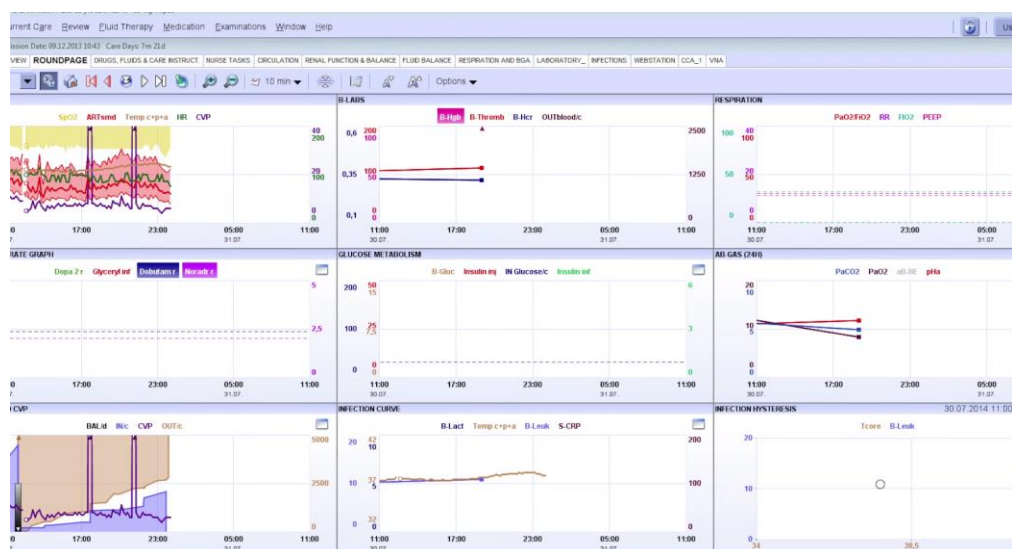


Figura 2: Pantalla principal de Centricity para la ayuda a la decisión [8]

Para la explotación de datos, Centricity cuenta con una integración con la herramienta Microsoft Dynamics CRM y Power BI [8] y puede apreciarse un ejemplo de dicha explotación en la figura 3:

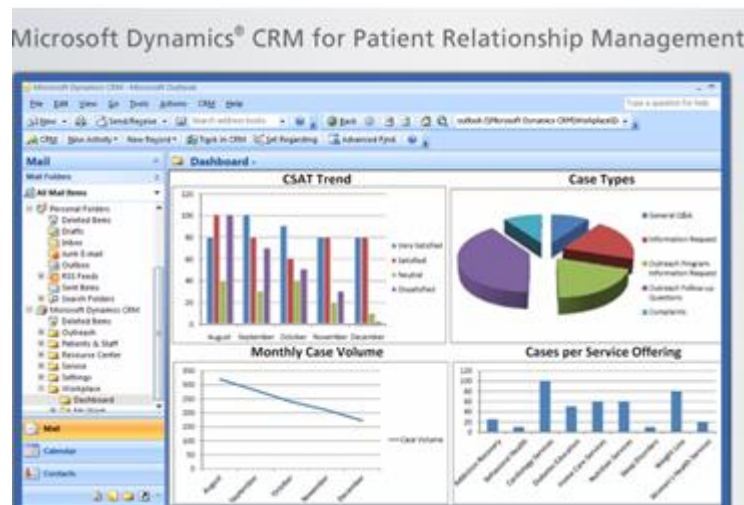


Figura 3: Dashboard de Centricity-Microsoft [6]

MetaVision ICU de iMDSOft.

La aplicación presenta un módulo propio de explotación de datos e informes, con el objetivo de evaluar potenciales pacientes para estudios clínicos y para extraer información para dichos estudios[9].

En la figura 4 se presenta el modo en el que se gestiona el sistema de alarmas en MetaVision.

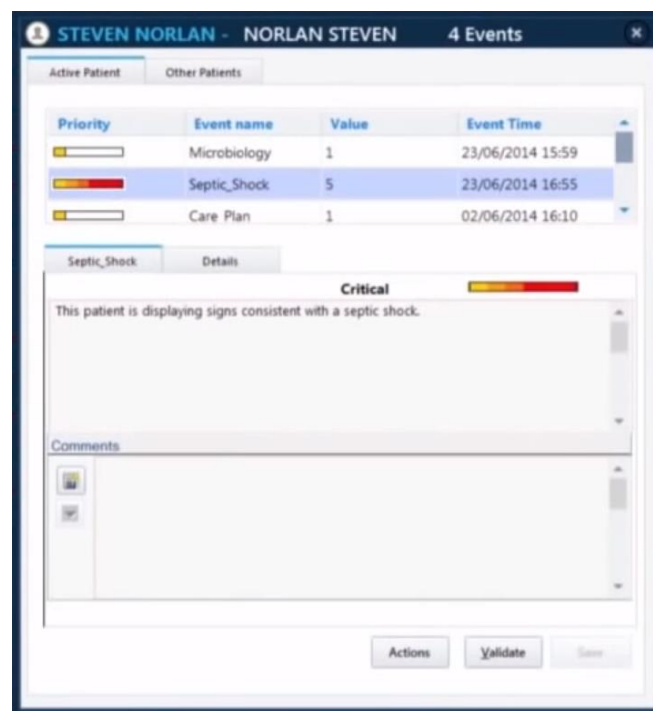


Figura 4: Pantalla de alarmas de Metavision [9]

ehCos de Everis Health

La aplicación ayuda a la toma de decisiones de diferentes maneras[10]:

1. Creación de patrones repetitivos que alimentan modelos predictivos.
2. Análisis retrospectivo de la historia del paciente.

En la figura 5 se expone un mapa de camas proporcionado por la aplicación, que permite conocer el estado de la unidad de una manera fácil y eficiente.

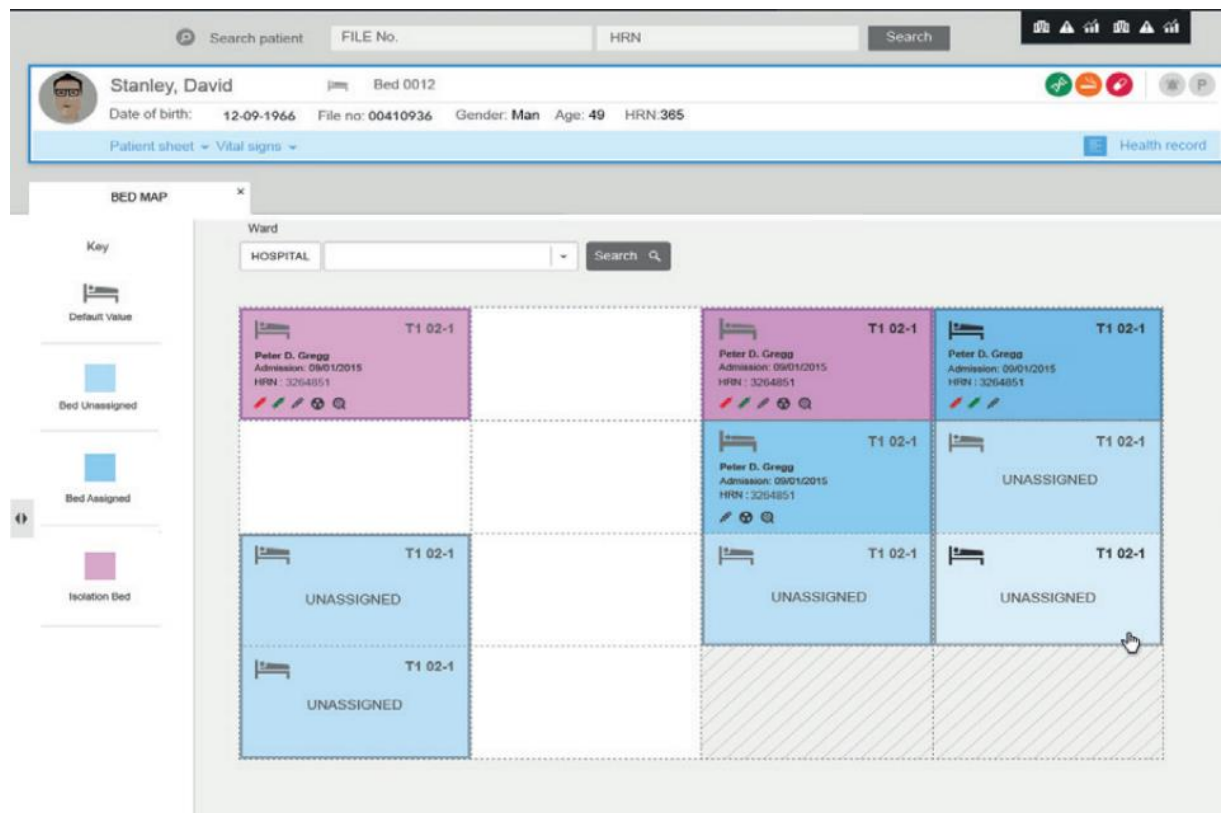


Figura 5: Mapa de camas de ehCos [10]

Mediante estas tecnologías y en colaboración con el Hospital Virgen del Rocío de Sevilla, han desarrollado un sistema de alarmas tempranas que permite predecir los episodios de shock séptico, hipotensión e hipoxia[11].

3. Objetivos concretos y metodología de trabajo

En este apartado se detallarán los objetivos tanto general como específicos que se pretenden alcanzar en el trabajo.

Más adelante se expondrá la metodología que se ha seguido para poder cumplir con los objetivos propuestos.

Objetivo general

El objetivo principal del proyecto es ayudar al personal clínico en la toma de decisiones en un entorno de cuidados críticos basándose en la predicción de la probabilidad de muerte y de la probabilidad de diagnósticos secundarios.

Cuando se conoce un posible diagnóstico, se pueden tomar medidas correctoras con anterioridad disminuyendo la probabilidad de efectos adversos.

Para ello se hará uso de herramientas de inteligencia artificial y de visualización.

Objetivos específicos

Los objetivos específicos en los que se divide este trabajo son los siguientes:

1. Elegir un conjunto de datos completo que permita crear un modelo de datos fiable.
2. Explorar el conjunto de datos con el fin de analizar qué campos son más relevantes para el estudio.
3. Crear una base de datos para el análisis y procesado de los datos.
4. Analizar herramientas y algoritmos de inteligencia artificial para elegir el más adecuado a las necesidades del proyecto.
5. Elaborar un modelo que ofrezca una alta fiabilidad.
6. Explorar herramientas y tecnologías de visualización de datos para crear el dashboard objetivo.
7. Implementar el dashboard que oriente al personal clínico en la toma de decisiones.
8. Evaluar la herramienta.

Metodología de trabajo

Para alcanzar los objetivos propuestos anteriormente se ha dividido el proyecto en diferentes fases, quedando reflejadas en los capítulos del 6 al 11.

A continuación se presentan brevemente las etapas del proyecto:

1. **Elección y tratamiento del conjunto de datos:** Durante esta fase se evaluarán distintas opciones de conjuntos, teniendo en cuenta que deben cumplir con los requisitos de anonimización adecuados para poder ser tratados.
2. **Elección de la herramienta e implementación del modelo:** En esta fase se desarrollará un modelo predictivo.
3. **Elección de la tecnología de visualización, diseño y ejecución de las visualizaciones:** A lo largo de esta etapa se decidirá la tecnología con la que se diseñará y llevará a cabo el dashboard objetivo.
4. **Evaluación, conclusión y análisis de líneas futuras:** Durante la última fase del proyecto se realizará una evaluación de la consecución de los objetivos propuestos al inicio de este capítulo y de los requisitos identificados, así como la elaboración de conclusiones del proyecto y posibles líneas de actuación futuras.

4. Identificación de requisitos

Un paciente en la UCI genera hasta 236 categorías de variables[12], pero los seres humanos son capaces de gestionar adecuadamente de 5 a 9 variables [13].

En base a esto, han ido apareciendo herramientas de gestión de la historia clínica, algunas de ellas expuestas en el punto anterior. Sin embargo, tal cantidad de datos no se está explotando con rigurosidad en la actualidad [1], a pesar de que el personal clínico considere que son datos valiosos tanto para el propio paciente como para estudios posteriores que puedan mejorar la calidad de la estancia de los pacientes en las unidades de cuidados críticos.

Desde la Semicyuc afirman [14] que:

“Los sistemas de información clínica (SIC) son herramientas necesarias para la gestión de la información generada por los pacientes en los Servicios de Medicina Intensiva. Son sistemas útiles para la mejora de la Seguridad del paciente, de la calidad asistencial y de la gestión clínica. Y son de gran ayuda en tareas como la investigación y la docencia. Dada la variedad de SIC disponibles en el mercado, es necesario disponer de unos estándares mínimos para que sean útiles y se adecuen a las necesidades de los pacientes y de los profesionales.”

La definición de dichos estándares mínimos puede encontrarse en el sitio web oficial de la sociedad [14], destacando por su relación con el objetivo de este trabajo indicador de calidad nº135:

“El SIC debe asegurar: [...]”

- Gestión de alarmas y metaalarmas
- Registro de CMBD incorporado, con códigos, e Informes
- Herramienta de búsqueda y extracción de datos de instalación obligada.”

Por lo tanto, un sistema de ayuda a la toma de decisión diseñado con el objetivo de ser visual, intuitivo e interactivo como el que se desarrolla en este proyecto, podría dar solución a algunos de los problemas detectados por el colectivo clínico y expuestos con anterioridad.

Para ello, el dashboard objetivo de este trabajo cumplirá con los requisitos definidos en las figuras 6 y 7:

Requisitos funcionales			
Id	Tipo	Descripción	Prioridad [0-5]
F1	Interfaz	El usuario deberá introducir todos los parámetros para realizar la predicción.	4
F2	Interfaz	Los parámetros de entrada para la predicción serán: edad, sexo, motivo de ingreso, frecuencia cardiaca, presión arterial, temperatura y presión venosa central	5
F3	Interfaz	Se proporcionará un botón para borrar todos los parámetros introducidos	2
F4	Interfaz	Se especificará el rango de cada parámetro	4
F5	Interfaz	Se especificará la unidad de medida de cada parámetro	5
F6	Interfaz	Se mostrará la hora de la predicción	4
F7	Proceso	La salida de una predicción será: estado del paciente al alta de la UCI y a los 90 días del alta de la UCI, probabilidad de diagnósticos asociados	5
F8	Proceso	Se guardará todas las predicciones para su posterior análisis	3
F9	Legal	El sistema no almacenará los datos de los pacientes, ni tendrá acceso a ellos de ninguna forma, solo podrá usar el modelo generado anteriormente	5

Figura 6: Requisitos funcionales

Requisitos no funcionales			
Id	Tipo	Descripción	Prioridad [0-5]
NF1	Usabilidad	Los usuarios podrán realizar una predicción en menos de 5 minutos	3
NF2	Usabilidad	Se le indicará al usuario si le falta un parámetro para realizar la predicción	3
NF3	Usabilidad	Se le indicará al usuario si ha introducido un dato con un tipo incorrecto	3
NF4	Usabilidad	Se podrá acceder al sistema de predicción desde los navegadores Chrome y Firefox	4
NF5	Carga	El cálculo de una predicción se realizará en un tiempo inferior a 30 segundos	4
NF6	Proceso	El modelo podrá recalcularse con los nuevos episodios asistenciales mediante un proceso manual	2
NF7	Legal	Los datos para la generación del modelo estarán completamente anonimizados	5
NF8	Carga	El servidor deberá soportar al menos 10 usuarios simultáneos	2

Figura 7: Requisitos no funcionales

5. Descripción de la herramienta software desarrollada

En este capítulo se describirá la arquitectura de la herramienta de software desarrollada mediante un diagrama que permita conocer las diferentes fases, tal y como refleja la figura 8.

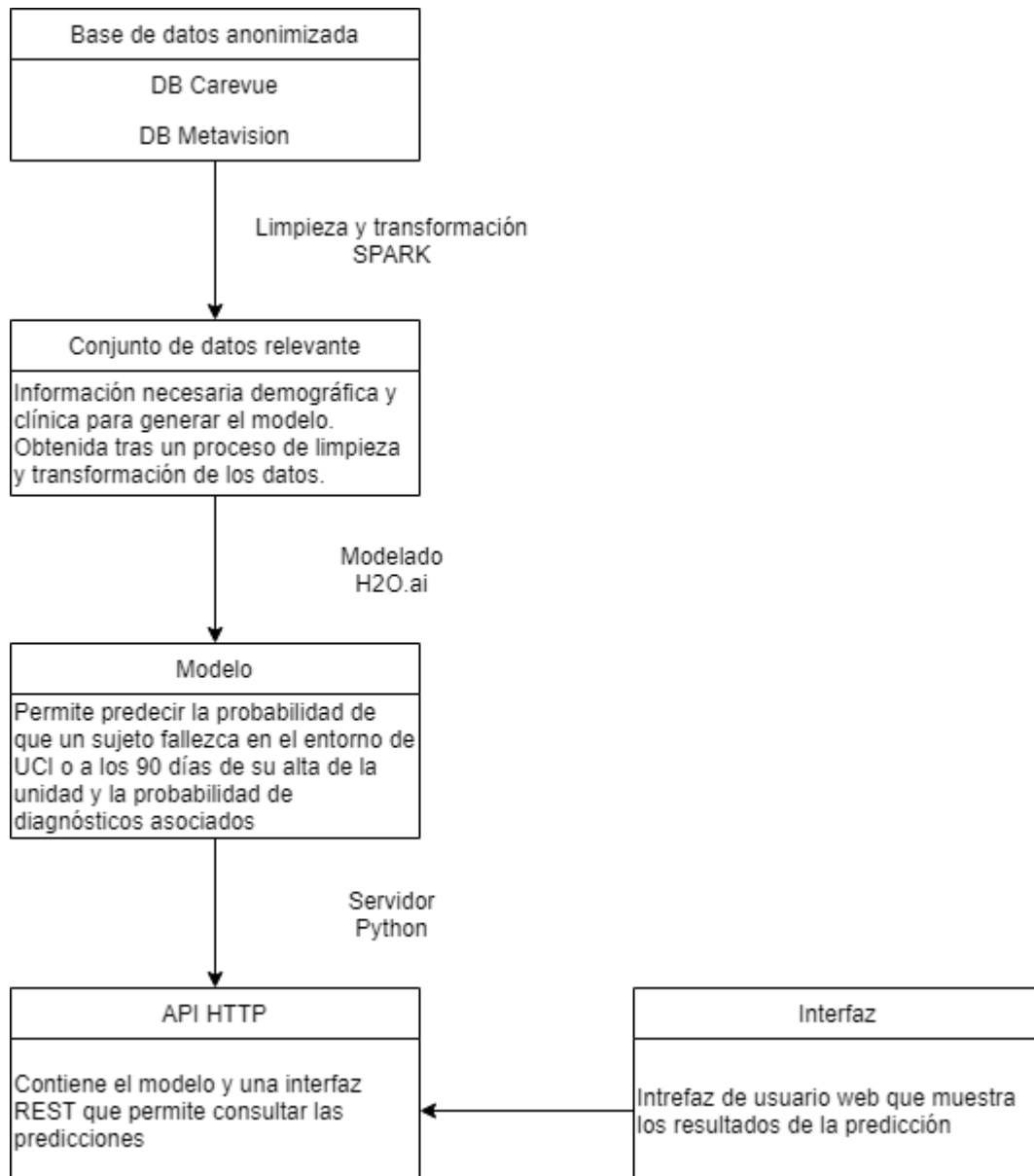


Figura 8: Arquitectura de la herramienta software desarrollada

6. Conjunto de datos a analizar

A lo largo de este capítulo se hará un análisis del conjunto de datos elegido para llevar a cabo el modelado.

Se ha decidido trabajar con el conjunto de datos MIMIC-III (**M**edical **I**nformation **M**art for Intensive **C**are III), un conjunto de datos elaborado por el *MIT Lab for Computational Physiology* que contiene información de más de 40.000 pacientes de cuidados críticos, incluyendo pacientes adultos y pacientes neonatales [15]. Se estudiará concretamente la

versión v1.4, publicada el 2 de septiembre de 2016. En ella se recoge información de dos softwares de información clínica diferentes y descritos anteriormente. Los primeros registros datan de 2001 y fueron registrados con el programa careVue de Philips. El nombre careVue hace referencia a versiones anteriores del producto ICCA descrito en el capítulo 2. En el año 2008 el hospital cambió de software a Metavision de iMDSOFT, por lo que los registros del conjunto de datos pertenecen a dos aplicaciones diferentes, hecho que se tendrá que tener en cuenta en la fase de análisis, limpieza y transformación de datos.

A lo largo de este apartado se detallará la motivación para la elección de esta base de datos frente a otras disponibles, la estructura que presenta y los pasos seguidos para obtener el acceso a los mismos.

Motivación para la elección del conjunto de datos

Después de analizar varios datasets, se decidió utilizar MIMIC-III por los siguientes motivos:

1. El laboratorio responsable define los datos del conjunto como **no identificables**: Este fue uno de los factores determinantes, ya que al tratarse de información sensible por ser datos médicos es imprescindible que no se pueda asociar dicha información a ningún individuo. A continuación se detallan las técnicas más relevantes del proceso de anonimización [16]:
 - 1.1. Alteración de las fechas: Mediante un proceso de adición de ruido.
 - 1.1.1. Todos los registros con formato fecha fueron modificados a una fecha futura.
 - 1.1.2. Se mantiene el incremento de fecha constante para cada paciente, variando entre pacientes, con el fin de ofrecer información fiable de la relación en los datos intra-hospitalarios.
 - 1.1.3. Se mantiene el día de la semana y la estación de año.
 - 1.2. Alteración en la edad de los pacientes: En este caso se lleva a cabo un proceso de supresión y un proceso de agregación.
 - 1.2.1. Aquellos pacientes que cumplieron los 90 años de edad durante el ingreso hospitalario, han sido suprimidos del estudio. Se trata de un grupo de pacientes muy pequeño y mantener esa información haría a los pacientes potencialmente identificables.
 - 1.2.2. Los pacientes con una edad de 89 años al ingreso sufrieron un proceso de agregación, unificando todas las edades bajo un valor de 200 años. De tal manera disminuye la posibilidad de re-identificación.
 - 1.3. Eliminación de número de historia clínica e identificador personal del personal sanitario. Tras un proceso de supresión, todos los identificadores únicos de persona, tanto paciente como equipo clínico, han sido eliminados.
2. **Gran volumen** de registros: El conjunto cuenta con datos de más de 40.000 pacientes.

3. **Consistencia:** Al tratarse de información registrada en un único hospital, el Beth Israel Deaconess Medical Center, se garantiza la consistencia de los datos. Se han rechazado otros conjuntos de datos como eICU porque, a pesar de incluir registros de aproximadamente 200.000 pacientes, pertenece a diferentes hospitales. En este proyecto se considera importante la unificación de criterios clínicos, y esto se obtiene en mayor grado trabajando con datos de un único hospital.
4. **Diversidad de tipología de variables:** El conjunto de datos elegido contiene información relevante de diferente tipología. En el siguiente apartado de este capítulo se detallará, destacando las siguientes:
 - 4.1. Datos demográficos
 - 4.2. Constantes vitales capturadas automáticamente de dispositivos de cabecera
 - 4.3. Resultados de laboratorio y microbiología
 - 4.4. Procedimientos y diagnósticos codificados
 - 4.5. Prescripción y registro de administración de medicaciones
 - 4.6. Estado del paciente al alta del hospital

Descripción de la base de datos

MIMIC-III es un conjunto de datos en el que la información está estructurada en tablas, que pueden unirse a través de identificadores únicos, definidos con el sufijo ID.

Cuenta con unas tablas de definiciones, identificadas con el prefijo “D_” cuyo objetivo es dar sentido a los numerosos “ID”.

Las tablas del conjunto pueden definirse siguiendo los siguientes grupos:

1. Descripción de pacientes: Éstas contienen la información necesaria para poder definir y agrupar pacientes, manteniendo siempre la privacidad, ya que se han excluido nombres y datos que puedan ser una potencial amenaza para la re-identificación.
 - 1.1. ADMISSIONS: En esta tabla se almacena la información básica demográfica del paciente junto con fecha y hora de admisión y de alta. Entre los datos demográficos más relevantes, se destaca el motivo o diagnóstico de ingreso y el estado al alta del paciente, tanto de la unidad de cuidados críticos como del hospital.
 - 1.2. CALLOUT: En esta tabla se registra la información de cuando un paciente recibe el alta médica y cuando finalmente se da de alta de la unidad. Este último dato es también conocido como “alta administrativa”, ya que dependerá de la disponibilidad de cama en el servicio de destino.
 - 1.3. ICUSTAYS: En esta tabla se encuentran los movimientos que ha tenido el paciente durante una estancia en las diferentes unidades. Otro dato relevante que se puede

analizar con la información de esta tabla es el sistema en el que se registraron los datos, carevue o metavision.

- 1.4. PATIENTS: En ésta, está registrada la fecha de nacimiento y sexo del paciente, así como desglosada con mayor detalle la fecha de muerte de un paciente, indicando si falleció dentro de la unidad, en el hospital o fuera de él.
 - 1.5. SERVICES: Se registra la información del servicio clínico al que está a cargo cada paciente. Esta información no será relevante para el estudio.
 - 1.6. TRANSFERS: En esta tabla se encuentran registradas las entradas y salidas de cada unidad de un paciente. Al igual que la tabla anterior, la información aquí almacenada no es relevante para el proyecto.
2. Información clínica recogida durante la estancia del paciente en la unidad de cuidados críticos:
 - 2.1. CAREGIVERS: Se almacena el rol y permisos de acceso a la aplicación de cada usuario.
 - 2.2. CHARTEVENTS: En esta tabla, se registra todos los datos aislados con formato número o cadena de caracteres, que hayan sido registrados por personal clínico para un paciente. Esta información será muy significativa para el estudio del proyecto.
 - 2.3. DATETIMEEVENTS: Se encuentra el mismo tipo de información que en la tabla anterior pero con formato fecha y hora.
 - 2.4. INPUTEVENTS_CV/INPUTEVENTS_MV: Se registran todos los volúmenes de entrada de los pacientes. En este caso, por tener diferente estructura los datos, se han separado en dos tablas diferentes, una para los pacientes cuyos registros se llevaron a cabo en carevue y otros para los pacientes de metavision.
 - 2.5. NOTEVENTS: Se trata de los informes tanto de médicos como de enfermería asociados a cada paciente. La información ha sido anonimizada. Esta tabla es muy valiosa por contener múltiple información clínica. Sin embargo, al tratarse de texto libre de gran longitud, el tiempo de análisis aumenta significativamente.
 - 2.6. OUTPUTEVENTS: Se trata de una tabla análoga a INPUTEVENTS, pero donde se registran todos los volúmenes asociados a pérdidas del paciente.
 - 2.7. PROCEDUREEVENTS_MV: En esta tabla están registrados los procedimientos intra-UCI de los pacientes con registro en Metavision. Esta información no se usará para el estudio, ya que no está recogida desde ambos sistemas.
 3. Información administrativa y clínica recogida en el entorno hospitalario, pero no exclusivamente en la unidad de cuidados críticos.
 - 3.1. CPTEVENTS: Se almacenan todos los códigos asociados a procedimientos realizados para cada paciente.

- 3.2. **DIAGNOSES_ICD**: Esta tabla contiene un registro por cada diagnóstico que se haya codificado para cada paciente siguiendo el código CIE-9. (Clasificación Internacional de Enfermedades 9ª revisión)
- 3.3. **DRGCODES**: En ella se encuentran los códigos diagnósticos usados por el hospital para facturación a los pacientes. No será de utilidad para el estudio.
- 3.4. **LABEVENTS**: En esta tabla se registran los resultados de los análisis de laboratorio de los pacientes.
- 3.5. **MICROBIOLOGYEVENTS**: Tabla análoga a la anterior, pero se registran los valores de laboratorio exclusivos de microbiología.
- 3.6. **PRESCRIPTIONS**: Se registran todas las órdenes que hayan sido prescritas por los médicos hayan sido o no administradas al paciente.
- 3.7. **PROCEDURES_ICD**: Esta tabla contiene un registro por cada procedimiento que se haya codificado para cada paciente siguiendo el código CIE-9.
- 4. **Tablas de definición**. En estas tablas podemos encontrar la relación entre los identificadores y el significante de los diagnósticos, procedimientos, datos de laboratorio y datos demográficos registrados en la historia clínica del paciente.

7. Herramientas y tecnologías para el análisis del conjunto de datos

En este capítulo se describirán las herramientas y tecnologías utilizadas tanto para la creación de la base de datos, como para su posterior análisis y explotación.

Docker

Docker [17] es un software libre cuyo objetivo es la automatización de despliegue de aplicaciones dentro de contenedores.

Actualmente es una de las plataformas de virtualización con mayor demanda e índice de uso, ya que a diferencia de las máquinas virtuales no es necesario contar con un kernel propio.

Además según estudios realizados, mejora los tiempos de ejecución en el análisis de grandes datasets [18].

Desde la publicación como código abierto en 2013, ha aumentado notablemente su uso, contando con contribuyentes importantes como Red Hat, IBM, Microsoft, Google o Cisco.

Su tecnología está basada en el sistema operativo Linux, contando con librerías propias.

En la figura 9, se presenta la arquitectura de Docker:

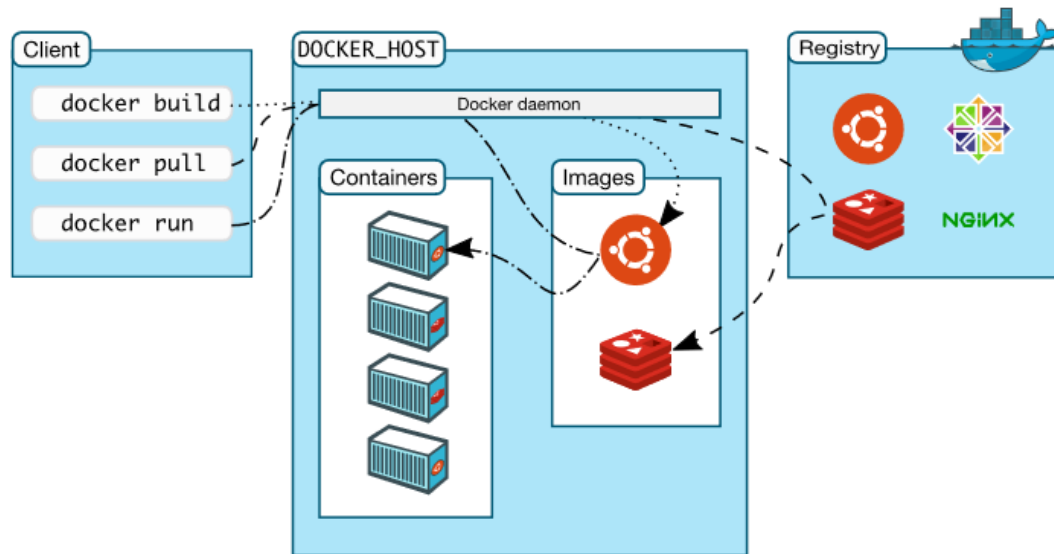


Figura 9: Arquitectura Docker [17]

Como puede observarse, existe un cliente que actúa como interfaz de usuario y un servidor compuesto por unas plantillas de solo lectura denominada imágenes y los contenedores, que proporcionan lo necesario para ejecutar la aplicación.

PostgreSQL

PostgreSQL[19] es una base de datos relacional de código libre. Como ventajas que propician la elección de esta base de datos frente a otra se encuentran las siguientes:

- Concurrencia y rendimiento: Destaca la paralelización en los procesos de lectura.
- Fiabilidad
- Buscador de texto

DBeaver

DBeaver[20] es una herramienta de base de datos de código libre cuya principal ventaja es la facilidad de uso. Además, utiliza la API de java JDBC para conectarse a las bases de datos pertinentes.

Por último, presenta una gran cantidad de plugins específicos para las bases de datos más usadas, entre ellas PostgreSQL.

Spark

Apache Spark [21] es una herramienta de código libre que proporciona diferentes APIs entre las que se incluye Python, Java, Scala y R con el objetivo de facilitar el procesamiento y preparación de datos masivos.

Además cuenta con herramientas de alto nivel como SPARK SQL, que permite el tratamiento de datos estructurados.

DataBricks

Databricks [22] es un servicio online desarrollado por los creadores de Apache Spark que proporciona las herramientas necesarias para llevar a cabo una completa preparación de los datos incluyendo limpieza y transformación de los mismos.

De forma gratuita, Databricks community ofrece un servidor con las siguientes características:

1. Memoria: 6GB
2. Scala versión: 2.11
3. Spark versión: 2.4.3
4. Python versión: 3

La figura 10 presenta Jupiter Notebook, proporcionado por Databricks para llevar a cabo la explotación de datos.

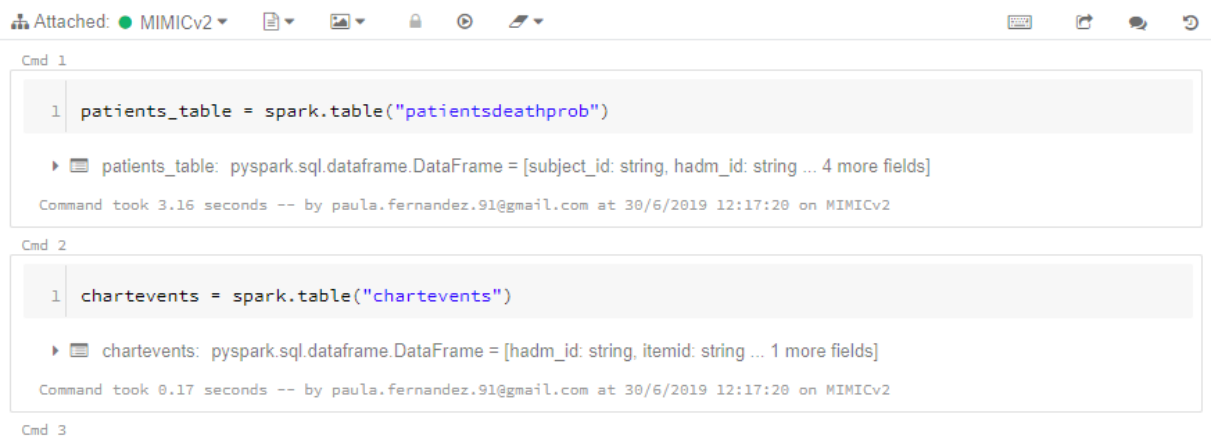


Figura 10. Jupiter Notebook

8. Herramientas y tecnologías para la creación del modelo

En este capítulo se describirá la herramienta utilizada para crear el modelo, exponiendo las ventajas que tiene el uso de esta herramienta y la motivación para su elección. Además, se explicarán las tecnologías aplicadas durante el proceso.

H2O

H2O [23] es una plataforma de código libre que incorpora diversos algoritmos de machine-learning y estadística con el fin de generar modelos predictivos usando tecnologías como el Deep-Learning.

El valor aportado por esta plataforma es su funcionalidad de AutoMachineLearning, que consiste en ejecutar diversos algoritmos con parámetros distintos en cada ejecución con el fin de obtener múltiples modelos junto con sus métricas y proporcionar al usuario un análisis de los mismos.

H2O incluye tanto algoritmos de aprendizaje no supervisado como supervisado entre los que destacan Random Forest o Generalized Low Rank Models [24].

Además, tras la realización de los experimentos, proporciona un informe completo con todos los datos relevantes para la comprensión y reproducción del modelo.

9. Herramientas de visualización para la presentación de los datos

GoogleCharts

Google charts [25] es una librería de código libre que garantiza compatibilidad con los distintos navegadores web.

Como punto clave ofrece una gran variedad de gráficos con una alta posibilidad de personalización, pudiendo adaptar cada visualización a las necesidades en función de la tipología de datos o la intención de la misma.

10. Implementación

En este capítulo, se detallarán los pasos seguidos desde la obtención del conjunto de datos, a la realización de la visualización.

Acceso al conjunto de datos

Al tratarse de un conjunto de datos con información sensible, el MIT-LCP exige que el usuario complete satisfactoriamente un curso CITI, con el título “Data for Specimens Only Research”.

Se trata de un curso online de 9 temas listado en la figura 11, cuyo objetivo es dar a conocer normativa de tratamiento de datos y ofrece diferentes casos de uso y situaciones comprometidas.

Modules	Completed
Belmont Report and Its Principles (ID 1127)	23-Apr-2019
History and Ethics of Human Subjects Research (ID 498)	23-Apr-2019
Basic Institutional Review Board (IRB) Regulations and Review Process (ID 2)	23-Apr-2019
Records-Based Research (ID 5)	23-Apr-2019
Genetic Research in Human Populations (ID 6)	23-Apr-2019
Populations in Research Requiring Additional Considerations and/or Protections (ID 16680)	23-Apr-2019
Research and HIPAA Privacy Protections (ID 14)	23-Apr-2019
Conflicts of Interest in Human Subjects Research (ID 17464)	23-Apr-2019
Massachusetts Institute of Technology (ID 1290)	23-Apr-2019

Figura 11: Contenido curso CITI “Data or Specimens Only Research”

Una vez completado el curso, un responsable del MIT valora la candidatura y proporciona acceso al conjunto de datos.

Importación de los datos

Para importar los datos y poder trabajar con ellos se usa una base de datos PostgreSQL en un contenedor de Docker.

El proceso de importación de los datos se resume en figura 12:

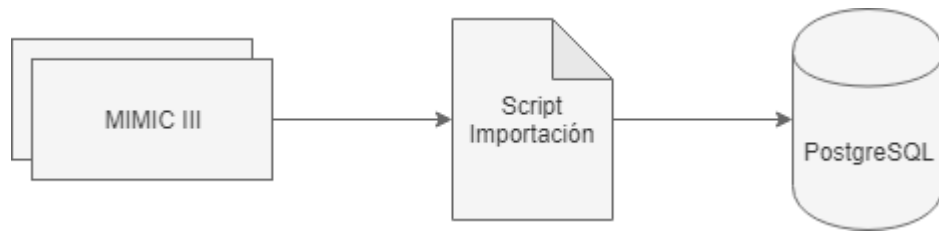


Figura 12: Importación de la base de datos MIMIC III a PostgreSQL

Para ello, el laboratorio MIT-LCP proporciona una imagen que contiene la base de datos PostgreSQL instalada y se deben seguir los siguientes pasos[26]:

Importación de la imagen

Para construir la imagen, ejecutamos el código fuente 1:

```
cd /mimic_code/buildmimic/docker
docker build -t postgres/mimic .
```

Código fuente 1: Construcción imagen Docker

La figura 13 muestra la confirmación de que el proceso se ha finalizado con éxito:

pau	fixe	run	media	mimic-code	buildmimic	docker	master	1\	\$	docker images				
REPOSITORY						TAG				IMAGE ID	CREATED		SIZE	
postgres/mimic						latest				8d5a5ee39abc	3 minutes ago		456 MB	

Figura 13: Salida comando docker images

Despliegue del contenedor

Para arrancar la base de datos PostgreSQL con Docker, se inicializa el contenedor con los parámetros especificados en el código fuente 2:

```
docker run
--name mimic
-p 5432:5432
-e BUILD_MIMIC=1
-e POSTGRES_PASSWORD=mimic
-e MIMIC_PASSWORD=mimic
-v /run/media/pau/Data/MIMICIII/uncompressed:/mimic_data:Z
-v /run/media/pau/Data/MIMICIII/pgdata:/var/lib/postgresql/data:Z
-d postgres/mimic
```

Código fuente 2: Despliegue contenedor Docker

La figura 14 muestra el proceso que se lleva a cabo durante la creación de la base de datos.

```
pau: fixe [0] run [0] media [0] pau [0] Data1 [0] MIMICIII $ docker logs -f mimic
The files belonging to this database system will be owned by user "postgres".
This user must also own the server process.
The database cluster will be initialized with locale "en_US.utf8".
The default database encoding has accordingly been set to "UTF8".
The default text search configuration will be set to "english".
Data page checksums are disabled.
fixing permissions on existing directory /var/lib/postgresql/data ... ok
creating subdirectories ... ok
selecting default max_connections ... 100
selecting default shared_buffers ... 128MB
selecting dynamic shared memory implementation ... posix
creating configuration files ... ok
running bootstrap script ... ok
performing post-bootstrap initialization ... ok
syncing data to disk ... ok
Success. You can now start the database server using:
   pg_ctl -D /var/lib/postgresql/data -l logfile start
WARNING: enabling "trust" authentication for local connections
You can change this by editing pg_hba.conf or using the option -A, or
--auth-local and --auth-host, the next time you run initdb.
waiting for server to start... 2019-06-08 15:19:43.495 UTC [44] LOG:  listening on Unix socket "/var/run/postgresql/.s.PGSQL.5432"
2019-06-08 15:19:43.676 UTC [45] LOG:  database system was shut down at 2019-06-08 15:19:41 UTC
2019-06-08 15:19:43.727 UTC [44] LOG:  database system is ready to accept connections
done
server started
```

Figura 14: Salida del comando `docker logs -f mimicdocker`, antes de terminar la creación de la base de datos

Se trata de un proceso largo de aproximadamente cuatro horas de duración.

Cuando todas las tablas estén correctamente creadas, obtendremos los mensajes que se observan en la figura 15 como parte de la salida del comando anterior:

```
/usr/local/bin/docker-entrypoint.sh: running postgres_checks.sql (all rows should return PASSED)
```

tbl	expected_count	observed_count	row_count_check
admissions	58976	58976	PASSED
callout	34499	34499	PASSED
caregivers	7567	7567	PASSED
chartevents	330712483	330712483	PASSED
cptevents	573146	573146	PASSED
datatimeevents	4485937	4485937	PASSED
d_cpt	134	134	PASSED
diagnoses_icd	651047	651047	PASSED
d_icd_diagnoses	14567	14567	PASSED
d_icd_procedures	3882	3882	PASSED
d_items	12487	12487	PASSED
d_labitems	753	753	PASSED
drgcodes	125557	125557	PASSED
icustays	61532	61532	PASSED
inputevents_cv	17527935	17527935	PASSED
inputevents_mv	3618991	3618991	PASSED
labevents	27854055	27854055	PASSED
microbiologyevents	631726	631726	PASSED
noteevents	2083180	2083180	PASSED
outputevents	4349218	4349218	PASSED
patients	46520	46520	PASSED
prescriptions	4156450	4156450	PASSED
procedureevents_mv	258066	258066	PASSED
procedures_icd	240095	240095	PASSED
services	73343	73343	PASSED
transfers	261897	261897	PASSED

(26 rows)

Figura 15: Salida del comando `docker logs -f mimicdocker`, confirmando la creación de la base de datos

Limpieza de los datos

Esta fase es una de las fases críticas del proyecto. Es imprescindible hacer una buena selección de los datos que se van a usar para la creación del modelo.

Por otro lado, es necesario definir cuáles van a ser las variables predictivas y las variables respuesta.

Variables predictivas

Las variables predictivas son aquellas que servirán de entrada al algoritmo y que permitirán clasificar el paciente. Podemos separarlas en demográficas y clínicas.

Variables predictivas demográficas: Edad, sexo y diagnóstico de ingreso.

A continuación, se expondrá cómo se han extraído dichas variables:

Edad: En este caso se ha calculado el campo en función de la fecha de nacimiento y la fecha de admisión del episodio. Esta información se encuentra almacenada en la tabla PATIENTS.

Motivo de ingreso: El motivo de ingreso está codificado en la tabla DIAGNOSES_ICD usando la nomenclatura CIE-9. En dicha tabla se encuentran todos los diagnósticos, por lo que filtramos por el seq_num = 1, ya que es éste el principal y por lo tanto el motivante del ingreso. Además, para evitar errores al registrar el código, eliminamos todos aquellos datos que contengan un carácter no numérico mediante una expresión regular.

Sexo: Esta información está registrada en la tabla PATIENTS, pudiendo tomar valores “M” para masculino y “F” para femenino.

Con las variables predictivas demográficas definidas, se procede a realizar la limpieza de los datos mediante el proceso reflejado en la figura 16:

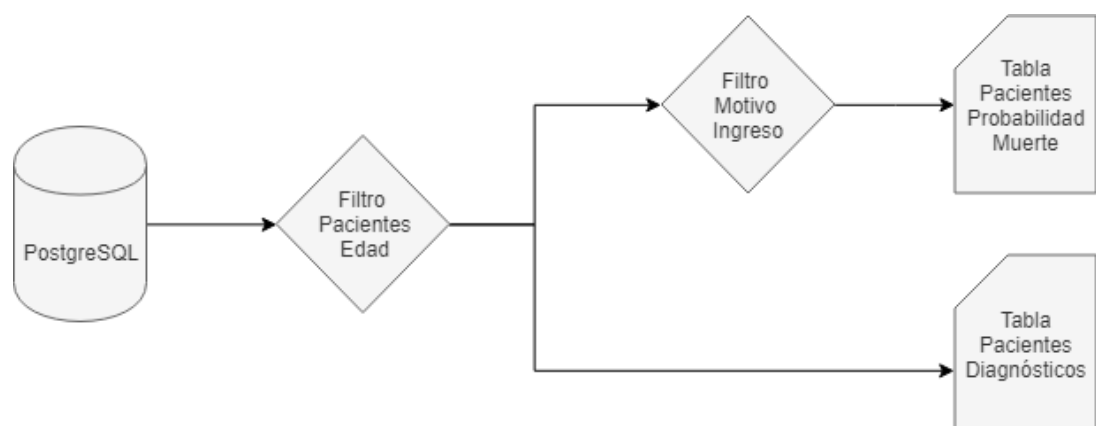


Figura 16: Proceso de limpieza de variables predictivas demográficas

En primer lugar, se excluirán del análisis los pacientes neonatales, ya que las patologías, evoluciones y tasas de mortalidad de este grupo de individuos varía considerablemente.

Además, se desecharán de igual modo todos aquellos pacientes que no cumplieran 15 años en el año del ingreso, por estar estos en la fase médica pediátrica.

Variables predictivas no demográficas: Estos datos corresponden a los eventos registrados o bien automáticamente desde los dispositivos de cabecera del paciente, o bien introducidos manualmente por parte del personal clínico.

Esta información está contenida en la tabla chartEvents, y el proceso de extracción y limpieza de los datos es el esquematizado en la figura 17:

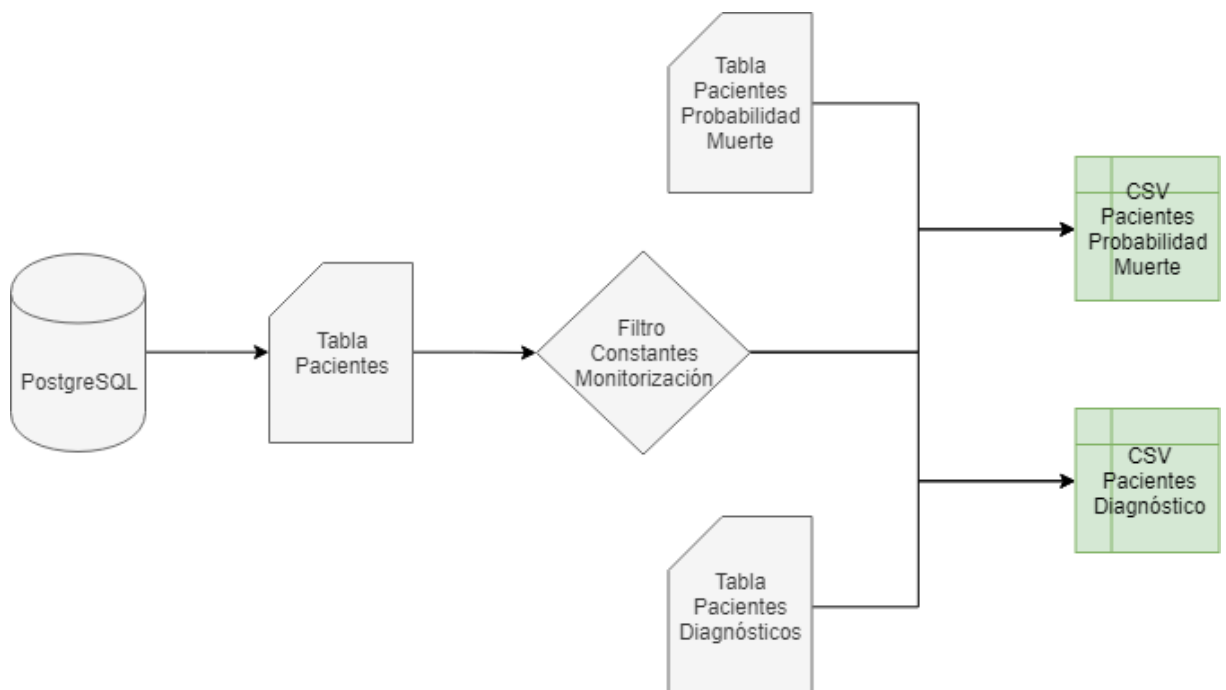


Figura 17: Proceso de limpieza de datos de variables predictivas clínicas

Puesto que la tabla chartevents, en la que se encuentran todos los datos de constantes médicas registrados para los pacientes tiene un tamaño de 33GB, ha sido necesario realizar una limpieza basada en los siguientes criterios.

Filtrado de columnas

Únicamente son necesarias hadm_id, itemid y valuenum. Para llevar a cabo esta reducción del fichero, se usa el comando del código fuente 3:

```
awk 'BEGIN { FS=","; OFS=","; } {print $2,$4,$8,$9}' CHARTEVENTS.csv > CHARTEVENTS_min.csv
```

Código fuente 3: Reducción columnas de la tabla de chartevents

Filtrado de columnas

Se decide usar como variables clínicas predictoras un conjunto de constantes de monitorización básicas.

Para obtener exclusivamente los datos que servirán como variables predictoras, se lleva a cabo un filtrado de filas dejando exclusivamente aquellas cuyo identificador está definido en la figura 18.

	Carevue		Metavision	
	itemid	label	itemid	label
Frecuencia Cardíaca	211	heart rate	220045	heart rate
	51	Arterial BP [Systolic]	220050	Arterial Blood Pressure systolic Non Invasive Blood Pressure
PA sistólica	455	NBP [Systolic]	220179	systolic
	3313	BP Cuff [Systolic]		
PA diastólica	8368	Arterial BP [Diastolic]	220051	Arterial Blood Pressure diastolic Non Invasive Blood Pressure
	8441	NBP [Diastolic]	220180	diastolic
PA media	8502	BP Cuff [Diastolic]		
	52	Arterial BP Mean	220052	Arterial Blood Pressure mean Non Invasive Blood Pressure
Temperatura	456	NBP Mean	220181	mean
	3312	BP Cuff [Mean]		
Presión Venosa Central	676	Temperature C	223761	Temperature Fahrenheit
	677	Temperature C (calc)	223762	Temperature Celsius
	113	CVP	220074	Central Venous Pressure

Figura 18: Tabla variables predictoras clínicas

Variables respuesta

Las variables a predecir serán el estado al alta del paciente y los diagnósticos del paciente durante el episodio.

Estado al alta del paciente

Según lo definido por la Semicyuc, el estado a alta de un paciente, es decir si está vivo o muerto, se puede asociar a un episodio asistencial concreto siempre y cuando dicha muerte se produzca dentro de la unidad o en los 90 días posteriores a su alta de la unidad de cuidados críticos.

Por lo tanto, en el caso de estudio, la variable puede tomar los siguientes valores:

Valor	Descripción
0	Paciente actualmente vivo o fallecido 90 días después del alta en la unidad de cuidados críticos. Según los indicadores de Semicyuc, no se debe asociar dicho fallecimiento al episodio asistencial.
1	Paciente fallecido en la unidad de cuidados críticos
2	Paciente fallecido en los 90 días posteriores al alta de la unidad de cuidados críticos

Diagnósticos durante el episodio asistencial

Durante el ingreso de un paciente se registran diversos diagnósticos. Por un lado, aquel que motiva el ingreso del paciente en la unidad y por otro, todos los asociados al paciente durante el ingreso.

Creación de tablas

Para la posterior creación del modelo es necesario generar una única tabla que incluya todas las variables predictoras y respuesta definidas con anterioridad.

En primer lugar, crearemos una tabla con todos los datos demográficos, independientemente de su naturaleza predictora o respuesta con el código fuente 4.

Toda relación entre los datos estará basada en el identificador `hadm_id`, ya que es éste el que identifica de manera única cada episodio clínico. No se debe confundir con `subject_id`, ya que este último está relacionado con un paciente y un paciente puede tener diversos episodios que deben ser tratados de manera independiente.

```

CREATE TABLE public.patientsInfo(
  subject_id numeric NOT NULL,
  hadm_id numeric NOT null,
  edadingreso numeric,
  diagIn varchar,
  deathloc numeric,
  sexo varchar
);

insert into patientsInfo
SELECT admissions.subject_id,
       admissions.hadm_id,
       ( SELECT date_part('year'::text, admissions.admittime) - date_part('year'::text, patients.dob)
         FROM patients
         WHERE patients.subject_id = admissions.subject_id) AS edadingreso,
       ( SELECT diagnoses_icd.icd9_code
         FROM diagnoses_icd
         WHERE diagnoses_icd.hadm_id = admissions.hadm_id AND diagnoses_icd.seq_num = 1 AND
         diagnoses_icd.icd9_code::text !~ '.*[a-zA-Z].*':text) AS diagIn,
       CASE
         WHEN ( SELECT patients.dod
                 FROM patients
                 WHERE patients.subject_id = admissions.subject_id) >= admissions.admittime AND ( SELECT patients.dod
                 FROM patients
                 WHERE patients.subject_id = admissions.subject_id) <= admissions.disctime THEN 1
         WHEN (( SELECT date_part('day'::text, (( SELECT patients_1.dod
                 FROM patients_1
                 WHERE patients_1.subject_id = admissions.subject_id)) - admissions.disctime) AS date_part
                 FROM patients
                 WHERE patients.subject_id = admissions.subject_id)) <= 90::double precision THEN 2
         ELSE 0
       END AS deathloc, ( SELECT patients.gender
                         FROM patients
                         WHERE patients.subject_id = admissions.subject_id) AS sexo
FROM admissions

```

Código fuente 4: Creación tabla patientsInfo

Para crear el modelo predictivo de la probabilidad de muerte, solo trabajaremos con aquellos pacientes catalogados clínicamente como adultos y para ello creamos una tabla nueva que incluya exclusivamente estos pacientes, como refleja el código fuente 5. Adicionalmente, en esta fase, eliminamos todos aquellos individuos cuya edad sea mayor de 99 años, ya que según lo especificado en la documentación del conjunto de datos usado la edad máxima es 99 y por lo tanto valores mayores se consideran errores.

```

CREATE TABLE public.cohortDeath(
  subject_id numeric NOT NULL,
  hadm_id numeric NOT null,
  edadingreso numeric,
  diagIn numeric,
  deathloc numeric,
  sexo varchar
);

insert into cohortDeath
SELECT subject_id, hadm_id, edadIngreso, cast(diagIn as integer), deathloc, sexo from patientsinfo where edadIngreso
>=15 and edadIngreso <=99

```

Código fuente 5: Creación tabla cohortDeath

Tras este proceso, se obtiene la table cohortDeath que incluye 48.250 admisiones diferentes.

Para el modelo predictivo de los diagnósticos asociados, se analizarán aquellos pacientes con edad comprendida entre 15 y 99 años y cuyo motivo de ingreso esté comprendido entre los diagnósticos asociados al sistema circulatorio, es decir con códigos CIE-9 entre 390 y 459.

Este filtrado de pacientes se realiza en el código fuente 6:

```
CREATE TABLE public.cohortDiagnostico(  
  subject_id numeric NOT NULL,  
  hadm_id numeric NOT null,  
  edadIngreso numeric,  
  diagIn numeric,  
  deathloc numeric,  
  sexo varchar  
);  
insert into cohortDiagnostico  
SELECT subject_id, hadm_id, edadIngreso, cast(diagIn as integer), deathloc, sexo from patientsinfo where edadIngreso  
>=15 and edadIngreso <=99 and ( cast(diagIn as Integer) between 390 and 459 )
```

Código fuente 6: Creación tabla cohortDiagnostico

Tras este proceso, se obtiene la tabla cohortDiagnostico, que contiene 1.517 admisiones diferentes.

A continuación, asociamos a los pacientes preseleccionados por el proceso anterior, todos los diagnósticos codificados durante su estancia en la unidad de cuidados críticos.

Para ello debemos crear en primer lugar una tabla que incluya tanto el código del CIE-9 almacenado en `diagnoses_icd`, como la etiqueta definida en la tabla de definición `D_icd_diagnoses`, ya que el usuario deberá obtener como resultado de la predicción la etiqueta y no el código. Esta acción se lleva a cabo con el código fuente 7.

```
CREATE TABLE public.patientsDiag(  
  hadm_id numeric NOT null,  
  codigo varchar,  
  etiqueta varchar  
);  
insert into patientsDiag  
select hadm_id, diagnoses_icd.icd9_code, long_title from diagnoses_icd join d_icd_diagnoses on  
diagnoses_icd.icd9_code = d_icd_diagnoses.icd9_code where seq_num != 1
```

Código fuente 7: Creación tabla patientsDiag

Posteriormente, con el código fuente 8 asociamos cada uno de los registros de diagnóstico codificado con el episodio asistencial al que pertenece y lo almacenaremos en una tabla con el nombre `patientsDiagnostico`.

```
CREATE TABLE public.patientsDiagnosticos(  
  hadm_id numeric NOT null,  
  edadingresso numeric,  
  diagin numeric,  
  sexo varchar,  
  codigo varchar,  
  etiqueta varchar  
);  
insert into patientsDiagnosticos  
select cohortdiagnostico.hadm_id, cohortdiagnostico.edadingresso, cohortdiagnostico.diagin, cohortdiagnostico.sexo,  
patientsDiag.codigo, patientsDiag.etiqueta from cohortdiagnostico join patientsDiag on cohortdiagnostico.hadm_id =  
patientsDiag.hadm_id
```

Código fuente 8: Creación tabla patientsDiagnosticos

Para extraer las variables clínicas, se ha recurrido a Databricks ya que el conjunto de datos era demasiado pesado para poder ser tratado con la herramienta DBeaver utilizada anteriormente.

El proceso llevado a cabo, está reflejado en el diagrama de la figura 19:

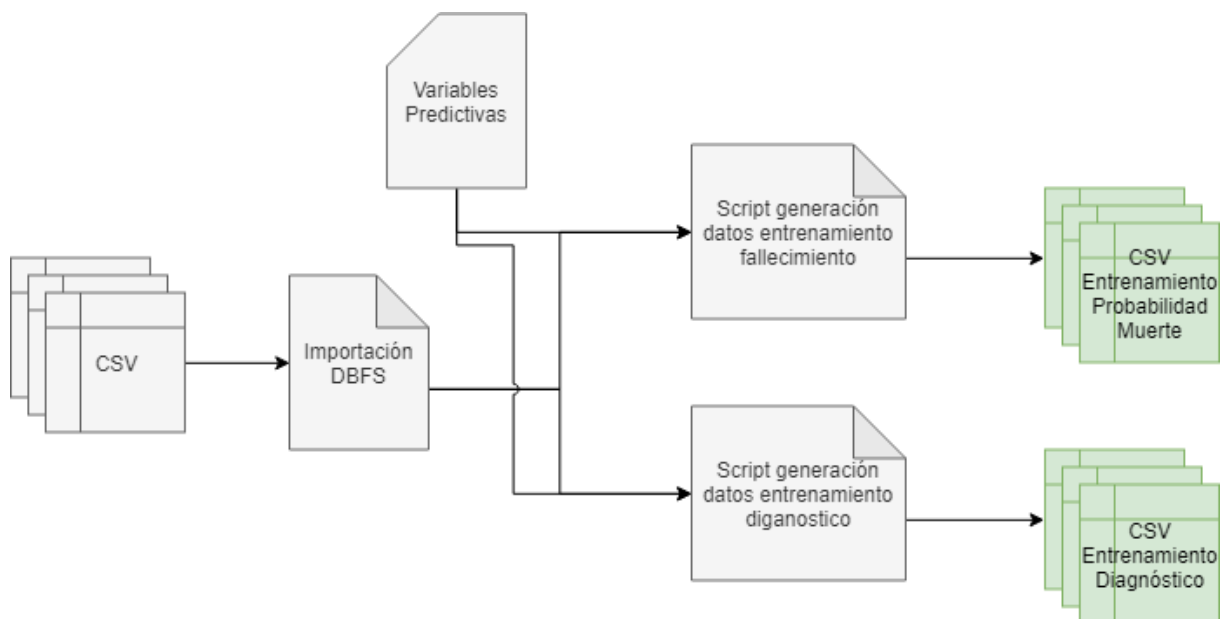


Figura 19: Generación datos de entrenamiento

En primer lugar, exportamos desde DBeaver a formato csv las tablas cohortDeath y patientsDiagnosticos, siguiendo el proceso descrito en la figura 20:

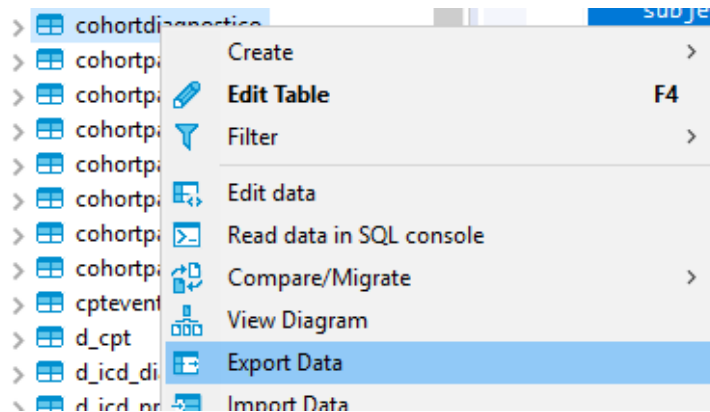


Figura 20. Exportación tablas

Para después importarlas en Databricks siguiendo los pasos de la figura 21, junto con la tabla chartevents que ha sido limpiada según los pasos descritos con anterioridad:

Create New Table (Python)

Data source: Upload File, S3, DBFS, Other Data Sources

Upload to DBFS: /FileStore/tables/ (optional) [Select]

File: cohortdeath.csv (1.3 MB) [Remove file]

File uploaded to /FileStore/tables/cohortdeath.csv

[Create Table with UI] [Create Table in Notebook]

Select a Cluster to Preview the Table

Choose a cluster with which you will read and preview the data.

Cluster: MIMICv2 (6 GB, Running, 5.4 (includes Apache Spark 2.4.0, ...))

[Preview Table]

Specify Table Attributes

Specify the Table Name, Database and Schema to add this to the data UI for other users to access

Table Name: cohortdeath_csv

Create in Database: default

File Type: CSV

Column Delimiter: ,

☒ First row is header

☐ Infer schema

☐ Multi-line

[Create Table]

[Create Table in Notebook]

Table Preview

	subject_id	hadm_id	edadingsreso	diagn	deathloc	sexo
	FLOAT	FLOAT	FLOAT	FLOAT	INT	STRING
22		165315	65	9678	0	F
23		152223	71	41401	0	M
23		124321	75	2252	0	M
24		161859	39	41041	0	M
25		129635	59	41071	0	M
26		197661	72	99604	0	M
28		162569	74	41401	0	M

Figura 21: Importación de las tablas en Databricks

A continuación, se detalla cómo se construye y la información que contiene el CSV que servirá para generar el modelo de predicción de muerte del paciente.

En primer lugar, se cargan las tablas importadas anteriormente.

Posteriormente declaramos las variables clínicas que serán usada como predictivas ejecutando el código fuente 9, estableciendo los límites inferiores y superiores establecidos según criterio clínico tal y como quedan recogidos en la figura 22:

	Mínimo	Bajo	Alto	Máximo
Frecuencia Cardíaca	10	50	120	300
PA sistólica	10	90	160	300
PA diastólica	5	50	90	180
PA media	10	70	110	200
Temperatura	30	35	41	45
PVC	-20	1	20	300

Figura 22. Límites variables predictivas clínicas

```
analysis = [
  {
    "name": "frec_cardiaca",
    "values": ["211", "51", "220045", "220050"],
    "min": 10,
    "max": 300.0
  },
  {
    "name": "pa_sistolica",
    "values": ["455", "3313", "8368", "220179", "220051"],
    "min": 10,
    "max": 300.0
  },
  {
    "name": "pa_diastolica",
    "values": ["8441", "8502", "52", "220180", "220052"],
    "min": 5,
    "max": 180.0
  },
  {
    "name": "pa_media",
    "values": ["456", "3312", "220181"],
    "min": 10,
    "max": 200.0
  },
  {
    "name": "temperature",
    "values": ["676", "677", "223761", "223762"],
    "min": 30.0,
    "max": 45.0
  },
  {
    "name": "pv_central",
    "values": ["113", "220074"],
    "min": -20.0,
    "max": 300.0
  }
]
```

Código fuente 9: Variables predictivas clínicas

Para después calcular mediante un script detallado en el Anexo II el valor mínimo, máximo y medio por episodio de cada una de las variables:

Finalmente se exporta la tabla generada en formato CSV.

Puesto que, para ser considerado como registro válido, debe contener todas las variables predictivas descritas anteriormente, esta tabla contará con 17.572 admisiones diferentes.

Procedemos de manera análoga, para construir la tabla que contenga la información necesaria para predecir los posibles diagnósticos.

Implementación del modelo

Con el archivo csv generado anteriormente, que contiene los datos relevantes limpios creamos el modelo predictivo con el uso de la herramienta H2O.

Las fases de la creación del modelo quedan reflejadas en la figura 23:



Figura 23: Fases de la creación del modelo

Entrada de datos: En esta fase reflejada en la figura 24, H2O detecta las columnas del csv como y seleccionamos la columna que queremos predecir y si queremos excluir alguna de las columnas.

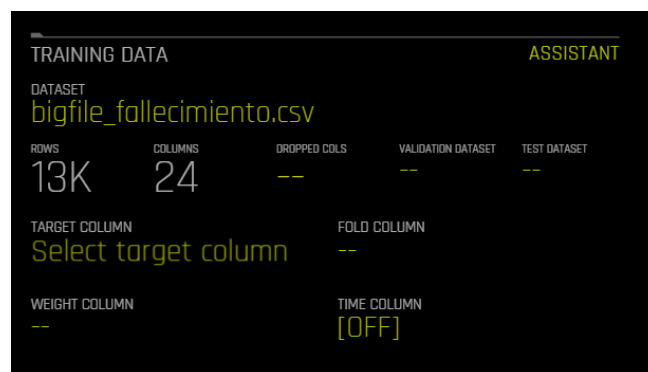


Figura 24: Entrada de datos H2O

En ambos casos, las columnas que excluimos son hadm_id y subject_id, ya que no es relevante ni el identificador de paciente ni el de episodio clínico.

Pre-procesado: Se realiza un pre-procesado automático con el fin de evaluar los datos y transformar aquellos que se pueda a valores numéricos.

Tuning: En esta fase, se realizan varias iteraciones para encontrar los parámetros óptimos. Para ello se genera un conjunto de modelos de entrenamiento con diferentes parámetros y determina como óptimos, aquellos cuyo valor de logLoss sea menor, ya que es éste el parámetro que hemos seleccionado como scorer.

El logLoss califica la calidad de un modelo penalizando los errores y otorgando un peso diferente a cada una de las clases. Por ello, a pesar de estar relacionado con la métrica “accuracy”, es más precisa al tener en cuenta más factores. Cuanto más cercano a cero sea el valor, más fiable será la predicción.

H2O permite evaluar los modelos en función de diversos parámetros reflejados en la figura 25:



Figura 25: Parámetros evaluación H2O

Evolution: En esta fase, se determina cuáles son los datos más representativos cuando se genera el modelo, y por lo tanto deben tener más peso en el algoritmo predictivo. Este proceso se lleva a cabo evaluando los diferentes modelos generados durante todas las iteraciones realizadas.

Model: En base a las iteraciones y pruebas, se determina cuál es mejor modelo predictivo.

Scoring Pipeline: Se crea y exporta el pipeline de Python.

La figura 26 representa el modelo de validación utilizado en ambas predicciones. Se trata de un modelo de validación cruzada con 5 iteraciones, con una relación 80-20 entre los datos de entrenamiento y los de validación:

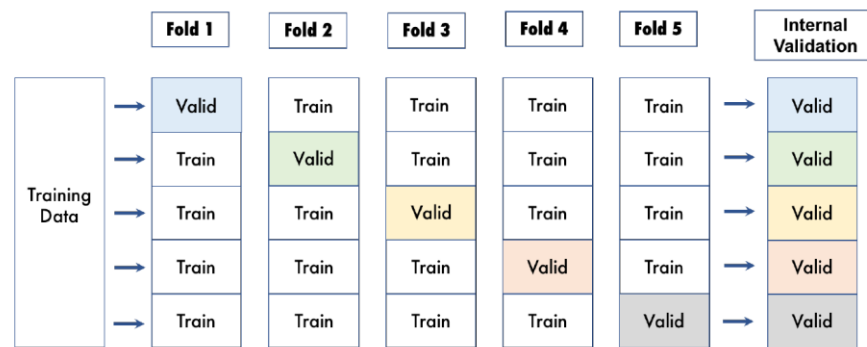


Figura 26: Validación cruzada 5-iteraciones

Predicción Probabilidad de Muerte

En esta sección se expondrán las particularidades del proceso definido anteriormente en el capítulo para la predicción de la probabilidad de muerte y se detallarán los resultados obtenidos.

En primer lugar, H2O hace una evaluación de la columna a predecir. En este primer modelo, el objetivo es la columna deathloc, que puede tener tres valores diferentes según lo reflejado en la sección “variables respuesta” de este capítulo. Esto es consistente con lo analizado por H2O, cuyos resultados muestra la figura 27:

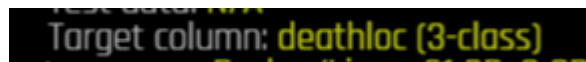
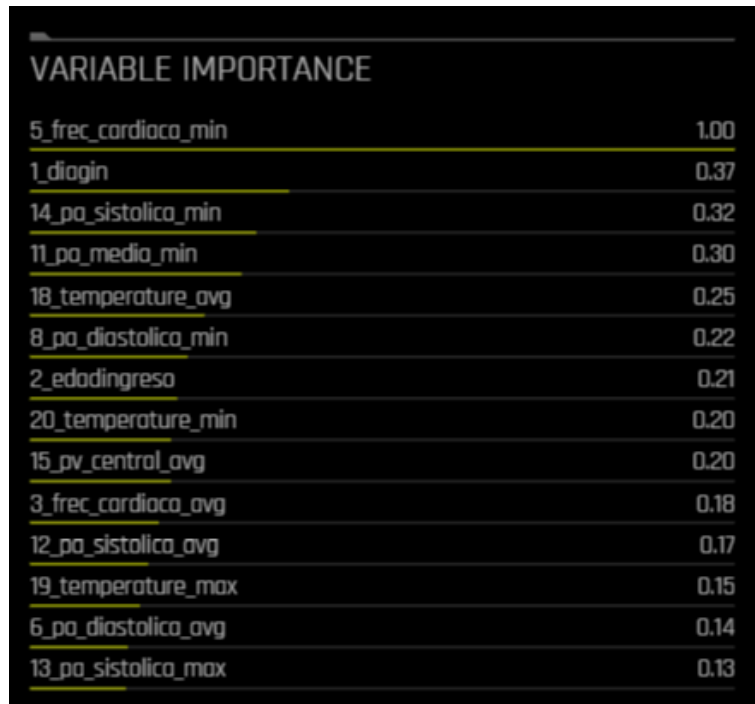


Figura 27: Evaluación columna a predecir

Una de las fases más determinantes de la predicción del modelo es “evolution” ya que la definición de los parámetros más relevantes es un punto crucial para luego poder integrar el modelo con una herramienta de software.

Como se aprecia en la figura 28, los valores mínimo de cada una de las constantes de monitorización básica seleccionadas tienen en todos los casos, excepto para la temperatura, más relevancia que las medias o máximos de los mismo.



VARIABLE IMPORTANCE	
5_frec_cardiaco_min	1.00
1_diagin	0.37
14_pa_sistolico_min	0.32
11_pa_medio_min	0.30
18_temperature_avg	0.25
8_pa_diastolico_min	0.22
2_edadingreso	0.21
20_temperature_min	0.20
15_pv_central_avg	0.20
3_frec_cardiaco_avg	0.18
12_pa_sistolico_avg	0.17
19_temperature_max	0.15
6_pa_diastolico_avg	0.14
13_pa_sistolico_max	0.13

Figura 28: Importancia de las variables en la predicción de probabilidad de muerte

Sin embargo, dado que la diferencia entre la importancia de máximo, mínimo y medias no es muy elevado, no podemos prescindir de ellos, ya que entonces el logLoss, cae drásticamente y el modelo generado no tiene valor.

Para llevar a cabo este análisis de importancia, se han realizado 40 experimentos con combinaciones diferentes y así lo refleja la memoria del experimento cuyo fragmento puede consultarse en la figura 29:

trained and scored 40 models to evaluate features and model parameters

Figura 29: Número de modelos para determinar la importancia de parámetros

Para la generación del modelo final, se realiza en la penúltima fase, una combinación de dos modelos GBMModels y un GLMModel.

El modelo GBM (Gradient Boosting Model) es un algoritmo de aprendizaje automático para llevar a cabo clasificaciones. Está basado en la construcción escalonada del modelo a partir de un conjunto de árboles de decisión y es un buen modelo para aquellos casos en los que las clases a predecir no están balanceadas [27], como es el caso que aplica a este trabajo.

El modelo GLM (Generalized Linear Model) es un algoritmo basado en la regresión lineal que está formado por una componente aleatoria de observaciones independientes, una componente sistemática que refleja la importancia que tiene cada variable predictora y una

función del valor predecido esperado. Este modelo permite llevar a cabo modelos con respuestas nominales [28] como el de este caso de estudio.

Para llegar hasta el modelo predictor, se han llevado a cabo 46 iteraciones, evaluando en cada una el score “logLoss” obteniendo los resultados de la figura 30:

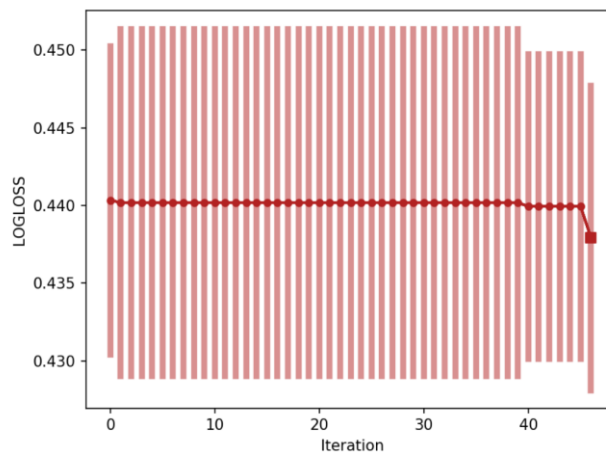


Figura 30: Valores de LogLoss para la predicción de probabilidad de muerte

Como se puede apreciar, los valores del score indican que el nivel de predicción correcta es elevado, ya que el logLoss es de 0.438.

La figura 31 presenta la matriz de confusión proporcionada por H2O tras el experimento.

MULTICLASS CONFUSION MATRIX					
		Predicted label			
		0	1	2	Error
Actual label	0	1947.6	45	2.8	0.0240
	1	160	193.2	3.2	0.4579
	2	152.4	24.4	4.2	0.9768
	Total	2260	262.6	10.2	0.1531

Figura 31: Matriz de confusión del algoritmo de probabilidad de muerte

Análisis de los resultados del modelo de predicción de muerte

Como puede verse en la matriz de confusión de la figura 30, el algoritmo predice con más de un 97% de acierto si el paciente fallece o no dentro de la unidad.

Según lo definido en la fase de limpieza y transformación de datos, el valor 0 significa que el paciente no fallece por causas asociadas al motivo que provoca su ingreso, por lo tanto, en base a los experimentos realizados, se puede determinar con tan solo un 2,4 % de probabilidad de error si el paciente fallece o no dentro de la unidad de cuidados críticos.

Sin embargo, dado que el número de registros para los valores 1 y 2 es reducido, la predicción de dichas clases no es buena.

Predicción Probabilidad de Diagnósticos Secundarios

En esta sección se expondrán las particularidades del proceso para la predicción de posibles diagnósticos asociados y se detallarán los resultados obtenidos.

En este caso la columna a predecir es el código CIE-9 de los diagnósticos. Por lo tanto además de excluir las columnas “hadm_id” y “subject_id”, excluimos también la columna “etiqueta”, ya que ésta representa el diagnóstico.

En primer lugar se eliminan todas las clases de la columna código que tenga una frecuencia de aparición menor de 5, ya que en este caso no hay datos suficientes para poder predecir una salida.

Para este modelo, las variables más relevantes para la predicción de la columna objetivo no son las columnas del csv proporcionado, si no nuevas columnas calculadas automáticamente a partir de las originales.

La figura 32 representa la importancia de cada una de las variables calculadas junto con las columnas que la forman y la transformación realizada.

Relative Importance	Feature	Description	Transformer
1.000	144_NumToCatWoE: frec_cardiaca_min: pa_media_min.2	Weight of Evidence for columns ['frec_cardiaca_min', 'pa_media_min']. Dim reduction into 5-D space applied, latent feature #2 (numeric columns are bucketed into 100 equally populated bins)	Numeric to Categorical Weight of Evidence
0.931	266_NumToCatWoE: pa_sistolica_max: temperature_max.2	Weight of Evidence for columns ['pa_sistolica_max', 'temperature_max']. Dim reduction into 5-D space applied, latent feature #2 (numeric columns are bucketed into 100 equally populated bins)	Numeric to Categorical Weight of Evidence
0.929	174_NumToCatWoE: pa_media_min: pa_sistolica_min.2	Weight of Evidence for columns ['pa_media_min', 'pa_sistolica_min']. Dim reduction into 5-D space applied, latent feature #2 (numeric columns are bucketed into 25 equally populated bins)	Numeric to Categorical Weight of Evidence
0.658	271_NumCatTE: pa_media_avg: sexo: temperature_avg: temperature_min.1	Out-of-fold mean of the response grouped by: ['pa_media_avg', 'sexo', 'temperature_avg', 'temperature_min'] using 5 folds. Dim reduction into 5-D space applied, latent feature 1. [internal:(10, 10, None)] (numeric columns are bucketed into 25 equally populated bins) [internal:(10, 10, None)]	TBD

Relative Importance	Feature	Description	Transformer
0.624	209_NumToCatWoE: frec_cardiaca_max: frec_cardiaca_min: pa_diastolica_avg: pa_diastolica_min: pa_media_min: temperature_avg.2	Weight of Evidence for columns ['frec_cardiaca_max', 'frec_cardiaca_min', 'pa_diastolica_avg', 'pa_diastolica_min', 'pa_media_min', 'temperature_avg']. Dim reduction into 5-D space applied, latent feature #2 (numeric columns are bucketed into 250 equally populated bins)	Numeric to Categorical Weight of Evidence
0.609	48_NumToCatWoE: pa_diastolica_avg: temperature_avg: temperature_min.2	Weight of Evidence for columns ['pa_diastolica_avg', 'temperature_avg', 'temperature_min']. Dim reduction into 5-D space applied, latent feature #2 (numeric columns are bucketed into 25 equally populated bins)	Numeric to Categorical Weight of Evidence
0.361	209_NumToCatWoE: frec_cardiaca_max: frec_cardiaca_min: pa_diastolica_avg: pa_diastolica_min: pa_media_min: temperature_avg.4	Weight of Evidence for columns ['frec_cardiaca_max', 'frec_cardiaca_min', 'pa_diastolica_avg', 'pa_diastolica_min', 'pa_media_min', 'temperature_avg']. Dim reduction into 5-D space applied, latent feature #4 (numeric columns are bucketed into 250 equally populated bins)	Numeric to Categorical Weight of Evidence
0.346	49_NumToCatTE: frec_cardiaca_max: pa_sistolica_max: pa_sistolica_min.0	Out-of-fold mean of the response grouped by: ['frec_cardiaca_max', 'pa_sistolica_max', 'pa_sistolica_min'] using 5 folds. Dim reduction into 5-D space applied, latent feature 0. [internal:(20, 10, 10)] (numeric columns are bucketed into 25 equally populated bins) [internal:(20, 10, 10)]	Numeric to Categorical Target Encoding
0.340	103_NumToCatTE: pa_diastolica_max: pa_media_avg: pa_media_max.3	Out-of-fold mean of the response grouped by: ['pa_diastolica_max', 'pa_media_avg', 'pa_media_max'] using 5 folds. Dim reduction into 5-D space applied, latent feature 3. [internal:(10, 1, 10)] (numeric columns are bucketed into 10 equally populated bins) [internal:(10, 1, 10)]	Numeric to Categorical Target Encoding
0.336	248_InteractionDiv: edadingreso: pa_sistolica_max	[edadingreso] / [pa_sistolica_max]	Interaction
0.299	147_NumToCatWoE: diagin: frec_cardiaca_avg: pa_diastolica_max: pa_media_max: pv_central_min.3	Weight of Evidence for columns ['diagin', 'frec_cardiaca_avg', 'pa_diastolica_max', 'pa_media_max', 'pv_central_min']. Dim reduction into 5-D space applied, latent feature #3 (numeric columns are bucketed into 100 equally populated bins)	Numeric to Categorical Weight of Evidence

Relative Importance	Feature	Description	Transformer
0.296	205_NumToCatTE: edadingreso: frec_cardiaca_max: pv_central_max: temperature_min.0	Out-of-fold mean of the response grouped by: ['edadingreso', 'frec_cardiaca_max', 'pv_central_max', 'temperature_min'] using 5 folds. Dim reduction into 5-D space applied, latent feature 0. [internal:(10, 3, None)] (numeric columns are bucketed into 25 equally populated bins) [internal:(10, 3, None)]	Numeric to Categorical Target Encoding
0.290	270_NumToCatWoE: frec_cardiaca_max: frec_cardiaca_min: pa_diastolica_avg: pa_media_max: pa_sistolica_min: pv_central_avg.4	Weight of Evidence for columns ['frec_cardiaca_max', 'frec_cardiaca_min', 'pa_diastolica_avg', 'pa_media_max', 'pa_sistolica_min', 'pv_central_avg']. Dim reduction into 5-D space applied, latent feature #4 (numeric columns are bucketed into 25 equally populated bins)	Numeric to Categorical Weight of Evidence
0.285	111_NumCatTE: diagin: pa_sistolica_avg: pa_sistolica_max: pv_central_avg: sexo: temperature_max.1	Out-of-fold mean of the response grouped by: ['diagin', 'pa_sistolica_avg', 'pa_sistolica_max', 'pv_central_avg', 'sexo', 'temperature_max'] using 5 folds. Dim reduction into 5-D space applied, latent feature 1. [internal:(10, 3, None)] (numeric columns are bucketed into 25 equally populated bins) [internal:(10, 3, None)]	TBD

Figura 32: Importancia de las variables en la predicción de diagnósticos secundarios

Al igual que para la predicción de la probabilidad de muerte, H2O general el modelo final, como combinación de dos modelos GBMModels y un GLMModel, pero en este caso durante 51 iteraciones.

La figura 33 muestra los resultados del score “logLoss”:

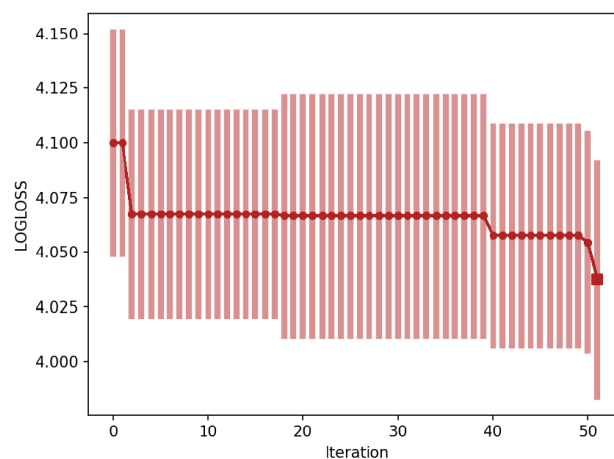


Figura 33: Valores de LogLoss para la predicción de diagnósticos secundarios

Análisis de los resultados del modelo de predicción de muerte

En este caso, a pesar de que el parámetro “accuracy” del modelo es de 0.7405, se puede ver en la figura 32 que el logLoss es significativamente mayor que en el modelo de predicción de probabilidad de muerte, teniendo un valor de 4.0375.

Esto es esperable, ya que el diagnóstico no solo depende de constante de monitorización básica, si no que debería contemplar otras variables como antecedentes médicos y medicación asociada entre otros.

Sin embargo, el algoritmo sí que determinará cuales son los diagnósticos más probables y orientará al personal médico.

Despliegue del modelo en el servidor

En la figura 34 se exponene los pasos seguidos para desplegar en un servidor el modelo obtenido en el paso anterior.

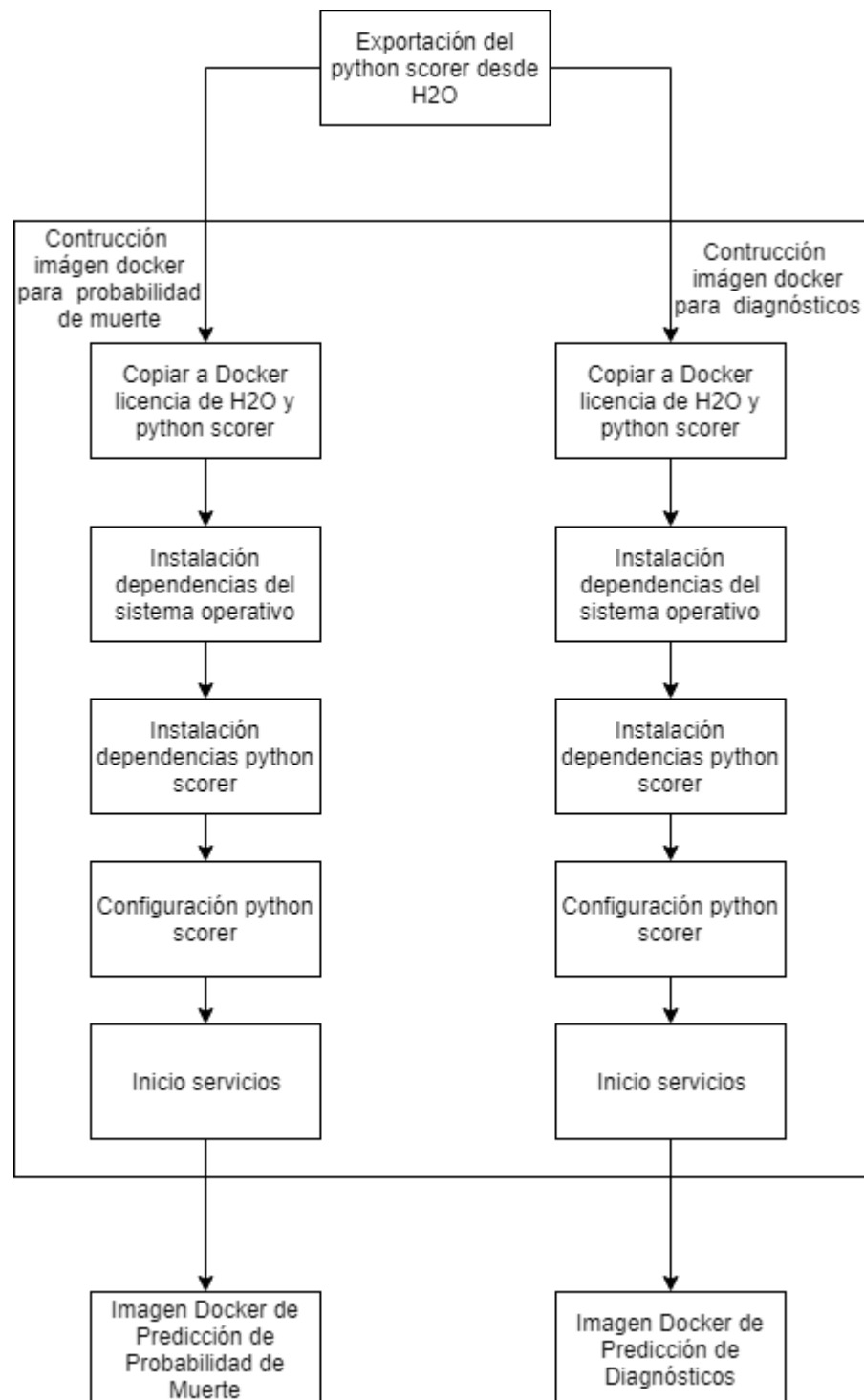


Figura 34: Pasos para el despliegue del modelo en el servidor

El código para llevar a cabo este despliegue puede consultarse en el anexo III.

Diseño de la interfaz

La interfaz de usuario debe ser clara y no debe necesitar explicaciones adicionales para su uso.

La figura 35 muestra la interfaz desarrollada. Se decide dividir la pantalla de tal manera que en la parte izquierda se presente el tablón con la información necesaria para realizar la predicción y los botones para ejecutar la prección y para borrar los datos introducidos. En la parte derecha, se presentan las visualizaciones con los resultados tanto de estado del paciente al alta como los 5 diagnósticos asociados más probables.

Predicción del estado al alta del paciente y diagnósticos asociados para el enfermo crítico

Paula Fernández Perez

Por favor, introduzca los datos solicitados y pulse calcular para obtener la predicción

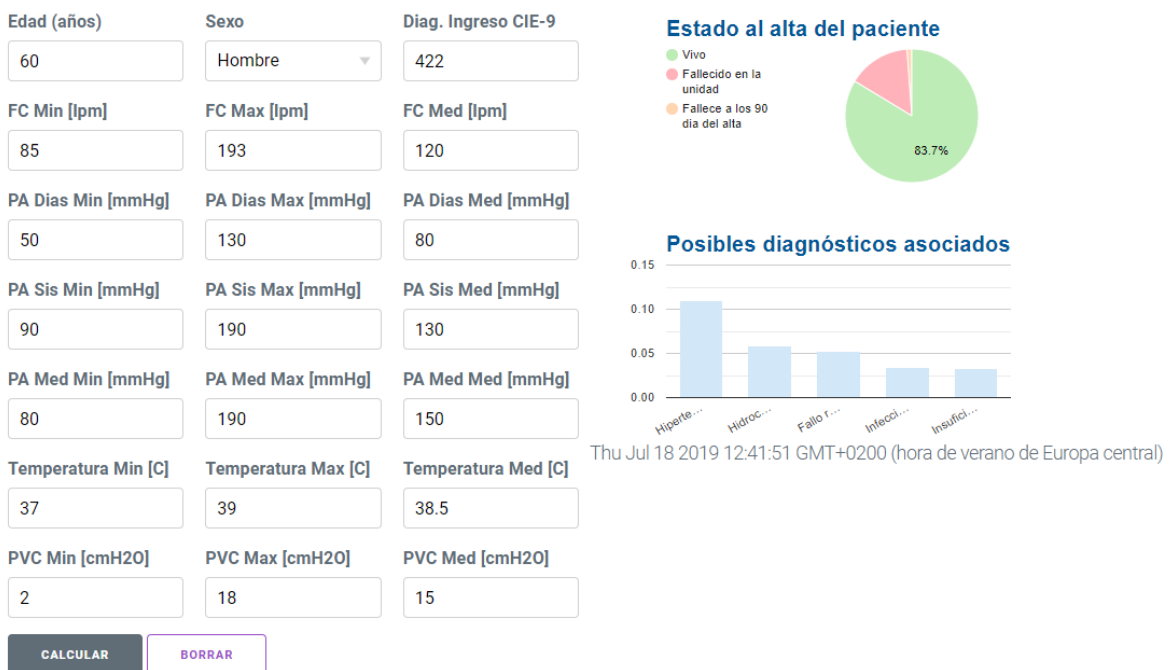


Figura 35: Interfaz

Se usa html y javascript para esta implementación. El código completo puede consultarse en el anexo IV.

11. Evaluación de la herramienta

En este capítulo se evaluarán tanto el cumplimiento de los requisitos definidos en el capítulo 4 como la experiencia de los usuarios de test.

Evaluación del cumplimiento de los requisitos

En la siguiente tabla se reflejará el cumplimiento o no de los requisitos definidos, referenciando en los casos que aplique, las figuras que reflejan el cumplimiento del mismo.

Requisitos funcionales			
Id	Descripción	Cumplimiento	Figura
F1	El usuario deberá introducir todos los parámetros para realizar la predicción.	Sí	36
F2	Los parámetros de entrada para la predicción serán: edad, sexo, motivo de ingreso, frecuencia cardiaca, presión arterial, temperatura y presión venosa central	Sí	37
F3	Se proporcionará un botón para borrar todos los parámetros introducidos	Sí	37
F4	Se especificará el rango de cada parámetro	No	N/A
F5	Se especificará la unidad de medida de cada parámetro	Sí	37
F6	Se mostrará la hora de la predicción	Sí	38
F7	La salida de una predicción será: estado del paciente al alta de la UCI y a los 90 días del alta de la UCI, probabilidad de diagnósticos asociados	Sí	38
F8	Se guardará todas las predicciones para su posterior análisis	Sí	N/A
F9	El sistema no almacenará los datos de los pacientes, ni tendrá acceso a ellos de ninguna forma, solo podrá usar el modelo generado anteriormente	Sí	N/A

Requisitos funcionales			
Id	Descripción	Cumplimiento	Figura
NF1	Los usuarios podrán realizar una predicción en menos de 5 minutos	Sí	N/A
NF2	Se le indicará al usuario si le falta un parámetro para realizar la predicción	Sí	36
NF3	Se le indicará al usuario si ha introducido un dato con un tipo incorrecto	Sí	N/A
NF4	Se podrá acceder al sistema de predicción desde los navegadores Chrome y Firefox	Sí	39
NF5	El cálculo de una predicción se realizará en un tiempo inferior a 30 segundos	Sí	N/A
NF6	El modelo podrá recalcularse con los nuevos episodios asistenciales mediante un proceso manual	Sí	N/A
NF7	Los datos para la generación del modelo estarán completamente anonimizados	Sí	N/A
NF8	El servidor deberá soportar al menos 10 usuarios simultáneos	Sí	N/A

Predicción del estado de salud asociado

f79b9cf1.ngrok.io dice
Edad (años) es obligatorio

Aceptar

Paula Fernández Perez

Por favor, introduzca los datos solicitados y pulse calcular para obtener la predicción

Edad (años) Sexo Diag. Ingreso CIE-9 Estado al alta de

Hombre 422

● Vivo
● Fallecido en la unidad
● Fallece a los 90 día del

Figura 36: Requisito F1

Edad (años)	Sexo	Diag. Ingreso CIE-9
	Hombre	
FC Min [lpm]	FC Max [lpm]	FC Med [lpm]
PA Dias Min [mmHg]	PA Dias Max [mmHg]	PA Dias Med [mmHg]
PA Sis Min [mmHg]	PA Sis Max [mmHg]	PA Sis Med [mmHg]
PA Med Min [mmHg]	PA Med Max [mmHg]	PA Med Med [mmHg]
Temperatura Min [C]	Temperatura Max [C]	Temperatura Med [C]
PVC Min [cmH2O]	PVC Max [cmH2O]	PVC Med [cmH2O]
CALCULAR BORRAR		

Figura 37: Requisitos F2, F3,F5

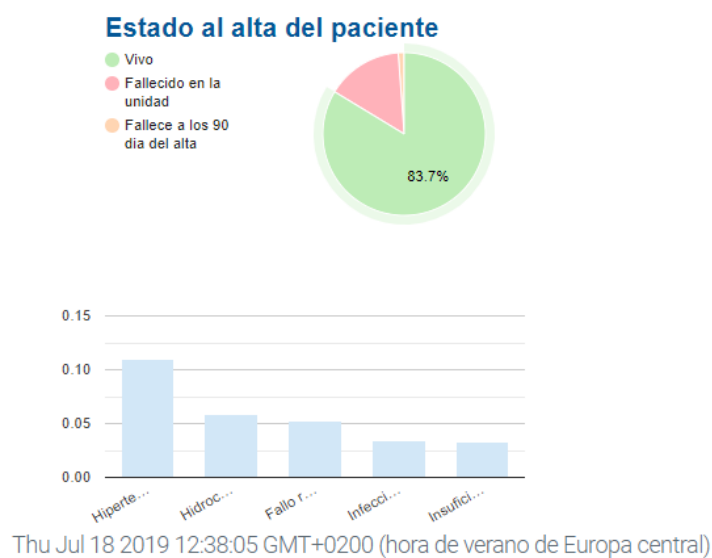


Figura 38: Requisitos F6, F7



Figura 39: requisito NF4

Evaluación de la usabilidad de la herramienta

La evaluación de usabilidad de la herramienta, se realiza siguiendo el estándar SUS (System Usability Scale) [29].

Esta escala fue creada por John Brooke en 1986 y permite evaluar aplicaciones o sitios web entre otros. Entre los beneficios que presenta este test se encuentran:

1. Facilidad de respuesta por parte del usuario.
2. Válido para muestras pequeñas de usuarios. Los resultados pueden considerarse relevantes a partir de 10 encuestados.
3. Permite diferenciar entre un sistema usable o no usable.

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es “Totalmente en desacuerdo” y 5 es “Totalmente de acuerdo”.

Pregunta	Respuesta
Creo que me gustaría utilizar este sistema frecuentemente	1-5
El sistema me resultó innecesariamente complejo	1-5
Creo que el sistema es bastante fácil de utilizar	1-5
Creo que necesitaría el soporte de un técnico para poder utilizar este sistema	1-5
Creo que las diferentes funciones del sistema se encuentran muy bien integradas	1-5
Opino que hubo demasiada inconsistencia en el sistema	1-5
Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente	1-5
Me sentí algo incómodo al utilizar este sistema	1-5
Me sentí muy seguro al utilizar este sistema	1-5
Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema	1-5

Para interpretar los resultados se procede del siguiente modo:

1. Pregunta impar: Restar un punto a la valoración del usuario
2. Pregunta par: Restar a 5, la valoración del usuario.
3. Multiplicar el resultado de cada pregunta por 2.5.
4. Sumar los resultados.

El resultado final de la escala será una puntuación entre 0 y 100. Valores por debajo de 50 se consideran sistemas no usables. Valores por encima de 50 se interpretan como sistemas usables y valores mayores de 80, se consideran sistemas sobresalientes desde el punto de vista de la usabilidad.

Análisis de los resultados del test de SUS

En base a las respuestas proporcionadas por 16 usuarios con un perfil clínico, y que pueden consultarse tanto individualmente como colectivo en el Anexo V, se concluye que todos los usuarios calificaron la herramienta como usable, siendo el valor mínimo de 70 y el máximo de 100, con una puntuación media de 89.21875.

Cabe destacar los siguientes resultados positivos:

1. El 62.5% de los usuarios manifestaron estar totalmente de acuerdo con que la herramienta es fácil de usar, y el 37.5 % estar muy de acuerdo, no otorgando ninguno una puntuación menor de 3.
2. Solo el 6.3% (2 usuarios), determinaron que necesitarían algo de soporte técnico para utilizar la herramienta, mientras que el 93,8% indicaron no necesitar soporte.
3. El 100% de los usuarios indicó no necesitar aprender muchas otras cosas antes de poder utilizar correctamente el sistema.

Sin embargo, es necesario identificar por qué algunos usuarios puntuaron con un 2 (12.5%) y un 3 (25%) la primera pregunta, que hace referencia al deseo de utilizar el sistema frecuentemente. Tras una conversación con los encuestados, se concluye que aquellos usuarios que no usarían el sistema frecuentemente no lo harían por una desconfianza en el software, prefieren usar los métodos actuales (escalas de predicción como el Apache II).

12. Conclusiones

Durante los últimos años, ha aumentado considerablemente la cantidad de datos en los entornos sanitarios, especialmente debido al auge de la implementación de sistemas de información clínica.

Estos datos son recopilados de forma automática y fiable de dispositivos de cabecera o de forma manual por el personal sanitario pero varios profesionales del sector médico y específicamente de los entornos de cuidado de paciente crítico, aseguran que la gestión del conocimiento debería estar dirigida a la capacidad de acción.

Es por esto que en este TFM se ha trabajado para desarrollar una herramienta que ayude al personal clínico en la toma de decisiones, incluyendo experiencias pasadas mediante el uso de tecnología MachineLearning. En la actualidad hay varios estudios y publicaciones orientadas al uso de MachineLearning en la medicina, pero no son muy populares en el día a día de las UCIs, donde siguen predominando escalas como el Apache II como sistemas de predicción de la mortalidad.

Para ello se han implementado dos modelos predictivos, el primero para proporcionar la probabilidad de que un paciente fallezca en el entorno intra-uci o intra-hospitalario y el segundo para orientar a los facultativos en la detección de posibles diagnósticos asociados.

Se concluye que la elección de las variables de predictoras del modelo es aceptable por los siguientes motivos:

1. En primer lugar, cumplen con el propósito de fiabilidad en el algoritmo, siendo el LogLoss para la predicción del estado del paciente al alta de 0.438 y de 4.0375 para el algoritmo. Aunque para el segundo caso no sea tan favorable, el algoritmo proporciona una orientación, que es el objetivo de la herramienta.
2. En segundo lugar, al realizar la evaluación de los usuarios, estos han referido sentirse cómodos con el uso de la herramienta, considerando que era fácil de usar y que no sería necesaria ninguna explicación adicional ni ayuda técnica para trabajar con ella. Este segundo punto es de vital importancia, ya que el sector médico necesita sentirse muy seguro con el uso de una herramienta para que esta tenga una larga vida de uso en el servicio.

Además, la herramienta desarrollada cumple 16 de los 17 requisitos funcionales, siendo el único no cumplido “Se especificará el rango de cada parámetro” dada la implicación visual que esto tendría en la interfaz final desarrollada.

13. Líneas futuras

Tras el estudio del contexto y la evaluación de las herramientas existentes y la desarrollada en este trabajo, se determina que hay varias líneas de trabajo futuras que ampliarían las posibilidades, efectividad y popularidad de la herramienta y son las siguientes:

1. Inclusión automática en base de datos de aprendizaje, todas las consultas realizadas por los usuarios. De esta manera, el modelo se podría ir actualizando con una cadencia de tiempo fijada y se podrían ir aumentando progresivamente el número de registros para mejorar los parámetros de efectividad.
2. Posibilitar a los usuarios incluir más variables predictoras para generar sus modelos. Esto permitiría mejorar los resultados en aquellos hospitales que dispongan de más medios, pero sin perjudicar la eficacia en aquellos entornos donde no se disponga de más información además de la monitorización básica.
3. Permitir descargar el modelo en local para poder hacer predicciones en cualquier entorno. Esta línea de trabajo abriría grandes posibilidades para el uso de la herramienta en entornos rurales de poca cobertura o incluso en países subdesarrollados.
4. Integración bidireccional de la herramienta con las bases de datos de los principales sistemas de información. De este modo no sería necesaria la introducción manual de los parámetros y podrían enviarse alertas al personal clínico cuando la probabilidad de supervivencia dentro de la unidad disminuya.

14. Referencias bibliográficas

- [1] R. Bravo, “La gestión del conocimiento en Medicina: a la búsqueda de la información perdida,” *An. Sist. Sanit. Navar.*, 2002.
- [2] O. Manor-Shulman, J. Beyene, H. Frndova, and C. S. Parshuram, “Quantifying the volume of documented clinical information in critical illness,” *J. Crit. Care*, 2008.
- [3] V. Gómez Tello *et al.*, “Estándares técnicos y funcionales, y proceso de implantación, de un sistema de información clínica en unidades de cuidados intensivos,” *Med Intensiva*, vol. 35, no. 8, pp. 484–496, 2011.
- [4] Ministerio de Sanidad y Política Social, “Unidad de cuidados intensivos Estándares y recomendaciones.” 2010.
- [5] W. A. . KNAUS, “APACHE II: a severity of disease classification system.,” *Crit. Care Med.*, vol. 13, no. 10, pp. 818–829, 1985.
- [6] Philips, “IntelliSpace Cuidados Críticos y Anestesia.” [Online]. Available: <https://www.philips.es/healthcare/product/HCNOCTN332/intellispace-para-cuidados-crticos-y-anestesia-icca>. [Accessed: 03-Jun-2019].
- [7] HealthIT and Health, “Hospital Analytics.” [Online]. Available: <http://healthitmhealth.com/wp-content/uploads/2015/03/Hospital-Analytics5.jpg>. [Accessed: 03-Jun-2019].
- [8] Microsoft, “Customer Stories.” [Online]. Available: <https://customers.microsoft.com/en-us/story/ge-healthcare-health-provider-power-bi>. [Accessed: 03-Jun-2019].
- [9] iMDsoft, “Products. Intensive Care.” [Online]. Available: <https://www.imdsoft.com/products/intensive-care>. [Accessed: 03-Jun-2019].
- [10] ehCOSSmartICU.Everis, “Health.” [Online]. Available: <https://www.ehcos.com/productos/ehcos-smarticu/>. [Accessed: 03-Jun-2019].
- [11] ehCOSSmartICU.Everis, “ehCOS SmartICU entra en su fase final con el pilotaje del sistema de alertas predictivas.” 2017.
- [12] F. Chen, Liddy M. MB, MSc; Martin, Claudio M. MSc, MD, FRCPC; Morrison, Teresa L. RN; Sibbald, William J. MD, FCCM, “Interobserver variability in data collection of the APACHE II score in teaching and community hospitals,” *Crit. Care Med.*, vol. 9, p. 15_523, 1999.

- [13] G. A. Miller, “The magical number seven, plus or minus two: some limits on our capacity for processing information,” *Psychol. Rev.*, 1956.
- [14] M. C. Martín *et al.*, “Indicadores de calidad en el enfermo crítico,” *Med. Intensiva.*, vol. 32, pp. 23–32, 2008.
- [15] A. E. W. Johnson *et al.*, “MIMIC-III, a freely accessible critical care database.,” *Sci. data*, 2016.
- [16] I. Neamatullah *et al.*, “Automated de-identification of free-text medical records,” *BMC Med. Inform. Decis. Mak.*, 2008.
- [17] “Docker.” [Online]. Available: <https://www.docker.com/>. [Accessed: 30-Jun-2019].
- [18] P. Di Tommaso, E. Palumbo, M. Chatzou, P. Prieto, M. L. Heuer, and C. Notredame, “The impact of Docker containers on the performance of genomic pipelines,” *PeerJ*, 2015.
- [19] “PostgreSQL.” [Online]. Available: <https://www.postgresql.org/about/>. [Accessed: 30-Jun-2019].
- [20] “DBeaver.” [Online]. Available: <https://dbeaver.io/about/>. [Accessed: 30-Jun-2019].
- [21] “Spark.” [Online]. Available: <https://spark.apache.org>. [Accessed: 30-Jun-2019].
- [22] “DataBricks.” [Online]. Available: <https://databricks.com/>. [Accessed: 30-Jun-2019].
- [23] “H2O.” [Online]. Available: <https://www.h2o.ai/>. [Accessed: 05-Jul-2019].
- [24] A. Candel, A., Parmar, V., LeDell, E., and Arora, *Deep Learning with H2O*, Sixth. Mountain View: H2O.ai, Inc., 2018.
- [25] Google, “Google Charts.” [Online]. Available: <https://developers.google.com/chart/>. [Accessed: 05-Jul-2019].
- [26] “Building the MIMIC database with Docker.” [Online]. Available: <https://github.com/MIT-LCP/mimic-code/tree/master/buildmimic/docker>. [Accessed: 30-Jun-2019].
- [27] C. Elvira, A. Ochoa, J. C. González, and F. Mochón, “Machine-Learning-Based No Show Prediction in Outpatient Visits,” *IJIMAI*, vol. 4, no. 7, pp. 29–34, 2018.
- [28] G. T. Ludwig Fahrmeir, “Models For Nominal Responses,” in *Multivariate Statistical Modelling Based on Generalized Linear Models*, Second., Springer, 2013, pp. 77–79.

- [29] B. W. Patrick W. Jordan, B. Thomas, Ian Lyall McClelland, "SUS: a 'quick and dirty' usability scale," in *Usability Evaluation in Industry*, 1996, pp. 189–194.

15. Bibliografía complementaria

- A. E. W. Johnson, D. J. Stone, L. A. Celi, and T. J. Pollard, "The MIMIC Code Repository: Enabling reproducibility in critical care research," *J. Am. Med. Informatics Assoc.*, 2018
- D. P. Ryan, B. J. Daley, K. Wong and X. Zhao, "Prediction of ICU in-hospital mortality using a deep Boltzmann machine and dropout neural net", *Biomedical Sciences and Engineering Conference* , 2013
- G. Meyfroidt, F. Güiza, J. Ramon, and M. Bruynooghe, "Machine learning techniques to examine large patient databases," *Best Practice and Research: Clinical Anaesthesiology*. 2009
- M. E. Charlson and F. L. Sax, "The therapeutic efficacy of critical care units from two perspectives: A traditional cohort approach vs a new case-control methodology," *J. Chronic Dis.*, 198
- M. Pal and P. M. Mather, "An assessment of the effectiveness of decision tree methods for land cover classification," *Remote Sens. Environ.*, 2003
- S. Purushotham, C. Meng, Z. Che, and Y. Liu, "Benchmark of Deep Learning Models on Large Healthcare MIMIC Datasets," *J. Biomed. Inform.*, 2017
- Z. Obermeyer and E. J. Emanuel, "Predicting the Future-Big Data, Machine Learning, and Clinical Medicine HHS Public Access," *N Engl J Med*, 2016

Anexo I. Extracción variables predictivas

Edad:

```
SELECT date_part('year'::text, admissions.admittime) - date_part('year'::text,
patients.dob) FROM patients
```

Motivo de ingreso

```
SELECT diagnoses_icd.icd9_code
      FROM diagnoses_icd
      WHERE diagnoses_icd.seq_num = 1 AND diagnoses_icd.icd9_code::text !~
'.*[a-zA-Z].*':text
```

Sexo

```
SELECT patients.gender FROM patients
```

Estado al alta del paciente

```
CASE
      WHEN ( SELECT patients.dod
            FROM patients
            WHERE patients.subject_id = admissions.subject_id) >=
admissions.admittime AND ( SELECT patients.dod
            FROM patients
            WHERE patients.subject_id = admissions.subject_id) <=
admissions.disctime THEN 1
      WHEN (( SELECT date_part('day'::text, (( SELECT patients_1.dod
            FROM patients patients_1
            WHERE patients_1.subject_id = admissions.subject_id)) -
admissions.disctime) AS date_part
            FROM patients
            WHERE patients.subject_id = admissions.subject_id)) <= 90::double
precision THEN 2
      ELSE 0
END
```

Anexo II. Script generación datos para el entrenamiento del modelo.

```
from pyspark.sql.functions import *

analysis_partials = []

for item in analysis:
    current_analysis_values = patients_table.join(charthevents,
patients_table.hadm_id == charthevents.hadm_id)\
    .filter(col("itemid").isin(item["values"]))

    current_analysis_values =
current_analysis_values.where(current_analysis_values.valuenum<item["max"]).where(
current_analysis_values.valuenum>item["min"])

    current_analysis = current_analysis_values\
    .groupBy(patients_table.hadm_id)\
    .agg(
        max("valuenum").alias(item["name"]+"_max"),
        min("valuenum").alias(item["name"]+"_min"),
        bround(avg("valuenum"),2).alias(item["name"]+"_avg")
    )

    analysis_partials.append(current_analysis)

patients_table.coalesce(1).write.format('com.databricks.spark.csv').save('predicMu
erte.csv', header = 'true')
```

Anexo III: Despliegue del modelo predictivo en el servidor

Creación de la imagen docker para la predicción de la probabilidad de muerte

```
FROM python:3.6-stretch

COPY license.sig /license.sig
COPY scorer_muerte.zip /

RUN apt-get update && apt-get -y install libopenblas-dev unzip
RUN unzip scorer_muerte.zip -d /app

WORKDIR /app/scoring-pipeline

RUN pip install -r requirements.txt

ENV DRIVERLESS_AI_LICENSE_FILE "/license.sig"
CMD ["python", "http_server.py", "--port=9090"]
EXPOSE 9090
```

Creación de la imagen docker para la predicción de diagnósticos

```
FROM python:3.6-stretch

COPY license.sig /license.sig
COPY scorer_diagnosticos.zip /

RUN apt-get update && apt-get -y install libopenblas-dev unzip
RUN unzip scorer_diagnosticos.zip -d /app

WORKDIR /app/scoring-pipeline

RUN pip install -r requirements.txt

ENV DRIVERLESS_AI_LICENSE_FILE "/license.sig"
CMD ["python", "http_server.py", "--port=9090"]
EXPOSE 9090
```

Creación de la imagen docker para la predicción de diagnósticos

```
version: "3.7"
services:
  scorer_diagnosticos:
    build:
      context: .
      dockerfile: Dockerfile_diagnosticos
    ports:
      - "8000:9090"

  scorer_muerte:
    build:
      context: .
      dockerfile: Dockerfile_muerte
    ports:
      - "8001:9090"
```

Anexo IV: Desarrollo Interfaz

```
<!DOCTYPE html>

<html>

<head>
  <script type="text/javascript"
src="https://www.gstatic.com/charts/loader.js"></script>
  <script src="https://cdnjs.cloudflare.com/ajax/libs/jquery/3.4.1/jquery.js"
    integrity="sha256-WpOohJOqMqqyKL9FccASB900KwACQJpFTUBLTYOVvVU="
crossorigin="anonymous"></script>

  <!--link a la fuente a usar -->
  <link href="https://fonts.googleapis.com/css?family=Roboto" rel="stylesheet">

  <link rel="stylesheet"
href="//fonts.googleapis.com/css?family=Roboto:300,300italic,700,700italic">
  <link rel="stylesheet"
href="//cdn.rawgit.com/necolas/normalize.css/master/normalize.css">
  <link rel="stylesheet"
href="//cdn.rawgit.com/milligram/milligram/master/dist/milligram.min.css">
</head>

<body>
  <div class="container">

    <div class="row">
      <div class="column">
        <h1 id=titulo>Predicción del estado al alta del paciente y
diagnósticos asociados para el enfermo
crítico </h1>
        <h2 id=sub> Paula Fernández Perez </h2>
        <h3 id=instrucciones> Por favor, introduzca los datos solicitados
y pulse calcular para obtener la
predicción </h2>
      </div>
    </div>

    <div class="row">
```

```

        <div class="column column-50">
            <div class="row">
                <div class="column column-33">
                    <label for="edad">Edad (años)</label>
                    <input required type="number" min="0" placeholder=""
name="edad" id="edad">
                </div>
                <div class="column column-33">
                    <label for="sexo">Sexo</label>
                    <select required id="sexo">
                        <option value=""></option>
                        <option value="M">Hombre</option>
                        <option value="F">Mujer</option>
                    </select>
                </div>
                <div class="column column-33">
                    <label for="diagnostico">Diag. Ingreso CIE-9</label>
                    <input required type="text" placeholder=""
name="diagnostico" id="diagnostico">
                </div>
            </div>
            <div class="row">
                <div class="column column-33">
                    <label for="fc_min">FC Min [lpm]</label>
                    <input type="number" min="0" placeholder="" name="fc_min"
id="fc_min">
                </div>
                <div class="column column-33">
                    <label for="fc_max">FC Max [lpm]</label>
                    <input type="number" min="0" placeholder="" name="fc_max"
id="fc_max">
                </div>
                <div class="column column-33">
                    <label for="fc_avg">FC Med [lpm]</label>
                    <input type="number" min="0" placeholder="" name="fc_avg"
id="fc_avg">
                </div>
            </div>
            <div class="row">
                <div class="column column-33">

```

```

        <label for="pad_min">PA Dias Min [mmHg]</label>
        <input type="number" min="0" placeholder="" name="pad_min"
id="pad_min">
    </div>
    <div class="column column-33">
        <label for="pad_max">PA Dias Max [mmHg]</label>
        <input type="number" min="0" placeholder="" name="pad_max"
id="pad_max">
    </div>
    <div class="column column-33">
        <label for="pad_avg">PA Dias Med [mmHg]</label>
        <input type="number" min="0" placeholder="" name="pad_avg"
id="pad_avg">
    </div>
</div>
<div class="row">
    <div class="column column-33">
        <label for="pas_min">PA Sis Min [mmHg]</label>
        <input type="number" min="0" placeholder="" name="pas_min"
id="pas_min">
    </div>
    <div class="column column-33">
        <label for="pas_max">PA Sis Max [mmHg]</label>
        <input type="number" min="0" placeholder="" name="pas_max"
id="pas_max">
    </div>
    <div class="column column-33">
        <label for="pas_avg">PA Sis Med [mmHg]</label>
        <input type="number" min="0" placeholder="" name="pas_avg"
id="pas_avg">
    </div>
</div>
<div class="row">
    <div class="column column-33">
        <label for="pam_min">PA Med Min [mmHg]</label>
        <input type="number" min="0" placeholder="" name="pam_min"
id="pam_min">
    </div>
    <div class="column column-33">
        <label for="pam_max">PA Med Max [mmHg]</label>

```

```

        <input type="number" min="0" placeholder="" name="pam_max"
id="pam_max">
    </div>
    <div class="column column-33">
        <label for="pam_avg">PA Med Med [mmHg]</label>
        <input type="number" min="0" placeholder="" name="pam_avg"
id="pam_avg">
    </div>
</div>
<div class="row">
    <div class="column column-33">
        <label for="tmp_min">Temperatura Min [C]</label>
        <input type="number" min="0" placeholder="" name="tmp_min"
id="tmp_min">
    </div>
    <div class="column column-33">
        <label for="tmp_max">Temperatura Max [C]</label>
        <input type="number" min="0" placeholder="" name="tmp_max"
id="tmp_max">
    </div>
    <div class="column column-33">
        <label for="tmp_avg">Temperatura Med [C]</label>
        <input type="number" min="0" placeholder="" name="tmp_avg"
id="tmp_avg">
    </div>
</div>
<div class="row">
    <div class="column column-33">
        <label for="pvc_min">PVC Min [cmH2O]</label>
        <input type="number" min="0" placeholder="" name="pvc_min"
id="pvc_min">
    </div>
    <div class="column column-33">
        <label for="pvc_max">PVC Max [cmH2O]</label>
        <input type="number" min="0" placeholder="" name="pvc_max"
id="pvc_max">
    </div>
    <div class="column column-33">
        <label for="pvc_avg">PVC Med [cmH2O]</label>

```

```

        <input type="number" min="0" placeholder="" name="pvc_avg"
id="pvc_avg">
    </div>
</div>
<div class="row">
    <div class="column">
        <input class="button-primary" id="submit" type="submit"
value="Calcular">
        <input class="button-outline" id="reset" type="reset"
value="Borrar">
    </div>
</div>
</div>
<div class="column">
    <div class="row">
        <div id='grafica1' class="grafica"> </div>
    </div>
    <div class="row">
        <div id='grafica2' class="grafica"> </div>
    </div>
</div>
</div>
</div>

<style>
    /* selector -->*/
    #titulo {
        color: #045797;
        text-align: center;
        font-family: 'Roboto', sans-serif
    }

    #sub {
        color: #61a5d9;
        text-align: center;
        font-family: 'Roboto', sans-serif
    }

    .grafica {

```

```
        float: left;
        width: 100%;
    }
</style>
<!-- sirve para hacer estructuras dentro de una web-->

<script>
    google.charts.load('current', { 'packages': ['corechart'] });

    function drawResponseMuerte(datos) {

        //creamos la tabla de datos
        var data = new google.visualization.DataTable();
        data.addColumn('string', 'Estado al alta');
        data.addColumn('number', 'Probabilidad');
        data.addRows([
            ["Vivo", datos[0]],
            ["Fallecido en la unidad", datos[1]],
            ["Fallece a los 90 dia del alta", datos[2]],
        ]);

        //creamos los gráficos
        var pie = new
google.visualization.PieChart(document.getElementById('grafica1'))

        //array de opciones para la gráfica de tarta.
        var options2 = {
            legend: { position: 'left' },
            title: 'Estado al alta del paciente',
            titleTextStyle: {
                color: '#045797',
                bold: true
            },
            titleFontSize: 20,
            colors: ['#bdec6', '#ffb2ba', '#ffd5b2'],
            //el nombre del grupo etiquetado en el gráfico.
            pieSliceText: 'value * 100',

            pieSliceTextStyle: {
                color: 'black',
```

```
    },  
  }  
  //pintamos las gráficas  
  pie.draw(data, options2)  
}  
  
function drawResponseDiagnostico(values) {  
  var labels = ['Infección Intestinal por Clostridium',  
    'Septicemia por Staphylococcus aureus susceptible a meticilina',  
    'Septicemia no especificada',  
    'Infección por Pseudomonas',  
    'Hepatitis C Cronica sin Coma hepático',  
    'Hipotiroidismo',  
    'Diabetes Mellitus tipo II o no especificada',  
    'Otros trastornos de neurohipófisis',  
    'Malnutrición inespecífica',  
    'Hipercolesterolemia pura',  
    'Hiperlipidemia no especificada',  
    'Hiperosmolalidad y / o hipernatremia',  
    'Hiposmolalidad y / o hiponatremia',  
    'Acidosis',  
    'Alcalosis',  
    'Trastorno del equilibrio ácido-base',  
    'Hipopotasemia',  
    'Obesidad',  
    'Anemia posthemorragia aguda',  
    'Anemia',  
    'Trombocitopenia',  
    'Abuso de alcohol',  
    'Hábito Tabaco',  
    'Trastorno depresivo',  
    'Meningitis por bacterias gramnegativas',  
    'Meningitis',  
    'Hidrocefalia comunicante',  
    'Hidrocefalia obstructiva',  
    'Hemiplegia',  
    'Migraña',  
    'Encefalopatía',  
    'Compresión cerebral',
```

'Edema cerebral',
'Hipertensión no especificada',
'Enfermedad renal crónica hipertensiva, no especificada, con enfermedad renal crónica en etapa V o enfermedad renal en etapa terminal',
'Infarto subendocárdico',
'Infarto de miocardio',
'Ateroesclerosis coronaria de la arteria coronaria nativa',
'Embolismo pulmonar e infarto',
'Trastorno de la válvula mitral',
'Fibrilación auricular',
'Aleteo auricular',
'Paro cardíaco',
'Otras disritmias cardíacas',
'Insuficiencia cardíaca congestiva',
'Hemorragia subhialoidea',
'Hemorragia intercerebral',
'Oclusión y estenosis de la arteria carótida sin mención de infarto cerebral',
'Oclusión de la arteria cerebral, no especificada con infarto cerebral',
'Otras isquemias cerebrales transitorias',
'Isquemia cerebral transitoria no especificada',
'Aneurisma cerebral, no fracturado',
'Embolismo venoso agudo y trombosis de los vasos profundos de la extremidad inferior proximal',
'Hipotensión, no especificada',
'Neumonía por pseudomona',
'Neumonía por haemophilus',
'Neumonía susceptible a meticilina por estafilococo áureo',
'Neumonía por otras bacterias gramnegativas',
'Neumonía, organismo no especificado',
'Asma, no especificada',
'Obstrucción crónica de las vías aéreas',
'Neumonitis por inhalación de comida o vómito',
'Derrame pleural, no especificado',
'Colapso pulmonar',
'Fallo respiratorio agudo',
'Otra insuficiencia pulmonar',
'Insuficiencia respiratoria aguda y crónica',
'Reflujo esofágico',

```

        'Insuficiencia renal aguda con lesión de necrosis tubular',
        'Insuficiencia renal aguda',
        'Infección tracto urinario',
        'Dermatitis por medicamento por vía interna',
        'Anomalías del sistema cerebrovascular',
        'Otras convulsiones',
        'Apnea del sueño, no especificada',
        'Afasia',
        'Shock séptico',
        'Disfagia, no especificada',
        'Otros niveles de enzimas séricas anormales, no específicas',
        'Bacteriemia',
        'Sepsis severa',
        'Complicación mecánica del dispositivo del sistema nervioso,
implante e injerto',
        'Infección y reacción inflamatoria debida a otro dispositivo
vascular, implante e injerto',
        'Infección y reacción inflamatoria debida al dispositivo del
sistema nervioso, implante e injerto',
        'Complicación cardíaca',
        'Neumonía asociada al respirador',
        'Hematoma que complica un procedimiento',
        'Otras no especificadas complicaciones médicas',
        'Accidente en institución residencial',
        'Otros procedimientos especificados como la causa de una reacción
anormal del paciente, o de una complicación posterior, sin mencionar la desventura
en el momento del procedimiento',
        'Infección con microorganismos resistentes a la penicilina',
        'Bypass aortocoronario',
        'Anticoagulantes de largo plazo',
        'Cuidados paliativos']

var result = [];
labels.forEach((label, i) => result.push([label, values[i]]));
result = result.sort((a, b) => b[1] - a[1]).slice(0,5)

//creamos la tabla de datos
var data = new google.visualization.DataTable();
data.addColumn('string', 'Estado al alta');
data.addColumn('number', 'Probabilidad');

```

```

        data.addRow(result);
        //creamos los gráficos
        var col = new
google.visualization.ColumnChart(document.getElementById('grafica2'))

        //array de opciones para la gráficas de barras, columnas y líneas.
        var options = {
            legend: { position: 'none' },
            titleTextStyle: {
                color: '#045797',
                bold: true
            },
            colors: ['#D2E7F7', '#98CDF6', '#61A5D9', '#338DD4', '#0C76C9',
'#045797']

        }

        //pintamos las gráficas
        col.draw(data, options)

    }

    $("#submit").click(function () {
        var allOk = true;

        $('input[type="number"], input[type="text"], select').each(function ()
{
            if ($(this).val()==""){
                alert($(this).parent().children()[0].textContent + " es
obligatorio");
                allOk = false;
                return false;
            }
        });

        if(allOk){
            var parameters = {
                "id": 1,
                "method": "score",

```

```

    "params": {
      "row": {
        "edadingreso": $('#edad').val(),
        "diagin": $('#diagnostico').val(),
        "sexo": $('#sexo').val(),
        "frec_cardiaca_max": $('#fc_max').val(),
        "frec_cardiaca_min": $('#fc_min').val(),
        "frec_cardiaca_avg": $('#fc_avg').val(),
        "pa_sistolica_max": $('#pas_max').val(),
        "pa_sistolica_min": $('#pas_min').val(),
        "pa_sistolica_avg": $('#pas_avg').val(),
        "pa_diastolica_max": $('#pad_max').val(),
        "pa_diastolica_min": $('#pad_min').val(),
        "pa_diastolica_avg": $('#pad_avg').val(),
        "pa_media_max": $('#pam_max').val(),
        "pa_media_min": $('#pam_min').val(),
        "pa_media_avg": $('#pam_avg').val(),
        "temperature_max": $('#tmp_max').val(),
        "temperature_min": $('#tmp_min').val(),
        "temperature_avg": $('#tmp_avg').val(),
        "pv_central_max": $('#pvc_max').val(),
        "pv_central_min": $('#pvc_min').val(),
        "pv_central_avg": $('#pvc_avg').val()
      }
    }
  }

  $("#grafica1").html("")
  $("#grafica2").html("")

  $.ajax({
    type: 'POST',
    url: '/scorer_diagnosticos/rpc',
    data: JSON.stringify (parameters),
    success: function (data) {
      drawResponseDiagnostico(data.result)
    },
    contentType: "application/json",
    dataType: 'json'
  })

```

```
});

$.ajax({
    type: 'POST',
    url: '/scorer_muerte/rpc',
    data: JSON.stringify(parameters),
    success: function (data) {
        $("#grafica1").html("")
        drawResponseMuerte(data.result)
    },
    contentType: "application/json",
    dataType: 'json'
});
}
});

$("#reset").click(function () {
    $('input[type="number"], input[type="text"], select').each(function ()
{
        var elt = $(this).val("");
    });
});
</script>
</body>

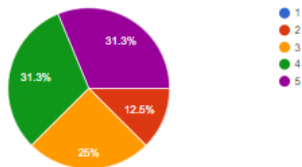
</html>
```

Anexo V: Respuestas de las encuestas de usabilidad

Resumen

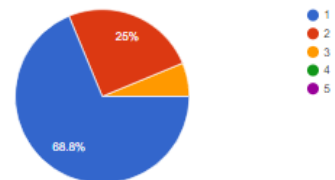
Creo que me gustaría utilizar este sistema frecuentemente

16 responses



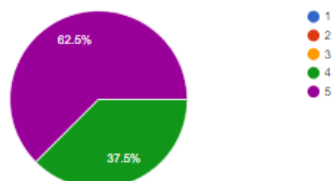
El sistema me resultó innecesariamente complejo

16 responses



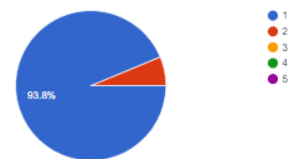
Creo que el sistema es bastante fácil de utilizar

16 responses



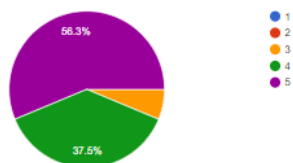
Creo que necesitaría el soporte de un técnico para poder utilizar este sistema

16 responses



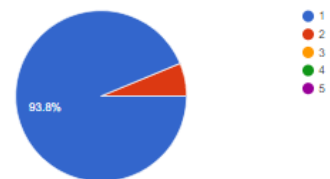
Creo que las diferentes funciones del sistema se encuentran muy bien integradas

16 responses



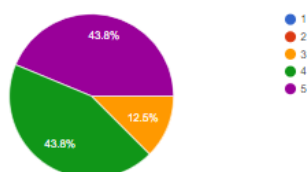
Opino que hubo demasiada inconsistencia en el sistema

16 responses



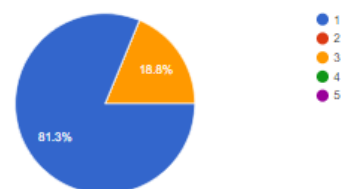
Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente

16 responses



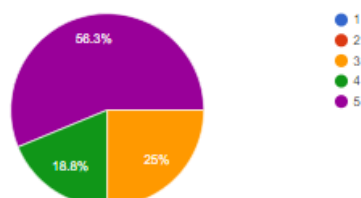
Me sentí algo incómodo al utilizar este sistema

16 responses



Me sentí muy seguro al utilizar este sistema

16 responses



Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema

16 responses



Usuario 1

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

4 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

4 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

4 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

5 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 2

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

5 ▼

El sistema me resultó innecesariamente complejo *

3 ▼

Creo que el sistema es bastante fácil de utilizar *

4 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

5 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

3 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

4 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 3

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

4 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

5 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

4 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

5 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 4

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

4 ▼

El sistema me resultó innecesariamente complejo *

2 ▼

Creo que el sistema es bastante fácil de utilizar *

4 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

4 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

4 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

5 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 5

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

4 ▼

El sistema me resultó innecesariamente complejo *

2 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

5 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

5 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

4 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 6

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

3 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

5 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

5 ▼

Me sentí algo incómodo al utilizar este sistema *

3 ▼

Me sentí muy seguro al utilizar este sistema *

3 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 7

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

5 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

5 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

5 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

5 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 8

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

2 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

4 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

3 ▼

Opino que hubo demasiada inconsistencia en el sistema *

2 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

4 ▼

Me sentí algo incómodo al utilizar este sistema *

3 ▼

Me sentí muy seguro al utilizar este sistema *

3 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 9

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

5 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

4 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

4 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

5 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

5 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 10

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

3 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

5 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

5 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

4 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 11

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

5 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

5 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

5 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

5 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 12

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

4 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

5 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

4 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

5 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 13

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

3 ▼

El sistema me resultó innecesariamente complejo *

2 ▼

Creo que el sistema es bastante fácil de utilizar *

4 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

2 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

4 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

4 ▼

Me sentí algo incómodo al utilizar este sistema *

3 ▼

Me sentí muy seguro al utilizar este sistema *

3 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 14

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

5 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

4 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

4 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

5 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 15

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

3 ▼

El sistema me resultó innecesariamente complejo *

1 ▼

Creo que el sistema es bastante fácil de utilizar *

5 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

4 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

5 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

5 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼

Usuario 16

Por favor, conteste a las siguientes preguntas con valores del 1 al 5, donde 1 es "Totalmente en desacuerdo" y 5 es "Totalmente de acuerdo".

* Required

Creo que me gustaría utilizar este sistema frecuentemente *

2 ▼

El sistema me resultó innecesariamente complejo *

2 ▼

Creo que el sistema es bastante fácil de utilizar *

4 ▼

Creo que necesitaría el soporte de un técnico para poder utilizar este sistema *

1 ▼

Creo que las diferentes funciones del sistema se encuentran muy bien integradas *

5 ▼

Opino que hubo demasiada inconsistencia en el sistema *

1 ▼

Imagino que la mayoría de las personas aprendería a utilizar el sistema rápidamente *

3 ▼

Me sentí algo incómodo al utilizar este sistema *

1 ▼

Me sentí muy seguro al utilizar este sistema *

3 ▼

Necesito aprender muchas otras cosas antes de poder utilizar correctamente el sistema *

1 ▼