

Sistemas de Información y Telemedicina II\*

# **Machine Learning para Mesotelioma Maligno**

Irene Estela García García

Ilán Francisco Carretero Juchnowicz

Ignacio Maria Amat Hernández

April 21, 2020

---

\*Grado en Ingeniería Biomédica, Escuela Técnica Superior de Ingenieros Industriales, Valencia, España.

# Contents

<b>1</b>	<b>Contexto y motivación</b>	<b>4</b>
<b>2</b>	<b>Objetivos</b>	<b>4</b>
<b>3</b>	<b>Materiales y métodos</b>	<b>5</b>
<b>4</b>	<b>Resultados y discusión</b>	<b>5</b>
4.1	Resolución ideal . . . . .	5
4.2	Resolución real . . . . .	6
4.2.1	Análisis exploratorio . . . . .	8
4.2.1.1	Tipos de datos . . . . .	8
4.2.1.2	Caja bigotes . . . . .	9
4.2.1.3	Histograma . . . . .	10
4.2.1.4	Kernel Density . . . . .	11
4.2.1.5	Cuantil - cuantil . . . . .	12
4.2.1.6	Correlaciones . . . . .	13
4.2.2	Extracción de Características . . . . .	14
4.2.2.1	Filter methods . . . . .	14
4.2.2.2	Wrapper methods . . . . .	15
4.2.2.3	PCA . . . . .	15
4.2.3	Modelos de Clasificación . . . . .	16
4.2.4	Interfaz . . . . .	16
4.2.4.1	Aplicación para médicos . . . . .	20
4.2.4.2	Aplicación para pacientes . . . . .	21
4.3	Servidores y conexiones y seguridad . . . . .	23
4.4	Proyecto ALERT . . . . .	24
<b>5</b>	<b>Conclusiones</b>	<b>24</b>
<b>6</b>	<b>Lecciones aprendidas y reflexiones</b>	<b>24</b>

## List of Figures

1	Esquema teórico de la solución ideal. . . . .	7
2	Caja bigotes. . . . .	9
3	Histograma. . . . .	10
4	Kernel Density. . . . .	11
5	Cuantil - cuantil. . . . .	12
6	Correlaciones. . . . .	13

7	Diagrama de pareto. . . . .	15
8	Ventana principal. . . . .	17
9	Selección de paciente. . . . .	17
10	Importación paciente. . . . .	18
11	Cargar variables. . . . .	18
12	Selección de modelos. . . . .	19
13	Selección de modelos. . . . .	19
14	Versión limitada para médicos. . . . .	20
15	PLACEHOLDER. . . . .	22

## Listings

1	Variables de la base de datos. . . . .	8
2	Selección de variables según <b>fscore</b> . . . . .	14
3	Mejores variables según <b>SFS</b> y <b>SBS</b> . . . . .	15

**1 Contexto y motivación**

**2 Objetivos**

### 3 Materiales y métodos

## 4 Resultados y discusión

### 4.1 Resolución ideal

Con el objetivo de ofrecer un servicio que sea de utilidad en múltiples vertientes en lo concerniente al tratamiento del mesotelioma maligno, se debe desarrollar dos productos que a su vez se hallen interconectados, pudiendo ofrecer grandes mejoras tanto en la eficiencia del hospital como en la calidad asistencial que el paciente recibe del centro médico.

En primer lugar, se implementaría un sistema de ayuda a la decisión médica en el cual, a partir de datos a los que previamente se les haya realizado un análisis exploratorio, se entrenarían varios modelos de predicción aplicando el aprendizaje supervisado como podrían ser modelos lineales, modelos cuadráticos, k-vecinos o random forest. Asimismo, una vez entrenados y analizadas sus métricas de acierto y error, se procedería a seleccionar el modelo que mejor resultados haya obtenido, siempre y cuando supere unos mínimos umbrales. Finalmente, se dotaría al sistema de ayuda a la decisión médica en cuestión de una interfaz amigable e intuitiva para el personal sanitario, de forma que pudiesen a partir de la carga de datos de un paciente, ejecutar el modelo y obtener en un tiempo reducido si el modelo considera en base a dichos datos si el paciente es enfermo o sano, ayudando de forma directa al diagnóstico que realizará el médico.

En segundo lugar, se realizaría una aplicación móvil, a partir de la cual mediante un sistema de usuarios y contraseñas pudiesen acceder profesionales sanitarios, enfermos e individuos con riesgo de contraer este cáncer. De esta forma, cada grupo de personas podría acceder a unos servicios u otros según su condición.

Así, aquellos pacientes que padezcan el cáncer dispondrían de un chat con otros enfermos con el fin de recibir apoyo y consejos unos de otros. También, tendrían la posibilidad de conversar con sus médicos, y a su vez estarían analizando su estado diariamente, ya que la aplicación enviaría notificaciones en las cuales se pediría a los pacientes que introdujesen datos que fuesen relevantes para el estudio de la situación del enfermo, como podría ser el dolor de pecho o la debilidad. Cabe destacar que, también estarían geolocalizados y se identificaría en un mapa aquellos puntos donde la exposición al asbesto es mayor, y al pasar cerca de éstos, la app emitiría una notificación alertando a tales enfermos, además de también alertar sobre las citas médicas y la ingesta de medicamentos.

Los pacientes en riesgo dispondrían de un chat para consultarle dudas a sus respectivos médicos, además de tener la misma monitorización que los pacientes enfermos, variando las notificaciones recibidas (un mayor número de recomendaciones sobre consejos y hábitos a realizar)

Finalmente, los médicos tendrían disponible un chat para poder resolver las cuestiones que tengan los pacientes de riesgo y los enfermos, e incluso poder preguntar a los pacientes que considere en que estado se encuentran, pudiendo realizar un primer estudio general vía chat. Cabe recalcar, que los tres grupos de usuarios, médicos, pacientes y personas en riesgo dispondrían de información sobre el amianto, donde se halla, que puede ocasionar, prevención, concienciación y buenas prácticas.

En síntesis, si bien es verdad que el servicio se compondría de dos productos, el objetivo principal es coordinación entre ambos, de forma que, los datos que se vayan recogiendo de los pacientes en la aplicación, se envíen y almacenen de forma segura en servidores a los que se pueda acceder desde el sistema de ayuda a la decisión clínica, de manera que a la hora de realizar distintas exploraciones y obtener nuevos análisis de los pacientes, los datos que haya podido introducir directamente desde la aplicación el paciente queden ya registrados, acelerando los flujos de trabajo del hospital. Por tanto, se promociona la condición de paciente activo, donde se hace partícipe al propio paciente del cuidado de su salud, ofreciéndole los medios necesarios para mejorar su calidad de vida mediante el tratamiento y prevención del mesotelioma maligno a partir del uso de las tecnologías de la información y la comunicación.

## 4.2 Resolución real

Si bien es verdad que las implementaciones realizadas servirían a modo de una primera aproximación de lo que podría ser un proyecto de esta envergadura, podrían considerarse de gran utilidad a la hora de considerarse como prototipados, de forma que la solución real acabara aproximándose al máximo a la solución ideal.

En lo concerniente al sistema de ayuda a la decisión clínica, en primer lugar, se ha realizado un análisis exploratorio de los datos para identificar datos anómalos y aquellas variables que podrían resultar más significativas. Así, pese a tener variables de relevancia con el mesotelioma maligno como podría ser la exposición al asbesto o la duración de ésta, analizando los datos disponibles en la base de datos de enfermos y sanos, observamos como los valores de unos y otros son muy similares en la mayoría de campos, por lo que se ha descartado la idea de reducir la dimensionalidad inicialmente. Muestra de ello puede ser los histogramas de las variables, donde se aprecia que los valores de para sanos y enfermos son muy similares en la mayoría de casos, o el gráfico de Pareto donde podemos visualizar como serían necesarias 17 componentes para explicar el 80% de la varianza de los datos, por lo que se concluye las componentes explican muy poca varianza individualmente.

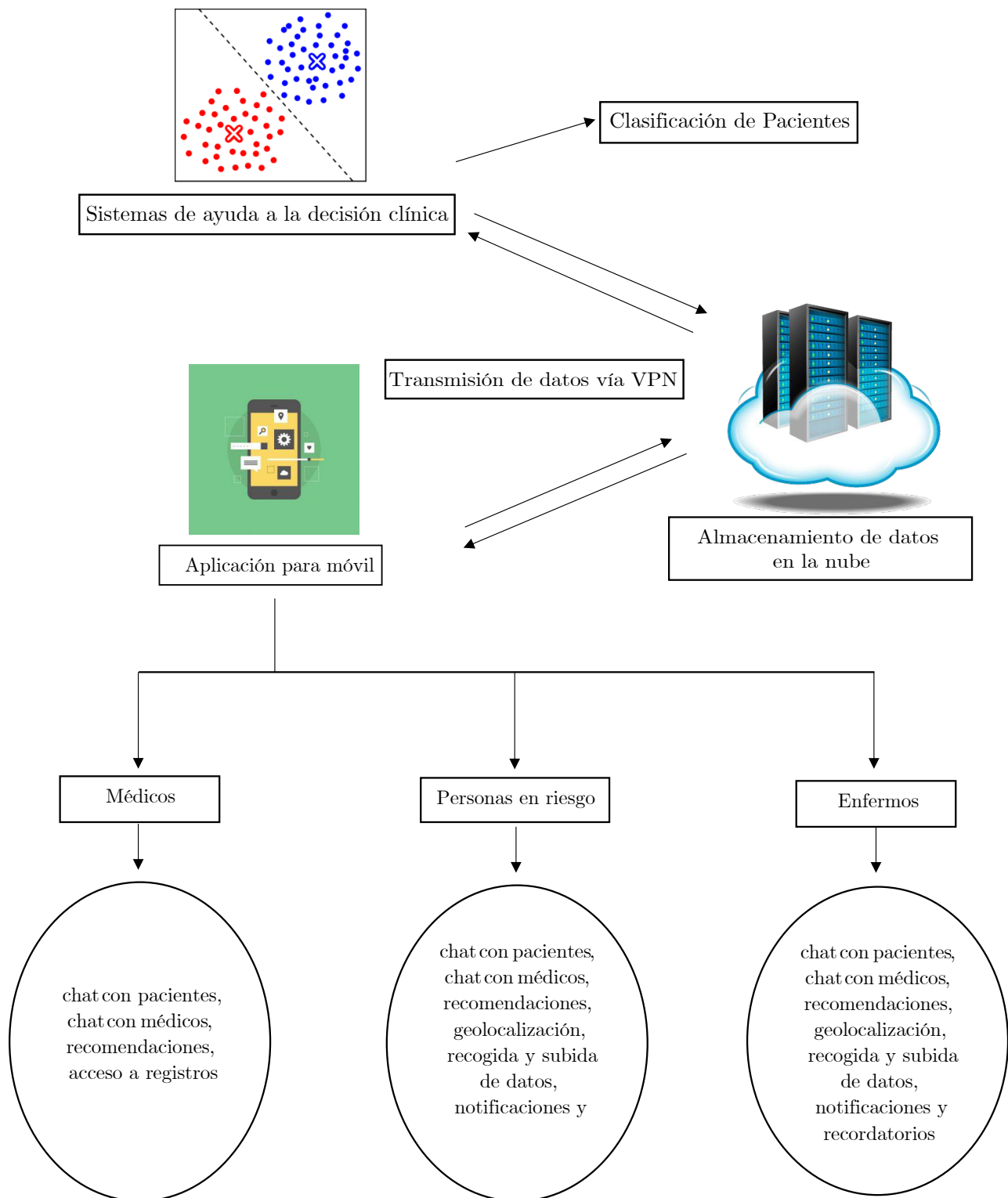


Fig. 1: Esquema teórico de la solución ideal.

## 4.2.1 Análisis exploratorio

### 4.2.1.1 Tipos de datos

Comenzamos analizando la base de datos, consta de las siguientes variables:

```
1 RangeIndex: 324 entries, 0 to 323
2 Data columns (total 30 columns):
3 #    Column                                Non-Null Count  Dtype
4 ---  -
5 0    age                                    324 non-null    float64
6 1    gender                                324 non-null    int64
7 2    city                                   324 non-null    int64
8 3    asbestos exposure                     324 non-null    int64
9 4    duration of asbestos exposure         324 non-null    float64
10 5    keep side                             324 non-null    int64
11 6    duration of symptoms                  324 non-null    float64
12 7    dyspnoea                              324 non-null    int64
13 8    ache on chest                         324 non-null    int64
14 9    weakness                              324 non-null    int64
15 10   habit of cigarette                   324 non-null    int64
16 11   performance status                  324 non-null    int64
17 12   white blood                          324 non-null    float64
18 13   cell count (WBC)                     324 non-null    int64
19 14   hemoglobin (HGB)                     324 non-null    int64
20 15   platelet count (PLT)                  324 non-null    float64
21 16   sedimentation                        324 non-null    float64
22 17   blood lactic dehydrogenise (LDH)     324 non-null    float64
23 18   alkaline phosphatise (ALP)           324 non-null    float64
24 19   total protein                         324 non-null    float64
25 20   albumin                              324 non-null    float64
26 21   glucose                              324 non-null    float64
27 22   pleural lactic dehydrogenise          324 non-null    float64
28 23   pleural protein                       324 non-null    float64
29 24   pleural albumin                       324 non-null    float64
30 25   pleural glucose                       324 non-null    float64
31 26   pleural effusion                      324 non-null    float64
32 27   pleural thickness on tomography       324 non-null    float64
33 28   pleural level of acidity (pH)         324 non-null    float64
34 29   C-reactive protein (CRP)             324 non-null    int64
35 dtypes: float64(18), int64(12)
36 memory usage: 76.1 KB
```

Listing 1: Variables de la base de datos.

Nuestra base de datos consta de 324 entradas, cada una con 30 variables. Todos los valores son `floats` e `ints`, además no tenemos ningún `NULL`.



#### 4.2.1.2 Caja bigotes

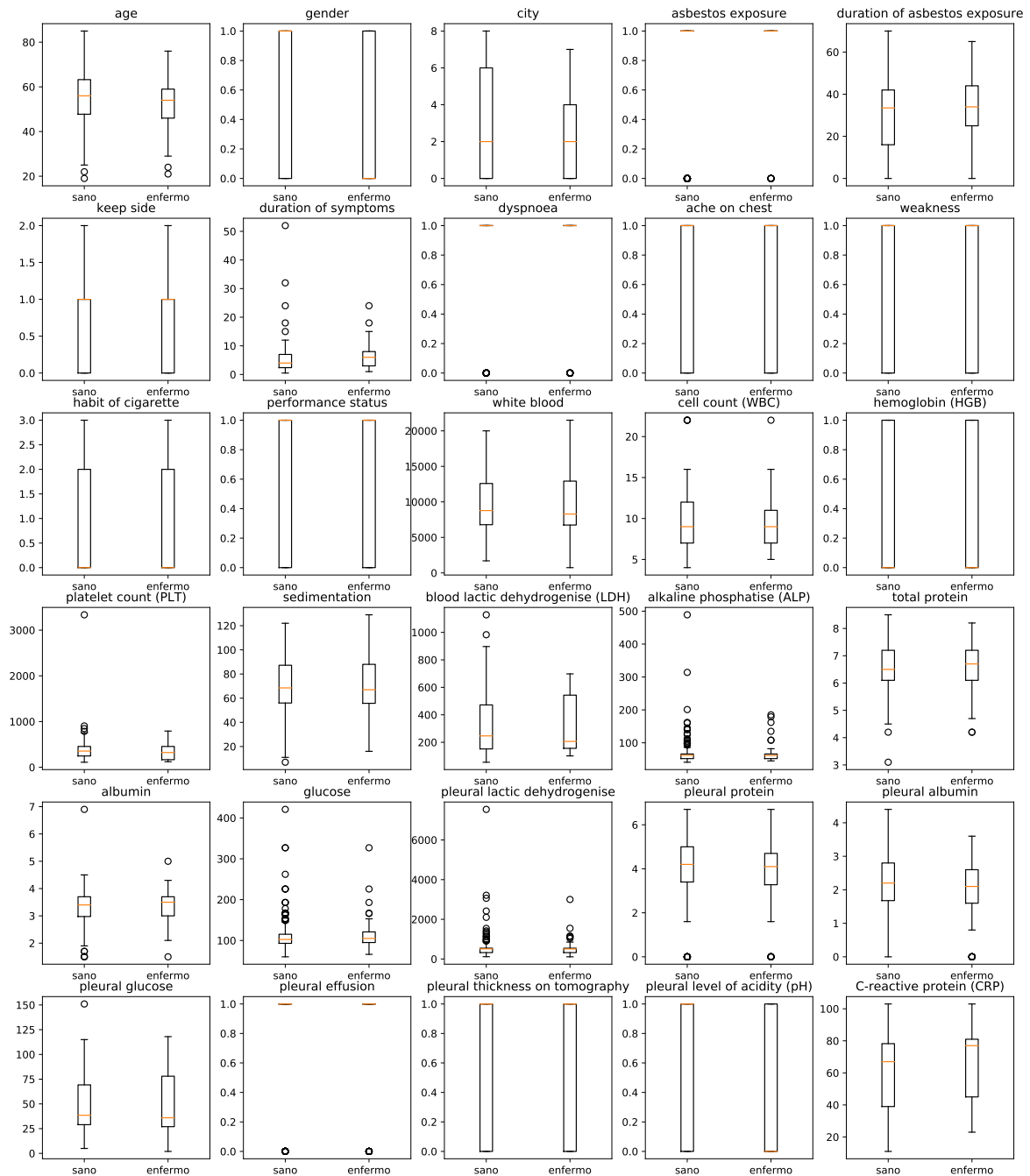


Fig. 2: Caja bigotes.

Vemos que no hay diferencias notables entre los diagramas de nuestra población de pacientes sanos y enfermos. Las variables PLT, PLT, glucosa y PLD parecen tener datos anómalos.

### 4.2.1.3 Histograma

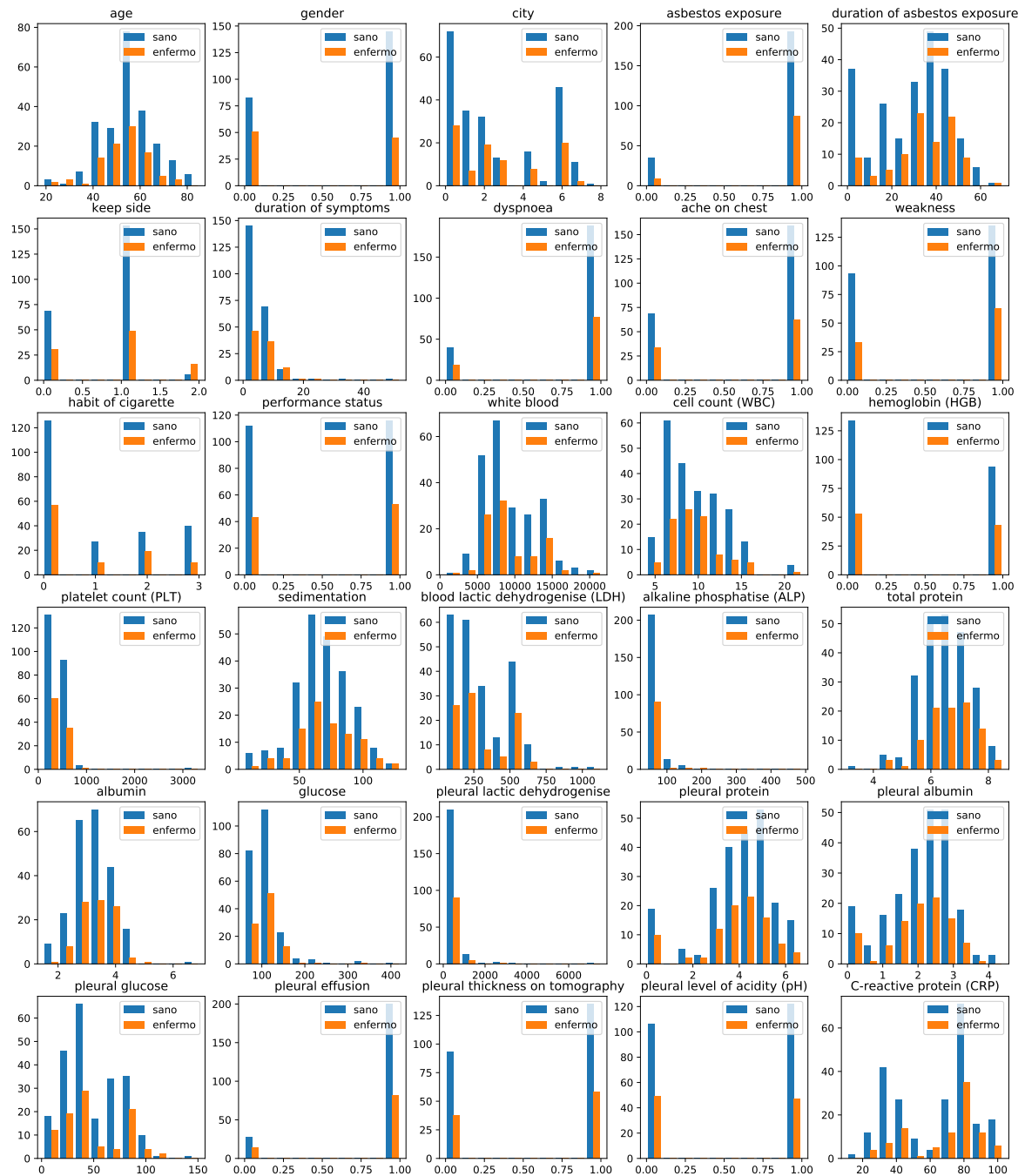


Fig. 3: Histograma.

Como es de esperar tras juzgar los diagramas de caja bigotes, los histogramas también tienen una distribución casi idéntica. Lo único en lo que se diferencian es en la cuentas totales, ello indica que tenemos más observaciones de pacientes sanos que de enfermos.

#### 4.2.1.4 Kernel Density

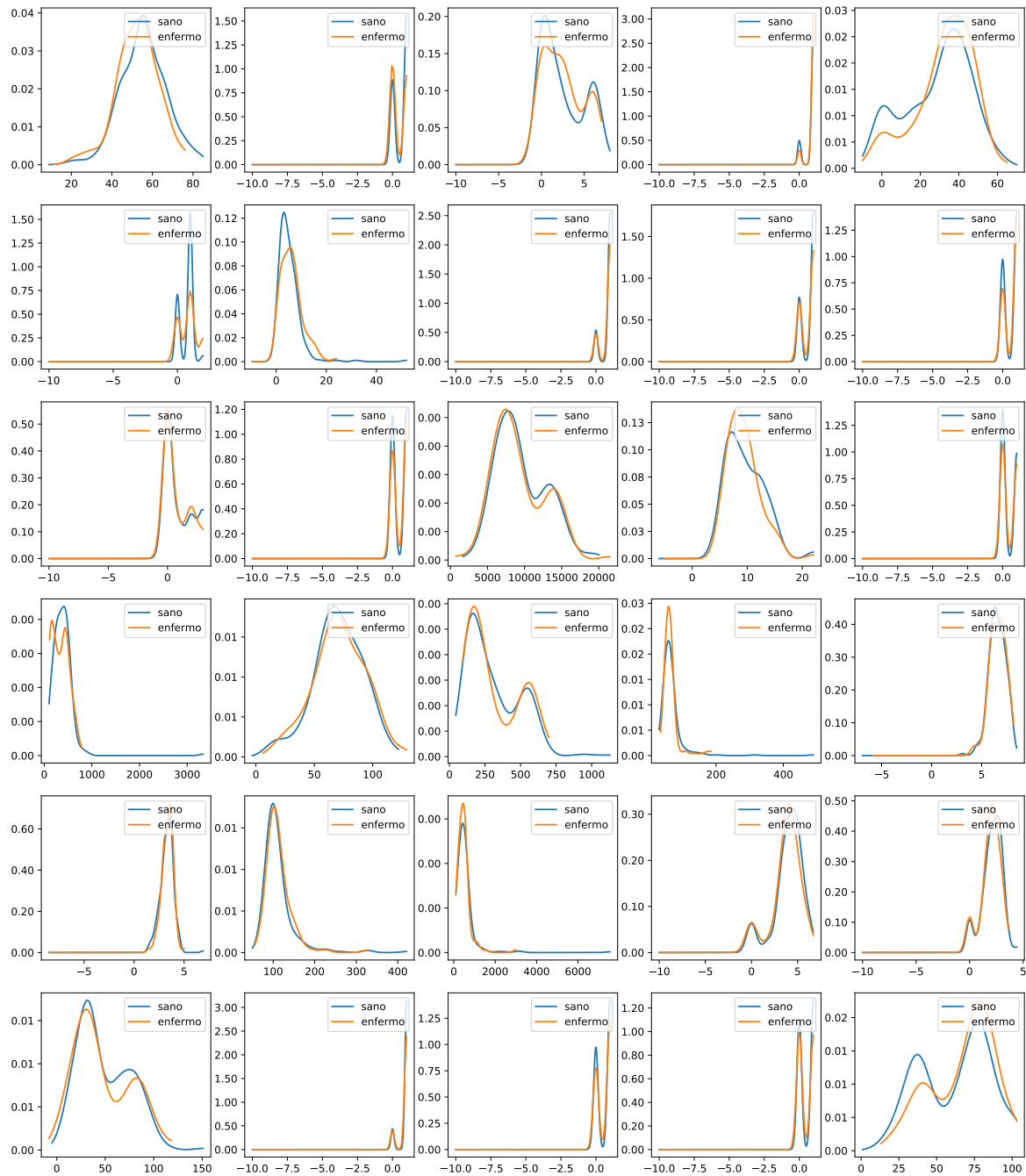


Fig. 4: Kernel Density.

A diferencia de los histogramas, al tratarse con densidades no se observan diferencias por la distinta cantidad de observaciones de sanos y enfermos. Únicamente vemos que las distribuciones son casi idénticas, como intuíamos de los histogramas.

#### 4.2.1.5 Cuantil - cuantil

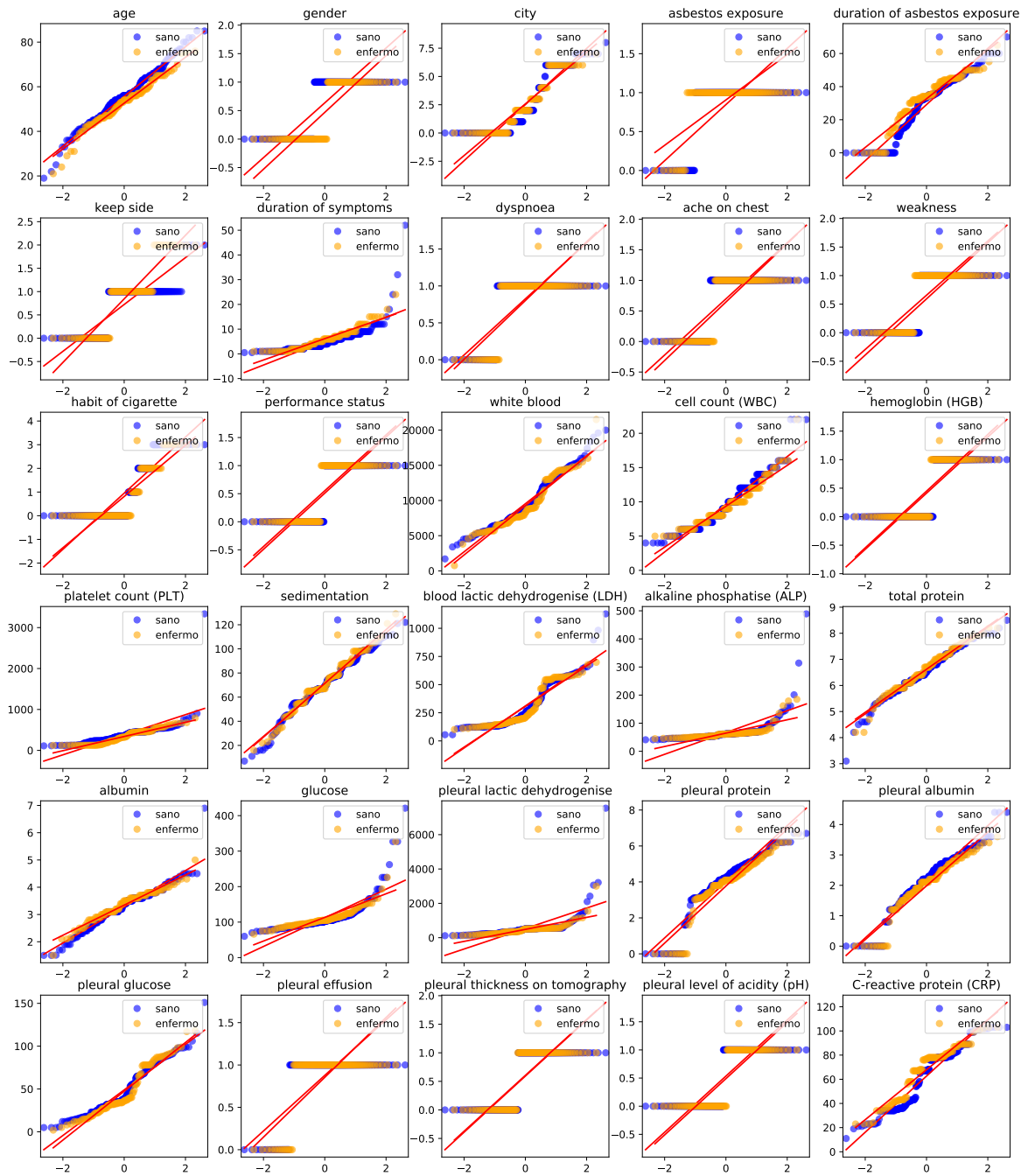


Fig. 5: Cuantil - cuantil.

4.2.1.6 Correlaciones

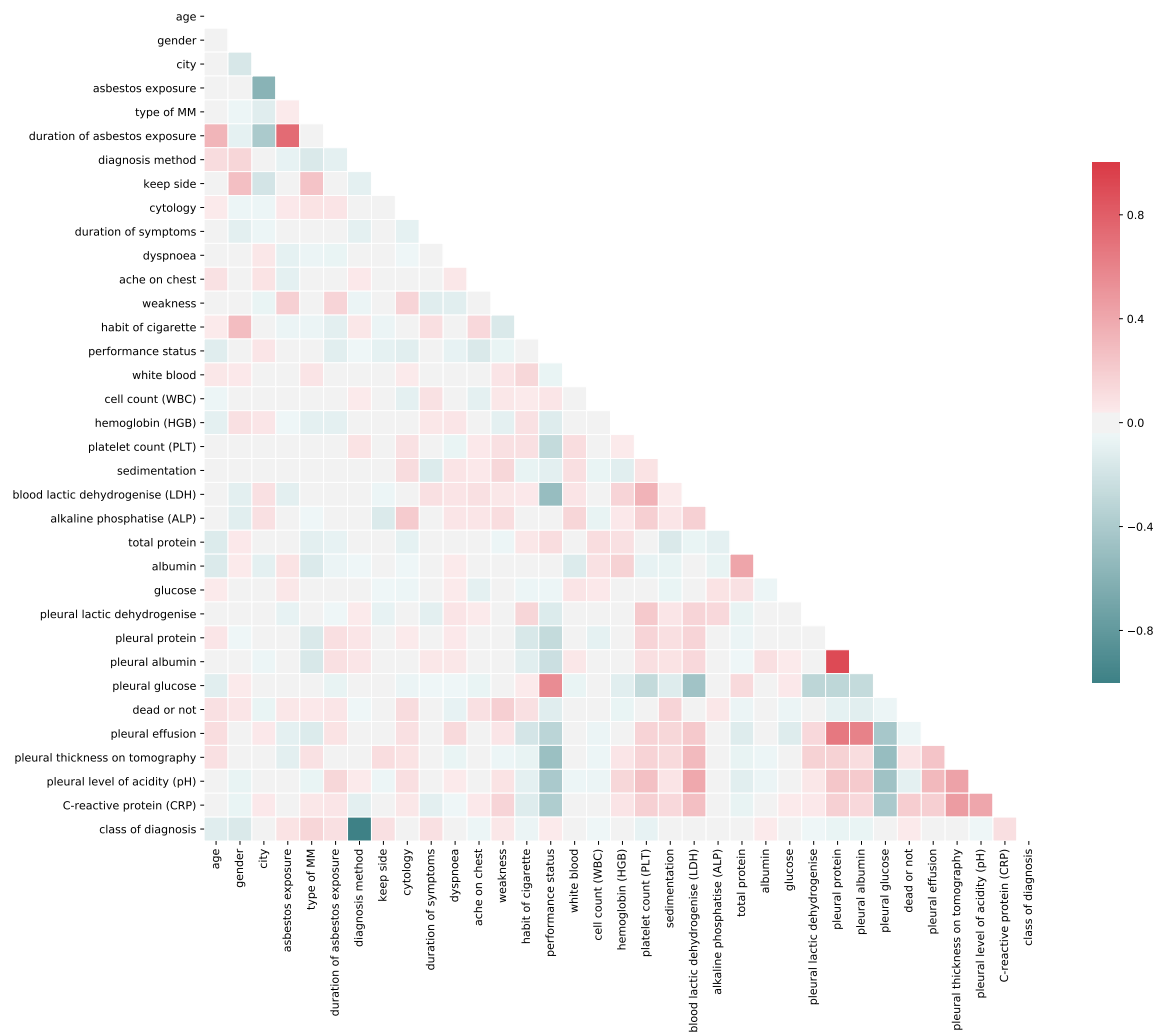


Fig. 6: Correlaciones.

## 4.2.2 Extracción de Características

### 4.2.2.1 Filter methods

1	Ranking	Variable
2	-----	-----
3	1	pleural level of acidity (pH)
4	2	C-reactive protein (CRP)
5	3	gender
6	4	pleural lactic dehydrogenise
7	5	pleural effusion
8	6	pleural glucose
9	7	pleural albumin
10	8	keep side
11	9	blood lactic dehydrogenise (LDH)
12	10	total protein
13	11	alkaline phosphatase (ALP)
14	12	white blood
15	13	performance status
16	14	cell count (WBC)
17	15	habit of cigarette
18	16	pleural protein
19	17	duration of asbestos exposure
20	18	city
21	19	dyspnoea
22	20	ache on chest
23	21	sedimentation
24	22	asbestos exposure
25	23	platelet count (PLT)
26	24	glucose
27	25	albumin
28	26	duration of symptoms
29	27	weakness
30	28	age
31	29	hemoglobin (HGB)
32	30	pleural thickness on tomography

Listing 2: Selección de variables según **f**score

Usando la puntuación de *Fisher* clasificamos las características de mayor a menor relevancia a la hora de resolver el problema de clasificación

#### 4.2.2.2 Wrapper methods

```
1 0.6782828282828283
2 Sequential Forward Selection ('asbestos exposure',
3                               'keep side',
4                               'weakness',
5                               'cell count (WBC)',
6                               'platelet count (PLT)',
7                               'alkaline phosphatase (ALP)',
8                               'glucose',
9                               'pleural protein',
10                              'pleural glucose',
11                              'C-reactive protein (CRP)')
12
13 0.5414285714285715
14 Sequential Backward Selection ('city',
15                               'asbestos exposure',
16                               'keep side',
17                               'duration of symptoms',
18                               'ache on chest',
19                               'performance status',
20                               'platelet count (PLT)',
21                               'alkaline phosphatase (ALP)',
22                               'pleural albumin',
23                               'C-reactive protein (CRP)')
```

Listing 3: Mejores variables según SFS y SBS

Una selección secuencial hacia adelante y hacia atrás con un modelo de **randomforest** con 100 estimadores y 10 parámetros calculamos una precisión del 67%. Los resultados son menores cuando hacemos la selección hacia atrás. También lo hemos intentado con un modelo **knn**, los resultados son peores en torno al 50%.

#### 4.2.2.3 PCA

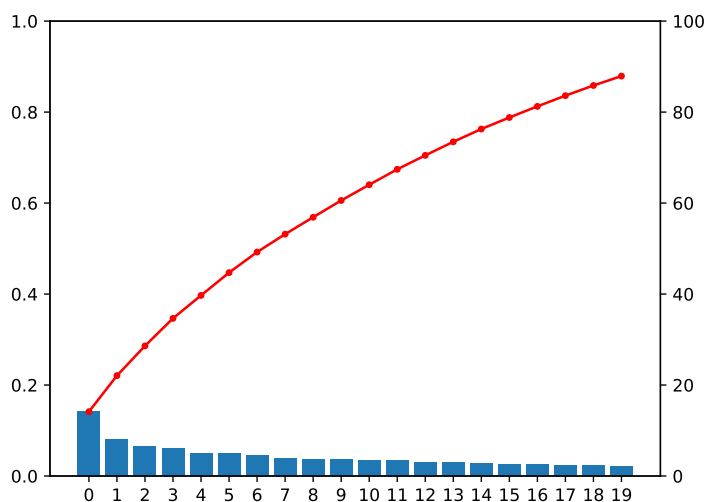


Fig. 7: Diagrama de Pareto.

Calculando el diagrama de Pareto vemos que necesitaríamos entorno a 17 componentes para explicar el 80% de la varianza de los datos. Con estos resultados vemos que será difícil reducir la dimensionalidad de nuestros datos.

### 4.2.3 Modelos de Clasificación

En segundo lugar, se ha procedido a programar varios modelos en Python a partir de las librerías disponibles para el desarrollo de tal implementación. Así se ha seleccionado el análisis discriminante lineal (LDA), el análisis discriminante cuadrático (QDA), k vecinos (KNN), random forest (RF), máquinas de soporte vectorial (SVM), una red neuronal probabilística (PNN) y un perceptrón multicapa (MLP). De esta forma, se han obtenido los siguientes resultados en lo relativo a la precisión, sensibilidad, especificidad y tiempo de ejecución de los respectivos modelos.

Model	TP	FP	FN	TN	Accuracy	Sensitivity	Specificity	Time
LDA	58.45	10.60	21.31	7.63	0.67	0.73	0.42	5.21
QDA	55.93	12.94	22.43	6.70	0.64	0.71	0.34	4.58
KNN	68.99	0.00	29.01	0.00	0.70	0.70	nan	9.74
FOREST	66.55	2.54	23.96	4.95	0.73	0.74	0.66	181.79
SVM	45.93	22.88	19.88	9.31	0.56	0.70	0.29	20.90
PNN	61.88	7.06	25.27	3.78	0.67	0.71	0.35	9.15
MLP	48.57	20.53	20.42	8.48	0.58	0.70	0.29	287.44

Table 1: Resultados agregados tras 1000 repeticiones.

### 4.2.4 Interfaz

Asimismo, con el fin de prestar una facilidad al usuario de cara a usar los modelos predictivos en cuestión, se han realizado dos interfaces simples para mostrar los resultados que se deseen. La primera interfaz desarrollada nos abre primeramente una ventana donde podemos introducir los datos del paciente a analizar mediante el sistema de ayuda a la decisión clínica. Asimismo, con el fin de facilitar la introducción de datos en el sistema la interfaz permite cargar un archivo (.csv) con todos los datos para proceder al estudio en cuestión. Una vez, se ha comprobado que los valores introducidos son correctos o modificado desde la propia interfaz aquellos que no lo sean, se indica con que modelo se pretende realizar la clasificación, mostrando en un tiempo reducido las métricas de evaluación del modelo para la base de datos y el diagnóstico particular para los datos introducidos (sano o enfermo). A continuación se muestra en detalladamente la funcionalidad y la navegación de nuestra la app.



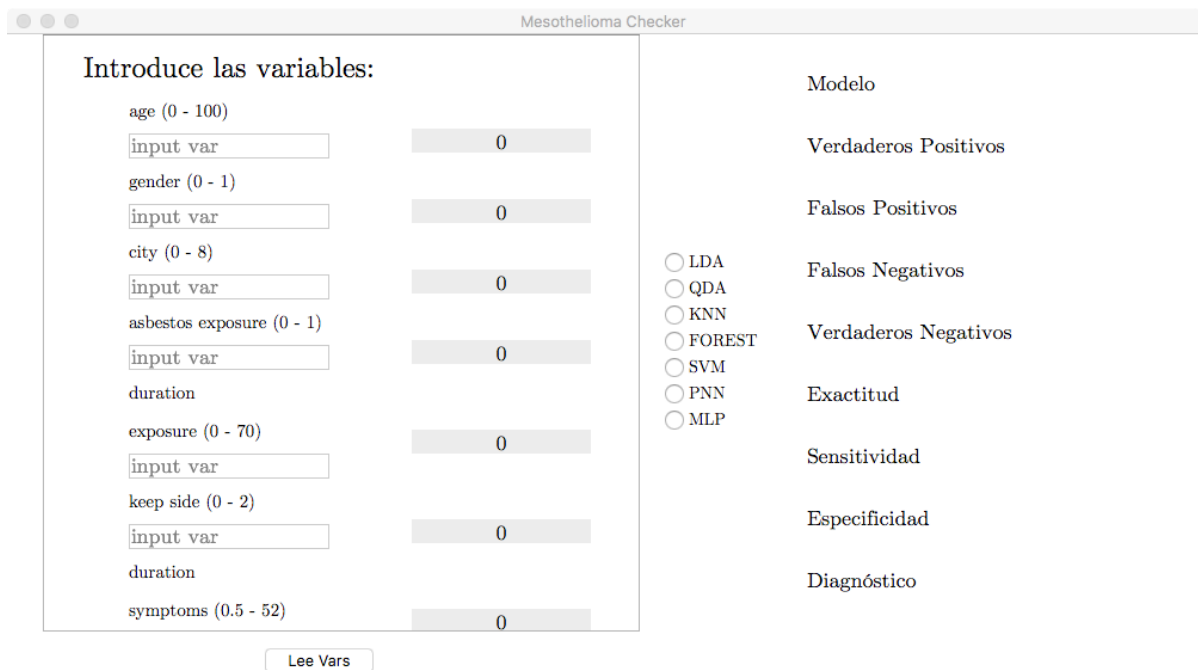


Fig. 8: Ventana principal.

Ventana principal de nuestra interfaz. A la izquierda campos de texto para introducir las variables, se muestra entre paréntesis el rango de dicha variable en nuestra base de datos.

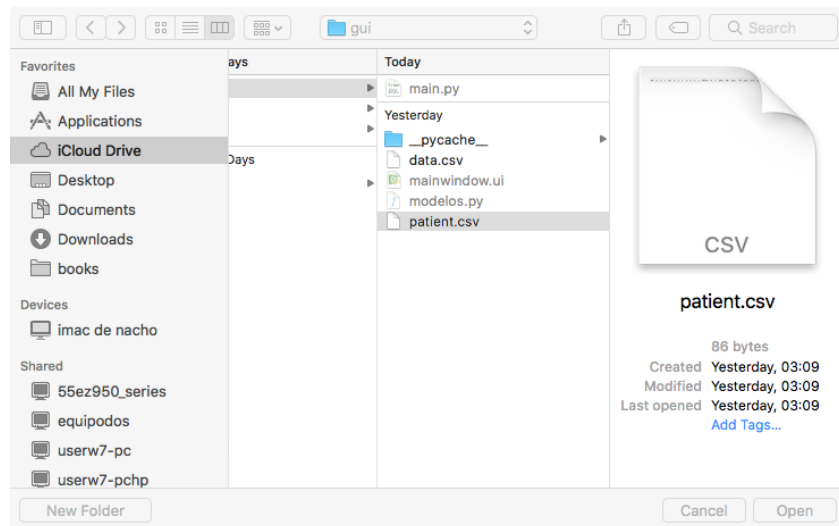


Fig. 9: Selección de paciente.

Para hacer mas cómodo su uso y evitar tener que introducir 30+ variables a mano, al presionar Cmd + O se abre una ventana para seleccionar un archivo .csv, contiene todos los valores para un paciente separados por comas.

**Mesothelioma Checker**

**Introduce las variables:**

age (0 - 100)	49	0
gender (0 - 1)	0	0
city (0 - 8)	2	0
asbestos exposure (0 - 1)	1	0
duration		
exposure (0 - 70)	20	0
keep side (0 - 2)	0	0
duration		
symptoms (0.5 - 52)		0

**Modelo**

☐ LDA

☐ QDA

☐ KNN

☐ FOREST

☐ SVM

☐ PNN

☐ MLP

Verdaderos Positivos

Falsos Positivos

Falsos Negativos

Verdaderos Negativos

Exactitud

Sensitividad

Especificidad

Diagnóstico

Fig. 10: Importación paciente.

Al importar el paciente se rellenan automáticamente las variables, se pueden comprobar y editar si fuera necesario.

**Mesothelioma Checker**

**Introduce las variables:**

age (0 - 100)	input var	49
gender (0 - 1)	input var	0
city (0 - 8)	input var	2
asbestos exposure (0 - 1)	input var	1
duration		
exposure (0 - 70)	input var	20
keep side (0 - 2)	input var	0
duration		
symptoms (0.5 - 52)		2

**Modelo**

☐ LDA

☐ QDA

☐ KNN

☐ FOREST

☐ SVM

☐ PNN

☐ MLP

Verdaderos Positivos

Falsos Positivos

Falsos Negativos

Verdaderos Negativos

Exactitud

Sensitividad

Especificidad

Diagnóstico

Fig. 11: Cargar variables.

Al presionar Enter con el teclado o el botón 'Leer Vars' con el ratón se guardan las variables en la memoria del programa. Para actualizar una variable no hay más que introducir un nuevo valor y pulsar enter o el ratón.

Para cargar el modelo se presiona Ctrl + M, o se elige desde el menú. Aparece una ventana para seleccionar el archivo .py con los modelos en python.

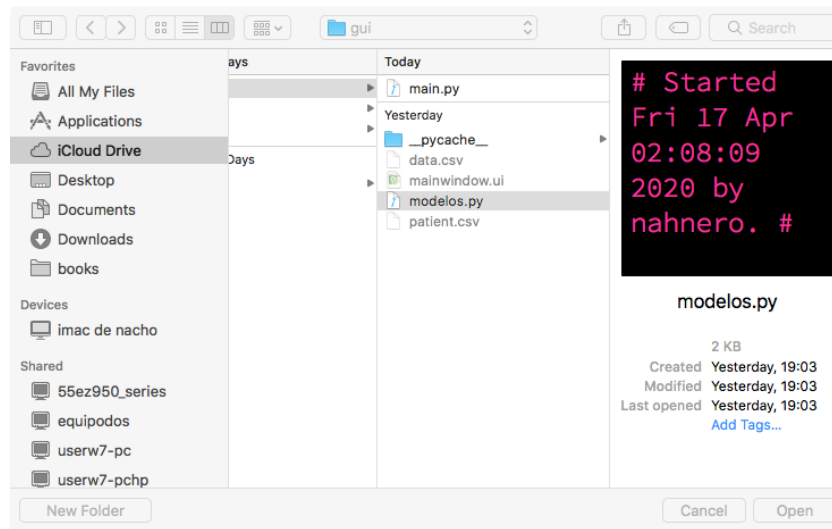


Fig. 12: Selección de modelos.

Una vez cargados los modelos se selecciona que modelo se desea ejecutar, los botones son mutuamente excluyentes. Al ejecutarse el modelo se generan las métricas de evaluación a la derecha y el diagnóstico.

The screenshot shows the 'Mesothelioma Checker' application. On the left, under 'Introduce las variables:', there are input fields for age (0-100), gender (0-1), city (0-8), asbestos exposure (0-1), duration, exposure (0-70), keep side (0-2), and symptoms (0.5-52), each with a corresponding value entered in a slider. Below these is a 'Lee Vars' button. In the center, there is a list of model selection options: LDA (selected with a blue dot), QDA, KNN, FOREST, SVM, PNN, and MLP. On the right, a table displays the results for the selected LDA model.

Modelo	LDA
Verdaderos Positivos	44.00
Falsos Positivos	5.00
Falsos Negativos	12.00
Verdaderos Negativos	4.00
Exactitud	0.74
Sensitividad	0.79
Especificidad	0.44
Diagnóstico	enfermo

Fig. 13: Selección de modelos.

Cuando se selecciona otro modelo se reevalúa al paciente, cada modelo tiene unas características diferentes en cuanto a exactitud, sensibilidad y especificidad. En función de lo que se busque se puede seleccionar un modelo adecuado.

Si bien es verdad que es un prototipado básico y que podría tener mejoras en múltiples aspectos (interfaz más amigable, más funciones, pruebas de usabilidad...) es importante recalcar el objetivo específico que se pretende con dicho interfaz en particular. En este caso, estaría destinado a una vertiente de carácter investigador, donde se pretenda comparar los múltiples modelos empleados, analizar sus resultados y determinar en que situaciones convendría utilizar un determinado modelo u otro según la distribución de los datos.

No obstante, a la hora de emplear un sistema de ayuda a la decisión clínica de forma continuada y habitual en centros hospitalarios, la opción de incluir varios de modelos de predicción no es relevante, ya que la finalidad en este caso es simple; obtener la clasificación del modelo con mejores métricas para ayudar al diagnóstico médico del mesotelioma maligno. Por tanto, se ha desarrollado también un segundo interfaz donde se anula la posibilidad de seleccionar un modelo en particular, y a partir de los datos introducidos el modelo clasifica al paciente como sano o enfermo.

#### 4.2.4.1 Aplicación para médicos

Esta versión restringe al usuario a usar un modelo predeterminado. Al pulsar el botón ejecutar, después de haber introducido los datos del paciente se calcula el modelo y se presentan los resultados.

Introduce las variables:		Modelo	FOREST
age (0 - 100)	49	Verdaderos Positivos	43.00
gender (0 - 1)	0	Falsos Positivos	3.00
city (0 - 8)	2	Falsos Negativos	17.00
asbestos exposure (0 - 1)	1	Verdaderos Negativos	2.00
duration exposure (0 - 70)	20	Exactitud	0.69
keep side (0 - 2)	0	Sensitividad	0.72
symptoms (0.5 - 52)	2	Especificidad	0.40
		Diagnóstico	enfermo

Fig. 14: Versión limitada para médicos.

Así, aunque sería más idóneo que el modelo clasificase el grado de enfermedad en vez de dar un resultado binario, se demuestra que perfeccionando los modelos de predicción y alimentando los modelos con más datos podría llegarse a un sistema de ayuda a la decisión clínica de gran utilidad para cualquier profesional médico, ya que en pocos segundos o minutos se obtendría una precisión a tener en cuenta. Asimismo, cabe destacar que también habría que consolidar una interfaz intuitiva y amigable de forma que al personal sanitario le suponga una mejora en su calidad y ritmo de trabajo y no un descenso en su rendimiento habitual.

El sistema de ayuda a la decisión clínica en este caso sería de gran utilidad para diagnosticar el mesotelioma maligno prematuramente, pudiendo ofrecer un mejor tratamiento desde un inicio. No obstante, tal y como se ha expuesto en la solución ideal, se busca una solución para el problema del mesotelioma maligno relación con

la exposición al amianto. Luego entonces, únicamente realizando un prototipado de un sistema de apoyo a la decisión clínica, si bien es cierto que nuestro proyecto sería útil, también se podría definir como un proyecto incompleto donde no se estaría optimizando los recursos existentes en la actualidad relacionados tanto con las TICS como con el propio mesotelioma maligno.

#### **4.2.4.2 Aplicación para pacientes**

De esta forma, intentando aportar una solución integral a la situación expuesta en lo referente al amianto y al mesotelioma maligno, se ha desarrollado también una primera aproximación de lo que sería una aplicación móvil para enfermos, personas en riesgo y médicos. Así, mediante la herramienta desarrollada por el instituto tecnológico de Massachusetts (MIT) App Inventor, se ha podido implementar un prototipado con varias de las funcionalidades de las que la aplicación en su fase definitiva debería ofrecer.

Inicialmente al abrir la app nos encontramos con una pantalla inicial en la cual podemos iniciar sesión o registrarnos como pacientes que padecen la enfermedad, individuos de riesgo o profesionales médicos. Así, aunque en este sistema de registro al final se termina accediendo al mismo menú en los tres casos debido a que en la base de datos se guardan sin distinción alguna entre enfermos, personas en riesgo y personal sanitario, se ha considerado relevante plasmar en el prototipado la distinción entre grupos distintos de usuarios. De esta forma, de cara a la consolidación de una aplicación sólida, habría que guardar cada usuario con la pertenencia correspondiente al grupo en el que se han registrado, de manera que los servicios ofrecidos en cada inicio de sesión serían distintos según seamos enfermos, personas con riesgo o médicos. Asimismo, la cuenta debería estar conectada con un identificador único, de manera que médicos y pacientes tuviesen la posibilidad de acceder a registros personales.

Una vez, llegados al menú principal, en el caso de primera versión de la aplicación se nos ofrecen cinco opciones: contactar con un médico o paciente mediante correo electrónico, visualizar nuestra ubicación y las zonas donde la exposición al amianto sea elevada, información y recomendaciones sobre el problema a tratar, las cuales pueden ser reproducidas vía audio, con un enlace para descargar una guía de prevención contra el amianto y un registro donde introducir variables referentes al paciente, las cuales pueden luego almacenarse como un archivo csv para luego poder enviarlas al centro médico. A su vez, pese a la precariedad del registro de datos implementado, ya que para enviar los datos habría que emplear otra aplicación en el caso del prototipado desarrollado, incluye varias funcionalidades como renombrar variables o eliminar y modificar registros. Finalmente, en el menú principal implementado también existe la opción de volver al inicio, lo cual sería equivalente en este desarrollo inicial a cerrar sesión.

Mesothelioma Checker

Introduce las variables:

age (0 - 100)

input var

0

gender (0 - 1)

input var

0

city (0 - 8)

input var

0

asbestos exposure (0 - 1)

input var

0

duration

exposure (0 - 70)

input var

0

keep side (0 - 2)

input var

0

duration

symptoms (0.5 - 52)

0

Lee Vars

Modelo

Verdaderos Positivos

Falsos Positivos

Falsos Negativos

Verdaderos Negativos

Exactitud

Sensitividad

Especificidad

Diagnóstico

☐ LDA

☐ QDA

☐ KNN

☐ FOREST

☐ SVM

☐ PNN

☐ MLP

Fig. 15: PLACEHOLDER.

Por tanto, ha quedado patente la versatilidad y eficacia que presenta App Inventor, ya que permite el desarrollo del prototipado de una aplicación de forma muy simple, intuitiva a partir de nociones básicas de programación que cualquiera puede aprender muy fácilmente. No obstante, pese a que se puede considerar una primera aproximación correcta, evidentemente es incompleta ya que carece de múltiples funcionalidades que requerirían sumergirnos en lenguajes de programación como C++, Java o Kotlin, para los cuales sería necesaria una inversión de tiempo mucho mayor. Así, para llegar a una aplicación donde verdaderamente se sobrepasase la fase de prototipado, habría que posiblemente emplear uno de estos lenguajes y añadir las siguientes funcionalidades principalmente:

- Conexión entre las sesiones y los registros médicos, de forma que médicos y pacientes pudiesen acceder a dichos registros
- Almacenar usuarios teniendo en cuenta si son profesionales sanitarios, personas de riesgo o enfermos, de forma que los servicios existentes en cada menú fuesen distintos
- Inclusión de un chat para todos los menús, de manera que los pacientes pudiesen contactar entre ellos y con sus respectivos médicos, pudiendo ayudarse unos a otros y disponiendo de un canal de comunicación fluido para contactar con el personal sanitario
- En los servicios disponibles para médicos, eliminar la opción de registrar datos y añadir un nuevo servicio que fuese para visualizar los registros de los pacientes
- Tanto en pacientes de riesgo como enfermos, posibilitar la opción de introducir todos los datos que precise

el médico para un diagnóstico inicial de los que pueda disponer el paciente, y subir los datos a un servidor a partir del cual el sistema de ayuda a la decisión clínica y otros servicios del hospital pudiesen acceder.

- Adición de notificaciones y recordatorios de citas, ingesta de medicamentos y medidas de seguridad para personas de riesgo y enfermos
- Monitorización de la posición de los pacientes, de forma que al estar a menos de una determinada distancia de una zona con una gran exposición al amianto, saltase una alarma.
- Inclusión en el servicio de información y recomendaciones enlaces a videos, noticias de periódico y más información que se fuese actualizando periódicamente.

### 4.3 Servidores y conexiones y seguridad

En lo concerniente a los servidores para almacenar la información, se ha decidido optar por la opción de un servidor en la nube. Así, el motivo de esta decisión reside en las ventajas que ofrece un servidor con estas características, ya que podríamos ampliar o reducir el espacio disponible de forma flexible aumentando la rentabilidad frente a almacenamientos compartidos o locales. A su vez, la idea del proyecto sería intentar expandir su alcance a nivel autonómico o nacional progresivamente, por lo que sería necesario almacenar la información de los pacientes en un espacio al que se pueda acceder fácilmente por todos los hospitales o centro sanitarios implicados en el proyecto sin comprometer por ello los recursos individuales de alguno de ellos. A su vez, habría que optar por una plataforma que ofreciese unos umbrales de seguridad aptos para el tratamiento de información sanitaria, de forma que realmente pudiese aprovecharse ventajas como la confiabilidad o escalabilidad de los servidores en la nube.

Asimismo, incidiendo en el aspecto de la seguridad, el proyecto en cuestión apuesta por un paciente activo el cual tiene la capacidad de participar activamente en su propia salud, en este caso mediante la información y recomendaciones ofrecidas en la aplicación y a partir de la consulta de registros médicos y envío de datos de su estado actual a los servidores mencionados anteriormente. También, sería mencionable que la aplicación dispondría no sólo de un chat para comunicar pacientes, sino también para contactar con profesionales sanitarios, por lo que se estaría transmitiendo continuamente información personal y confidencial. Por tanto, debido al flujo de datos que se estaría dando y a la condición de éstos, sería necesario el uso de una red privada virtual (VPN). Así, el tráfico de datos generado viajaría cifrado, dificultando que agentes externos puedan robar la información confidencial transmitida. A su vez, todos los usuarios de la aplicación podrían emplearla haciendo uso de redes Wi-Fi locales o públicas con la tranquilidad de que se están empleando recursos para proteger la integridad y confidencialidad de la información. Por tanto, podríamos emplear protocolos como Ipsec para mejorar la seguridad de nuestras conexiones y transmisiones de datos.

## 4.4 Proyecto ALERT

Con el fin de localizar aquellas regiones en la que la exposición al asbesto fuese mayor, se emplearía el dispositivo ALERT. Este sistema desarrollado como un proyecto europeo en el que han colaborado varios centros de investigación entre los que podríamos destacar el Instituto de Biomecánica de Valencia (IBV), permite detectar la presencia de fibras de amianto en el aire analizando los patrones de luz de las partículas suspendidas en el entorno. Así, gracias a su pequeño tamaño, lo que hace que se considere un dispositivo portable, junto con la capacidad de analizar en tiempo real la cantidad de microfibras de amianto en el espacio que lo rodea, permitiría identificar aquellas industrias donde el riesgo fuese mayor pudiendo obtener información espacial sobre las zonas donde el nivel de riesgo debido a la manipulación del amianto es mayor.

## 5 Conclusiones

## 6 Lecciones aprendidas y reflexiones

Mediante la realización del trabajo en cuestión hemos podido comprobar la dificultad que supone no sólo plantear una idea, sino también llevarla a la práctica. Es preciso y fundamental iniciar cualquier proyecto comprendiendo correctamente el problema identificado, para poder luego desarrollar una solución que realmente sea útil y pueda aplicarse en la resolución de aquello que se pretende alcanzar. Es precisamente por ello por lo que se ha pretendido abordar el tratamiento del mesotelioma maligno desde la perspectiva de las tecnologías de la información y la comunicación mediante un enfoque global. Con esto nos referimos a que se ha intentado prototipar de un proyecto en el que se pretende ayudar tanto a pacientes e individuos de riesgo como a profesionales sanitarios, con el fin de poder consolidar una solución integral al problema en cuestión.

No obstante, si dicho proyecto se fuese a realizar para finalmente obtener una solución verdaderamente aplicable a los hospitales y a la sociedad, es evidente que se necesitaría un equipo multidisciplinar en el que deberían colaborar médicos, ingenieros y pacientes para realmente identificar aquellos puntos del problema donde la utilización de recursos tecnológicos e informáticos permitiesen una mejora directa en la calidad de vida del paciente. Asimismo, durante el desarrollo del trabajo expuesto se ha incidido en el concepto de paciente activo, ya que mediante el uso de las tecnologías de la información y comunicación dicho paciente, en un pasado pasivo, puede adquirir un rol principal en el cuidado de su salud a partir de su colaboración con el personal sanitario mediante el uso por ejemplo de aplicaciones móviles.

A su vez, es de vital importancia mencionar los controles que deberían pasar todas y cada una de las partes que integran el proyecto, tanto el sistema de ayuda a la decisión clínica y la aplicación móvil, como la seguridad en la transmisión y el almacenamiento de datos durante el flujo de trabajo de la solución desarrollada.



Las tecnologías de la información y la comunicación pueden considerarse como uno de los valores más significativos a nivel actual, no sólo para el sector industrial, sino también para el mercado sanitario. Así, a partir del acoplamiento de múltiples herramientas que colaboren y actúen en sintonía unas con otras existe la posibilidad de mejorar y optimizar los recursos clínicos y médicos, lo que conlleva de forma directa una mejora en la calidad asistencial del protagonista del campo sanitario: el paciente. Finalmente, destacar que durante la realización del trabajo se ha ido aplicando los conocimientos adquiridos durante la carrera, y específicamente, los desarrollados en las asignaturas de sistemas de información y telemedicina, por lo que se ha podido aplicar a la resolución de un problema (al menos, en el planteamiento de la solución y el prototipado de la misma) demostrando utilidad y aplicabilidad de los contenidos académicos a la realidad.