



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



ESCUELA TÉCNICA
SUPERIOR INGENIEROS
INDUSTRIALES VALENCIA

TRABAJO FIN DE GRADO EN INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES

**DESARROLLO E IMPLEMENTACIÓN DE UNA
APLICACIÓN PARA LA
CARACTERIZACIÓN, EL PROCESADO DE Y
ANÁLISIS DE SEÑALES
ELECTROHISTEROGRÁFICAS EN ENTORNO
MATLAB.**

AUTOR: Alejandro Golfe San Martín

TUTORA: Yiyao Ye Lin

COTUTORES: Gema Prats Boluda
Javier Mas Cabo

Curso Académico: 2017-18



RESUMEN

La electrohisterografía (EHG) es el registro de la actividad mioeléctrica uterina en superficie. Estudios anteriores han demostrado que se puede utilizar como herramienta alternativa de monitorización no invasiva de la dinámica uterina. En comparación con el registro de toconodinometría convencional, los registros de EHG aportan información adicional sobre el estado electrofisiológico del útero, lo cual es de especial relevancia para la detección precoz del parto prematuro. Éste es uno de las complicaciones comunes en la obstetricia, representa el 7% de todos los nacimientos y el 85% de las muertes perinatales. Aunque se conocen que la amplitud de la señal EHG aumenta conforme se acerca el parto, y se produce un desplazamiento del contenido espectral hacia altas frecuencias, no existen estudios de la evolución temporal de los parámetros de no-linealidad y orden durante la gestación. El objetivo de este trabajo es desarrollar herramientas Matlab que permitan caracterizar la evolución de la electrofisiología uterina durante la gestación en base a los registros de EHG. Para llevar a cabo el estudio planteado se empleará la base de datos pública de registros de EHG (TPEHG) disponible.

Palabras clave: Electrohisterografía parto_prematuro herramientas_Matlab análisis_espectral parámetros_no_lineales análisis_estadístico análisis_roc



RESUM

L'electrohisterografia (EHG) és el registre de l'activitat mioelèctrica uterina en superfície. Estudis anteriors han demostrat que es pot utilitzar com a ferrament alternativa de monitorització no invasiva de la dinàmica uterina. En comparació amb el registre de toconodinamometria convencional, els registres d'EKG aporten informació addicional sobre l'estat electrofisiològic de l'úter, la qual cosa és d'especial rellevància per a la detecció precoç del part prematur. Este és un de les complicacions comunes en l'obstetrícia, representa el 7% de tots els naixements i el 85% de les morts perinatals. Encara que es coneixen que l'amplitud del senyal EHG augmenta conforme s'acosta el part, i es produïx un desplaçament del contingut espectral cap a altes freqüències, no hi ha estudis de l'evolució temporal dels paràmetres de no-linealitat i orde durant la gestació. L'objectiu d'este treball és desenrotllar ferramentes Matlab que permeten caracteritzar l'evolució de l'electrofisiologia uterina durant la gestació basant-se en els registres d'EKG. Per a dur a terme l'estudi plantejat s'emprarà la base de dades públiques de registres d'EKG (TPEKG) disponible.

Paraules clau: Electrohisterografia part_prematur ferramentes_Matlab anàlisis_espectral paràmetres_no_lineals anàlisis_estadístic anàlisis_roc



ABSTRACT

Electrohysterography (EHG) is the register of the uterine myoelectric activity on the surface. Previous studies have demonstrated that it can be used as a non invasive alternative tool to monitor the dynamic of the uterus. In comparison to the register of conventional toconodinamometry, the EHG register gives additional information about the electrophysiological state of the uterus, which is particularly relevant for the early detection of premature births. This is one of the main complications in obstetrics, representing 7% of all births and 85% of perinatal deaths. Although it is known that the EHG amplitude signal increases as the birth approaches, and there is a displacement of the spectral content towards big frequencies, there are no studies of the temporal evolution of the non-linear parameters and order during gestation. The goal of this project is to develop Matlab tools that allow to characterize the evolution of the uterine electrophysiology during gestation in relation to the EHG registers. In order to carry out the proposed study, the available public EHG register database (TPDHG) will be used

Key words: electrohysterography preterm_delivery Matlab_tools spectral_analysis non_linear_parameters statistic_analisis ROC_analysis



DOCUMENTOS CONTENIDOS EN EL TRABAJO

- MEMORIA
- MANUAL DE USUARIO
- PRESUPUESTO



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Desarrollo e implementación de una aplicación para la
caracterización, el procesado de y análisis de señales
electrohisterográficas en entorno MATLAB



ESCUELA TÉCNICA
SUPERIOR INGENIEROS
INDUSTRIALES VALENCIA

PORTADA DE LA MEMORIA



Índice de la Memoria

1. INTRODUCCIÓN.....	2
1.1 Parto prematuro	5
1.2 Técnicas para la detección trabajo de parto.....	6
1.3 EHG.....	7
2. JUSTIFICACIÓN Y OBJETIVOS DEL TRABAJO.....	9
3. DESARROLLO ENTORNO GRÁFICO GUIDE	10
3.1 Descripción de la herramienta.....	10
3.2 Especificaciones de diseño	11
3.3 Descripción funciones entorno gráfico.....	14
3.3.1 Pantalla principal.....	14
3.3.2 Visualización.....	16
3.3.3 Resegmentar	17
3.3.4 Análisis de señal EHG.....	19
3.3.5 Análisis estadístico.....	23
3.3.6 Informe estadístico.....	27
3.3.7 Sistema predictor	29
3.3.8 Evaluación del sistema predictor	38
4. RESULTADOS.....	40
4.1 Visualización señal.....	40
4.2 Visualización de la segmentación.....	43
4.3 Análisis de los parámetros.....	46
4.4 Análisis estadístico	48
4.5 Informe estadístico.....	52
4.6. Modelos predictivos	55
5. CONCLUSIÓN	63
6. BIBLIOGRAFÍA	64



Índice de figuras

Figura 1. Ejemplo señal EHG para un ancho de banda de 0.1 a 4 Hz.....	7
Figura 2. Configuración de los electrodos para la obtención de registros[10].....	7
Figura 3. Descripción teórica del electromiograma uterino (EMG). IUP, presión intrauterina[13]....	8
Figura 4. Entorno de desarrollo GUIDE de Matlab	10
Figura 5. Flujograma explicativo del proceso de razonamiento de la interfaz.....	13
Figura 6. Pantalla principal.	14
Figura 7. Directorio funciones pantalla principal.....	15
Figura 8. Pantalla de visualización	16
Figura 9. .Directorio funciones ventana visualización	16
Figura 10. Pantalla resegmentar.....	17
Figura 11. Directorio funciones ventana resegmentar.	18
Figura 12. Ventana de análisis.....	19
Figura 13. Registro EHG de 1200 segundos en el que se pueden identificar cinco contracciones, resaltadas en verde.....	20
Figura 14. Esquema de la división de la señal en ventanas para el análisis de registro completo...21	
Figura 15. Esquema de cálculo de los parámetros espectrales.	21
Figura 16. Directorio funciones ventana análisis.....	22
Figura 17. Ventana distribución variables.	23
Figura 18. Directorio funciones distribución de variables.....	24
Figura 19. Ventana habilitada para realizar el análisis estadístico de los datos.....	25
Figura 20. Directorio funciones ventana análisis estadístico	26
Figura 21. Ventana informe estadístico	27
Figura 22. Directorio funciones ventana informe estadístico.	27
Figura 23. Ventana de generación de modelos predictivos	29
Figura 24. Panel mostrado para la introducción de los porcentajes de datos destinados al conjunto de entrenamiento, validación y test.	30
Figura 25. Esquema proceso de generación de sistemas predictores	30
Figura 26. Esquema red neuronal antes del entrenamiento.....	34
Figura 27. Conexiones con distinto peso sináptico que convergen en la misma neurona.	34
Figura 28. Ejemplo separación categorías en dos dimensiones.	35



Figura 29. Esquema estructura árbol de decisión.36

Figura 30. Directorio funciones ventana algoritmo37

Figura 31. Directorio funciones curva ROC.....39

Figura 32. Formato de almacenamiento de un registro en planta en directorio codificado por fecha.Figura 33. Formato de almacenamiento de un registro de guardia en directorio codificado por fecha y su conversión en el directorio CONEX.40

Figura 34. Creación de la carpeta de sesión 1 asociada a los datos crudos del directorio F11_06_2015H8_20.....41

Figura 35. Señal obtenida para el paciente con N° 2368679. Ancho de banda de 0.1 a 4 Hz.....42

Figura 36. Señal obtenida para el paciente con N° 2368679. Ancho de banda de 0.2 a 0.34 Hz....42

Figura 37. Ejemplo contenido archivo de segmentación de registro completo del paciente 1552644 sesión 143

Figura 38. Ejemplo de la visualización de la segmentación de la señal EHG adquirida en la paciente 2263442 para el análisis de registro completo.....44

Figura 39. Ejemplo de la visualización de la segmentación de la señal EHG adquirida en la paciente 2263442 para el análisis de contracciones.....44

Figura 40. Ventana de modificación de la segmentación mostrada anteriormente. En amarillo el tramo seleccionado para su tratamiento en la interfaz.....45

Figura 41. Directorio de análisis dónde se almacena el análisis de la base de datos completa y los análisis de pacientes específicos46

Figura 42. Ejemplo de nomenclatura almacenamiento archivo extensión .mat de un análisis específico de un paciente.....47

Figura 43. Fichero Excel generado después del cómputo de la base de datos completa.47

Figura 44. Visualización de la distribución del parámetro valor pico a pico.48

Figura 45. Diagrama Caja-Bigotes generado para grupo de clasificación a término y pretérmino. .49

Figura 46. Diagrama Caja-Bigotes generado para grupo de clasificación menor de 7 días y mayor de 7 días de tiempo al parto.....49

Figura 47. Diagrama Caja-Bigotes generado para grupo de clasificación menor de 14 días y mayor de 14 días de tiempo al parto.....50

Figura 48. Diagrama Caja-Bigotes generado para grupo de clasificación menor de 30 días y mayor de 30 días de tiempo al parto.....50

Figura 49. Ejemplo informe estadístico para los parámetros seleccionados anteriormente52

Figura 50. Ejemplo de exportación a Excel del informe estadístico53

Figura 51. Directorios generados para cada grupo de clasificación seleccionado en el análisis estadístico.54

Figura 52. Informe estadístico exportado y p-valor almacenados para el grupo de clasificación pretérmino y a término.....	54
Figura 53. Ventana selección para la generación de modelos predictivos.	55
Figura 54. Directorio raíz con los distintos grupos de clasificación	56
Figura 55. Ejemplo directorio de la reducción de dimensión para el grupo de clasificación menor de 14 vs mayor de 14.	56
Figura 56. Ejemplo directorios que contienen la información de los modelos generados.....	56
Figura 57. Ejemplo almacenamiento particiones en regresión lineal.....	57
Figura 58. Ejemplo partición generada	57
Figura 59. Resultados obtenidos en la ventana gráfica de análisis ROC mediante regresión logística para grupo de clasificación TTD<7 vs TTD>7, con reducción de dimensión mediante el algoritmo PCA.....	58
Figura 60. Ejemplo directorio almacenamiento redes generadas por la interfaz	60
Figura 61. Ejemplo almacenamiento redes generadas en su directorio correspondiente	60
Figura 62. Resultados obtenidos en la ventana gráfica de análisis ROC mediante redes neuronales para grupo de clasificación TTD<7 vs TTD>7 con reducción de dimensión mediante el algoritmo PCA.....	61

Índice de tablas

Tabla 1. Especificaciones técnicas necesarias para la ejecución del entorno matemático.	12
Tabla 2. Esquema de cálculo de parámetros para la señal EHG mostrada en la Figura 13.....	20
Tabla 3. Elementos que componen la matriz de confusión de un modelo predictivo generado.	39
Tabla 4. Matriz de confusión obtenida para el método de clasificación regresión logística, grupo de clasificación TTD<7 vs TTD>7 con reducción de dimensión mediante el algoritmo PCA.	58
Tabla 5. Resultados obtenidos en el análisis ROC para método de clasificación regresión logística, grupo de clasificación TTD<7 vs TTD>7 y reducción de dimensión mediante algoritmo PCA.....	59
Tabla 6. Matriz de confusión obtenida para el método de clasificación redes neuronales, grupo de clasificación TTD<7 vs TTD>7, con reducción de dimensión mediante el algoritmo PCA.....	61
Tabla 7. Cuadro resumen de los resultados obtenidos en el análisis ROC para método de clasificación redes neuronales, grupo de clasificación TTD<7 vs TTD>7 y reducción de dimensión mediante algoritmo PCA.	62



1. INTRODUCCIÓN

1.1 Parto prematuro

El parto humano es la culminación del embarazo hasta el periodo de salida del bebé del útero materno. Por otra parte, el parto prematuro o parto pretérmino se define como el parto ocurrido antes de las 37 semanas de gestación, en oposición a la mayoría de embarazos que duran más de 37 semanas[1]. El parto puede prolongarse hasta la semana 41 o 42 de gestación, considerándose a término a partir de la semana 37. Según la edad gestacional en el momento del parto un recién nacido puede ser prematuro, a término y a posttérmino. El parto prematuro se ha convertido en la primera causa de mortalidad en niños a nivel mundial y se estima que al año nacen 15 millones de niños prematuros, cifra que está en aumento [2].

El parto prematuro lleva asociado un mayor riesgo para el niño de desarrollar complicaciones en el desarrollo, además de una mayor tasa de mortalidad. Dichas tasas de mortalidad presentan una gran disparidad entre los distintos países del mundo, debido principalmente a su nivel de desarrollo.

Muchos de los niños prematuros que sobreviven sufren algún tipo de discapacidad, normalmente relacionadas con el aprendizaje y problemas visuales y auditivos [2]. Según la OMS, más de una tercera parte de los bebés prematuros pueden evitar sufrir las complicaciones citadas anteriormente con una atención sencilla y costoeficaz, basada en servicios sanitarios esenciales durante el parto y periodo postnatal.

Las complicaciones citadas anteriormente llevan asociado un coste económico para la sociedad. El parto prematuro provoca un mayor número de hospitalizaciones, si a ello se le suma que los tiempos de ingreso son mayores, todo ello se traduce en un coste más de 5 veces superior frente a casos de parto a término. A todo ello hay que añadir las posteriores visitas a las distintas especialidades necesarias durante los primeros 10 años de vida. Estos hechos han sido probados por estudios sobre el impacto económico del parto prematuro en EEUU [3].

1.2 Técnicas para la detección trabajo de parto

Se han estudiado diversos factores y su viabilidad para ser usados como herramientas para la predicción del comienzo del parto. Se citan a continuación los más importantes:

- Longitud cervical: medición mediante técnicas de ultrasonidos de la longitud cervical.
- Índice Bishop: también conocido como test de Bishop, mediante una puntuación de 0 a 13 valora el cuello uterino en el trabajo de parto[4].
- Múltiples síntomas del parto (MPTLS): se recogen diversos síntomas y signos, como presencia de contracciones y dilataciones, para realizar una estimación subjetiva a cada paciente.
- Fibronectina fetal: la fibronectina fetal es una proteína considerada el adhesivo que une el saco fetal con el revestimiento del útero[5]. Mediante una prueba de detección se determina la concentración de dicha proteína.
- LIF cervical: mediante técnicas de fluorescencia se determina la composición del cuello del útero y se establece el riesgo de parto prematuro en base a los niveles de colágeno. Es de carácter no invasivo [8].
- TOCO: se emplea un transductor de presión que registra los cambios de presión debido a las contracciones uterinas.

Las técnicas mencionadas anteriormente tienen una baja capacidad para detectar el trabajo de parto prematuro. Este trabajo se centra en la señal obtenida mediante la Electromiografía uterina (EMG), que consiste en registrar la actividad eléctrica sobre el abdomen o en el propio útero. Dicha actividad es generada por las diferencias de potencial originadas por el intercambio iónico entre las membranas celulares[6], [7].

1.3 EHG

El electrohisterograma (EHG) es un método empleado para la monitorización de la dinámica uterina mediante el registro de la actividad eléctrica, dicha señal tiene el aspecto que se muestra en la Figura 1. Este método registra una señal asociada a la propagación de potenciales de acción por parte del miometrio de una mujer embarazada[9]. Esta técnica es de carácter no invasivo, por tanto, lo único que precisa es la limpieza previa de la superficie abdominal dónde se situarán los electrodos para el registro de dicha señal eléctrica, como se aprecia en la Figura 2.

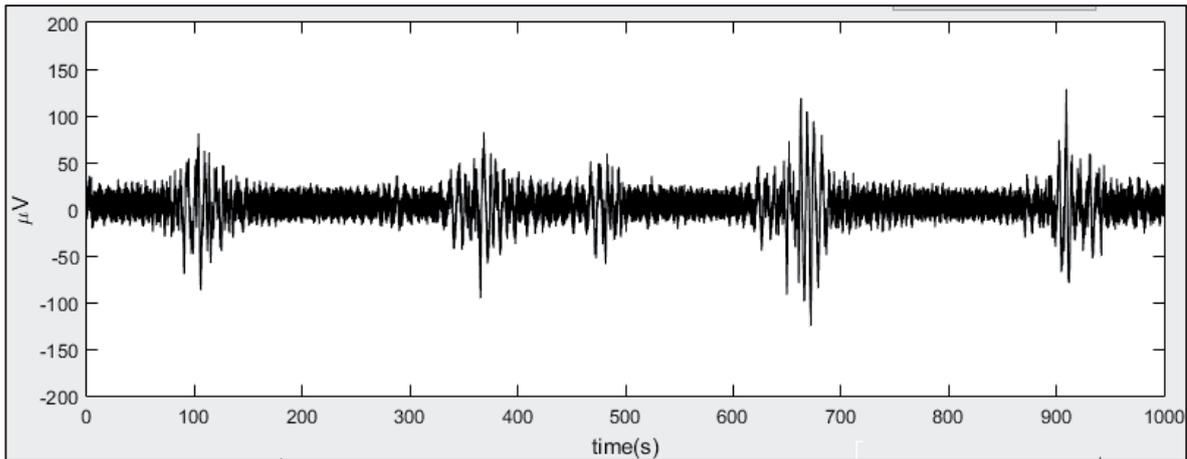


Figura 1. Ejemplo señal EHG para un ancho de banda de 0.1 a 4 Hz.

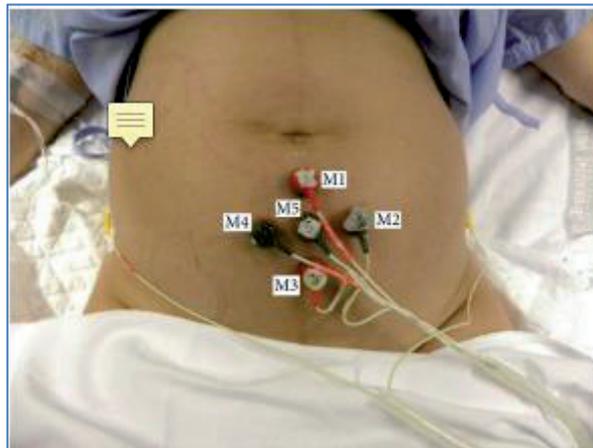


Figura 2. Configuración de los electrodos para la obtención de registros[10].

Estudios previos determinaron que la energía de la señal EHG se distribuye en un rango de frecuencia entre 0.1 y 4 Hz[11]. La señal EHG está compuesta por dos tipos de onda: la onda lenta (Slow wave) y onda rápida (Fast wave). Las ondas rápidas representan el disparo de las señales eléctricas individuales, mientras que las ondas lentas corresponden a las contracciones mecánicas resultantes[12]. En la Figura 3 se muestra un ejemplo de las componentes de la señal EHG junto con la señal de presión intrauterina.

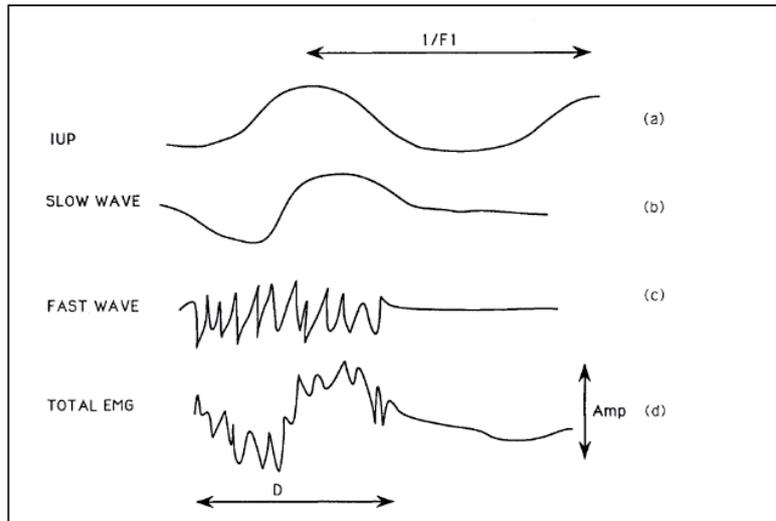


Figura 3. Descripción teórica del electromiograma uterino (EMG). IUP, presión intrauterina[13].

Se dice que la onda lenta no tiene significado fisiológico debido a que su ancho de banda se solapa con las fluctuaciones debidas a estiramientos de la piel, movimientos abdominales y de los electrodos, lo cual hace difícil su interpretación[14]. Por esta razón los estudios se centran en el análisis de la onda rápida. A su vez, la onda rápida tiene dos componentes: fast wave low (FWL) y fast wave high (FWH) [15]. La FWL está relacionada con la propagación de la actividad eléctrica y la FWH a la excitabilidad del útero[16].

En lo que respecta al análisis de la señal EHG, la gran mayoría de los estudios han optado por centrarse en el registro bipolar de dicha señal. Este registro se obtiene mediante el empleo de electrodos monopolares dispuestos sobre la superficie abdominal [15], de esta manera se obtienen los registros que van a ser tratados en el presente trabajo.

2. JUSTIFICACIÓN Y OBJETIVOS DEL TRABAJO

Como se ha expuesto en los apartados anteriores, el parto prematuro constituye un problema tanto económico como social que además carece de herramientas específicas para ser diagnosticado.

Este trabajo ha sido realizado dentro del grupo de investigación del Centro de Investigación e Innovación en Bioingeniería (Ci2b). Dicho grupo de investigación acumula varios años de experiencia trabajando en análisis de EHG para el diagnóstico del parto prematuro con la colaboración del hospital La Fe de Valencia.

El análisis de la señal EHG comprende desde la obtención de parámetros hasta su posterior análisis estadístico para la obtención de modelos eficaces que permitan diagnosticar con antelación el parto prematuro. De todo ello surge la necesidad de crear una herramienta que integre todo el proceso en una interfaz de usuario, siendo la finalidad de este trabajo la creación de la misma. Por tanto, el objetivo de esta interfaz es poner al alcance de cualquier usuario sin conocimientos de lenguaje de programación la posibilidad de analizar las señales EHG. La herramienta debe proporcionar al usuario un entorno gráfico versátil que facilite la tarea del análisis e interpretación de resultados. Las necesidades que debe cubrir la herramienta que se va a desarrollar son:

- Filtrado de la señal EHG en diferentes anchos de banda.
- Visualización de la señal filtrada en distintos anchos de banda junto con la segmentación (registro completo o contracción) asociada en caso de que se haya definido previamente. Si no se ha definido, debe ofrecer la posibilidad de crear una segmentación nueva. La herramienta desarrollada debe permitir que registro o conjunto de registro se desea visualizar y parametrizar.
- Realizar un análisis de parámetros lineales, no lineales y espectrales de la señal. El usuario debe poder definir los parámetros que desea obtener en dicho análisis.
- Visualizar la distribución de cada parámetro en la población analizada, que en este caso se trata de parto a término y parto pretérmino.
- Permitir realizar un estudio de la distribución de los datos y un estudio estadístico en función de los grupos de clasificación definidos, que deben ser parto a término y pretérmino, $TTD < 7$ vs $TTD > 7$, $TTD < 14$ vs $TTD > 14$ y $TTD < 30$ vs $TTD > 30$ días.
- Elaboración de modelos predictivos mediante regresión logística o redes neuronales según los grupos mencionados anteriormente.

3. DESARROLLO ENTORNO GRÁFICO GUIDE

3.1 Descripción de la herramienta

El desarrollo de la interfaz gráfica con las características descritas en apartados anteriores se ha realizado mediante el software matemático Matlab. Dicho software ofrece un entorno de desarrollo con lenguaje de programación propio. El presente trabajo se ha desarrollado en este entorno debido a que el grupo de investigación ha desarrollado sus herramientas de análisis en él.

Puesto que la herramienta a diseñar consiste en un entorno gráfico que implemente la programación necesaria en una interfaz gráfica, para ello se ha empleado la herramienta de desarrollo GUIDE de Matlab (ver Figura 4). A continuación se describe brevemente el funcionamiento del entorno de desarrollo de la interfaz con el fin de facilitar la comprensión al lector.

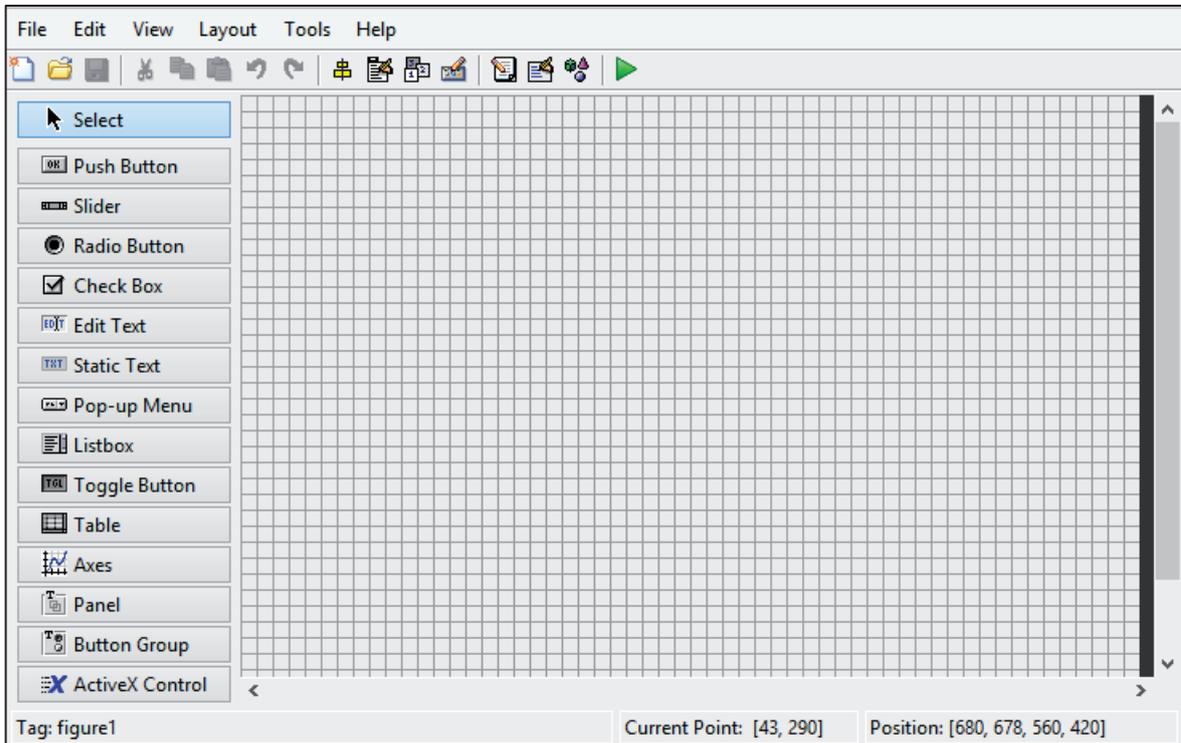


Figura 4. Entorno de desarrollo GUIDE de Matlab

Para el desarrollo de la interfaz se han empleado algunos de los elementos que se pueden apreciar en el panel izquierdo de la Figura 4. Dichos elementos permiten la creación de botones, desplegables, listados de cadenas de caracteres y la representación de gráficas, entre otras funciones. Este entorno de programación genera para ventana de la interfaz un archivo de extensión FIG y otro de extensión M con el código de programación asociado a dicha ventana.

3.2 Especificaciones de diseño

Para el correcto desarrollo de la aplicación, se deben establecer una serie de especificaciones de diseño. Estas especificaciones se han basado en los objetivos explicados en el apartado 2.

Estas especificaciones se enuncian a continuación:

- La pantalla principal debe tener habilitada una opción de búsqueda de paciente.
- La visualización debe ser en el eje de abscisas segundos y en el eje de ordenadas la amplitud del parámetro. Se tiene que poder regular los límites de visualización de los ejes de manera manual e intuitiva.
- Para seleccionar cada ancho de banda que se desea visualizar debe emplearse un desplegable con las distintas opciones disponibles.
- La ventana para modificar la segmentación debe permitir al usuario cargar segmentaciones existentes, que en este caso pueden ser tanto de registro completo como de contracción. Además, debe permitir modificarlas o crear nuevas en función de si ya existían en la base de datos.
- La ventana de análisis debe permitir seleccionar cada parámetro mediante *checkbox* y seleccionar mediante un desplegable si se desea analizar un paciente en particular o la base de datos completa.
- La ventana gráfica de selección del análisis estadístico debe poseer botones específicos para seleccionar cada grupo de clasificación de entre los disponibles y permitir seleccionar los parámetros que se desean analizar de manera intuitiva.
- El informe estadístico debe mostrar una tabla con las medianas y los rangos intercuartílicos para cada población dependiendo del grupo de clasificación escogido, junto con su p-valor.
- La interfaz debe ofrecer la posibilidad de seleccionar entre los distintos tipos de algoritmos de clasificación entre los grupos de clasificación junto con la posibilidad de introducir manualmente los porcentajes de registro para entrenamiento, validación y test. Con todo ello, se deben representar los resultados del sistema predictor desarrollado mediante la representación gráfica de la curva ROC y de la matriz de confusión.

Puesto que la interfaz gráfica ha sido diseñada en el entorno matemático Matlab, se detalla a continuación los requisitos mínimos que debe tener el ordenador para poder soportar la ejecución del entorno.

Sistema operativo

Windows	Windows 7 SP1 en adelante, Windows Server 2008 SP2
Mac	macOS 10.10 - 10.11
Linux	Kernel 2.6 o superior, glibc 2.11 o superior.
Procesador	Intel o AMD x86-64 con soporte de instrucciones AVX2
Disco	2 GB solo para MATLAB, 4-6 GB para una instalación típica.
RAM	1GB mínimo, 4 GB recomendado.
Tarjeta gráfica	Soporte para Open GL 3.3 recomendado con 1 GB en GPU.

Tabla 1. Especificaciones técnicas necesarias para la ejecución del entorno matemático.

En la página siguiente, se muestra un flujograma que describe el proceso lógico de utilización del programa y sus posibilidades.

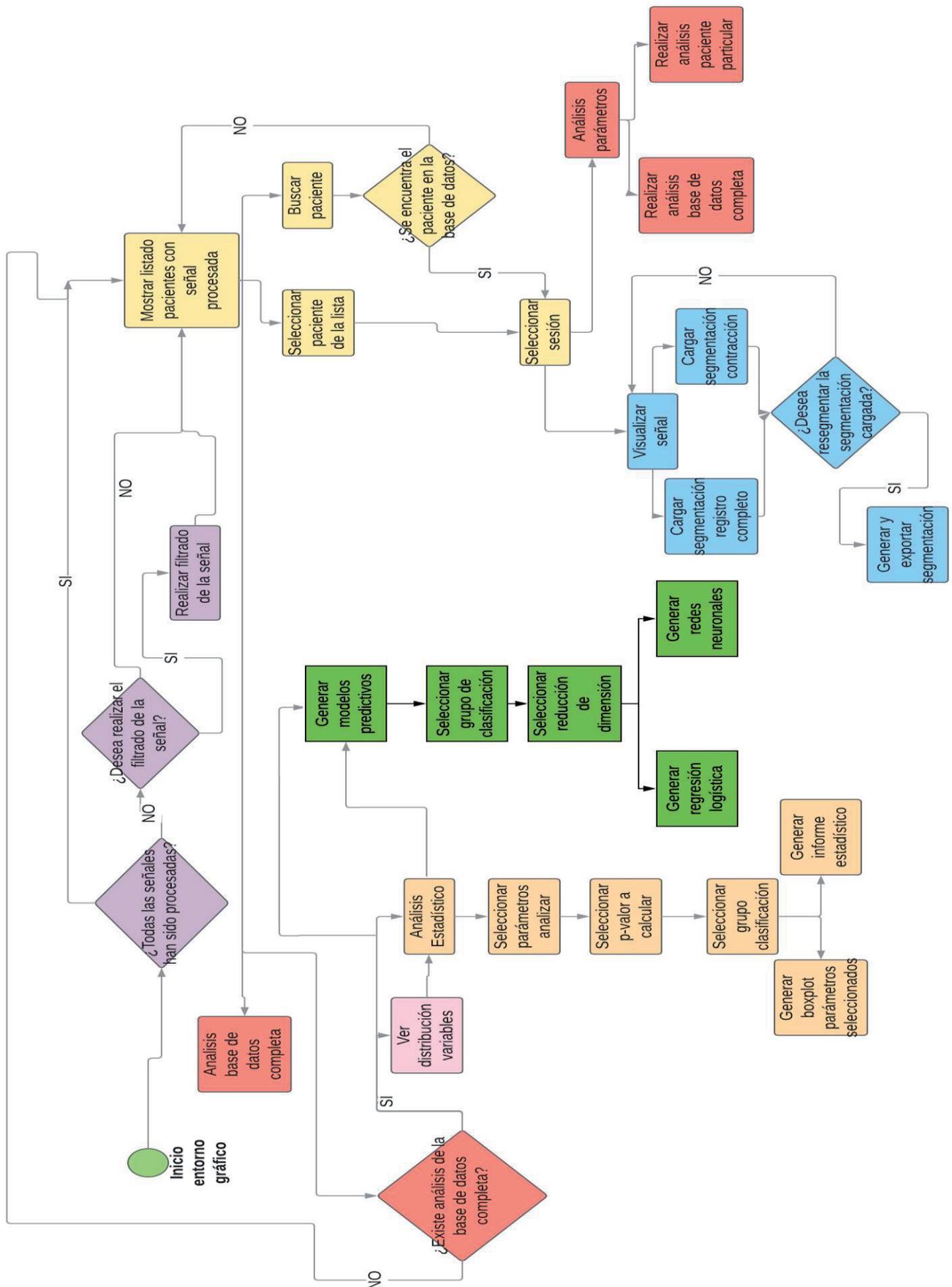


Figura 5. Flujo de trabajo explicativo del proceso de razonamiento de la interfaz

3.3 Descripción funciones entorno gráfico

La finalidad de este apartado es proporcionar al lector idea básica de las posibilidades que ofrece cada ventana gráfica de la interfaz, así como una breve contextualización teórica de lo que se está realizando. Para una descripción más detallada de la utilización de la interfaz se encuentra el documento adjunto Manual de usuario.

3.3.1 Pantalla principal

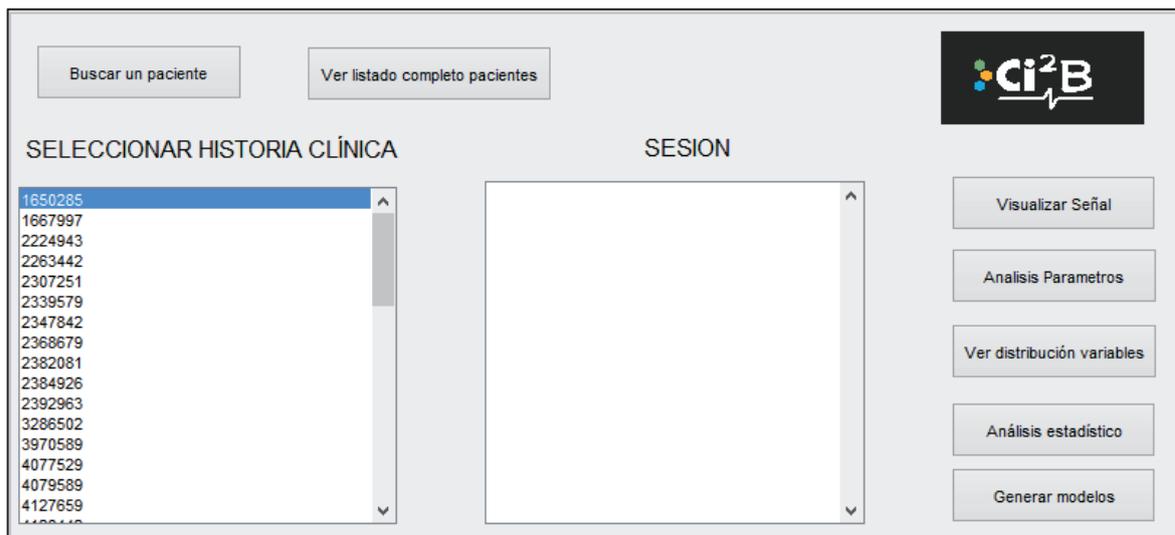


Figura 6. Pantalla principal.

En la ventana mostrada en la Figura 6 se muestra la pantalla principal del programa. Esta ventana permite la selección de los pacientes contenidos en la base de datos y la búsqueda de un paciente específico. A partir de la ventana mostrada se permite al usuario acceder a las diferentes funcionalidades del programa, como son la visualización de la señal, el análisis de parámetros, la visualización de la distribución de las variables, el análisis estadístico y la generación de modelos predictivos. En la Figura 7 se observan las funciones que conforman dicha ventana gráfica, cuyo funcionamiento se detalla a continuación.

Este equipo ▸ Documentos ▸ TFG ▸ programa ▸ pantalla_principal

Nombre	Fecha de modifica...	Tipo	Tamaño
 comprobar_existe_filtrado.m	02/06/2018 16:32	Archivo M	1 KB
 distinguir_guardia.m	02/06/2018 18:16	Archivo M	1 KB
 leer_directorio.m	07/06/2018 17:04	Archivo M	2 KB
 Montaje_Guardia_GUI.m	06/06/2018 10:17	Archivo M	3 KB
 Montaje_GUI.m	03/06/2018 11:13	Archivo M	4 KB
 Montar_Guardar_Registros_GUI.m	06/06/2018 11:44	Archivo M	3 KB
 pantalla_principal.fig	19/06/2018 9:48	Archivo FIG	26 KB
 pantalla_principal.m	19/06/2018 9:48	Archivo M	12 KB

Figura 7. Directorio funciones pantalla principal

- Pantalla_principal.fig: contiene la distribución gráfica de la pantalla principal.
- Pantalla_principal.m: contiene el código completo asociado a la ventana pantalla principal. Esta función es la principal de esta ventana y es la encargada de llamar al resto de funciones.
- Comprobar_existe_filtrado.m: esta función comprueba al iniciar la pantalla principal, si existe algún registro que no haya sido filtrado.
- Leer_directorio: función encargada de leer los directorios correspondientes y obtener las cadenas con sus nombres.
- Montaje_Guardia_GUI.m: función que forma parte de las necesarias para el montaje de los registros en caso de que no hayan sido filtrados previamente. Esta función es la encargada de realizar esta operación cuando se trata de un registro de guardia.
- Montaje_GUI: función que forma parte de las necesarias para el montaje de registros, como la anterior. Esta función monta los registros que se han realizado en planta.
- Montar_Guardar_Registros_GUI.m: función principal para el montaje de registros. Se encarga de distinguir si se trata de un registro en planta o de guardia y de llamar a las dos funciones anteriores dependiendo de qué caso se trate.

3.3.2 Visualización

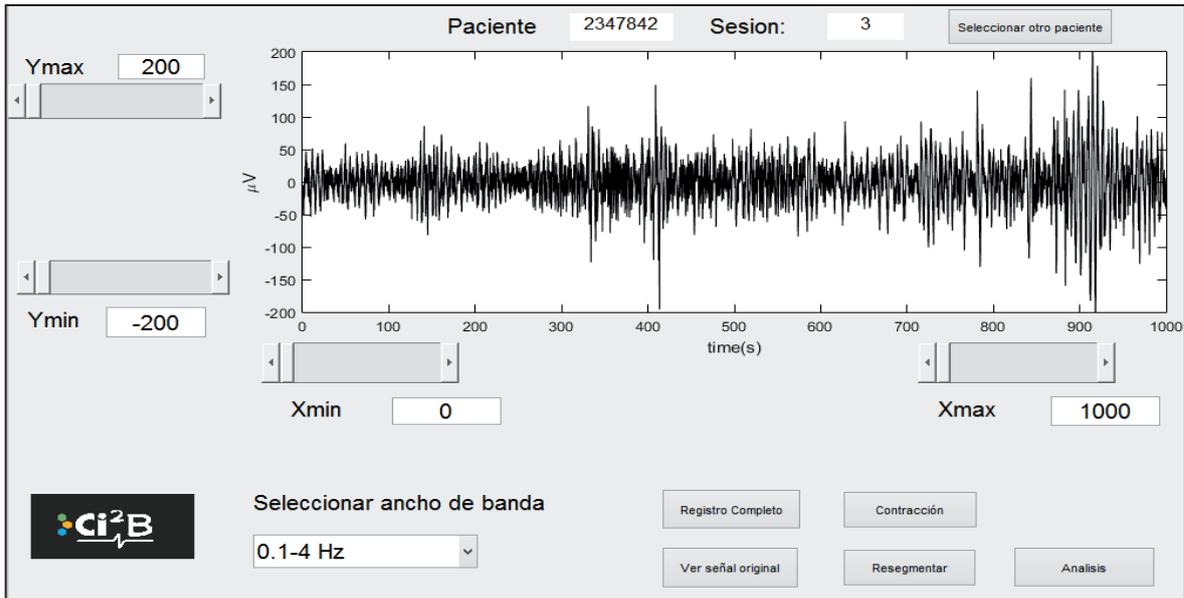


Figura 8. Pantalla de visualización

La ventana mostrada en la Figura 8 permite al usuario visualizar la señal procesada en distintos anchos de banda para el paciente seleccionado en la pantalla principal. Se dispone de las herramientas necesarias para el ajuste de los ejes de visualización, visualización tanto de registro completo como de contracción, continuar al análisis de parámetros y un botón de acceso a la ventana resegmentar cuyo funcionamiento se explicará en el apartado siguiente. En la Figura 9 se observan las funciones que conforman dicha ventana gráfica, cuyo funcionamiento se detalla a continuación.

Nombre	Fecha de modifica...	Tipo	Tamaño
leer_CT.m	07/06/2018 17:13	Archivo M	1 KB
leer_RC.m	07/06/2018 17:13	Archivo M	1 KB
representar anchos_banda.m	09/06/2018 11:04	Archivo M	1 KB
visualizacion.fig	19/06/2018 10:34	Archivo FIG	44 KB
visualizacion.m	19/06/2018 10:34	Archivo M	20 KB

Figura 9. .Directorio funciones venta visualización

- Visualizacion.fig: contiene la distribución gráfica de la ventana de visualización.
- Visualizacion.m: contiene el código completo asociado a la ventana gráfica de análisis. Esta función es la principal de esta ventana y es la encargada de llamar al resto de funciones.
- Leer_CT.m: función encargada de leer la segmentación de contracción para su visualización en el entorno gráfico.

- Leer_RC.m: función encargada de leer la segmentación de registro completo para su visualización en el entorno gráfico.
- Representar anchos_banda.m: esta función se encarga de representar la señal en el entorno gráfico en función del ancho de banda que seleccione el usuario.

3.3.3 Resegmentar

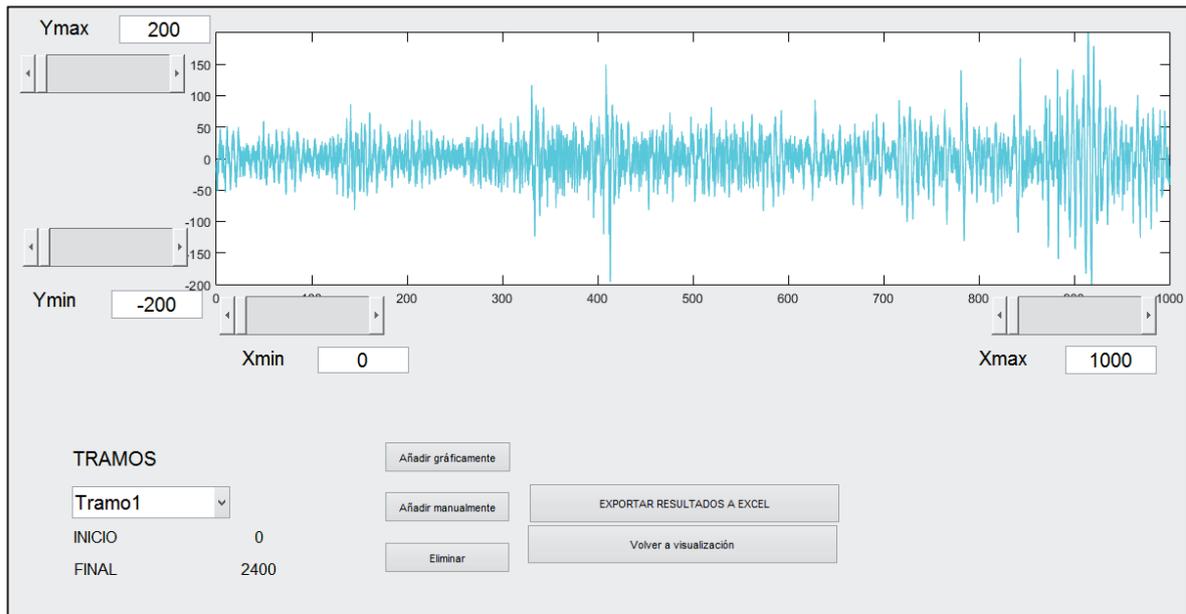


Figura 10. Pantalla resegmentar.

La ventana mostrada en la Figura 10 permite al usuario modificar o generar una nueva segmentación para el paciente seleccionado en la pantalla principal. Permite modificar tanto la segmentación de registro completo como la de contracción, para ello el usuario debe haber seleccionado en la pantalla de visualización la segmentación que desea modificar en esta ventana.

El sistema de ajuste de los ejes de visualización es idéntico al de la pantalla de visualización. Mediante el panel mostrado en la parte inferior de la Figura 10 se permite al usuario la creación, selección y eliminación de tramos de la segmentación. Finalmente, para almacenar la segmentación creada o modificada es imprescindible presionar el botón “Exportar resultados a Excel”. En la Figura 11 se observan las funciones que conforman dicha ventana gráfica, cuyo funcionamiento se detalla a continuación.

Este equipo ▸ Documentos ▸ TFG ▸ programa ▸ resegmentar

Nombre	Fecha de modifica...	Tipo	Tamaño
 almacenar_segmentacion_generada.m	09/06/2018 13:04	Archivo M	2 KB
 buscar_tramos_iguales.m	27/05/2018 13:00	Archivo M	1 KB
 generar_tramos.m	13/05/2018 14:02	Archivo M	1 KB
 resegmentar.fig	27/05/2018 11:51	Archivo FIG	31 KB
 resegmentar.m	09/06/2018 12:43	Archivo M	19 KB

Figura 11. Directorio funciones ventana resegmentar.

- Resegmentar.fig: contiene la distribución gráfica de la ventana de resegmentar.
- Resegmentar.m: contiene el código completo asociado a la ventana gráfica de resegmentar.
- Almacenar_segmentacióngenerada.m: función que realiza la exportación de la segmentación al fichero Excel correspondiente. Para ello se debe presionar el botón “Exportar a Excel” que se aprecia en la Figura 10.
- Buscar_tramos_iguales.m: en caso de que el usuario defina un tramo en la segmentación que ya existía, esta función lo detecta e informa del error al usuario.
- Generar_tramos: función que genera los tramos introducidos por el usuario, almacenando su valor y representándolos sobre la gráfica de la Figura 10. Estos tramos se generan cuando se añade tanto de forma manual como gráfica en la ventana correspondiente.

3.3.4 Análisis de señal EHG

Espectrales

- Frecuencia Media
- Frecuencia Mediana
- DF1
- DF2
- NE1
- NE2
- NE3
- H_L Ratio
- Dec1
- Dec2
- Dec3
- Dec4
- Dec5
- Dec6
- Dec7
- Dec8
- Dec9
- Spectral Moment Ratio

Parámetros Lineales

- Amplitud pico a pico
- RMS
- AUC
- Combinado 1 (EHG_LC)
- Combinado 3 (EHG_LC)
- Teager
- Primer Momento
- Curtosis
- Varianza Coef. Wavelet

Parámetros No Lineales

- Lempel-Ziv Binario
- Lempel-Ziv Multiestado (6)
- Sample Entropy
- Fuzzy entropy
- Spectral Entropy
- Time Reversibility
- PonCaire

Rangos No Lineales

- 0.1-4 Hz
- 0.2-0.34 Hz
- 0.34-4 Hz
- 0.34-1 Hz
- 1-4 Hz

Paciente: 2347842
Sesion: 3
Analizar el paciente
 Seleccionar todos los parametros

Ejecutar Analisis
Visualizar señal paciente
Seleccionar otro paciente

Figura 12. Ventana de análisis.

La ventana mostrada en la Figura 12 se emplea para realizar el análisis de los parámetros de la señal que el usuario desea computar. Los parámetros disponibles se agrupan en lineales y no lineales, siendo los lineales el conjunto formado por los parámetros temporales y espectrales. Para la selección de dichos parámetros la ventana gráfica cuenta con una serie de *checkbox* asociados al parámetro que se indica a su derecha, siendo necesario en el caso de los no lineales especificar el ancho de banda en el que se desean calcular.

Según el tipo de registro se va a distinguir entre dos tipos de análisis:

1. *Análisis de registro de contracción*: en este análisis la segmentación está compuesta por los tramos correspondientes a las contracciones visibles en la señal, como se aprecia en la Figura 13. Para el cálculo de parámetros, se computa el cálculo de todos los parámetros para cada contracción y su resultado final es el valor mediana de todas las contracciones para cada parámetro (ver Tabla 2).

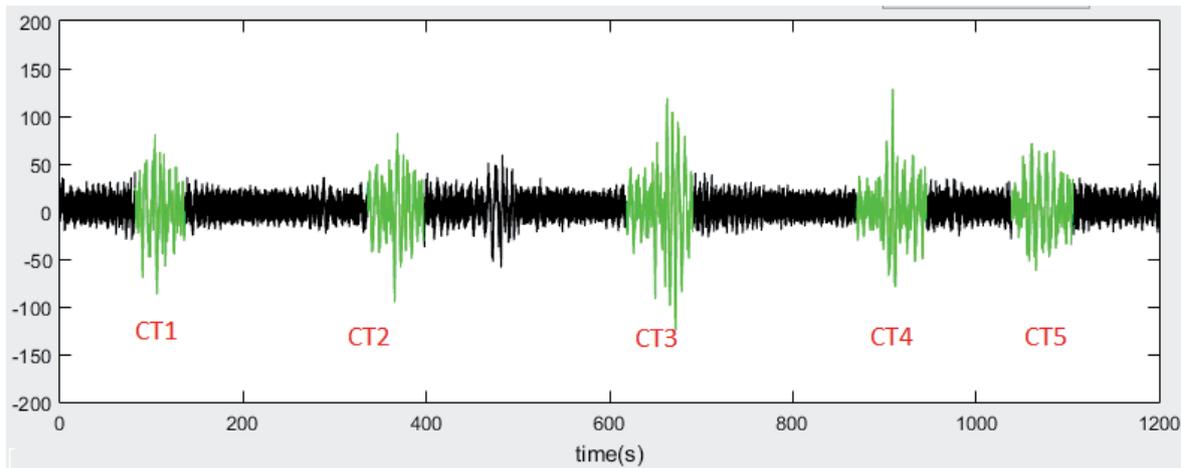


Figura 13. Registro EHG de 1200 segundos en el que se pueden identificar cinco contracciones, resaltadas en verde.

	Parámetro 1	Parámetro 2	Parámetro N
CT1	Valor_CT1_P1	Valor_CT1_P2	Valor_CT1_PN
CT2	Valor_CT2_P1	Valor_CT2_P2	Valor_CT2_PN
CT3	Valor_CT3_P1	Valor_CT3_P2	Valor_CT3_PN
CT4	Valor_CT4_P1	Valor_CT4_P2	Valor_CT4_PN
CT5	Valor_CT5_P1	Valor_CT5_P2	Valor_CT5_PN
Resultado	Mediana(Valor_CTi_P1)	Mediana(Valor_CTi_P2)	Mediana(Valor_CTi_PN)

Tabla 2. Esquema de cálculo de parámetros para la señal EHG mostrada en la Figura 13.

2. Análisis de registro completo: para llevar a cabo el análisis de la señal de registro completo se emplea un sistema distinto al expuesto anteriormente. En este caso, se define una ventana móvil que se desplazará sobre el registro de EHG. Los registros a analizar se dividen en ventanas de 120 segundos con un solapamiento del 50%, es decir, 60 segundos como se muestra en la Figura 14. Para el cálculo de los parámetros temporales y no lineales, se computan dichos parámetros para cada ventana y el resultado se obtiene mediante los valores promedio de los parámetros obtenidos en cada ventana de análisis. Por otra parte, para el cálculo de los parámetros espectrales se obtiene la PSD para cada una de las ventanas y finalmente se toma la media de ésta de todas las ventanas para el cálculo de parámetros espectrales, dicho proceso se resume en la Figura 15.

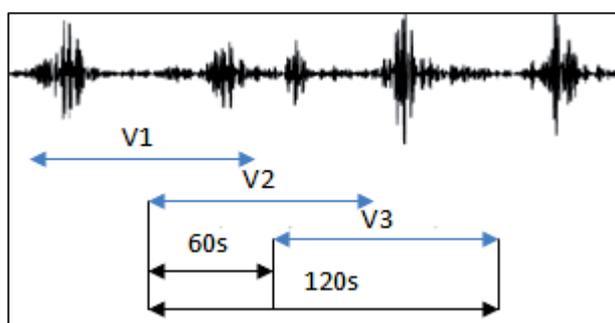


Figura 14. Esquema de la división de la señal en ventanas para el análisis de registro completo.

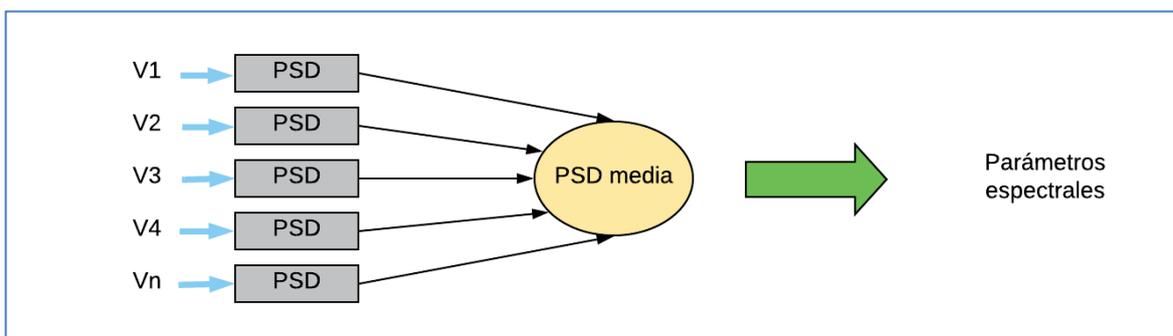


Figura 15. Esquema de cálculo de los parámetros espectrales.

Esta ventana permite tanto el análisis de la base de datos completa como el de un paciente específico. Para elegir una opción u otra se cuenta con el desplegable que se aprecia en la Figura 12. En la Figura 16 se observan las funciones que conforman dicha ventana gráfica, cuyo funcionamiento se detalla a continuación.

Este equipo ▸ Documentos ▸ TFG ▸ programa ▸ analisis

Nombre	Fecha de modifica...	Tipo	Tamaño
 analisis.fig	11/06/2018 12:07	Archivo FIG	46 KB
 analisis.m	14/06/2018 18:10	Archivo M	29 KB
 Calcula_Parametros_selectivo.m	15/05/2018 11:46	Archivo M	5 KB
 escribir_excel_resultados.m	12/06/2018 13:11	Archivo M	4 KB
 escribir_excel_resultados_particular.m	12/06/2018 13:15	Archivo M	4 KB
 espectrales.m	16/05/2018 10:52	Archivo M	5 KB
 gestion_parametros.m	09/06/2018 13:28	Archivo M	2 KB
 lineales_and_nlineales.m	17/05/2018 18:15	Archivo M	3 KB
 quitar_ceros.m	21/05/2018 16:53	Archivo M	1 KB

Figura 16. Directorio funciones ventana análisis.

- Analisis.fig: contiene la distribución gráfica de la ventana análisis.
- Analisis.m: contiene el código completo asociado a la ventana gráfica de análisis. Esta función es la principal de esta ventana y es la encargada de llamar al resto de funciones.
- Calcula_Parametros_selectivo.m: esta función se encuentra dentro de las necesarias para el cálculo de parámetros lineales. Su finalidad es recibir los parámetros que ha seleccionado el usuario en la ventana de la Figura 12 y calcular estos únicamente.
- Escribir_excel_resultados.m: función que implementa la exportación de los parámetros calculados al fichero Excel correspondiente cuando se selecciona el análisis de la base de datos completa. Dicha función crea el fichero Excel y le da el formato en caso de que no exista. Su uso se activa cuando se presiona el botón “Ejecutar análisis”.
- Escribir_excel_resultados_particular.m: tiene la misma función que la anterior, se emplea cuando se selecciona el análisis de un paciente en particular.
- Espectrales.m: función empleada para el cálculo de los parámetros espectrales, solamente calcula los que haya seleccionado el usuario.
- Gestión_parametros.m: función que devuelve el resultado de todos los parámetros calculados. Gestiona los parámetros seleccionados por el usuario y si por alguna razón el paciente no es analizable, devuelve un indicador para que no sea tenido en cuenta.
- Lineales_and_nlineales.m: función empleada para el cálculo de los parámetros temporales y no lineales, únicamente los que haya seleccionado el usuario en la ventana de análisis.
- Quitar_ceros.m: elimina las filas nulas de la matriz de análisis, normalmente estas filas nulas se deben a que hay pacientes no analizables por tratarse de registros de baja calidad y/o artefactados.

3.3.5 Análisis estadístico

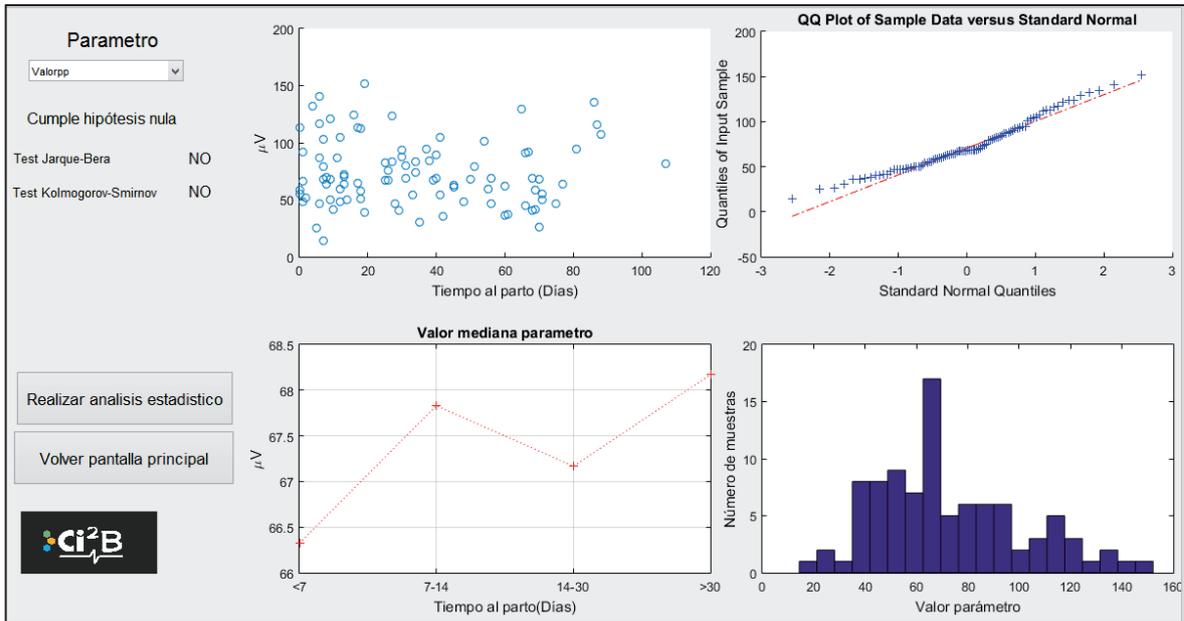


Figura 17. Ventana distribución variables.

La ventana mostrada en la Figura 17 forma parte del conjunto de ventanas destinadas para la realización del análisis estadístico. Esta ventana permite al usuario la visualización de la distribución del parámetro mediante los cuatro gráficos mostrados. Los gráficos empleados son los siguientes:

- **Gráfico 1:** Conocido como diagrama de dispersión, se emplea para descubrir y mostrar la relación entre dos conjuntos de datos. En este caso podría emplearse para identificar visualmente si existe alguna tendencia de la amplitud de la señal EHG conforme se acerca al parto. En el eje de abscisas se muestra el tiempo al parto en días y en el eje de ordenadas, el valor del parámetro en las unidades que se expresan en el gráfico.
- **Gráfico 2:** Este tipo de gráficos están implementados en la función de Matlab *qqplot* y se emplea para determinar si la distribución del parámetro seleccionado por el usuario es normal, en caso de que la distribución sea normal el parámetro se aprecia visualmente lineal en la representación.
- **Gráfico 3:** Este gráfico nos muestra los valores mediana de los parámetros para los grupos de clasificación definidos en el eje de abscisas, en función del tiempo al parto en días. En el eje de ordenadas se muestra la magnitud del parámetro en sus unidades correspondientes.
- **Gráfico 4:** se trata de un histograma que representa para cada parámetro el número de muestras que presentan un valor determinado.

En la Figura 18 se observan las funciones que conforman dicha ventana gráfica, cuyo funcionamiento se detalla a continuación.

Este equipo ▸ Documentos ▸ TFG ▸ programa ▸ distribucion_variables

Nombre	Fecha de modifica...	Tipo	Tamaño
deteccion_correccion_outliers.m	11/06/2018 17:56	Archivo M	1 KB
distribucion_variables.fig	11/06/2018 10:29	Archivo FIG	69 KB
distribucion_variables.m	11/06/2018 18:05	Archivo M	8 KB
grafica_mediana.m	11/06/2018 17:29	Archivo M	1 KB
grafica_qqplot.m	18/05/2018 13:00	Archivo M	1 KB
grafica_scatter_plot.m	04/06/2018 18:27	Archivo M	1 KB
isoutlier.m	11/06/2018 17:43	Archivo M	1 KB
parametros_disponibles.m	09/06/2018 10:54	Archivo M	2 KB
test_hipotesis.m	22/05/2018 11:29	Archivo M	1 KB

Figura 18. Directorio funciones distribución de variables

- Distribucion_variables.fig: contiene la distribución gráfica de la ventana distribución de variables.
- Distribucion_variables.m: contiene el código completo asociado a la ventana gráfica de distribución de variables. Esta función es la principal de esta ventana y es la encargada de llamar al resto de funciones.
- Detección_corrección_outliers.m: función encargada de gestionar la eliminación de los outliers detectados por la función Isoutlier.
- Grafica_mediana.m: representa la gráfica con el valor mediana de los datos en función del tiempo al parto.
- Grafica_qqplot.m: representa el diagrama Q-Q de los datos para el parámetro seleccionado.
- Grafica_scatter_plot.m: representa el diagrama de dispersión de los datos para el parámetro seleccionado.
- Isoutlier.m: detecta los outliers para que la función deteccion_correccion_outliers los pueda eliminar.

La ventana mostrada en la Figura 19 permite al usuario realizar un análisis estadístico de los datos en función del grupo de clasificación seleccionado. Dichos grupos de clasificación son cuatro y se dividen en término y pretérmino, TTD<7 vs TTD>7, TTD<14 vs TTD>14 y TTD<30 vs TTD>30. Para la selección de cada grupo de clasificación se habilitan los correspondientes botones que pueden verse en la parte inferior derecha de la Figura 19.

Figura 19. Ventana habilitada para realizar el análisis estadístico de los datos.

El usuario tiene la opción de seleccionar los parámetros que desea analizar mediante el listado mostrado en la parte superior izquierda de la Figura 19. Mediante el empleo de los botones “Añadir”, “Eliminar” o “Añadir todos” se seleccionan los parámetros deseados para continuar con el análisis. Por último, cabe mencionar que la ventana mostrada permite calcular tanto el pvalor de Wilcoxon como Tstudent, en los casos que proceda su cálculo.

Una vez realizado el análisis, a modo de informe se genera la ventana mostrada en el siguiente apartado. Dicho informe puede exportarse a un fichero Excel para su almacenamiento, como se comentará en el apartado siguiente.

En la Figura 20 se observan las funciones que conforman dicha ventana gráfica, cuyo funcionamiento se detalla a continuación.

Este equipo ▸ Documentos ▸ TFG ▸ programa ▸ analisis_estadistico

Nombre	Fecha de modifica...	Tipo	Tamaño
 almacenar_pvalor.m	03/06/2018 18:08	Archivo M	1 KB
 analisis_estadistico.fig	11/06/2018 12:07	Archivo FIG	113 KB
 analisis_estadistico.m	18/06/2018 12:22	Archivo M	13 KB
 Boxplot_Test.m	31/05/2018 11:29	Archivo M	2 KB
 generar_cadenas_pvalor.m	22/05/2018 12:14	Archivo M	1 KB
 term_preterm.m	31/05/2018 12:01	Archivo M	1 KB
 tiempo_parto.m	31/05/2018 12:02	Archivo M	1 KB

Figura 20. Directorio funciones ventana análisis estadístico

- Analisis_estadistico.fig: contiene la distribución gráfica de la ventana análisis estadístico.
- Analisis_estadistico.m: contiene el código completo asociado a la ventana gráfica de análisis estadístico. Esta función es la principal de esta ventana y es la encargada de llamar al resto de funciones.
- Almacenar_pvalor.m: función encargada de almacenar en un fichero los pvalor calculados para la distribución de los parámetros.
- Boxplot_Test.m: función que integra los comandos necesarios para la representación de los diagramas de Caja-Bigotes de los parámetros seleccionados por el usuario.
- Generar_cadenas_pvalor: función que genera las cadenas con los pvalor que acompañarán a cada parámetro en su diagrama Caja-Bigotes.
- Term_preterm.m: función que lleva a cabo el análisis estadístico a término y pretérmino al pulsar sobre el botón correspondiente.
- Tiempo_parto.m: función que lleva a cabo el análisis estadístico en función del tiempo al parto al pulsar sobre el botón correspondiente en la ventana.

3.3.6 Informe estadístico

Informe estadístico						Exportar a Excel
	A término		Pretérmino			
	Media	Rango Intercuartílico	Media	Rango Intercuartílico	P-valor	
Time Reversibility(0.1-4 Hz)	2.8223	1.8188	3.653	1.713	0.0023026	^
EHG_LC3	85596.0677	70148.245	117394.9681	74716.8918	0.0094206	
EHG_LC1	11502.3307	8729.2705	13917.3221	9482.4688	0.093015	
Ziv1(0.1-4 Hz)	0.44291	0.11559	0.41893	0.11504	0.098401	
DF2	0.36082	0.021973	0.36441	0.019531	0.12779	
AUC	182682.9714	166188.6304	230919.8569	212287.9418	0.16052	
RMS	8.6802	4.9289	9.4551	5.6078	0.22621	
PON3	17.3271	11.7833	19.2553	10.7453	0.23803	
PON2	24.2477	16.496	26.9875	15.2179	0.25028	
Valorpp	70.2529	40.034	76.243	36.6855	0.28964	
NE2	0.28765	0.07609	0.30283	0.081534	0.30721	
Dec6	0.33132	0.026855	0.33463	0.026855	0.32276	
PON4	0.13672	0.056531	0.13153	0.050887	0.32547	
PON1	3.0559	1.8357	3.4062	2.4245	0.33675	
Fuzzy entropy(0.1-4 Hz)	0.31257	0.11005	0.30078	0.069111	0.39904	
Spectral Moment Ratio	3.4788	0.56155	3.3499	0.86525	0.40969	
Dec2	0.23432	0.013428	0.2366	0.01709	0.43682	
Dec5	0.30416	0.029297	0.30592	0.029297	0.44654	
Primer momento	6499569.3723	6497275.1438	7404461.1622	5084605.9658	0.45371	
Dec8	0.429	0.054932	0.42465	0.061035	0.48399	
Curtosis	2.2056	0.35421	2.1864	0.32306	0.49114	v

Figura 21. Ventana informe estadístico

En la ventana mostrada en la Figura 21 se muestra el informe generado por la ventana gráfica anterior. Los parámetros mostrados son los que han sido seleccionados en la ventana anterior y en el informe se ordenan de menor a mayor en función de su p-valor. Para el grupo de clasificación seleccionado, muestra de cada población su mediana y su rango intercuartílico.

Si una vez generado el informe el usuario desea almacenarlo, la ventana gráfica mostrada posee la opción de exportar dicho informe a un fichero Excel presionando sobre el botón “Exportar a Excel”. En caso de que exista un informe almacenado previamente, el programa solicita al usuario la confirmación de que desea sobrescribir los datos.

Este equipo > Documentos > TFG > programa > informe_estadístico

Nombre	Fecha de modifica...	Tipo	Tamaño
almacenar_informe_estadistico.m	18/06/2018 13:04	Archivo M	2 KB
generar_tabla_informe_estadistico.m	18/06/2018 11:48	Archivo M	1 KB
informe_estadistico.fig	15/06/2018 11:08	Archivo FIG	7 KB
informe_estadistico.m	18/06/2018 12:33	Archivo M	5 KB
ordenar_pvalor.m	18/06/2018 11:27	Archivo M	2 KB

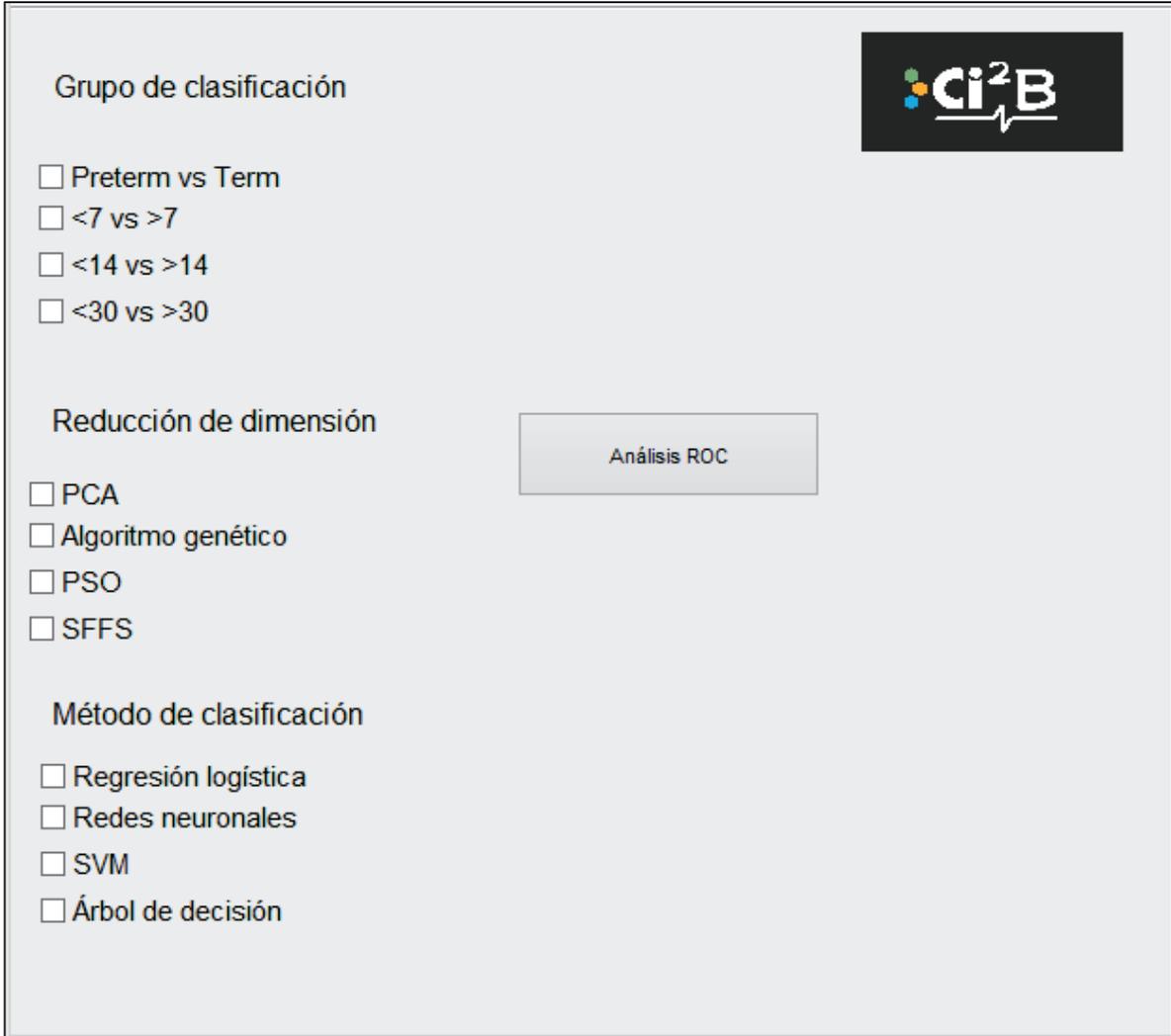
Figura 22. Directorio funciones ventana informe estadístico.

- Informe_estadístico.fig: contiene la distribución gráfica de la ventana análisis estadístico.
- Informe_estadístico.m: contiene el código completo asociado a la ventana gráfica de informe estadístico. Esta función es la principal de esta ventana y es la encargada de llamar al resto de funciones.



- Almacenar_informe_estadístico.m: función que exporta el análisis estadístico generado a un fichero Excel, si el usuario lo desea. Se emplea cuando se presiona el botón “Exportar a Excel” de la Figura 21.
- Generar_tabla_informe_estadístico.m: da el formato adecuado a los datos calculados para mostrar la tabla correspondiente.
- Ordenar_pvalor.m: puesto que los resultados se muestran mediante el pvalor ordenado de menor a mayor, esta función es la encargada de establecer dicho orden mediante el método de la burbuja.

3.3.7 Sistema predictor



Grupo de clasificación

- Preterm vs Term
- <7 vs >7
- <14 vs >14
- <30 vs >30

Reducción de dimensión

- PCA
- Algoritmo genético
- PSO
- SFFS

Método de clasificación

- Regresión logística
- Redes neuronales
- SVM
- Árbol de decisión

Análisis ROC

Figura 23. Ventana de generación de modelos predictivos

La ventana que se muestra en la Figura 23 permite al usuario generar los modelos predictivos deseados en función del grupo de clasificación, su reducción de dimensión y el método de clasificación.

Para la generación de los modelos predictivos deseados se ha empleado la técnica de validación cruzada para el diseño y validación del modelo. Mediante el panel mostrado en la Figura 24 se asignan los valores de los porcentajes de los conjuntos de entrenamiento, validación y test. El procedimiento se repite 30 veces para minimizar el sesgo de la selección aleatoria de las muestras. Para determinar la capacidad de generalización del modelo, se reserva el porcentaje de datos del grupo de test para probar su rendimiento.

Método de clasificación	Especificar porcentaje		
	Entrenamiento	Validación	test
<input type="checkbox"/> Regresión logística			
<input checked="" type="checkbox"/> Redes neuronales	30	30	
<input type="checkbox"/> SVM			
<input type="checkbox"/> Árbol de decisión			
	Topología neuronas		
	1		

Figura 24. Panel mostrado para la introducción de los porcentajes de datos destinados al conjunto de entrenamiento, validación y test.

Los grupos de clasificación disponibles son los mismos que en apartados anteriores, a término y pretérmino y $TTD < 7/14/30$ vs $TTD > 7/14/30$ días. La generación del sistema predictor consiste en recibir una serie de características, en este caso los parámetros de la señal EHG, llevar a cabo una reducción de dimensión de los datos y finalmente obtener un método de clasificación que nos permita distinguir si una muestra pertenece a la categoría A o B, dependiendo de la categoría escogida en el grupo de clasificación. Tanto los métodos de reducción de dimensión como los modelos predictivos se explican a continuación en los siguientes apartados, el diagrama mostrado en la Figura 25 resume el proceso de generación de sistemas predictores.

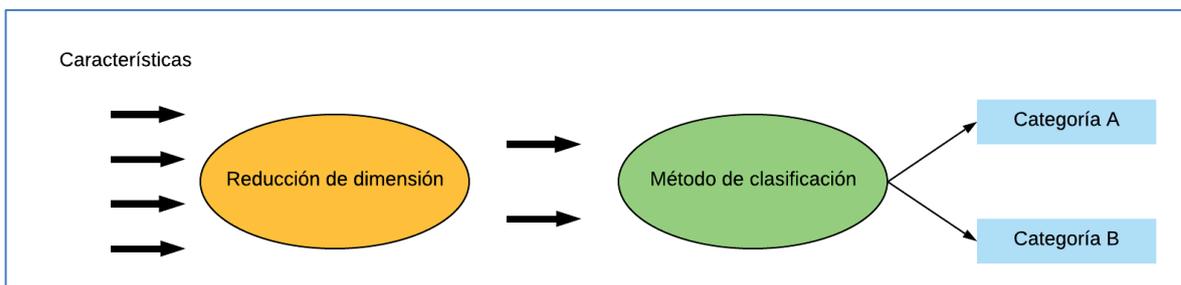


Figura 25. Esquema proceso de generación de sistemas predictores

3.3.7.1 Reducción de dimensión

Para la construcción de modelos predictivos es recomendable realizar una reducción de dimensión para evitar el sobreentrenamiento. La interfaz diseñada ha sido pensada para poseer cuatro tipos de métodos de reducción de dimensionalidad, como son el algoritmo PCA, el algoritmo genético, la optimización por enjambre de partículas (PSO) y el método de la selección secuencial flotante (SFFS).

El algoritmo PCA (Principal Component Analysis) se emplea para describir un conjunto de datos en términos de nuevas variables no correlacionadas. En este caso, para la ordenación de los datos se ha empleado un umbral de selección de un 98% para la salida del PCA. Como resultado del método PCA se obtiene una matriz no correlacionada.

Otro método de reducción de dimensión es el algoritmo genético, cuyo nombre está inspirado en la evolución biológica y su base genético-molecular. Estos algoritmos funcionan sobre un conjunto de soluciones llamado fenotipo y el conjunto de individuos de la población, codificando normalmente de forma binaria la información en cada cadena, que recibe el nombre de cromosoma. Dichos cromosomas evolucionan en cada iteración del algoritmo, estas iteraciones reciben el nombre de generaciones.

Estos algoritmos son aplicables a cualquier tipo de función pero son particularmente eficaces a la hora de calcular funciones no derivables. En cualquier caso, debe tenerse en cuenta una serie de consideraciones a la hora de aplicar los algoritmos genéticos:

- En caso de que la función cuente con un gran número de máximos y mínimos locales, se requerirán más iteraciones para garantizar la obtención del máximo o mínimo global.
- Si la función posee varios valores cercanos al valor óptimo, el algoritmo solamente garantiza que encontrará uno de ellos, no necesariamente el óptimo.

El funcionamiento básico de este tipo de algoritmos se explica a continuación:

- Inicialización: se genera de manera aleatoria la población inicial. Tal población está compuesta por un conjunto de cromosomas que representan las posibles soluciones al problema. Es importante que la población inicial tenga diversidad estructural para ser lo más representativa posible.
- Evaluación: a cada cromosoma de la población inicial se le aplica la función de aptitud que determina lo óptima que es la solución que codifica.
- Condición de término: el algoritmo genético debe detenerse cuando se alcance la solución óptima. Puesto que dicha solución se desconoce, se establecen unos criterios de detención del algoritmo. Normalmente o bien se establece un número máximo de iteraciones o se detiene cuando no haya cambios en la población. Mientras que no se produzca la detención, el algoritmo selecciona los cromosomas que serán cruzados en la siguiente operación, los recombina y muta los cromosomas para alcanzar soluciones que no estaban cubiertas por la población inicial. Una vez aplicado el algoritmo, se seleccionan los individuos más óptimos que conformarán la población de la siguiente generación.

En cuanto a las desventajas de este tipo de algoritmos cabe mencionar que cuando el problema a tratar es de gran complejidad, es posible que implique un elevado coste de tiempo y recursos. En algunos casos dependiendo de los parámetros empleados, la solución del algoritmo puede converger a soluciones que no sean óptimas.

El algoritmo **PSO**, también conocido como algoritmo de optimización por nube de partículas, posee eficacia probada para problemas de optimización. El funcionamiento de este algoritmo consiste en la definición de una población de soluciones candidatas, a las que denomina partículas. Estas partículas describen su movimiento a través del espacio acorde a diferentes reglas matemáticas sencillas. El movimiento de cada partícula está condicionado por su mejor posición y por la mejor posición global encontrada en todo el espacio de búsqueda. En el momento que se determinan nuevas mejores posiciones, éstas pasan a orientar el movimiento de las partículas. Este proceso se repite hasta encontrar una solución lo suficientemente satisfactoria.

En la PSO, es determinante llevar a cabo una elección de parámetros correcta para favorecer un buen rendimiento del algoritmo. De manera gráfica, puede entenderse que el problema a resolver genera una hipersuperficie de dimensionalidad equivalente al número de parámetros. La irregularidad de la hipersuperficie dependerá del grado de complejidad del problema.

Uno de los problemas de este tipo de algoritmos es la convergencia prematura, ya que puede incurrir fácilmente en óptimos locales.

Por último se encuentra el algoritmo **SFFS**, también conocido como algoritmo de selección secuencial flotante. Su funcionamiento consiste en definir un conjunto de variables y un criterio para maximizar. En cada iteración se añaden nuevas características que maximicen el criterio definido y se elimina la característica que más reduce el criterio. Este proceso se repite mientras el valor del criterio definido no decrezca, y el algoritmo termina cuando la adición de nuevas características no mejora el criterio definido.

3.3.7.2 Modelos predictivos

La interfaz se ha desarrollado para soportar la implementación de los cuatro métodos de clasificación que pueden observarse tanto en la Figura 23 y Figura 24.

La **regresión logística** es una particularización del modelo lineal generalizado, mediante la aplicación de una función no lineal a una regresión múltiple [17]. En este caso se define una variable de tipo binario para diferenciar de si se trata de un parto a término o pretérmino.

$$p_n = b_0 + b_1 * \lambda_{1n} + b_2 * \lambda_{2n} + \dots + b_p * \lambda_{pn}$$

Ecuación 1

La transformación que hay que aplicar a p_n es la siguiente:

$$\ln\left(\frac{y_n}{1-y_n}\right)$$

Ecuación 2

y_n representa la probabilidad de que un suceso ocurra, mientras que $1-y_n$ representa la probabilidad de que dicho suceso no ocurra.

Realizando la siguiente transformación sobre la Ecuación 1

$$\ln\left(\frac{y_n}{1-y_n}\right) \rightarrow = b_0 + b_1 * X_{1n} + b_2 * X_{2n} + \dots + b_p * X_{pn} = \ln\left(\frac{y_n}{1-y_n}\right)$$

Ecuación 3

Finalmente se obtiene que la formulación equivalente del modelo de regresión logística es:

$$Y_n = \frac{1}{1 + e^{-[b_0 + b_1 * X_{1n} + b_2 * X_{2n} + \dots + b_p * X_{pn}]}}$$

Ecuación 4

Por otra parte, la generación de modelos predictivos mediante **redes neuronales** consiste en crear un modelo computacional que se basa en un conjunto de unidades neuronales artificiales simples que se comporta de forma análoga a los axones de neuronas de cerebros orgánicos. Las neuronas se encuentran conectadas entre sí y sus enlaces pueden provocar la activación o inhibición de las neuronas adyacentes. Estos sistemas no se programan de forma explícita, ya que se entrenan y se forman a sí mismos.

La estructura de capas empleada en este trabajo es la siguiente:

- Capa de entradas: capa que recibe la información del exterior, en este caso los parámetros calculados para cada paciente.
- Capas ocultas: se trata de unas capas intermedias mediante las cuales se realiza el trabajo de la red.
- Capa de salidas: proporciona el resultado del trabajo al exterior.

Mediante la Figura 26 se ilustra la distribución de capas y las relaciones que presentan. En el ejemplo siguiente se diferencia entre el grupo a término y pretérmino, pero la interfaz desarrollada permite además la diferenciación entre $TTD < 7/14/30$ vs. $TTD > 7/14/30$.

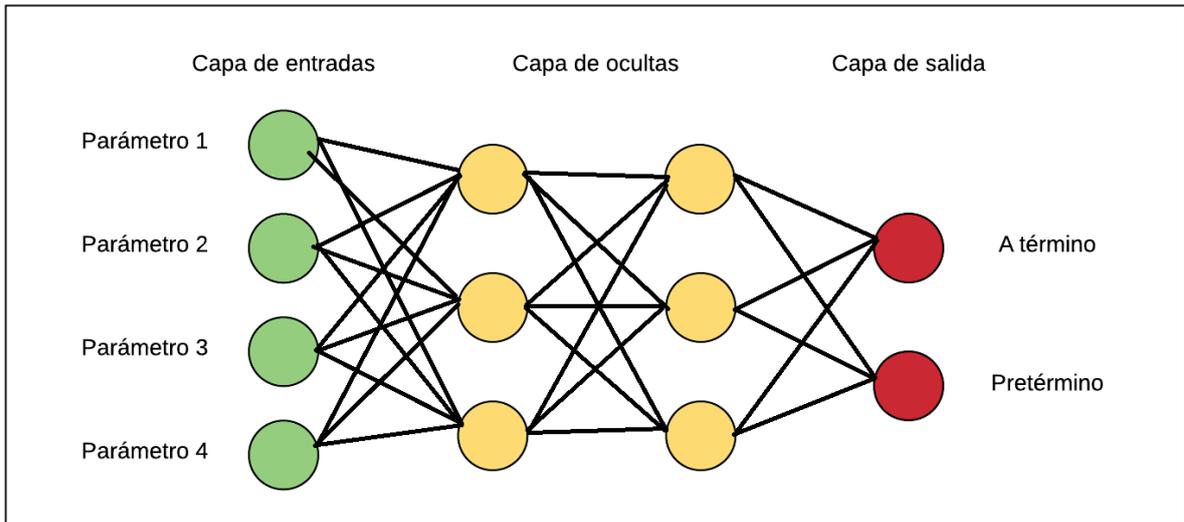


Figura 26. Esquema red neuronal antes del entrenamiento.

Las distintas neuronas que conforman la red son unidades de procesamiento de información, es decir, reciben la información de la neurona anterior, la procesan y emite el resultado a través de sus conexiones a la neurona siguiente en caso de que se supere el valor "umbral". El valor umbral determina la activación o no de la neurona, en caso de activarse la neurona emite un valor (1) y en caso contrario, emitirá otro valor (0).

Cuando las redes neuronales ya se han entrenado, las conexiones entre ellas llevan asociado un determinado peso sináptico, el cual define la importancia de la información que se transmite por dicha conexión. A modo ilustrativo, se muestra en la Figura 27 las conexiones entre neuronas con su peso sináptico W_i .

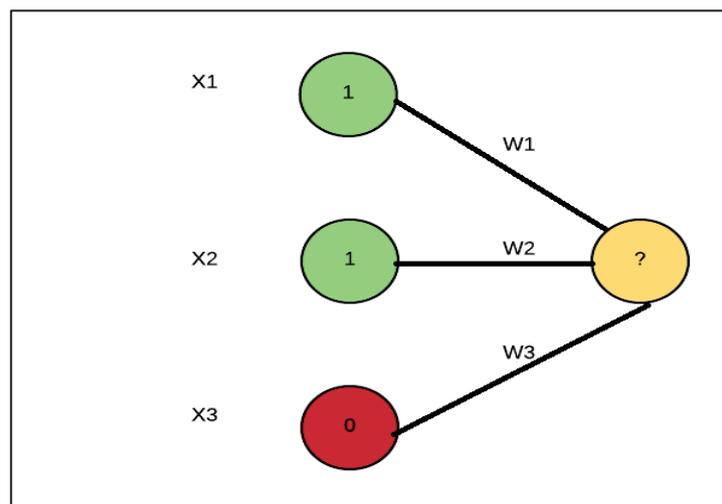


Figura 27. Conexiones con distinto peso sináptico que convergen en la misma neurona.

El procesamiento de la información por parte de cada neurona de la red consiste en una función (F) que trabaja con los valores obtenidos de las neuronas que la preceden. Dicha información suele ser 0 o 1, y para su procesamiento se tiene en cuenta el peso sináptico de la conexión por la que se ha recibido dicha información.

El entrenamiento de las redes consiste en la modificación del peso sináptico y el valor umbral de las conexiones con el fin de que el resultado obtenido coincida con el valor esperado o se aproxime.

Al final del proceso, las capas de salida generan un resultado que puede coincidir o no con el que se desea asociar el estímulo. Esta precisión dependerá de la información con la que se haya entrenado la red, dicha precisión es objeto de estudio de este trabajo como se expondrá en el apartado Resultados.

Por otra parte, otro de los métodos de clasificación para los que se ha pensado la interfaz es la **máquina de vectores de soporte**. Las máquinas de vectores de soporte son algoritmos de aprendizaje supervisado y pensados para la resolución de problemas de clasificación y regresión. Mediante un conjunto de entrenamiento se definen las clases y se entrena la máquina de vectores de soporte para construir un modelo predictivo que analice y determine la clase a la que pertenece una nueva muestra.

El funcionamiento se resume en que la SVM busca un hiperplano que diferencie de la forma más óptima los puntos de clases diferentes. La característica fundamental de estos algoritmos reside en el concepto de “separación óptima”, que consiste en buscar el hiperplano que tenga mayor distancia con los puntos más cercanos. De esta manera, los puntos del vector que pertenezcan a una categoría se encontrarán a un lado del hiperplano y los que pertenezcan a la otra categoría al otro.

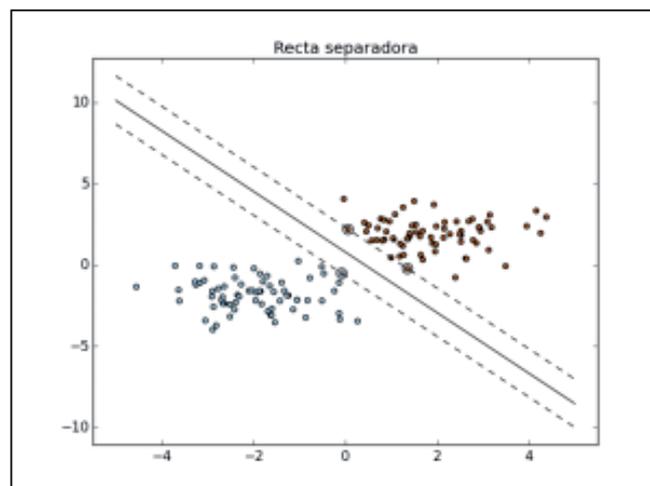


Figura 28. Ejemplo separación categorías en dos dimensiones.

Una de las principales ventajas de las SVM es que pueden ser empleadas para la resolución de problemas tanto lineales como no lineales.

Finalmente, el último método de clasificación para el que ha sido pensada la interfaz son los árboles de decisión. Los **árboles de decisión** son un modelo predictivo que puede utilizarse en el ámbito de la inteligencia artificial. Mediante un conjunto de datos dado se generan diagramas de construcciones lógicas que se utilizan para representar y categorizar unas condiciones que ocurren sucesivamente, para la resolución de un problema.

Para la construcción de un árbol de decisión, se dispone de tres elementos:

- Nodos: momento en el que se debe tomar una decisión entre varias posibles, a mayor número de nodos existirá mayor número de posibles finales.
- Vectores de números: contienen la solución para cada una de las diversas posibilidades que se tiene.
- Flechas: son las encargadas de unir los nodos, cabe mencionar que un nodo solo puede ser apuntado por una flecha, lo cual implica que hay un único camino para llegar del nodo inicial a cada uno de los nodos del árbol de decisión.

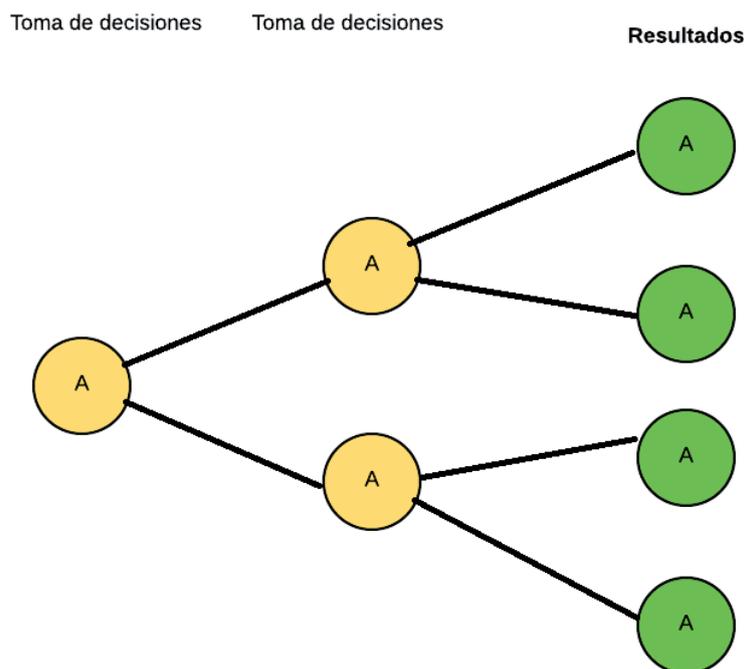


Figura 29. Esquema estructura árbol de decisión.

Los árboles de decisión en el diseño de aplicaciones informáticas consisten en una representación similar a la mostrada en la Figura 29 cuyas ramas se bifurcan dependiendo del valor tomado por las variables y que resultan en una acción concreta.

En la Figura 30 se observan las funciones que conforman dicha ventana gráfica, cuyo funcionamiento se detalla a continuación.

Nombre	Fecha de modifica...	Tipo	Tamaño
algoritmo.fig	15/06/2018 10:35	Archivo FIG	32 KB
algoritmo.m	15/06/2018 10:37	Archivo M	17 KB
analiza_redes_gui.m	29/05/2018 18:20	Archivo M	10 KB
directorio_particiones.m	24/05/2018 16:22	Archivo M	1 KB
entrena_redes_gui.m	04/06/2018 10:39	Archivo M	3 KB
Generar_Particiones_Auto_A.m	24/05/2018 11:03	Archivo M	2 KB
Generar_Particiones_Auto_B.m	14/06/2018 12:21	Archivo M	2 KB
Lanzar_PCA98_Particiones.m	27/03/2018 18:05	Archivo M	2 KB
Lanzar_PCA98_ParticionesB.m	29/03/2018 15:06	Archivo M	3 KB
regresion_logistica.m	14/06/2018 10:37	Archivo M	4 KB
seleccionar_var_clasificadora.m	17/06/2018 12:34	Archivo M	2 KB

Figura 30. Directorio funciones ventana algoritmo

- Algoritmo.fig: contiene la distribución gráfica de la ventana algoritmo.
- Algoritmo.m: contiene el código completo asociado a la ventana gráfica de algoritmo. Esta función es la principal de esta ventana y es la encargada de llamar al resto de funciones.
- Analiza_redes_gui.m: función encargada de analizar las redes generadas por el programa. Calcula lo necesario para la representación de la curva ROC.
- Directorio_particiones.m: función que gestiona el directorio dónde se guardan los modelos generados en función de lo que seleccione el usuario.
- Entrena_redes_gui.m: función que entrena las redes generadas por el programa.
- Generar_Particiones_Auto_A: genera las particiones necesarias, en este caso treinta, para el modelo de regresión logística.
- Genera_Particiones_Auto_B.m: función análoga a la anterior particularizada para redes neuronales.
- Lanzar_PCA98_Particiones.m: función encargada de la reducción de dimensión de los datos para regresión logística.
- Lanzar_PCA98_ParticionesB.m: función encargada de la reducción de dimensión de los datos para redes neuronales.
- Regresion_logistica.m: función encargada de realizar la regresión lineal de los datos. Devuelve los valores necesarios para la representación de la curva ROC.
- Seleccionar_var_clasificadora.m: función que gestiona la elección de la variable clasificadora por parte del usuario.

3.3.8 Evaluación del sistema predictor

Para llevar a cabo la evaluación del sistema predictor, se va a realizar la representación y análisis de la curva ROC. La curva ROC representa la sensibilidad frente a la especificidad de un sistema clasificador binario variando el umbral de discriminación. La representación de la curva ROC junto con los parámetros expuestos a continuación proporcionan al usuario las herramientas necesarias para valorar si se trata de un modelo válido o no.

Para evaluar la validez del modelo predictivo, se han definido los siguientes parámetros:

- Sensibilidad: es un indicador de la capacidad de nuestro estimador de determinar que un caso es positivo cuando realmente se trata de un parto pretérmino. En definitiva, se trata de una medida de la capacidad de la prueba de detectar pacientes que presentarán un parto pretérmino.
- Especificidad: al contrario que la sensibilidad, la especificidad es un indicador de la capacidad de determinar casos negativos cuando realmente lo son, es decir, que se va a producir el parto a término. Por tanto, se trata de una medida de la capacidad de la prueba de detectar pacientes que presentarán un parto a término.
- AUC: se trata del área bajo la Curva ROC y se emplea para evaluar la capacidad del modelo de discriminar si una muestra pertenece a una población u otra. Si la prueba de diagnóstico es perfecta, su valor será 1.
- Valor predictivo positivo (VPP): es un dato, normalmente en tanto por uno, que mide la probabilidad de que un caso sea positivo en caso de que la prueba diagnóstica sea positiva. Particularizando para este caso, la probabilidad de que se presente un parto pretérmino si la prueba es positiva.
- Valor predictivo negativo (VPN): al contrario que el valor predictivo positivo, este valor mide la probabilidad de ser un caso negativo en caso de que la prueba diagnóstica sea negativa.

Con ayuda de la Tabla 3 se definen los valores que van a componer lo que se conoce como la matriz de confusión. Cabe mencionar que el programa, al realizar varias particiones, a la hora de obtener los valores de la matriz de confusión realiza la media de los obtenidos en las distintas particiones. Esta media es orientativa para que el usuario interprete la validez del modelo generado.

<u>Tipos de diagnósticos</u>		<u>Parto pretérmino</u>	
		<u>Ausente</u>	<u>Presente</u>
<u>Prueba diagnóstica</u>	<u>Negativa</u>	Verdadero negativo (VN)	Falso negativo (FN)
	<u>Positiva</u>	Falso positivo (FP)	Verdadero positivo (VP)

Tabla 3. Elementos que componen la matriz de confusión de un modelo predictivo generado.

En la Figura 31 se observan las funciones que conforman la ventana gráfica de análisis ROC, cuyo funcionamiento se detalla a continuación.

Este equipo > Documentos > TFG > programa > curva ROC			
Nombre	Fecha de modifica...	Tipo	Tamaño
 crear_cadenas_matriz_confusion.m	28/05/2018 12:51	Archivo M	1 KB
 curva_troc.fig	11/06/2018 12:07	Archivo FIG	30 KB
 curva_troc.m	11/06/2018 12:07	Archivo M	7 KB

Figura 31. Directorio funciones curva ROC.

- Curva_troc.fig: contiene la distribución gráfica de la ventana análisis estadístico.
- Curva_troc.m: contiene el código completo asociado a la ventana gráfica de informe estadístico. Esta función es la principal de esta ventana y es la encargada de llamar al resto de funciones.
- Crear_cadenas_matriz_confusion.m: da el formato necesario para visualizar los datos en la ventana correspondiente.

4. RESULTADOS

4.1 Visualización señal

Como se ha comentado anteriormente, la señal EHG se registra mediante la colocación de unos electrodos sobre la superficie abdominal de la mujer embarazada y acondicionada con un filtro paso banda entre 0.1 y 150 Hz y adquirida con una frecuencia de muestreo de 500 Hz. No obstante, en la literatura se ha demostrado que la señal EHG distribuye su energía entre 0.1 y 4 Hz, por lo que se necesita realizar un preprocesado de la señal para la visualización y análisis de la misma. Los registros que se tratan en este trabajo son de dos tipos:

- **Registro en planta:** son los registros tomados a pacientes que se encuentran ingresados en planta. . Además de la señal EHG que se encuentra almacenada en la carpeta 'ehg', se ha adquirido simultáneamente el registro de TOCO y se almacena en la carpeta 'monitor_externo'. La estructura de almacenamiento de dichos registros se puede ver en la Figura 32.
- **Registro de guardia:** son los registros tomados a pacientes que acuden al servicio de urgencia por amenaza del parto prematuro, los cuales sólo contienen el registro de EHG. La estructura de almacenamiento de dichos registros se puede ver en la Figura 33.

Este equipo > Documentos > TFG > RegistrosAPP_Matlab-completos > 1552644 > F26_09_2016H9_07

Nombre	Fecha de modifica...	Tipo	Tamaño
Datas propio	01/06/2018 13:16	Carpeta de archivos	
ehg	01/06/2018 13:16	Carpeta de archivos	
monitor_externo	01/06/2018 13:16	Carpeta de archivos	
monitor_interno	01/06/2018 13:16	Carpeta de archivos	
existe	13/06/2018 11:09	Documento de tex...	0 KB
info	26/09/2016 9:07	Documento de tex...	1 KB

Figura 32. Formato de almacenamiento de un registro en planta en directorio codificado por fecha.

Este equipo > Documentos > TFG > RegistrosAPP_Matlab-completos > 6683399 > F_23_04_2017 > CONEX00

Nombre	Fecha de modifica...	Tipo	Tamaño
000	01/06/2018 13:18	Carpeta de archivos	
000_CONV	12/06/2018 18:36	Carpeta de archivos	

Figura 33. Formato de almacenamiento de un registro de guardia en directorio codificado por fecha y su conversión en el directorio CONEX.

El programa primero detecta automáticamente si existen las señales de EHG preprocesadas o no mediante la existencia de un fichero de texto llamado “existe.txt”. En caso de que no existiera, el programa permite realizar el preprocesado de la señal en la ventana principal y su resultado se almacena en la carpeta con el número de sesión correspondiente (ver Figura 34). De esta manera, se dispone de la señal preprocesada para futuras ejecuciones del programa.

Nombre	Fecha de modifica...	Tipo	Tamaño
1	29/05/2018 18:33	Carpeta de archivos	
F11_06_2015H8_20	12/06/2018 18:32	Carpeta de archivos	
MontadoMatlab	12/06/2018 18:32	Carpeta de archivos	

Figura 34. Creación de la carpeta de sesión 1 asociada a los datos crudos del directorio
F11_06_2015H8_20.

Para filtrar la señal se ha empleado un filtro anti-aliasing en cascada con un filtro pasa bajo, y otro pasa alto. En este trabajo se ha empleado el filtro Butterworth de orden 5 de fase cero. Los anchos de banda disponibles son: 0.1 a 4 Hz, 0.2 a 0.34 Hz, 0.34 a 4 Hz, 0.34 a 1 Hz, 1 a 4Hz, 4 a 30 Hz, aunque el último carece de interés. Puesto que la base de datos es amplia y para cada paciente y sesión se poseen 6 anchos de banda en los que visualizar la señal, a modo representativo vamos a visualizar la señal para los dos primeros anchos de banda (Ver Figura 35 y Figura 36) en un paciente de la base de datos. Así mismo, en la ventana mostrada en la Figura 35 se muestra en la parte superior un indicador del paciente y sesión seleccionada. En caso de querer visualizar un fragmento de señal con más detalle, el programa incorpora una función para el ajuste de ejes tanto manual como con el cursor.

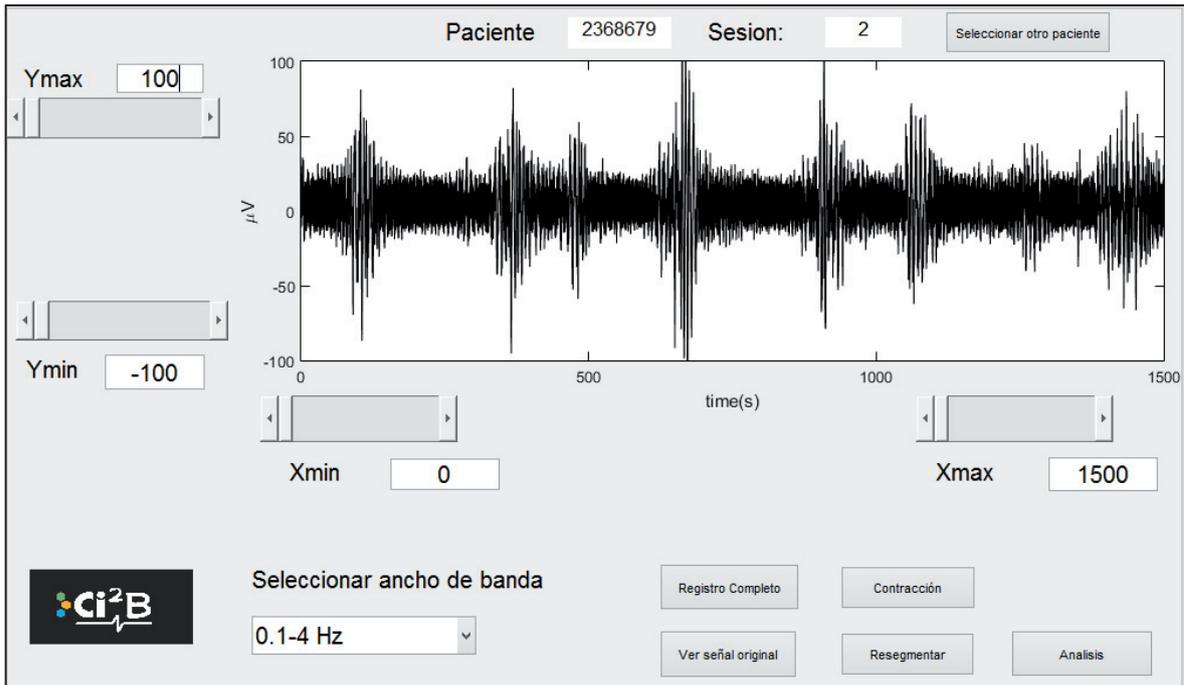


Figura 35. Señal obtenida para el paciente con Nº 2368679. Ancho de banda de 0.1 a 4 Hz

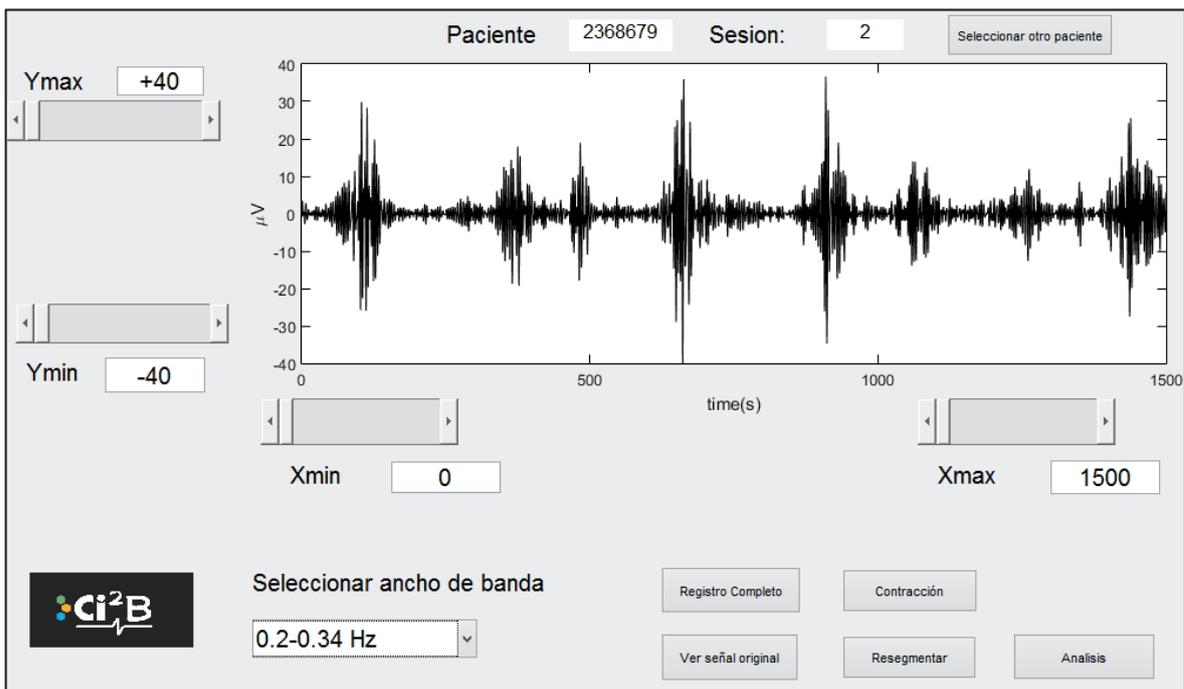


Figura 36. Señal obtenida para el paciente con Nº 2368679. Ancho de banda de 0.2 a 0.34 Hz

4.2 Visualización de la segmentación

Igual que otros registros electrofisiológicos, la señal de EHG podría estar embebida de una serie de interferencias fisiológicas como respiración, fluctuación de la línea basal, y artefactos de movimiento materno-fetal. Por lo que en general no se suele analizar todo el registro de EHG. En cuanto al análisis de la señal EHG, tradicionalmente los investigadores han recurrido el análisis de los EHG-bursts asociados a las contracciones uterinas, ignorando la información de la actividad basal cuando el útero se encuentra en reposo[18]. En cambio, estudios recientes indican que el análisis de registro completo de señal EHG que incluye no sólo las contracciones uterinas sino también la actividad basal podría utilizarse para la predicción del parto prematuro[19]. En este aspecto, la segmentación de la señal consiste en identificar los EHG-bursts asociadas a contracciones uterinas y descartar aquellos tramos de señal asociados a la respiración y/o artefactos de movimientos mediante la inspección visual en el análisis de contracciones y análisis de registro completo respectivamente.

En este caso, se cuenta con una base de datos creada por el Grupo de Investigación que ha definido previamente los tramos a analizar de algunos pacientes. Dicha segmentación asociada a cada paciente se encuentra en formato Excel, como se muestra en la siguiente figura:

	A	B	C	D	E
1		Registro completo			
2		Inicio	Fin	Duración(min)	
3	TR1	200	560	6.00	
4	TR2	600	1260	11.00	
5					

Figura 37. Ejemplo contenido archivo de segmentación de registro completo del paciente
1552644 sesión 1

Como se puede observar en la Figura 37, además de la segmentación almacenada en la pestaña Bipolar, se dispone de otra información adicional como son datos obstétricos del paciente y se utilizará en el análisis estadístico que se comentará más adelante.

Una de las funciones del entorno gráfico es cargar la segmentación, en caso de que exista, para permitir al usuario su visualización y si lo desea, su modificación.

En la siguiente figura ilustramos mediante un ejemplo de la señal EHG de la paciente 2263442 con su segmentación correspondiente, tanto para registro completo como para contracción.

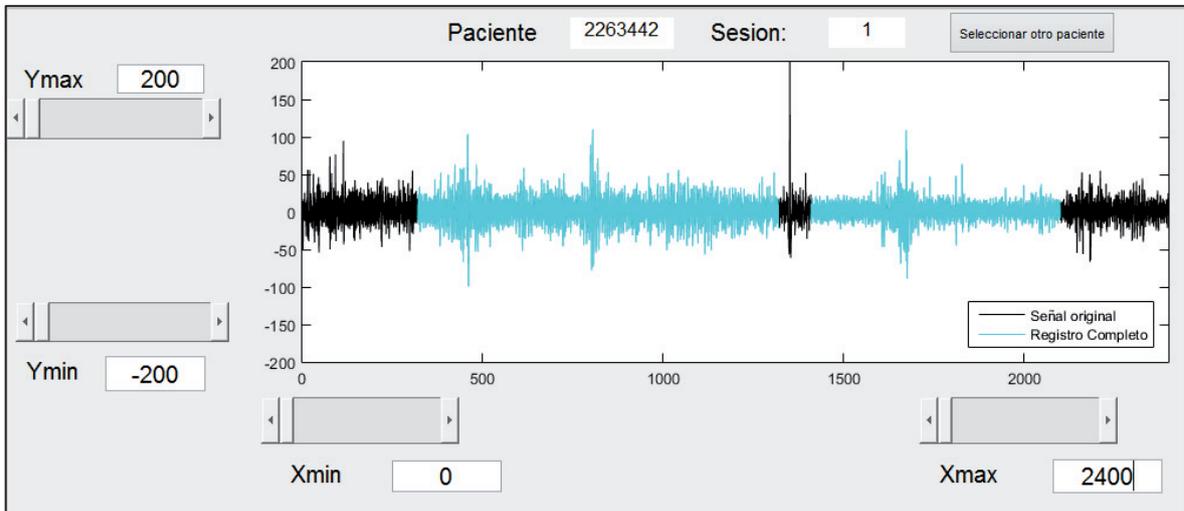


Figura 38. Ejemplo de la visualización de la segmentación de la señal EHG adquirida en la paciente 2263442 para el análisis de registro completo.

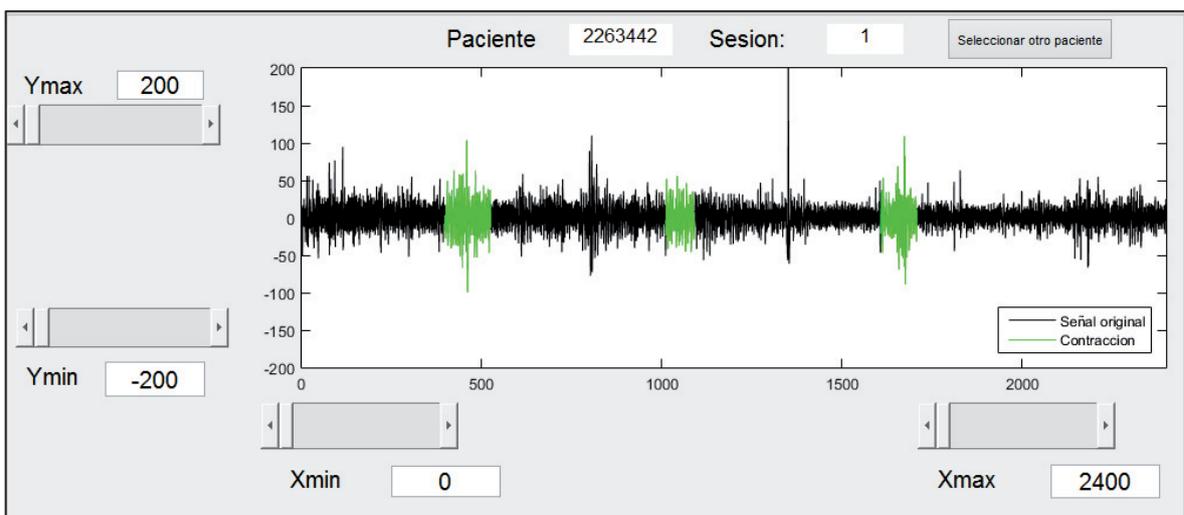


Figura 39. Ejemplo de la visualización de la segmentación de la señal EHG adquirida en la paciente 2263442 para el análisis de contracciones.

En ocasiones, puede que no exista segmentación previa al tratarse de un registro nuevo y el usuario considere oportuno realizar la segmentación de la señal o quiera modificar la segmentación ya existente bajo su criterio. En cualquier caso se puede modificar o generar la segmentación y su almacenamiento en la base de datos si el usuario lo desea en la siguiente ventana.

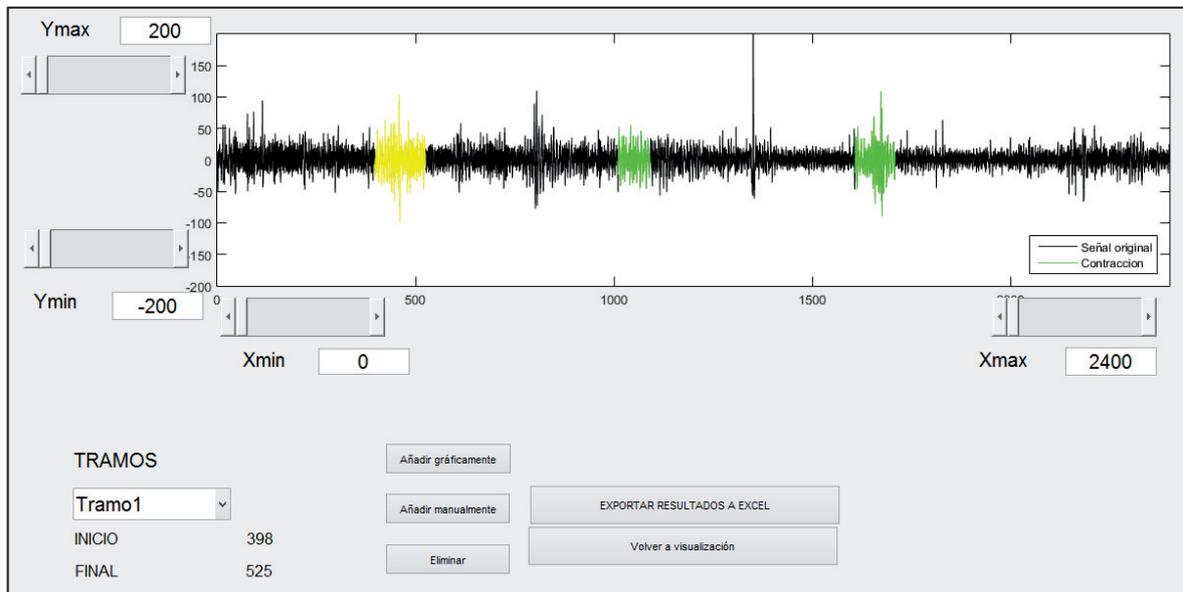


Figura 40. Ventana de modificación de la segmentación mostrada anteriormente. En amarillo el tramo seleccionado para su tratamiento en la interfaz.

Para la modificación de la segmentación, como puede verse en la Figura 40, puede realizarse la introducción de tramos de manera manual o gráficamente mediante el cursor o la eliminación de tramos mediante su selección en el desplegable y pulsando el botón “eliminar”. Para almacenar la segmentación generada se debe pulsar el botón “exportar resultados a Excel”. En caso de que exista segmentación previa, el programa pide al usuario la confirmación de que desea sobrescribir sobre la segmentación anterior. Finalmente, se obtiene un archivo Excel que mantiene el mismo formato que el resto de segmentaciones y cuyo título sirve como identificador para asociarlo al paciente correspondiente.

4.3 Análisis de los parámetros

Como se ha explicado en el apartado anterior, la segmentación asociada a cada paciente se emplea para realizar el análisis de la señal. Mediante dicho análisis se obtienen una serie de parámetros y se almacena junto a ellos una serie de datos obstétricos, que permiten su diferenciación a la hora de realizar un análisis estadístico.

El entorno gráfico permite realizar el análisis de toda la base de datos o de un paciente en concreto.

En caso de seleccionar un paciente en concreto se obtiene como resultado lo siguiente:

- Un directorio cuyo nombre será el número de historia clínica del paciente y que contendrá un fichero de extensión .mat con los parámetros calculados (Ver Figura 42).
- Dichos parámetros calculados para el paciente serán exportados a un fichero Excel cuyo nombre es *Analisis_particulares* (Ver Figura 41). Los parámetros calculados se almacenarán en el fichero Excel junto a los que ya existían.

Por otra parte, si se selecciona el análisis de parámetros de la base de datos completa mediante el desplegable habilitado en la Figura 12, se obtiene lo siguiente:

- Un fichero denominado Analisis.mat que contiene una matriz cuyas filas corresponden a cada paciente y su sesión. La razón de almacenar este fichero en esta extensión es que el programa trabaja mucho más rápido con los datos en tal formato (Ver Figura 41).
- Un archivo Excel (ver Figura 43) que contiene los parámetros calculados para toda la base de datos, el formato de presentación de los datos es el mismo que para el análisis de un paciente particular.

Nombre	Fecha de modifica...	Tipo	Tamaño
2307251	15/06/2018 10:29	Carpeta de archivos	
2339579	12/06/2018 13:18	Carpeta de archivos	
Analisis	15/06/2018 11:59	Microsoft Access ...	19 KB
Analisis_particulares	15/06/2018 10:29	Hoja de cálculo d...	10 KB
Analisisistotal	14/06/2018 19:26	Hoja de cálculo d...	26 KB

Figura 41. Directorio de análisis dónde se almacena el análisis de la base de datos completa y los análisis de pacientes específicos

Este equipo > Documentos > TFG > Analisis > 2307251

Nombre	Fecha de modifica...	Tipo	Tamaño
Analisis_2307251_sesion_2	17/06/2018 11:07	Microsoft Access ...	1 KB

Figura 42. Ejemplo de nomenclatura almacenamiento archivo extensión .mat de un análisis específico de un paciente.

Analisisistotal - Microsoft Excel

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista

Calibri 11 General Formato condicional Insertar Σ

Pegar Fuente Alineación Número Estilos Estilos de celda Eliminar Ordenar y filtrar Buscar y seleccionar

Portapap... Celdas Modificar

	A	B	C	D	E	F	G	H	I	J
1										
2		Sesion	Valorpp	RMS	AUC	EHG_LC1	EHG_LC3	Teager	Primer mom	Curtosis
3	Paciente									
4	1650285	1	80.325742	10.7009169	281842.895	7226.74089	21016.5369	141900.523	9997476.47	2.0742709
5	1650285	2	87.8454622	10.561342	281349.193	7815.25537	50969.7692	115936.332	9101485.06	2.29370921
6	1667997	1	68.170204	8.75634004	185431.603	13198.0962	101889.03	73278.2872	6565933.74	2.21713276
7	2224943	1	40.3924795	4.54211246	52500.1981	3088.24695	22749.3043	19220.6041	2294691.77	2.21713276
8	2224943	2	46.7661401	5.71328056	80541.3241	5369.4216	51210.7898	30250.7114	3564730.95	2.50632822
9	2263442	1	67.377024	8.78849362	187938.716	7830.77983	28947.6972	72601.7031	6578818.92	2.16562948
10	2263442	2	84.3101303	8.22440348	167320.07	4921.17853	82048.6936	70435.4984	6313115.44	2.21588213
11	2307251	1	92.205241	10.4082278	260853.14	28983.6822	243044.538	87980.6018	8043147.53	1.953977
12	2307251	2	55.5197135	6.41379391	102788.126	11420.9029	287982.095	40311.1556	3577571.96	2.21546525
13	2339579	1	74.3024831	7.46481054	134156.45	14906.2722	125198.39	55371.9322	4639849.64	2.45398495
14	2347842	1	91.8086527	13.0953764	420585.517	15577.2414	61615.9977	202323.189	14684494.2	1.91810953
15	2347842	2	90.6288484	12.9949801	436911.216	16181.8969	64356.7354	155318.365	12407794.7	1.83779278
16	2347842	3	129.306072	16.8068182	185431.603	30604.1507	77522.1879	73278.2872	6565933.74	1.99775288
17	2368679	2	67.167263	9.60548334	224732.141	7491.07138	46866.8069	180672.74	10550398.8	2.21713276
18	2382081	1	44.7137107	5.15683732	66189.9794	5515.83162	65802.2832	28440.5914	2341415.79	2.36905095
19	2384926	2	68.6355021	9.21358342	207597.257	5463.08571	27684.4319	96921.1367	9117704.99	2.62877449
20	2392963	2	75.4385348	9.79163867	236616.285	8763.56612	64101.3766	93988.8587	11572744.8	1.82905705
21	3286502	1	113.283663	11.3453401	360858.249	32805.2953	101889.03	62503.5148	11568072.8	2.21713276
22	3970589	1	105.101059	12.8218018	406338.354	16930.7648	198274.222	123068.343	11790950.8	1.90557967
23	4077529	1	82.9827002	11.0396993	302433.096	12601.379	69738.9383	115142.272	10354984.7	2.44624958
24	4077529	2	54.3475274	8.01620543	157468.177	6561.17404	18270.8353	90866.9969	6553048.55	2.37427553
25	4079589	1	68.767734	8.56070384	177778.226	11111.1391	76909.0067	97743.8044	6185245.36	2.11414833

Hoja1 Hoja2 Hoja3

Figura 43. Fichero Excel generado después del cómputo de la base de datos completa.

4.4 Análisis estadístico

Una vez realizado el análisis de la base de datos completa, la ventana mostrada en la Figura 44 permite al usuario visualizar la distribución cada parámetro. Cabe mencionar que para poder visualizar la distribución de un parámetro concreto, éste debe estar computado previamente tal como se comentó en el apartado anterior.

El programa realiza dos test estadísticos del parámetro seleccionado por el usuario Jarque-bera y Kolmogorov. Tanto el test de Kolmogorov como el de Jarque-Bera sirven para determinar si la población proviene de una distribución normal o no.

Como ejemplo en la Figura 44, visualizamos el parámetro de valor pico a pico de la señal EHG.

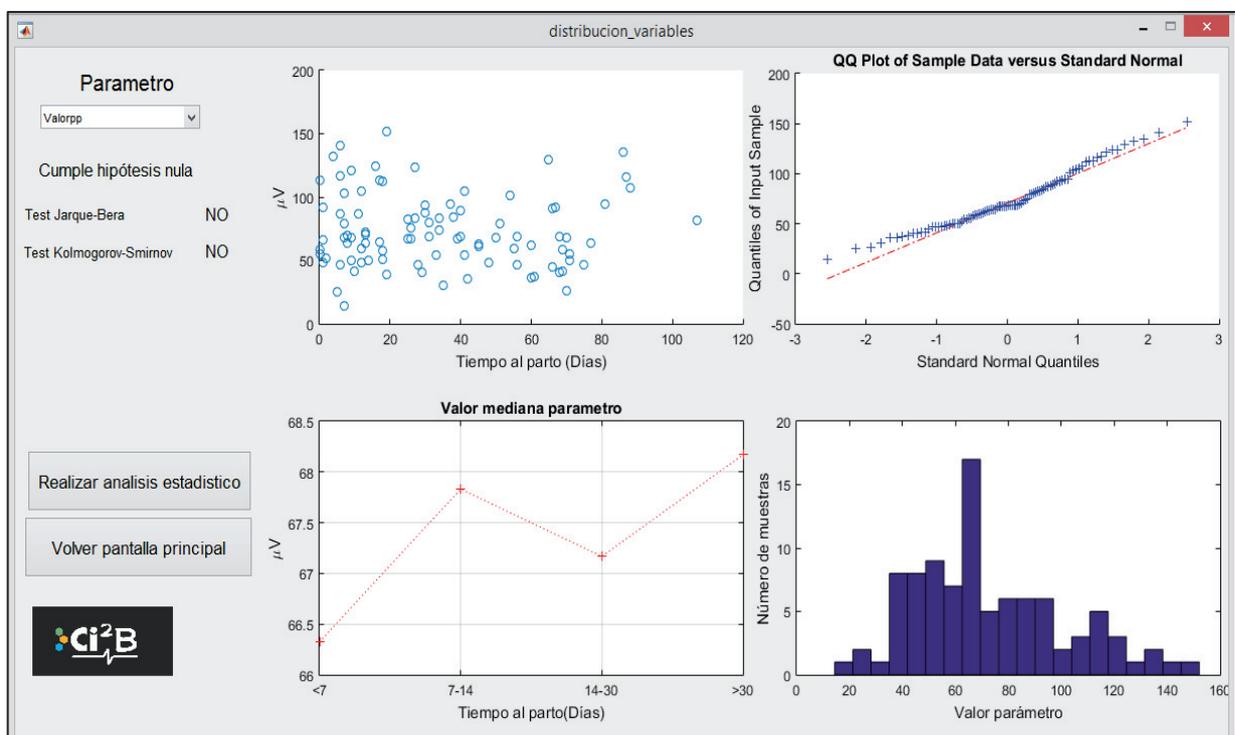


Figura 44. Visualización de la distribución del parámetro valor pico a pico.

Una vez realizado el análisis de los parámetros de la base de datos completa, mediante la ventana mostrada en la Figura 19 se lleva a cabo el análisis estadístico de los datos. Para ello se da la opción al usuario de seleccionar entre cuatro grupos de clasificación, uno que distingue entre pacientes a término y pretérmino y otros tres que distinguen a los pacientes en función del tiempo al parto en días que seleccione el usuario, que pueden ser 7 , 14 o 30. Estos datos provienen de la lectura de los datos obstétricos de cada paciente.

Los parámetros analizados, como se ha expuesto en apartados anteriores, se dividen en lineales, no lineales y espectrales. A modo representativo, se han seleccionado dos parámetros de cada familia y se muestra el resultado obtenido para los cuatro grupos de clasificación.

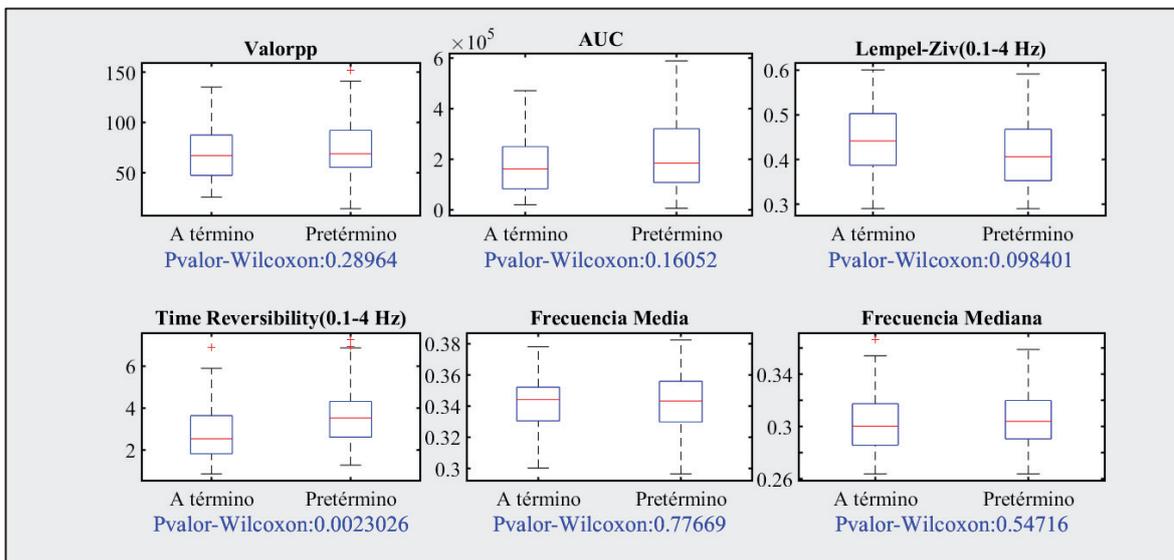


Figura 45. Diagrama Caja-Bigotes generado para grupo de clasificación a término y pretérmino.

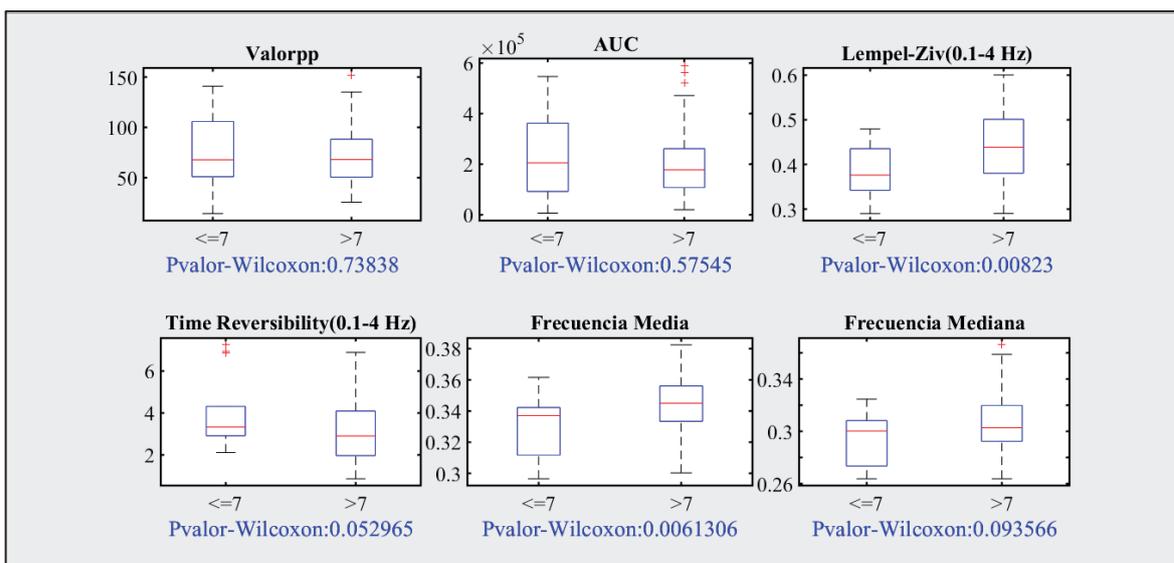


Figura 46. Diagrama Caja-Bigotes generado para grupo de clasificación menor de 7 días y mayor de 7 días de tiempo al parto.

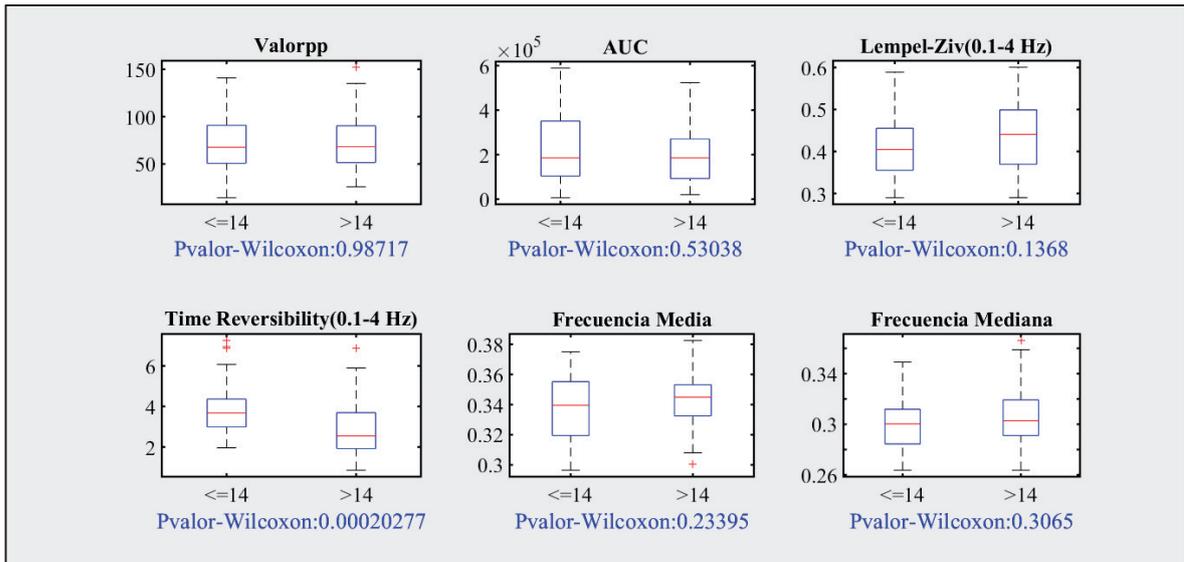


Figura 47. Diagrama Caja-Bigotes generado para grupo de clasificación menor de 14 días y mayor de 14 días de tiempo al parto.

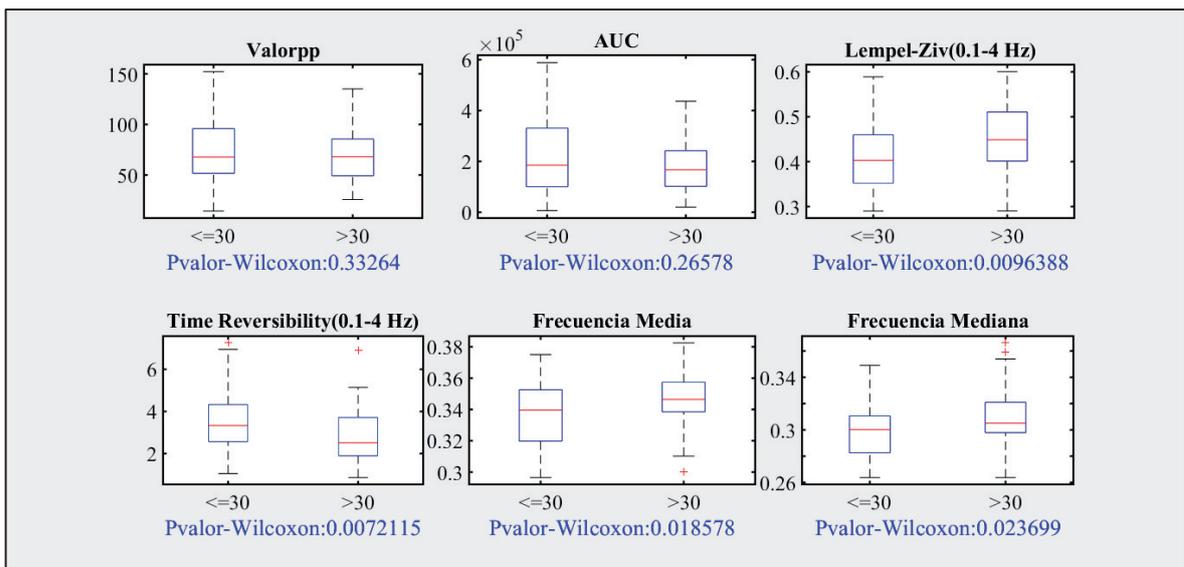


Figura 48. Diagrama Caja-Bigotes generado para grupo de clasificación menor de 30 días y mayor de 30 días de tiempo al parto.

De los seis parámetros expuestos, solamente la reversibilidad en el tiempo (Time Reversibility) presenta diferencias estadísticamente significativas entre el grupo de pretérmino y a término y TTD<14 y TTD>14. Se observa que la reversibilidad en el tiempo es mayor para el grupo pretérmino, lo cual coincide con la observación de otros autores [20].

Por otra parte, Lempel-Ziv y frecuencia media del grupo TTD<7 es significativamente inferior al de TTD>7, lo cual indica que se reduce la complejidad de la señal conforme se acerca el parto. No obstante, la frecuencia media disminuye conforme se acerca el parto lo cual contradice a la tendencia del desplazamiento del contenido espectral hacia altas frecuencias conforme se acerca



el parto. Esto probablemente sea debido a que los registros de la base de datos empleada en este trabajo fueron tomados en pacientes con amenaza de parto prematuro bajo distintas fases de terapia tocolítica. En la literatura se ha demostrado que los agentes tocolíticos reducen la excitabilidad de las células miométricas modificando la distribución espectral de la señal EHG[21].

En cambio, tanto Lempel-Ziv, Time reversibility, como Frecuencia media y Spectral Moment Ratio presentan diferencias estadísticamente significativas entre TTD<30 y TTD>30 días. Del mismo modo, se presenta una tendencia de disminución de la complejidad de la señal y aumento de la no-linealidad de la señal Time reversibility, así como un desplazamiento del contenido espectral hacia baja frecuencia conforme se acerca el parto, así como un desplazamiento del contenido espectral hacia baja frecuencia.

El hecho de que se hayan encontrado más o menos número de parámetros con diferencias estadísticamente significativas entre los distintos grupos se debe a la diferencia del tamaño de la muestra.

4.5 Informe estadístico

Como resumen del análisis estadístico de los distintos parámetros entre los distintos grupos, el programa permite crear un informe de dicho análisis. Para ello, se muestra para cada parámetro su mediana y su rango intercuartílico en cada grupo de clasificación. Los valores se encuentran ordenados de menor a mayor según el p-valor, para observar si existen resultados con significancia estadística.

A continuación, se muestra como ejemplo en la Figura 49 el resultado obtenido para el grupo de clasificación a término y pretérmino para todos los parámetros que se encuentran disponibles en el análisis, junto con el fichero Excel con los datos exportados. Estos resultados indican que tan sólo el parámetro reversibilidad en el tiempo y EHG/LC3 presentan diferencias estadísticamente significativas entre el grupo de pretérmino y a término.

Informe estadístico						Exportar a Excel
	A término		Pretérmino		P-valor	
	Mediana	Rango Intercuartílico	Mediana	Rango Intercuartílico		
Time Reversibility(0.1-4 Hz)	2.5346	1.8188	3.5221	1.713	0.0023026	^
EHG_LC3	64205.608	70148.245	101507.0658	74716.8918	0.0094206	
EHG_LC1	11269.3781	8729.2705	13043.4037	9482.4688	0.093015	
Ziv1(0.1-4 Hz)	0.44157	0.11559	0.40629	0.11504	0.098401	
DF2	0.35645	0.021973	0.36133	0.019531	0.12779	
AUC	162394.1235	166188.6304	185431.6029	212287.9418	0.16052	
RMS	8.1203	4.9289	8.769	5.6078	0.22621	
PON3	15.6439	11.7833	18.1589	10.7453	0.23803	
PON2	21.7786	16.496	25.3897	15.2179	0.25028	
Valorpp	67.1069	40.034	68.8458	36.6855	0.28964	
NE2	0.29433	0.07609	0.29781	0.081534	0.30721	
Dec6	0.33325	0.026855	0.33813	0.026855	0.32276	
PON4	0.13177	0.056531	0.12641	0.050887	0.32547	
PON1	3.0335	1.8357	2.9475	2.4245	0.33675	
Fuzzy entropy(0.1-4 Hz)	0.31455	0.11005	0.29599	0.069111	0.39904	
Spectral Moment Ratio	3.5338	0.56155	3.5243	0.86525	0.40969	
Dec2	0.23438	0.013428	0.2356	0.01709	0.43682	
Dec5	0.30273	0.029297	0.30518	0.029297	0.44654	
Primer momento	6550370.7932	6497275.1438	6557880.4941	5084605.9658	0.45371	
Dec8	0.43945	0.054932	0.42603	0.061035	0.48399	v

Figura 49. Ejemplo informe estadístico para los parámetros seleccionados anteriormente

Como se aprecia en la figura anterior, existe una opción que permite al usuario exportar los resultados obtenidos a un fichero Excel. En la Figura 50 se muestra dicho fichero generado.

	A	B	C	D	E	F
1		A término		Pretérmino		
2	Parametros	Mediana	Rango Inter cuartílico	Mediana	Rango Inter cuartílico	Pvalor
3	Time Revers	2.53455399	1.818757033	3.522056	1.712950466	0.00230257
4	EHG_LC3	64205.608	70148.24497	101507.066	74716.89181	0.00942057
5	EHG_LC1	11269.3781	8729.270462	13043.4037	9482.468763	0.0930148
6	Lempel-Ziv(0.44157156	0.115589252	0.40628964	0.115038301	0.09840102
7	DF2	0.35644531	0.021972656	0.36132813	0.01953125	0.12779164
8	AUC	162394.123	166188.6304	185431.603	212287.9418	0.16051724
9	RMS	8.12030445	4.928880646	8.76896636	5.60777254	0.2262089
10	PON3	15.6439239	11.78334716	18.15889	10.74530933	0.23802891
11	PON2	21.7786239	16.49601265	25.3897081	15.21791427	0.25027846
12	Valorpp	67.106896	40.03399965	68.8457844	36.68552749	0.28964095
13	NE2	0.29433117	0.07609006	0.29780816	0.081534139	0.30721113
14	Dec6	0.33325195	0.026855469	0.33813477	0.026855469	0.32275591
15	PON4	0.13176649	0.05653128	0.12640611	0.050886638	0.32547124
16	PON1	3.03353194	1.835727003	2.94754187	2.424528345	0.33674653
17	Fuzzy entrop	0.31455459	0.110054931	0.29598894	0.069110541	0.39903674
18	Spectral Mor	3.53376169	0.56155004	3.52426096	0.865247102	0.40969143
19	Dec2	0.234375	0.013427734	0.2355957	0.017089844	0.43682176
20	Dec5	0.30273438	0.029296875	0.30517578	0.029296875	0.44654438
21	Primer mom	6550370.79	6497275.144	6557880.49	5084605.966	0.45370993
22	Dec8	0.43945313	0.054931641	0.42602539	0.061035156	0.48398679
23	Curtosis	2.21671588	0.35421411	2.15930318	0.323060972	0.49113785
24	Lempel_mul	0.3086208	0.103867489	0.28894395	0.077433395	0.52274144
25	Frecuencia M	0.30029297	0.031738281	0.30395508	0.029296875	0.5471573

Figura 50. Ejemplo de exportación a Excel del informe estadístico

El programa genera para cada grupo de clasificación un directorio dentro del cual se ubican dos ficheros:

- Fichero Excel con la exportación del informe: dicho fichero se almacenará exclusivamente si el usuario lo desea, mediante el botón habilitado en la interfaz.
- Fichero de datos con los p-valor y sus parámetros asociados.

A continuación se muestra el sistema de almacenamiento de dichos ficheros en las figuras X y X. Aunque en el ejemplo mostrado anteriormente se ha seleccionado el grupo de clasificación a término y pretérmino, es importante señalar que el programa permite seleccionar los cuatro grupos que se muestran en la Figura 51.

Este equipo > Documentos > TFG > Analisis_estadistico

Nombre	Fecha de modifica...	Tipo	Tamaño
TermPreterm	18/06/2018 12:17	Carpeta de archivos	
30	18/06/2018 12:35	Carpeta de archivos	
14	18/06/2018 12:24	Carpeta de archivos	
7	18/06/2018 12:33	Carpeta de archivos	

Figura 51. Directorios generados para cada grupo de clasificación seleccionado en el análisis estadístico.

Este equipo > Documentos > TFG > Analisis_estadistico > TermPreterm

Nombre	Fecha de modifica...	Tipo	Tamaño
Informe_estadistico	18/06/2018 12:03	Hoja de cálculo d...	14 KB
Pvalor_almacenados	18/06/2018 12:00	Microsoft Access ...	2 KB

Figura 52. Informe estadístico exportado y p-valor almacenados para el grupo de clasificación pretérmino y a término.

4.6. Modelos predictivos

Tras el análisis estadístico de los parámetros de EHG, el usuario puede proceder al diseño y validación del modelo predictivo. En la ventana que se muestra a continuación en la Figura 53, seleccionamos el grupo de clasificación, el método de reducción de dimensión y el método de clasificación. En este caso se han definido cuatro grupos, de los cuales tres están basados en el tiempo al parto en días y uno en término o pretérmino.

Por limitaciones de tiempo y desarrollo, no se han implementado todas las funciones que se aprecian en la Figura 53. En concreto, para la reducción de dimensión se encuentra disponible el algoritmo PCA y en cuanto a los métodos de clasificación, es posible elegir entre regresión logística y redes neuronales. El resto de métodos se han dejado propuestos para ser implementados en versiones futuras de la interfaz.

Grupo de clasificación

- Preterm vs Term
- <7 vs >7
- <14 vs >14
- <30 vs >30

Reducción de dimensión

- PCA
- Algoritmo genético
- PSO
- SFFS

Método de clasificación

- Regresión logística
- Redes neuronales
- SVM
- Árbol de decisión

Análisis ROC

Figura 53. Ventana selección para la generación de modelos predictivos.

El almacenamiento de la información generada se realiza sistemáticamente según las opciones seleccionadas. La carpeta que raíz con los modelos recibe el nombre *Modelos*, dentro de ella se encuentran los distintos grupos de clasificación. A su vez, dentro de ellos encontraremos una carpeta con el tipo de reducción de dimensión seleccionado. Finalmente, dentro de la carpeta de reducción de dimensión encontramos la carpeta con el modelo seleccionado, que contendrá toda la información generada por el programa. Mediante las Figura 54, Figura 55 y Figura 56 ilustramos el sistema de almacenamiento de los modelos.

Este equipo > Documentos > TFG > Modelos			
Nombre	Fecha de modifica...	Tipo	Tamaño
7	14/06/2018 10:40	Carpeta de archivos	
14	14/06/2018 10:42	Carpeta de archivos	
30	14/06/2018 10:40	Carpeta de archivos	
Term-preterm	14/06/2018 10:43	Carpeta de archivos	

Figura 54. Directorio raíz con los distintos grupos de clasificación

Este equipo > Documentos > TFG > Modelos > 14			
Nombre	Fecha de modifica...	Tipo	Tamaño
PCA	14/06/2018 10:42	Carpeta de archivos	

Figura 55. Ejemplo directorio de la reducción de dimensión para el grupo de clasificación menor de 14 vs mayor de 14.

Este equipo > Documentos > TFG > Modelos > Term-preterm > PCA			
Nombre	Fecha de modifica...	Tipo	Tamaño
Redes neuronales	14/06/2018 11:05	Carpeta de archivos	
Regresion lineal	14/06/2018 10:43	Carpeta de archivos	

Figura 56. Ejemplo directorios que contienen la información de los modelos generados.

Cuando se realiza la **regresión logística**, el programa solicita al usuario el porcentaje de datos que se destinan al grupo de entrenamiento y el porcentaje de datos que se destinan al grupo de validación. Si el usuario decide introducir una proporción no válida, es decir, que no sume el 100%, el programa mostrará un aviso y no dejará continuar hasta que se introduzca

correctamente. Para garantizar que los resultados obtenidos tienen cierta validez estadística, se generan 30 particiones de los datos. Estas particiones consisten en separar los datos aleatoriamente según los porcentajes indicados y se almacenan como se muestra en la Figura 57.

Nombre	Fecha de modifica...	Tipo	Tamaño
P1	14/06/2018 10:42	Carpeta de archivos	
P2	14/06/2018 10:42	Carpeta de archivos	
P3	14/06/2018 10:42	Carpeta de archivos	
P4	14/06/2018 10:42	Carpeta de archivos	
P5	14/06/2018 10:42	Carpeta de archivos	
P6	14/06/2018 10:42	Carpeta de archivos	
P7	14/06/2018 10:42	Carpeta de archivos	
P8	14/06/2018 10:42	Carpeta de archivos	
P9	14/06/2018 10:42	Carpeta de archivos	
P10	14/06/2018 10:42	Carpeta de archivos	
P11	14/06/2018 10:42	Carpeta de archivos	
P12	14/06/2018 10:42	Carpeta de archivos	
P13	14/06/2018 10:42	Carpeta de archivos	
P14	14/06/2018 10:42	Carpeta de archivos	
P15	14/06/2018 10:42	Carpeta de archivos	

Figura 57. Ejemplo almacenamiento particiones en regresión lineal

Para cada partición el programa almacena los datos y su reducción de dimensión, como se muestra en la Figura 58. Los datos separados aleatoriamente se encuentran en el fichero *Datos* y cuando se aplica la reducción de dimensión se almacena el fichero *Datos98*. La reducción de dimensión se empleará a continuación para llevar a cabo la regresión logística y mostrar los resultados obtenidos.

Nombre	Fecha de modifica...	Tipo	Tamaño
Datos	14/06/2018 10:42	Microsoft Access ...	19 KB
Datos98	14/06/2018 10:42	Microsoft Access ...	14 KB

Figura 58. Ejemplo partición generada

Finalmente, una vez realizado el análisis ROC obtenemos la ventana siguiente con los resultados (ver Figura 59). Para el ejemplo mostrado se ha seleccionado como grupo de clasificación a término y pretérmino, la reducción de dimensión mediante PCA y la regresión logística.

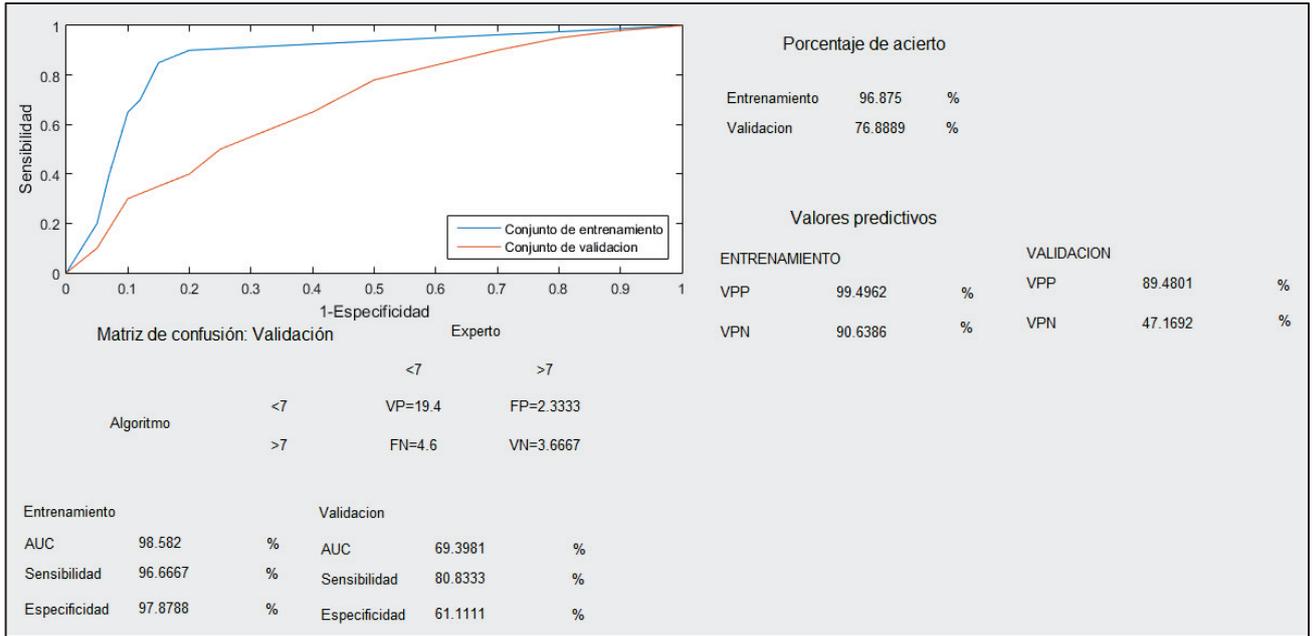


Figura 59. Resultados obtenidos en la ventana gráfica de análisis ROC mediante regresión logística para grupo de clasificación TTD<7 vs TTD>7, con reducción de dimensión mediante el algoritmo PCA.

A modo de resumen, se elaboran tablas con los datos obtenidos del análisis que se muestra en la Tabla 4 y Tabla 5.

<u>Matriz de confusión</u>	
<u>VP</u>	19.4
<u>FP</u>	2.3333
<u>FN</u>	4.6
<u>VN</u>	3.6667

Tabla 4. Matriz de confusión obtenida para el método de clasificación regresión logística, grupo de clasificación TTD<7 vs TTD>7 con reducción de dimensión mediante el algoritmo PCA.

	<u>Entrenamiento</u>	<u>Validación</u>
<u>Exactitud</u>	96.875%	76.8889
<u>AUC</u>	98.582%	69.3981%
<u>Sensibilidad</u>	97.8788%	80.8333%
<u>Especificidad</u>	97.8788%	61.1111%
<u>VPP</u>	99.4962%	89.4801%
<u>VPN</u>	90.6386%	47.1692%

Tabla 5. Resultados obtenidos en el análisis ROC para método de clasificación regresión logística, grupo de clasificación TTD<7 vs TTD>7 y reducción de dimensión mediante algoritmo PCA.

En base a los resultados obtenidos, se evalúa la capacidad del método de clasificación para diagnosticar el parto prematuro. En este caso, los inconvenientes que puede presentar un falso positivo son menores que los de un falso negativo, ya que la falta de diagnóstico de parto prematuro podría conllevar consecuencias negativas para el bebé. Por tanto, para este tipo de pruebas es importante presentar valores de sensibilidad elevados y en este caso los valores del modelo son lo suficientemente válidos. Respecto al AUC, valores mayores de 0.9 se asocian a pruebas con exactitud alta mientras que valores comprendidos entre 0.5-0.7 son de exactitud baja. En este caso, el conjunto de entrenamiento presenta un valor de AUC lo suficientemente elevado como para decir que posee una exactitud elevada aunque el conjunto de validación presenta una notable disminución de rendimiento, que lleva a categorizarlo como un conjunto exactitud baja. El conjunto de entrenamiento posee valores de sensibilidad y especificidad elevados, lo cual conlleva que tanto el valor predictivo positivo como el negativo posean valores elevados. Por otra parte, mientras que la sensibilidad del grupo de validación es alta, la especificidad posee un valor bajo por lo que presenta un valor predictivo positivo alto y un valor predictivo negativo bajo. En base a los resultados obtenidos, la disminución de rendimiento del grupo de entrenamiento al grupo de validación puede deberse a que se ha producido sobreentrenamiento como consecuencia del bajo número de muestras y elevado número de parámetros. Por tanto, no se puede afirmar con certeza que el modelo obtenido es válido para el diagnóstico de pacientes que no se encuentran en la base de datos.

Como se ha comentado anteriormente, el programa posee la capacidad de generar un método de clasificación mediante regresión logística y redes neuronales. A continuación se exponen los resultados obtenidos para la selección de **redes neuronales**.

El sistema de almacenamiento de archivos de redes neuronales por parte del programa es similar al expuesto en el apartado anterior, con la única diferencia de que en el directorio de particiones se almacenarán junto a ellas las redes generadas por la interfaz, como se muestra en la Figura 60 y Figura 61.

	P24	14/06/2018 11:05	Carpeta de archivos
	P25	14/06/2018 11:05	Carpeta de archivos
	P26	14/06/2018 11:05	Carpeta de archivos
	P27	14/06/2018 11:05	Carpeta de archivos
	P28	14/06/2018 11:05	Carpeta de archivos
	P29	14/06/2018 11:05	Carpeta de archivos
	P30	14/06/2018 11:05	Carpeta de archivos
	Redes generadas	14/06/2018 11:55	Carpeta de archivos

Figura 60. Ejemplo directorio almacenamiento redes generadas por la interfaz

Nombre	Fecha de modifica...	Tipo	Tamaño
red_1	14/06/2018 12:23	Microsoft Access ...	19 KB
red_2	14/06/2018 12:24	Microsoft Access ...	12 KB
red_3	14/06/2018 12:24	Microsoft Access ...	12 KB
red_4	14/06/2018 12:24	Microsoft Access ...	18 KB
red_5	14/06/2018 12:24	Microsoft Access ...	12 KB
red_6	14/06/2018 12:24	Microsoft Access ...	12 KB
red_7	14/06/2018 12:25	Microsoft Access ...	13 KB
red_8	14/06/2018 12:25	Microsoft Access ...	12 KB
red_9	14/06/2018 12:25	Microsoft Access ...	12 KB
red_10	14/06/2018 12:25	Microsoft Access ...	12 KB

Figura 61. Ejemplo almacenamiento redes generadas en su directorio correspondiente

En este caso, se generan diez redes por cada partición, lo cual genera en total trescientas redes neuronales. A modo de muestra se observa en la Figura 61 diez de las trescientas redes generadas.

Finalmente, una vez realizado el análisis ROC obtenemos la ventana siguiente con los resultados (ver Figura 62). Para el ejemplo mostrado se ha seleccionado como grupo de clasificación a término y pretérmino, la reducción de dimensión mediante PCA y la regresión logística.

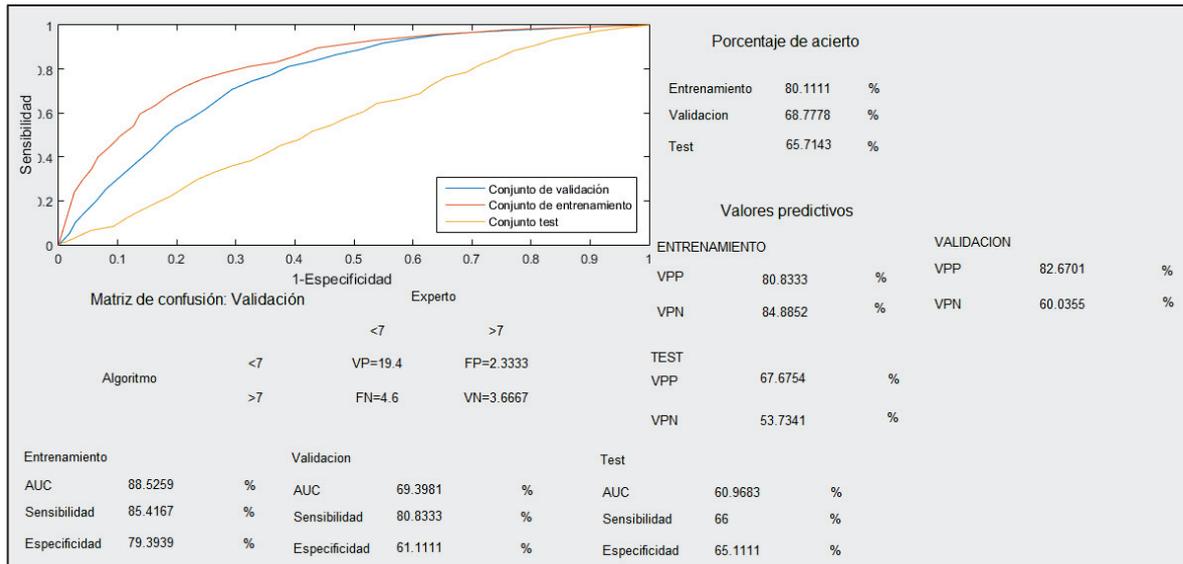


Figura 62. Resultados obtenidos en la ventana gráfica de análisis ROC mediante redes neuronales para grupo de clasificación TTD<7 vs TTD>7 con reducción de dimensión mediante el algoritmo PCA.

A modo de resumen, se elabora las siguientes tablas con los datos obtenidos del análisis que se muestran en la Tabla 6 y Tabla 7.

<u>Matriz de confusión</u>	
<u>VP</u>	19.4
<u>FP</u>	2.3333
<u>FN</u>	4.6
<u>VN</u>	3.6667

Tabla 6. Matriz de confusión obtenida para el método de clasificación redes neuronales, grupo de clasificación TTD<7 vs TTD>7, con reducción de dimensión mediante el algoritmo PCA.

	<u>Entrenamiento</u>	<u>Validación</u>	<u>Test</u>
<u>Exactitud</u>	80.1111%	68.7778%	65.7143%
<u>AUC</u>	88.5259%	69.3981%	60.9683%
<u>Sensibilidad</u>	85.4167%	80.8333%	66%
<u>Especificidad</u>	79.3939%	61.1111%	65.1111%
<u>VPP</u>	80.8333%	82.6701%	67.6754%
<u>VPN</u>	84.8852%	60.0355%	53.7341%

Tabla 7. Cuadro resumen de los resultados obtenidos en el análisis ROC para método de clasificación redes neuronales, grupo de clasificación TTD<7 vs TTD>7 y reducción de dimensión mediante algoritmo PCA.

Al igual que para regresión logística, con los resultados obtenidos se evalúa la validez del modelo predictivo generado. Como se ha comentado anteriormente para este tipo de pruebas es importante presentar valores de sensibilidad elevados, por tanto, tanto el grupo de entrenamiento como el de validación presentan valores razonables mientras que en el grupo de test se observa un decremento notable de sensibilidad. La especificidad sigue un comportamiento similar aunque con valores de rendimiento menores que la sensibilidad. En base a los resultados obtenidos, al igual que en regresión logística, es posible que la disminución de rendimiento del grupo de validación y test pueda deberse a un sobreentrenamiento de la red neuronal. Por tanto, no queda garantizado que dicho modelo predictivo generado sea válido para pacientes que no se encuentren en la base de datos.



5. CONCLUSIÓN

Finalmente, se ha desarrollado una interfaz gráfica que pone a disposición del usuario de manera versátil e intuitiva las distintas herramientas requeridas por el equipo de investigación.

La herramienta diseñada permite al usuario tanto la visualización de la señal EHG de los pacientes en la base de datos como la segmentación y/o resegmentación del registro EHG para el análisis de contracciones y/o el análisis de registro completo. La segmentación de las contracciones uterinas y/o tramos de señal analizables sirven como base para la obtención de los parámetros característicos representativos del registro. En este respecto, la interfaz gráfica permite computar una serie de parámetros lineales y no-lineales de un determinado paciente seleccionado por el usuario o del total de base de datos disponible.

Asimismo la interfaz gráfica permite realizar el análisis estadístico de los parámetros característicos de EHG: test de normalidad de los datos, gráficas de visualización, y test estadístico para determinar si un determinado parámetro presenta diferencia estadísticamente significativa entre distintos grupos de interés (pretérmino vs. a término, $TTD < 7/14/30$ vs. $TTD > 7/14/30$). Con ello, el usuario puede determinar fácilmente aquellos parámetros que contienen información relevante para diferenciar los distintos grupos.

Finalmente la herramienta diseñada permite al usuario no experto diseñar y validar los modelos matemáticos para la predicción del parto prematuro y/o parto inminente empleando distintas técnicas de reducción de dimensiones y métodos de clasificación.

6. BIBLIOGRAFÍA

- [1] J. González-Merlo and J. R. del. Sol, *Obstetricia*, 5ta ed. Elsevier, 2006.
- [2] “Nacimientos prematuros.” [Online]. Available: <http://www.who.int/es/news-room/fact-sheets/detail/preterm-birth>. [Accessed: 01-Jul-2018].
- [3] S. Petrou, Z. Mehta, C. Hockley, P. Cook-Mozaffari, J. Henderson, and M. Goldacre, “The impact of preterm birth on hospital inpatient admissions and costs during the first 5 years of life.,” *Pediatrics*, vol. 112, no. 6 Pt 1, pp. 1290–7, Dec. 2003.
- [4] E. H. BISHOP and E. H. Bishop, “PELVIC SCORING FOR ELECTIVE INDUCTION.,” *Obstet. Gynecol.*, vol. 24, no. 2, pp. 266–8, Aug. 1964.
- [5] “Prueba de fibronectina fetal - Mayo Clinic.” [Online]. Available: <https://www.mayoclinic.org/es-es/tests-procedures/fetal-fibronectin/about/pac-20384676>. [Accessed: 01-Jul-2018].
- [6] R. E. Garfield and W. L. Maner, “Physiology and electrical activity of uterine contractions,” *Semin. Cell Dev. Biol.*, vol. 18, no. 3, pp. 289–295, Jun. 2007.
- [7] M. P. G. C. Vinken, C. Rabotti, M. Mischi, and S. G. Oei, “Accuracy of Frequency-Related Parameters of the Electrohysterogram for Predicting Preterm Delivery,” *Obstet. Gynecol. Surv.*, vol. 64, no. 8, pp. 529–541, Aug. 2009.
- [8] R. E. Garfield *et al.*, “Uterine Electromyography and Light-Induced Fluorescence in the Management of Term and Preterm Labor,” *J. Soc. Gynecol. Investig.*, vol. 9, no. 5, pp. 265–275, Sep. 2002.
- [9] M. Hassan, J. Terrien, A. Alexandersson, C. Marque, and B. Karlsson, “Improving the classification rate of labor vs. normal pregnancy contractions by using EHG multichannel recordings,” *2010 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC’10*, pp. 4642–4645, 2010.
- [10] Y. Ye-Lin, J. Garcia-Casado, G. Prats-Boluda, J. Alberola-Rubio, and A. Perales, “Automatic identification of motion artifacts in EHG recording for robust analysis of uterine contractions,” *Comput. Math. Methods Med.*, vol. 2014, 2014.
- [11] H. Maul, W. L. Maner, G. Olson, G. R. Saade, and R. E. Garfield, “Non-invasive transabdominal uterine electromyography correlates with the strength of intrauterine pressure and is predictive of labor and delivery.,” *J. Matern. Neonatal Med.*, vol. 15, pp. 297–301, 2004.
- [12] P. Fergus, P. Cheung, A. Hussain, D. Al-Jumeily, C. Dobbins, and S. Iram, “Prediction of Preterm Deliveries from EHG Signals Using Machine Learning,” *PLoS One*, vol. 8, no. 10, 2013.
- [13] D. Devedeux, C. Marque, S. Mansour, G. Germain, and J. Duchêne, “Uterine electromyography: A critical review,” *Am. J. Obstet. Gynecol.*, vol. 169, no. 6, pp. 1636–1653, 1993.
- [14] J. Garcia-Casado, Y. Ye-Lin, G. Prats-Boluda, J. Mas-Cabo, J. Alberola-Rubio, and A. Perales, “Electrohysterography in the diagnosis of preterm birth: a review,” *Physiol. Meas.*, 2018.



- [15] B. Karlsson, J. Terrien, V. Gudmundsson, T. Steingrimsdottir, and C. Marque, "Abdominal EHG on a 4 by 4 grid: mapping and presenting the propagation of uterine contractions," *11th Mediterr. Conf. Med. Biomed. Eng. Comput.*, pp. 139–143, 2007.
- [16] H. de Lau, C. Rabotti, H. P. Oosterbaan, M. Mischi, and G. S. Oei, "Study protocol: PoPE-Prediction of Preterm delivery by Electrohysterography," *BMC Pregnancy Childbirth*, vol. 14, no. 1, p. 192, 2014.
- [17] R. E. Garfield and W. L. Maner, "Physiology and electrical activity of uterine contractions," *Semin. Cell Dev. Biol.*, vol. 18, no. 3, pp. 289–295, 2007.
- [18] D. Alamedine, A. Diab, C. Muszynski, B. Karlsson, M. Khalil, and C. Marque, "Selection algorithm for parameters to characterize uterine EHG signals for the detection of preterm labor," *Signal, Image Video Process.*, vol. 8, no. 6, pp. 1169–1178, 2014.
- [19] G. Fele-Žorž, G. Kavšek, Ž. Novak-Antolič, and F. Jager, "A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups," *Med. Biol. Eng. Comput.*, vol. 46, no. 9, pp. 911–922, 2008.
- [20] M. Hassan, J. Terrien, C. Marque, and B. Karlsson, "Comparison between approximate entropy, correntropy and time reversibility: Application to uterine electromyogram signals," *Med. Eng. Phys.*, vol. 33, no. 8, pp. 980–986, 2011.
- [21] M. P. G. C. Vinken, C. Rabotti, M. Mischi, J. O. E. H. van Laar, and S. G. Oei, "Nifedipine-Induced Changes in the Electrohysterogram of Preterm Contractions: Feasibility in Clinical Practice," *Obstet. Gynecol. Int.*, 2010.



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Desarrollo e implementación de una aplicación para la
caracterización, el procesado de y análisis de señales
electrohisterográficas en entorno MATLAB



ESCUELA TÉCNICA
SUPERIOR INGENIEROS
INDUSTRIALES VALENCIA

PORTADA DEL MANUAL DE USUARIO



Índice del manual de usuario

1. PERFIL DE USUARIO.....	5
2. SISTEMA DE ALMACENAMIENTO.....	6
2.1 Directorio principal.....	6
2.2 Análisis de parámetros y análisis estadístico.....	6
2.3 Modelos.....	7
2.4 Registros APP.....	7
2.5 Almacenamiento de la segmentación.....	8
3. DESCRIPCIÓN USO ENTORNO GRÁFICO.....	9
3.1 Pantalla principal.....	9
3.1.1 Búsqueda y selección de pacientes.....	9
3.1.2 Visualizar señal.....	11
3.1.3 Análisis de los parámetros.....	12
3.1.4 Ver distribución de variables.....	13
3.1.5 Análisis estadístico.....	13
3.1.6 Generar modelos.....	14
3.2 Visualización.....	14
3.2.1 Selección del ancho de banda a visualizar.....	15
3.2.2 Ajuste de ejes de visualización.....	15
3.2.3 Visualizar segmentación.....	16
3.2.4 Resegmentar.....	17
3.2.5 Análisis.....	18
3.3 Resegmentar.....	18
3.3.1 Edición y creación de la segmentación.....	18
3.3.2 Exportación de la segmentación modificada.....	21
3.4 Análisis.....	21
3.4.1 Análisis de un paciente específico.....	22
3.4.2 Análisis de la base de datos completa.....	23
3.4.3 Almacenamiento análisis parámetros.....	23
3.5 <i>Distribución de las variables</i>	24



3.6 Análisis estadístico	25
3.6.1 Selección de parámetros a analizar	25
3.6.2 Realizar análisis estadístico	26
3.6.3 Exportación informe estadístico y p-valor calculado.....	28
3.7 Generación de modelos predictivos.....	29
3.7.1 Regresión logística.....	30
3.7.2 Redes neuronales	31
3.7.3 Continuar análisis ROC.....	31
3.8 Análisis ROC	32

Índice de figuras

Figura 1. Directorio principal de la interfaz.....	6
Figura 2. Sistema de almacenamiento de los modelos predictivos según el grupo de clasificación..	7
Figura 3. Contenido directorio RegistrosAPP_Matlab	7
Figura 4. Directorio de almacenamiento de la segmentación de contracción.....	8
Figura 5. Directorio de almacenamiento de la segmentación de registro completo.	8
Figura 6. Ventana principal.....	9
Figura 7. Ventana emergente que solicita al usuario si desea procesar la señal de los pacientes que no tienen la señal procesada.....	9
Figura 8. Ventana emergente para la búsqueda de pacientes por número de historia clínica.....	10
Figura 9. Mensaje de error en caso de no encontrar el paciente buscado en la base de datos.	10
Figura 10. Ejemplo búsqueda paciente N°1650285.	10
Figura 11. Ejemplo paciente seleccionado N°6479638 junto con su sesiones disponibles.	11
Figura 12. Mensaje de advertencia en caso de no seleccionar el paciente correctamente.	12
Figura 13. Ejemplo cuadro superior derecho de la ventana Análisis en caso de no seleccionar ningún paciente.	12
Figura 14. Ejemplo cuadro superior derecho de la ventana Análisis en caso de seleccionar paciente N°6479638 y la sesión 3.....	13
Figura 15. Mensaje error por no encontrar un análisis de la base de datos completa.	13
Figura 16. Ventana de visualización de la señal	14
Figura 17. Anchos de banda disponibles para una señal filtrada por el programa.	15



Figura 18. Cuadro de texto para la introducción del valor.....	15
Figura 19. Ejemplo slider ajuste de valor límite de ejes.....	16
Figura 20. Ejemplo visualización segmentación de contracción	16
Figura 21. Ejemplo visualización segmentación de registro completo.	17
Figura 22. Mensaje de advertencia cuando no existe segmentación asociada.....	17
Figura 23. Selección tramo 2 mediante el desplegable para su posterior edición.	19
Figura 24. Ejemplo cursor para la definición gráfica de tramos en la segmentación.	19
Figura 25. Ventana emergente introducción inicio del tramo.	20
Figura 26. Ventana emergente introducción final de tramo.....	20
Figura 27. . Mensaje de advertencia si el tramo introducido no es válido o ya se encontraba definido.	20
Figura 28. Solicitud de confirmación por parte del usuario de que desea sobrescribir la segmentación.	21
Figura 29. Ventana de análisis de los parámetros.....	21
Figura 30. Opción seleccionada en el desplegable para realizar el análisis de un paciente específico.....	22
Figura 31. Barra de progreso de análisis del paciente Nº6250758 sesión 1.	22
Figura 32. Ventana emergente proceso finalizado.	22
Figura 33. Opción seleccionada en el desplegable para realizar el análisis de la base de datos completa.....	23
Figura 34. Barra de progreso de análisis de la base de datos completa.	23
Figura 35. Almacenamiento interno de los análisis calculados.	23
Figura 36. Ventana visualización de distribución de las variables.	24
Figura 37. Desplegable con los parámetros disponibles para su visualización.	24
Figura 38. Ventana análisis estadístico.	25
Figura 39. Mensaje de advertencia en caso de no haber seleccionado ningún parámetro.	26
Figura 40. Ejemplo diagrama caja-bigotes generado.....	27
Figura 41. Ventana informe estadístico generado.....	27
Figura 42. Ejemplo almacenamiento informe y pvalor calculados para grupo de clasificación término y pretérmino.	28
Figura 43. Ventana para la generación de modelos predictivos.	29
Figura 44. Mensaje de advertencia para los métodos no implementados.	30



Figura 45. Panel mostrado para la introducción de los porcentajes de datos de entrenamiento y validación en regresión logística.....	30
Figura 46. Mensaje de advertencia por introducción de proporciones incorrectas.....	31
Figura 47. Panel mostrado para la introducción de los porcentajes de datos de entrenamiento, validación y test en redes neuronales.....	31
Figura 48. Ejemplo neuronales.ventana análisis ROC para redes	32
Figura 49. Ejemplo ventana análisis ROC para redes neuronales.....	32



1. PERFIL DE USUARIO

La interfaz diseñada ha sido desarrollada en el ámbito académico ya que se trata de un trabajo de final de grado. Puesto que ha sido pensada para el grupo de investigación del Centro de Investigación e Innovación en Bioingeniería (Ci2b), se espera que el perfil de usuario que va a utilizar esta interfaz sea el de una persona con conocimientos suficientes sobre la señal EHG y su análisis. La finalidad de esta interfaz es proporcionar al usuario un entorno gráfico que permita cumplir las funciones necesarias dadas sin necesidad de introducción de comandos. Los conocimientos del usuario junto con el presente manual y la robustez del programa proporcionan un entorno sólido y versátil de trabajo.

2. SISTEMA DE ALMACENAMIENTO

Es importante señalar que la interfaz se ha desarrollado para una base de datos interna, y por tanto va condicionada a una distribución de carpetas fija.

2.1 Directorio principal

Es el directorio que contiene todo el programa, es imprescindible añadir al Path de Matlab este directorio para que el programa funcione. Puesto que el lenguaje de programación busca el nombre del directorio específicamente, es imprescindible que los directorios que se muestran en la Figura 1 tengan exactamente los mismos nombres, en caso contrario el programa presentará errores de lectura de datos. La interfaz completa se encuentra contenida en la carpeta “programa” que puede verse en la Figura 1.

Nombre	Fecha de modifica...	Tipo	Tamaño
Analisis	23/06/2018 18:25	Carpeta de archivos	
Analisis_estadistico	18/06/2018 12:35	Carpeta de archivos	
Modelos	17/06/2018 12:21	Carpeta de archivos	
programa	21/06/2018 12:23	Carpeta de archivos	
RegistrosAPP_Matlab	14/06/2018 18:03	Carpeta de archivos	
Segmentacion_CT	22/06/2018 17:47	Carpeta de archivos	
Segmentacion_RC	15/06/2018 17:55	Carpeta de archivos	

Figura 1. Directorio principal de la interfaz

En los siguientes apartados se explica el contenido de cada directorio específicamente y su nomenclatura condicionada por el lenguaje de programación.

2.2 Análisis de parámetros y análisis estadístico

El análisis de parámetros se almacena en el directorio “Análisis” que puede apreciarse en la Figura 1. A su vez, el análisis estadístico se almacena en la carpeta “Análisis_estadístico” que se puede ver en la Figura 1.

2.3 Modelos

Este directorio es el destinado a almacenar todos los modelos generados por la interfaz. Es necesaria la definición de este directorio y de las subcarpetas que se encuentran en su interior. Las subcarpetas de su interior son los directorios correspondientes a cada grupo de clasificación, y deben ser definidas exactamente como se muestra en la Figura 2.

Este equipo > Documentos > TFG > Modelos

Nombre	Fecha de modifica...	Tipo	Tamaño
7	14/06/2018 10:40	Carpeta de archivos	
14	14/06/2018 10:42	Carpeta de archivos	
30	14/06/2018 10:40	Carpeta de archivos	
Term-preterm	14/06/2018 10:43	Carpeta de archivos	

Figura 2. Sistema de almacenamiento de los modelos predictivos según el grupo de clasificación.

2.4 Registros APP

Los registros de amenaza de parto prematuro deben estar contenidos en la carpeta “RegistrosAPP_Matlab”, únicamente con dicho nombre. Esta carpeta contendrá diferentes directorios nombrados por el número de historia clínica del paciente, que contendrán los registros filtrados o no. Se puede ver en la Figura 3 el contenido de este directorio.

Este equipo > Documentos > TFG > RegistrosAPP_Matlab

Nombre	Fecha de modifica...	Tipo	Tamaño
1650285	14/06/2018 18:01	Carpeta de archivos	
1667997	14/06/2018 18:01	Carpeta de archivos	
2224943	14/06/2018 18:01	Carpeta de archivos	
2263442	14/06/2018 18:01	Carpeta de archivos	
2307251	14/06/2018 18:01	Carpeta de archivos	
2339579	14/06/2018 18:01	Carpeta de archivos	
2347842	14/06/2018 18:01	Carpeta de archivos	
2368679	14/06/2018 18:01	Carpeta de archivos	
2382081	14/06/2018 18:01	Carpeta de archivos	
2384926	14/06/2018 18:01	Carpeta de archivos	
2392963	14/06/2018 18:01	Carpeta de archivos	

Figura 3. Contenido directorio RegistrosAPP_Matlab

2.5 Almacenamiento de la segmentación

Como puede apreciarse en la Figura 1, existen dos directorios que contienen la segmentación:

- Segmentacion CT: este directorio contiene la segmentación de contracción. Es imprescindible que esta segmentación se almacene en el directorio con este nombre. En su interior, contiene los ficheros Excel que se nombran de la siguiente forma “Nº historia clínica paciente_sesion_SegCT” como se muestra en la Figura 4.

Nombre	Fecha de modifica...	Tipo	Tamaño
1650285_1_SegCT	02/05/2016 18:01	Hoja de cálculo d...	18 KB
1650285_2_SegCT	02/05/2016 18:02	Hoja de cálculo d...	18 KB
1667997_1_SegCT	20/03/2017 18:43	Hoja de cálculo d...	17 KB
2224943_1_SegCT	07/11/2016 16:36	Hoja de cálculo d...	17 KB
2224943_2_SegCT	07/11/2016 16:37	Hoja de cálculo d...	16 KB

Figura 4. Directorio de almacenamiento de la segmentación de contracción.

- Segmentación RC: este directorio contiene la segmentación de registro completo. Es imprescindible que esta segmentación se almacene en el directorio con este nombre. En su interior, contiene los ficheros Excel que se nombran de la siguiente forma “Nº historia clínica paciente_sesion_Completo” como se muestra en la Figura 5.

Nombre	Fecha de modifica...	Tipo	Tamaño
1243183_1_Completo	14/05/2018 19:11	Hoja de cálculo d...	17 KB
1552644_1_Completo	14/05/2018 19:12	Hoja de cálculo d...	17 KB
1650285_1_Completo	14/05/2018 19:08	Hoja de cálculo d...	17 KB
1650285_2_Completo	14/05/2018 19:08	Hoja de cálculo d...	17 KB
1663696_1_Completo	14/05/2018 19:11	Hoja de cálculo d...	67 KB

Figura 5. Directorio de almacenamiento de la segmentación de registro completo.

3. DESCRIPCIÓN USO ENTORNO GRÁFICO.

3.1 Pantalla principal

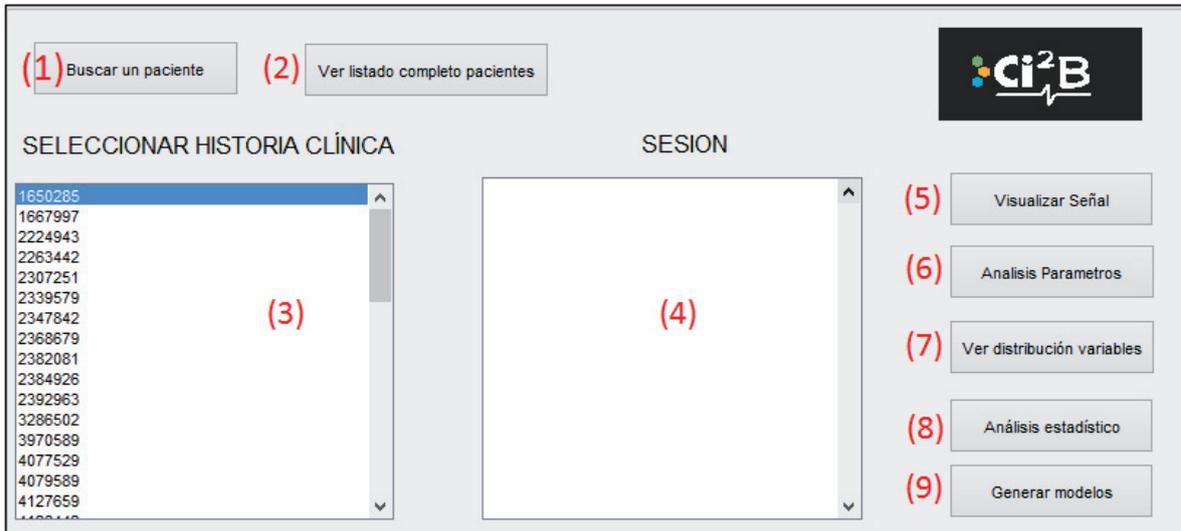


Figura 6. Ventana principal.

La ventana mostrada en la Figura 6 es la principal del programa, ya que a través de ésta se permite acceder a todas las opciones que la interfaz contiene. Al iniciar esta ventana el programa detecta si existen registros de pacientes sin procesar y en ese caso pregunta al usuario si desea procesarlos en ese momento, como se muestra en la Figura 7. A continuación se detalla las funciones de la ventana correspondiente.

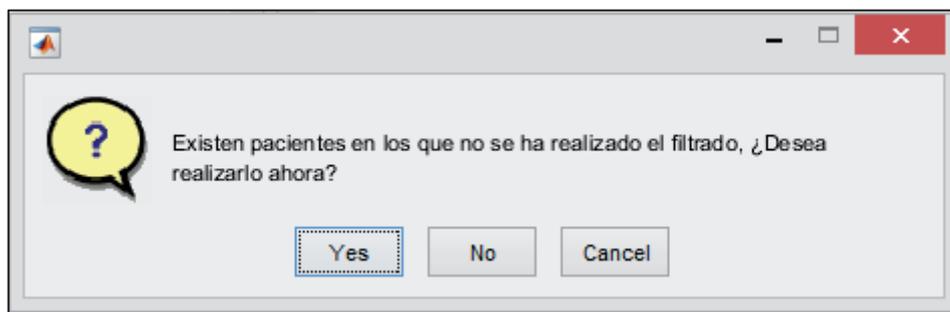


Figura 7. Ventana emergente que solicita al usuario si desea procesar la señal de los pacientes que no tienen la señal procesada.

3.1.1 Búsqueda y selección de pacientes

Para la búsqueda de pacientes se encuentra la opción de la marca 1 de la Figura 6, que mediante la ventana emergente mostrada en la Figura 8 nos permite buscar un paciente que se encuentre en la base de datos mediante la introducción de su número de historia clínica. En caso de que dicho paciente se encuentre en ella, se mostrará en el listado inferior (marca 3 Figura 6). Si el paciente no se encuentra en la base de datos, aparecerá un mensaje de advertencia para el usuario como se aprecia en la Figura 9.

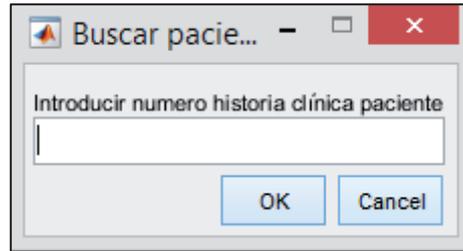


Figura 8. Ventana emergente para la búsqueda de pacientes por número de historia clínica.

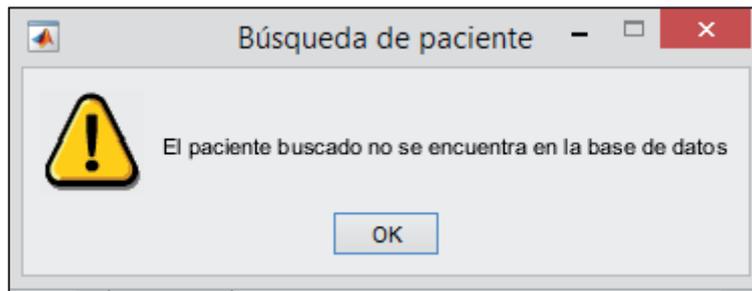


Figura 9. Mensaje de error en caso de no encontrar el paciente buscado en la base de datos.

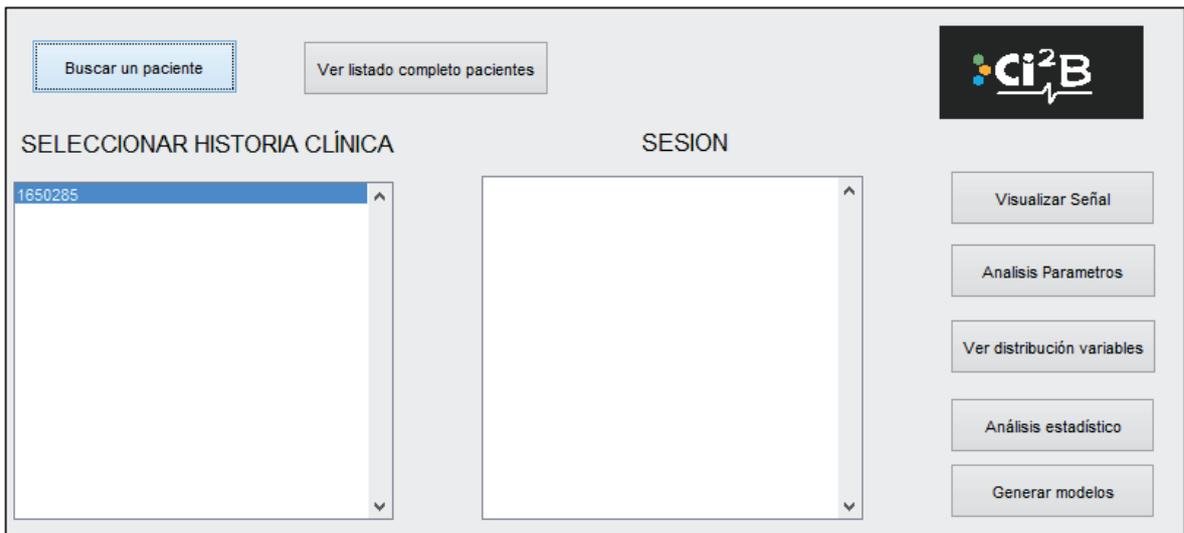


Figura 10. Ejemplo búsqueda paciente Nº1650285.

En caso de haber seleccionado un paciente para su búsqueda y querer visualizar nuevamente la lista de la base de datos completa, la opción “ver listado completo pacientes” que se muestra en la marca 2 de la Figura 6 permite realizarlo.

Para la selección de un paciente en concreto se cuenta con las herramientas mostradas en la Figura 6 mediante la marca 3 y 4. Dichas herramientas consisten en listados que muestran las opciones disponibles para seleccionar, que varían dependiendo del listado escogido. A continuación se explica la función de cada listado:

- Seleccionar historia clínica: corresponde a la marca 3 de la Figura 6, su función es la de mostrar el listado de pacientes completo de la base de datos con su número de historia clínica, o el paciente buscado si se encuentra en la base de datos, como se ha comentado en el apartado anterior.
- Sesión: corresponde a la marca 4 de la Figura 6, una vez se ha seleccionado el paciente en el listado anterior, este listado muestra automáticamente las sesiones disponibles como se muestra en la Figura 11. A demás, permite la selección de la sesión deseada para ser tratada en el programa.

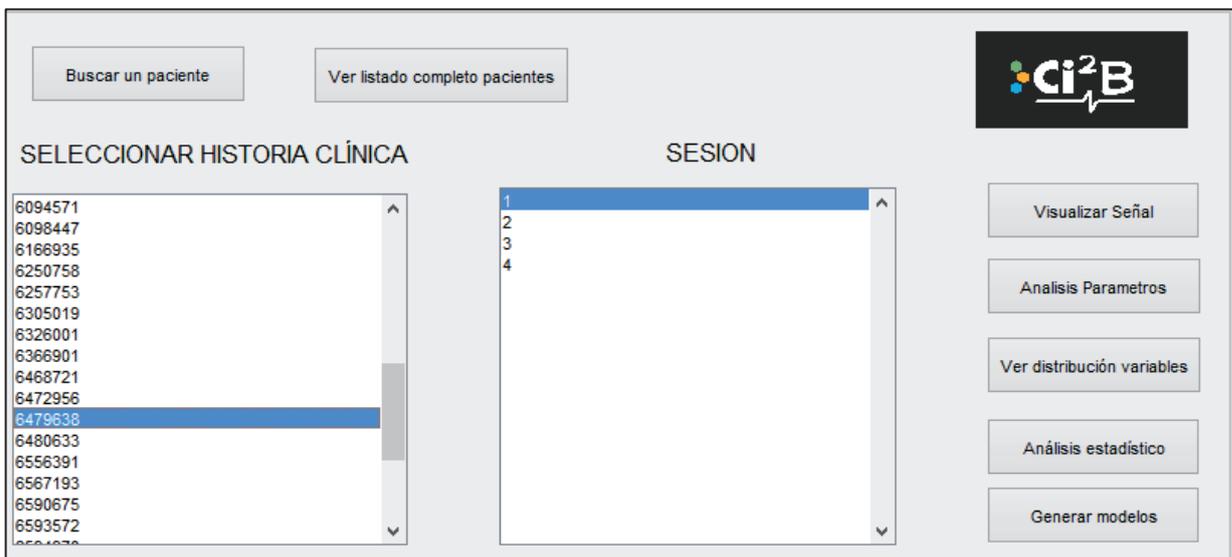


Figura 11. Ejemplo paciente seleccionado N°6479638 junto con su sesiones disponibles.

Se entiende que el paciente ha sido seleccionado cuando se selecciona tanto el número de historia clínica como la sesión que se desea tratar. Una vez seleccionado, se puede proceder a la visualización de la señal o al análisis de los parámetros del paciente.

3.1.2 Visualizar señal

Para visualizar la señal se dispone del botón habilitado para dicha función como puede apreciarse en la marca 5 de la Figura 6. Esta opción nos permite avanzar a la ventana siguiente que permite la visualización de la señal en los distintos anchos de banda que ha sido filtrada. Es importante señalar que para poder visualizar la señal, el paciente debe haber sido correctamente seleccionado. Para ello, como se ha expuesto anteriormente, debe haber sido seleccionado tanto el paciente como la sesión correspondiente. En caso de no realizar esta operación correctamente, aparecerá el mensaje de advertencia mostrado en la Figura 12.

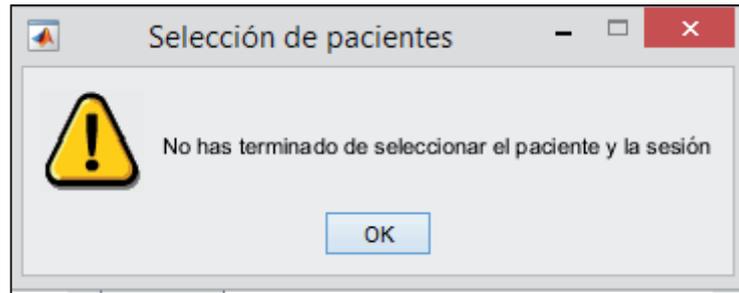


Figura 12. Mensaje de advertencia en caso de no seleccionar el paciente correctamente.

3.1.3 Análisis de los parámetros

Para continuar a la ventana que permite el análisis de los parámetros de la señal, se habilita el botón mostrado en la marca 6 de la Figura 6. Al pulsar este botón pueden darse dos situaciones:

- No se ha seleccionado ningún paciente en concreto: en este caso, posteriormente en la ventana de análisis, el programa solamente permitirá el análisis de los parámetros de la base de datos completa y en el indicador de paciente y sesión seleccionados de la ventana de análisis se mostrará que no se ha seleccionado ningún paciente, como puede apreciarse en la Figura 13.

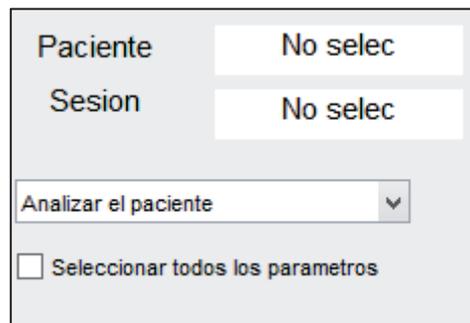


Figura 13. Ejemplo cuadro superior derecho de la ventana Análisis en caso de no seleccionar ningún paciente.

- Se ha seleccionado un paciente: en tal caso, en el indicador de la ventana de análisis mencionado anteriormente aparecerá el número del paciente y la sesión seleccionada, como se muestra en la Figura 14. Es importante señalar que el análisis de un paciente en concreto sólo será posible si se ha realizado esta operación correctamente.

Paciente	6479638
Sesion	3
Analizar el paciente	▼
<input type="checkbox"/> Seleccionar todos los parametros	

Figura 14. Ejemplo cuadro superior derecho de la ventana Análisis en caso de seleccionar paciente N°6479638 y la sesión 3.

3.1.4 Ver distribución de variables

Para continuar a la ventana que permite visualizar la distribución de las variables, se habilita el botón mostrado en la marca 7 de la Figura 6. Es importante saber que dicha visualización solamente es posible si se cuenta con un análisis de la base de datos completa en la base de datos del programa. En caso de no contar con dicho análisis, el programa mostrará el mensaje de advertencia que puede verse en la Figura 15 y no permitirá acceder a dicha ventana hasta que el usuario realice el análisis de la base de datos completa.

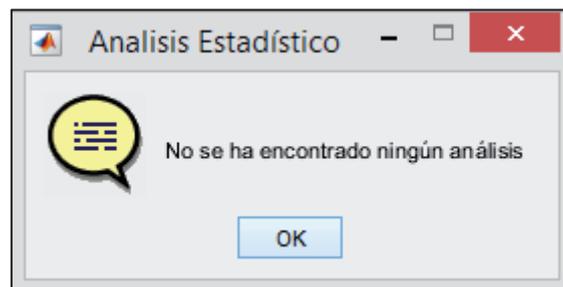


Figura 15. Mensaje error por no encontrar un análisis de la base de datos completa.

3.1.5 Análisis estadístico

Para continuar a la ventana que permite realizar un análisis estadístico de los datos obtenidos, se habilita el botón mostrado en la marca 8 de la Figura 6. Para acceder a esta opción también es imprescindible contar con un análisis de la base de datos completa y en caso no encontrarse dicho análisis, aparecerá una advertencia como la mostrada en la Figura 15. Por tanto, al igual que en el apartado anterior, no se permitirá el acceso a dicha ventana hasta que el usuario realice el análisis de la base de datos completa.

3.1.6 Generar modelos

Para continuar a la ventana que permite la creación de modelos predictivos basándonos en los datos, se habilita el botón mostrado en la marca 9 de la Figura 6. Así como para ver la distribución de las variables como para realizar el análisis estadístico de los datos, es indispensable contar con un análisis de la base de datos completa para poder continuar a la ventana siguiente. En caso de no contar con dicho análisis, aparecerá un mensaje de advertencia como el mostrado en la Figura 15.

3.2 Visualización

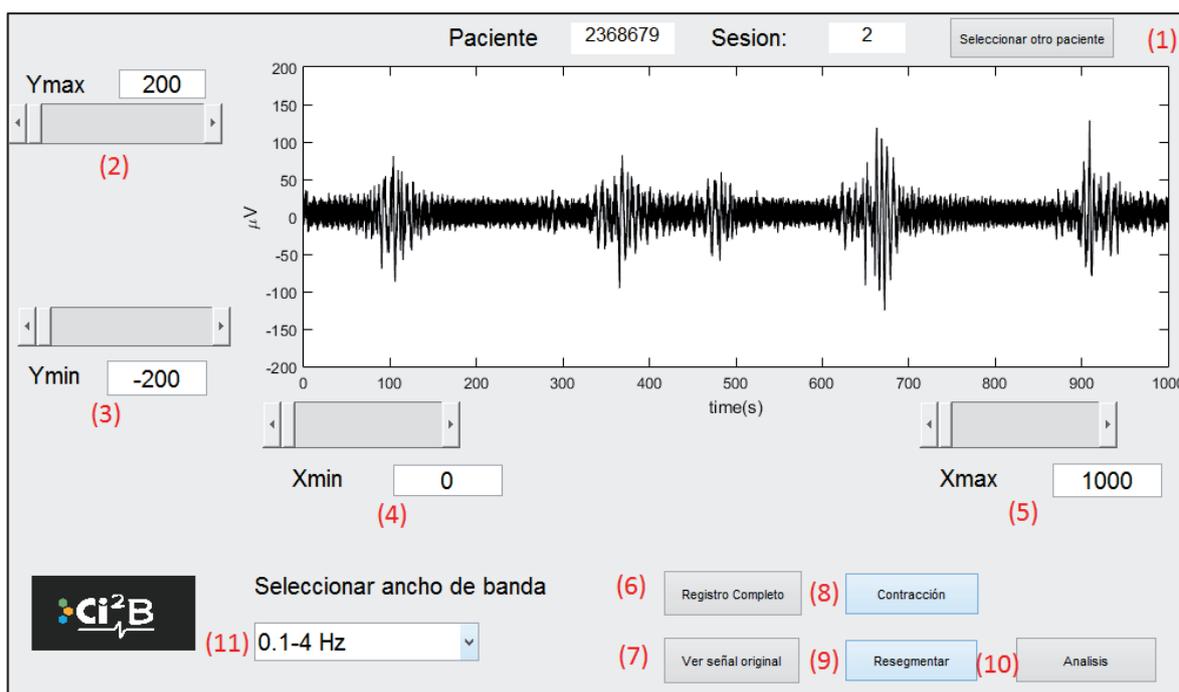


Figura 16. Ventana de visualización de la señal

La ventana mostrada en la Figura 16 permite al usuario visualizar, para el paciente y sesión seleccionados en la ventana anterior, la señal en los distintos anchos de banda que se filtra la señal. Además, la visualización permite cargar tanto la segmentación de registro completo como la de contracción. Junto con estas opciones se tiene la posibilidad de modificar o crear la segmentación asociada al paciente seleccionado y de analizar los parámetros para el mismo.

Como puede apreciarse en la marca 1 de la Figura 16, existe un botón que permite al usuario volver a la ventana principal de selección de pacientes, para acceder a otras opciones de la herramienta o seleccionar otro paciente distinto para su visualización. Para no olvidar que paciente se ha seleccionado, éste aparecerá en la parte superior como se aprecia en la Figura 16.

A continuación, se explican las distintas funcionalidades de la ventana mostrada.

3.2.1 Selección del ancho de banda a visualizar

Tal como ha sido comentado en apartados anteriores, una vez procesada la señal original se obtiene la señal filtrada en distintos anchos de banda. Los distintos anchos de banda posibles se pueden seleccionar en el desplegable que se activa al pulsar sobre la marca 11 de la Figura 6. A modo ilustrativo, se muestra mediante la Figura 17 los distintos anchos de banda que permite seleccionar el programa.

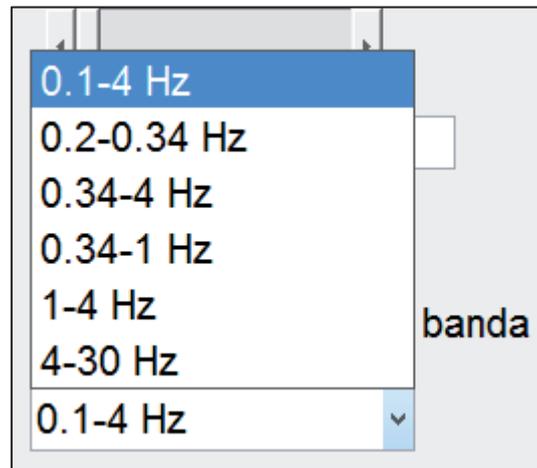


Figura 17. Anchos de banda disponibles para una señal filtrada por el programa.

3.2.2 Ajuste de ejes de visualización

Mediante las herramientas mostradas por las marcas 2, 3, 4 y 5 de la Figura 16, se permite el ajuste de los límites de visualización de los ejes de ordenadas y abscisas. El hecho de que se trate de cuatro herramientas es porque se tiene que poder ajustar tanto límite superior como inferior de ambos ejes. Los valores definidos por defecto son 200 y -200 para el límite superior e inferior respectivamente del eje de ordenadas, y 0 y 1500 para los límites del eje de abscisas. Existen dos maneras de llevar a cabo el ajuste de los límites de los ejes:

Introducción del valor límite de los ejes de forma manual: para ello se encuentran habilitados unos cuadros de texto que permiten la introducción del valor que sea desea fijar. Dichos cuadros de texto se muestran a continuación en la Figura 18.



Figura 18. Cuadro de texto para la introducción del valor

Ajuste de ejes mediante slider: mediante el slider habilitado para cada límite, empleando el cursor puede ajustarse de forma gráfica el valor límite de los ejes. Esta función tiene como inconveniente que no se puede fijar un valor determinado, pero para solventar esto se encuentra disponible la opción anterior. En la Figura 19 se muestra un slider como ejemplo.



Figura 19. Ejemplo slider ajuste de valor límite de ejes

3.2.3 Visualizar segmentación

Para la visualización asociada a la segmentación de cada paciente se habilitan los botones mostrados mediante la marca 6 y 8 de la Figura 16. Dichos botones permiten visualizar sobre la gráfica original la segmentación de registro completo y de contracción, en caso de que exista en la base de datos. La segmentación de registro completo se visualiza de color azul, mientras que la de contracción se hace de color verde y con los fragmentos de señal artefactada en color rojo.

A modo de ejemplo, se muestra mediante las Figura 20 y Figura 21 para cada opción seleccionada la visualización de la segmentación obtenida.

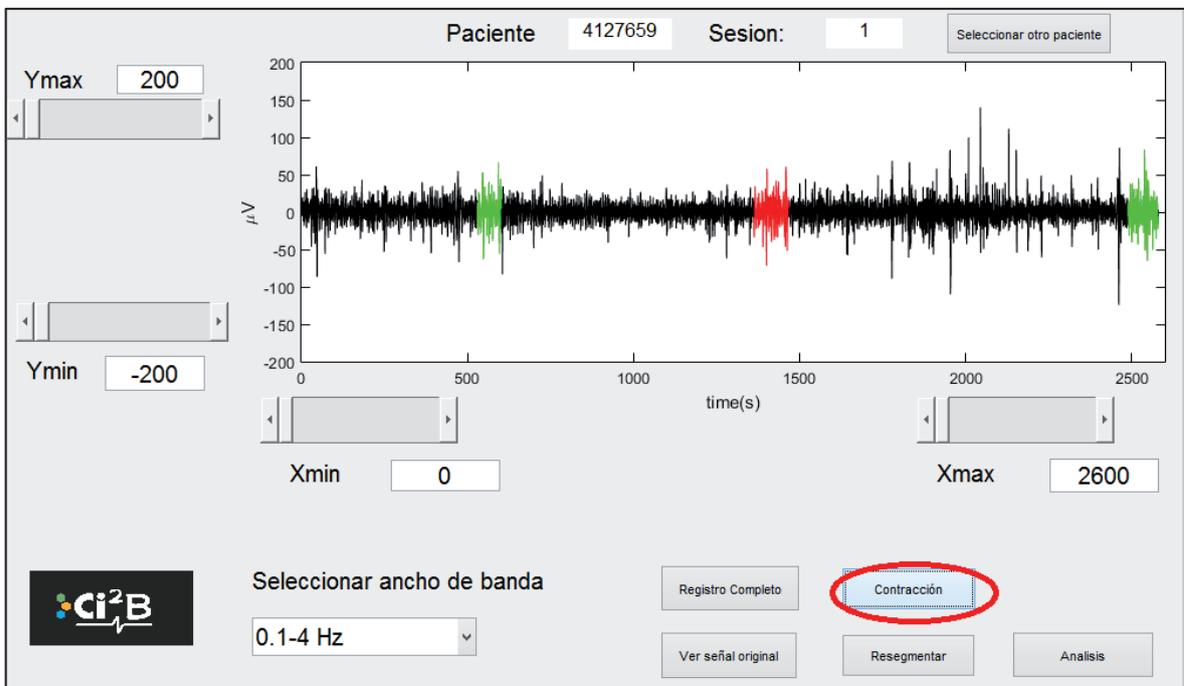


Figura 20. Ejemplo visualización segmentación de contracción

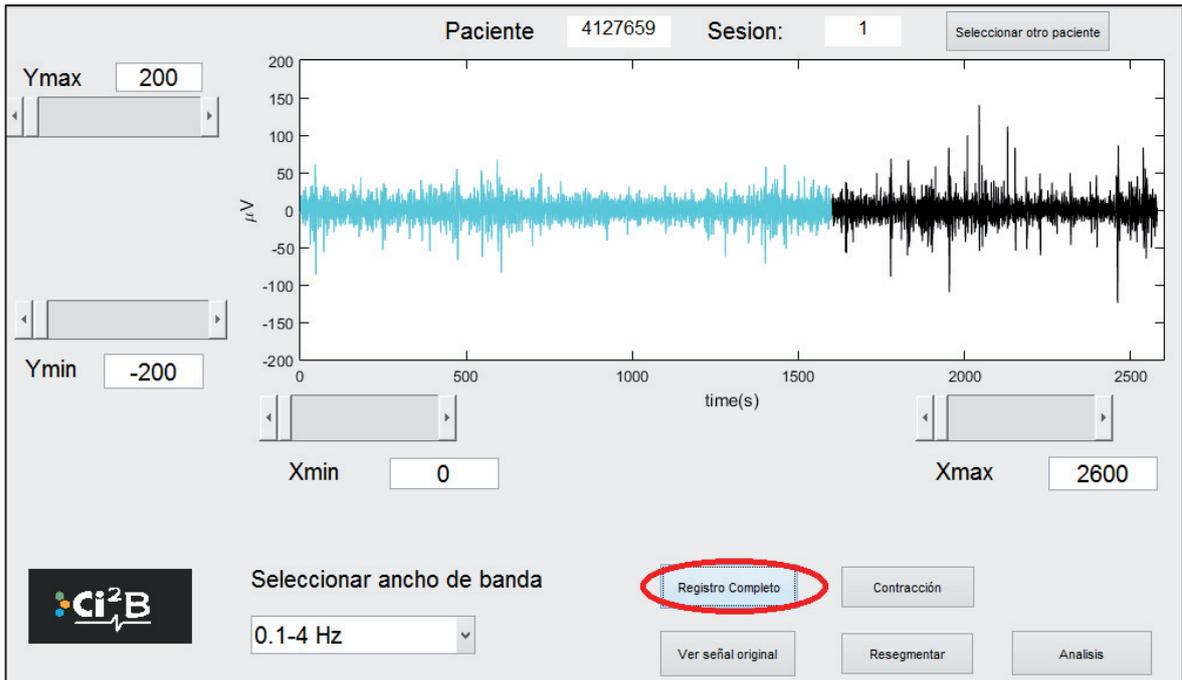


Figura 21. Ejemplo visualización segmentación de registro completo.

En caso de que se seleccione alguna de las dos opciones anteriores y no exista una segmentación asociada, aparecerá el mensaje de error mostrado en la Figura 22 como advertencia.



Figura 22. Mensaje de advertencia cuando no existe segmentación asociada.

Si se desea volver a ver la señal original sin cargar la segmentación, se pulsa el botón indicado en la marca 7 de la Figura 6 “Ver señal original”.

3.2.4 Resegmentar

Esta opción permite al usuario modificar la segmentación existente o crear una nueva en caso de que no exista. Primero debe seleccionar si se desea modificar la segmentación de contracción o de registro completo, mediante los botones mostrados en las marcas 6 y 8 de la Figura 6. A continuación, al pulsar sobre el botón “resegmentar” cuya marca es la 9 de la Figura 6, se continuará a ventana que se describe en el apartado 3 de este manual.

3.2.5 Análisis

Desde la ventana de visualización se puede continuar a la ventana de análisis mediante el botón habilitado que puede apreciarse en la marca 10 de la Figura 6. Puesto que para la visualización se ha seleccionado un paciente en concreto, a la hora de efectuar el análisis aparecerá seleccionado el mismo paciente que para la visualización.

3.3 Resegmentar

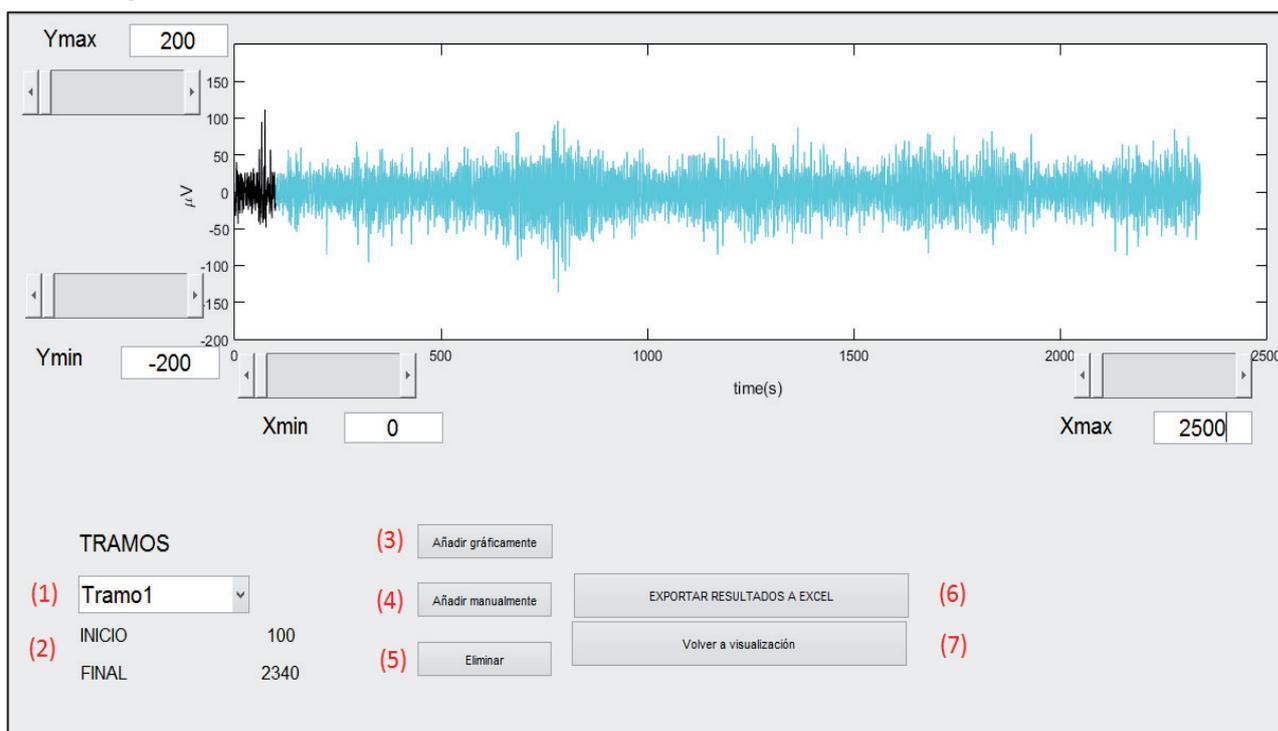


Figura 16. Ventana resegmentar.

La ventana mostrada en la Figura 22 se emplea para modificar la segmentación existente o crear una nueva en caso de que no exista. Para ello, se dispone de las opciones de marca 1, 2, 3, 4 y 5 de la Figura 22 que permiten al usuario la modificación de la segmentación. Para la visualización de la señal se cuenta con las herramientas de ajuste de los límites de los ejes, cuyo funcionamiento ha sido explicado en el apartado 2.2 de este documento.

A continuación se describe las funcionalidades de la ventana y de sus distintos elementos.

3.3.1 Edición y creación de la segmentación

Para editar un tramo es imprescindible su selección previa en la interfaz, para ello se habilita el desplegable de la marca 1 de la Figura 22. Dicho desplegable muestra una lista con los tramos definidos en la segmentación, numerados de izquierda a derecha. El tramo seleccionado se muestra de color amarillo para su distinción. Todo lo mencionado anteriormente se ilustra mediante el ejemplo de la Figura 23.

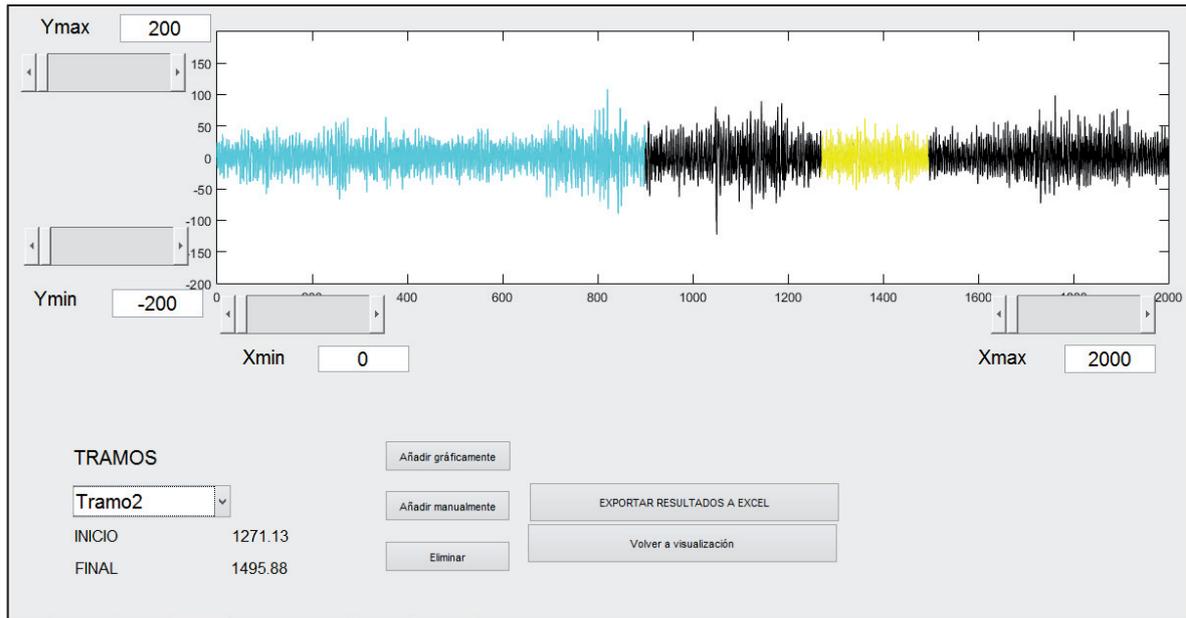


Figura 23. Selección tramo 2 mediante el desplegable para su posterior edición.

Para llevar a cabo la edición de los tramos de la segmentación, se cuenta con diversas herramientas que permiten realizar lo siguiente:

- ***Añadir tramos gráficamente:*** al pulsar sobre el botón que se muestra en la marca 3 de la Figura 22, se habilita un cursor que permite al usuario seleccionar el inicio y el final del tramo deseado sobre la gráfica pulsando el ratón. Dicho proceso se muestra en la Figura 24.

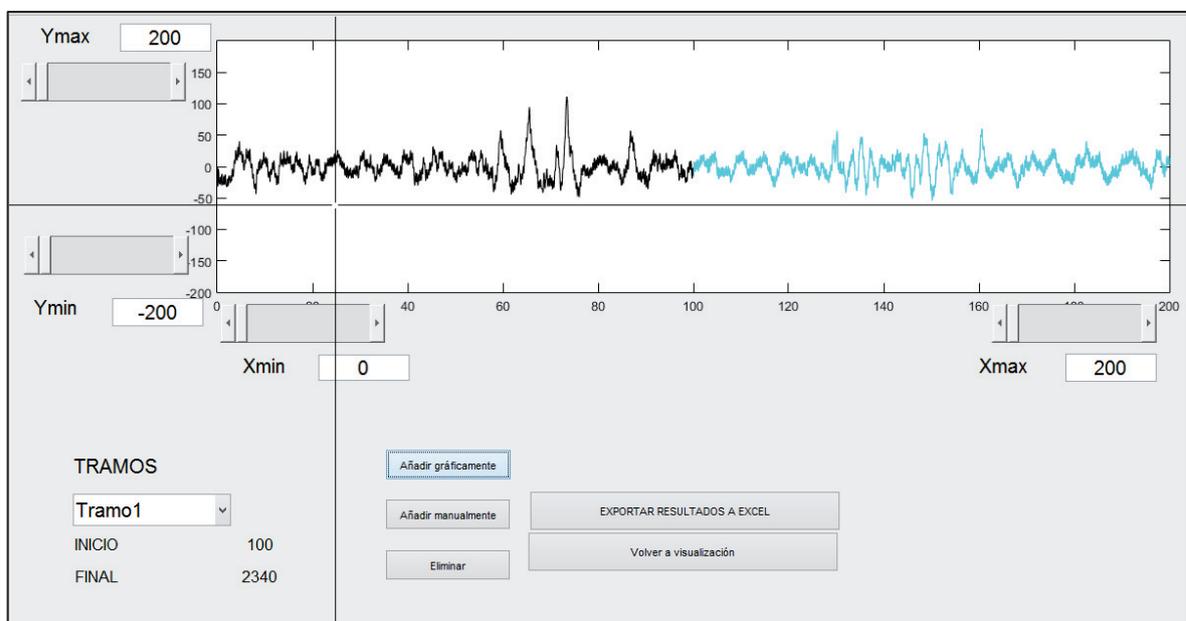


Figura 24. Ejemplo cursor para la definición gráfica de tramos en la segmentación.

- Añadir tramos manualmente: en ocasiones es necesario introducir los tramos de manera manual debido a que la introducción gráfica es imprecisa. Para ello, al pulsar sobre el botón de marca 4 de la Figura 22, aparecen dos ventanas emergentes que permiten la introducción del inicio y del final del tramo, como se muestra en la Figura 25 y Figura 26. Inmediatamente después de su introducción se mostrará el nuevo tramo sobre la gráfica con su color distintivo correspondiente. En caso de introducir un tramo que no sea válido, aparecerá un mensaje de advertencia como se muestra en la Figura 27.

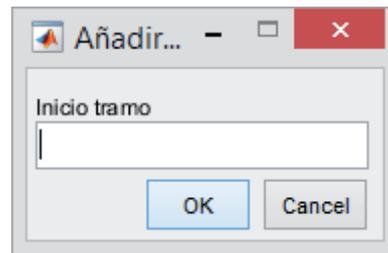


Figura 25. Ventana emergente introducción inicio del tramo.

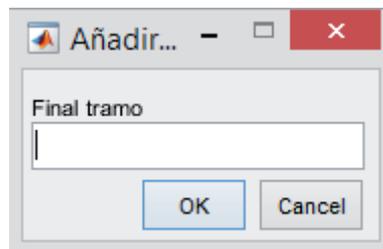


Figura 26. Ventana emergente introducción final de tramo.

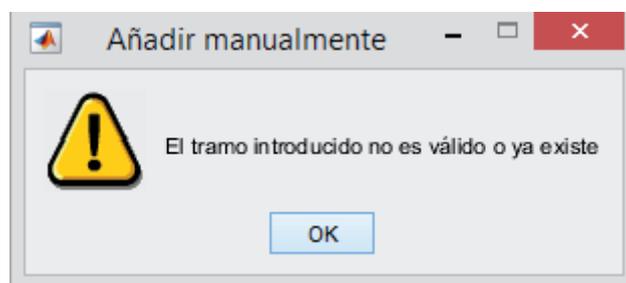


Figura 27. . Mensaje de advertencia si el tramo introducido no es válido o ya se encontraba definido.

- Eliminar tramos: el usuario tiene la opción de eliminar tramos de la segmentación mediante el botón de marca 5 de la Figura 22. Para eliminar un tramo, únicamente es necesario seleccionarlo mediante el desplegable y apretar el botón “Eliminar”.
- Volver a la pantalla anterior: mediante el botón “Volver a visualización” con marca 7 de la Figura 22, se permite al usuario volver a la pantalla anterior para la visualización de la señal.

3.3.2 Exportación de la segmentación modificada

Una vez definida mediante las herramientas anteriores la nueva segmentación, para guardarla se dispone del botón “EXPORTAR RESULTADOS A EXCEL”, como se aprecia en la marca 6 de la Figura 22, que almacena en el fichero Excel correspondiente al paciente la segmentación modificada o creada.

Para evitar la exportación accidental de una segmentación no deseada, en caso de que exista segmentación previamente el programa preguntará al usuario si desea sobrescribir los datos, como puede verse en la Figura 28.

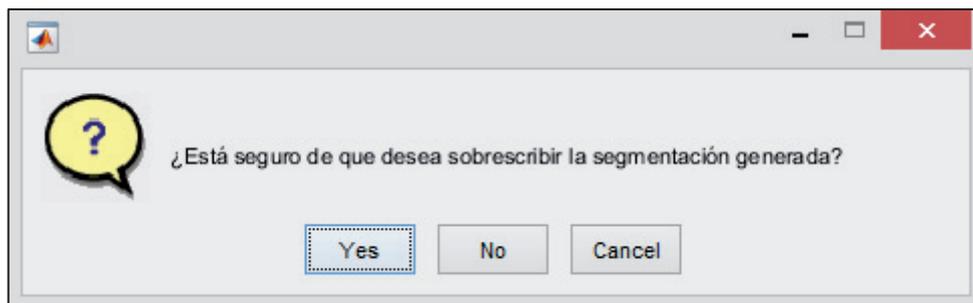


Figura 28. Solicitud de confirmación por parte del usuario de que desea sobrescribir la segmentación.

3.4 Análisis

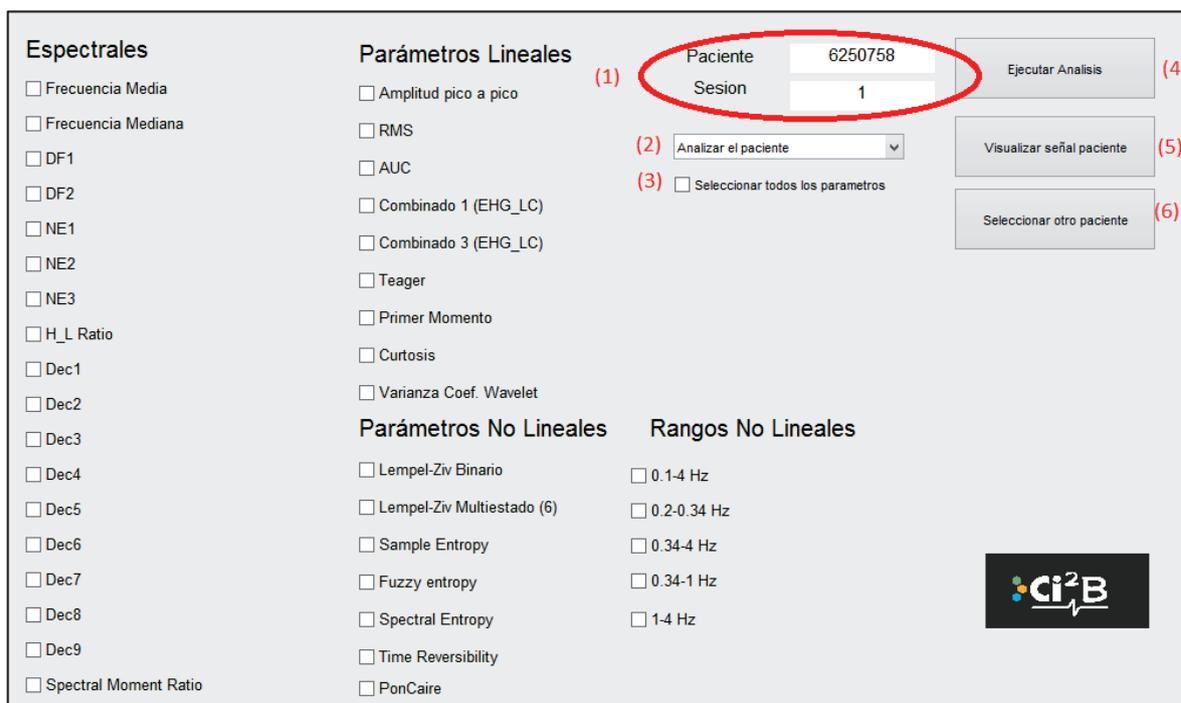


Figura 29. Ventana de análisis de los parámetros

La ventana mostrada en la Figura 29 permite al usuario realizar el análisis de los parámetros deseados y almacenarlos en la base de datos. Para la selección de los parámetros que se desean analizar se cuenta con diversos *checkbox* para su selección. En caso de que se desee marcar toda la lista de parámetros, se debe marcar el *checkbox* correspondiente a la marca 3 de la Figura 29 “Seleccionar todos los parámetros”. Por último, cabe mencionar que los rangos de frecuencia para los que se va a calcular los parámetros no lineales deben ser especificados manualmente.

Mediante esta ventana se puede realizar tanto el análisis de los parámetros de un paciente en concreto como de la base de datos completa.

3.4.1 Análisis de un paciente específico

Es imprescindible haber seleccionado el paciente correctamente en la ventana principal, como se explica en el apartado 1.4 de este documento. Dicho paciente y la sesión seleccionada se muestran en el panel con marca 1 de la Figura 29. En el desplegable de marca 2 de la Figura 29, debe aparecer la opción que se muestra en la Figura 30 “Analizar el paciente”.



Figura 30. Opción seleccionada en el desplegable para realizar el análisis de un paciente específico.

Una vez realizado lo anterior, para ejecutar el análisis debe pulsarse el botón “Ejecutar análisis” que se muestra en la Figura 29 con marca 4. Mientras se ejecuta el análisis puede observarse una barra de progreso como la que se muestra en la Figura 31 y cuando finaliza muestra una ventana emergente como en la Figura 32.

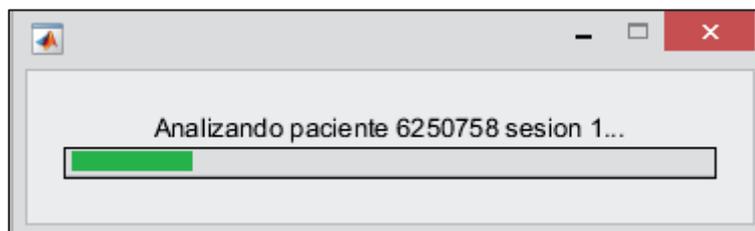


Figura 31. Barra de progreso de análisis del paciente N°6250758 sesión 1.

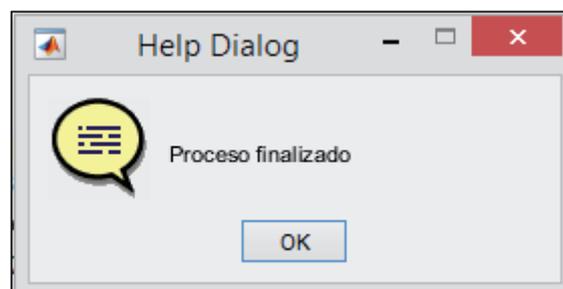


Figura 32. Ventana emergente proceso finalizado.

3.4.2 Análisis de la base de datos completa.

Para llevar a cabo esta operación no es necesario haber seleccionado previamente ningún paciente, únicamente se tiene que seleccionar en el desplegable con marca 2 de la Figura 29, debe aparecer la opción que se muestra en la Figura 33 “Analizar base de datos completa”.

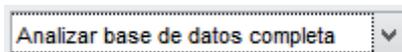


Figura 33. Opción seleccionada en el desplegable para realizar el análisis de la base de datos completa.

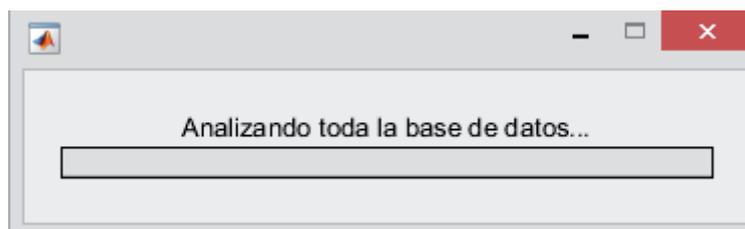


Figura 34. Barra de progreso de análisis de la base de datos completa.

3.4.3 Almacenamiento análisis parámetros.

Dependiendo del tipo de análisis realizado se almacenarán los resultados de una forma u otra. En la Figura 35 mostramos el directorio de almacenamiento de los distintos análisis, para cada opción se genera un fichero Excel y un fichero de extensión .mat dónde se almacenan los resultados.

Nombre	Fecha de modifica...	Tipo	Tamaño
2307251	17/06/2018 11:07	Carpeta de archivos	
2392963	17/06/2018 11:09	Carpeta de archivos	
6250758	23/06/2018 18:25	Carpeta de archivos	
Analisis	24/06/2018 11:56	Microsoft Access ...	19 KB
Analisis_particulares	23/06/2018 18:25	Hoja de cálculo d...	11 KB
Analisisistotal	21/06/2018 16:17	Hoja de cálculo d...	63 KB

Figura 35. Almacenamiento interno de los análisis calculados.

Los directorios cuyo nombre corresponde al número de historia clínica del paciente contienen el archivo de extensión .mat con los parámetros calculados para dicho paciente. Los datos se vuelcan además al fichero Excel de “Análisis_particulares”.

Por otra parte, el fichero “Análisis” y “Análisisistotal” almacenan los parámetros calculados para toda la base de datos.

3.5 Distribución de las variables

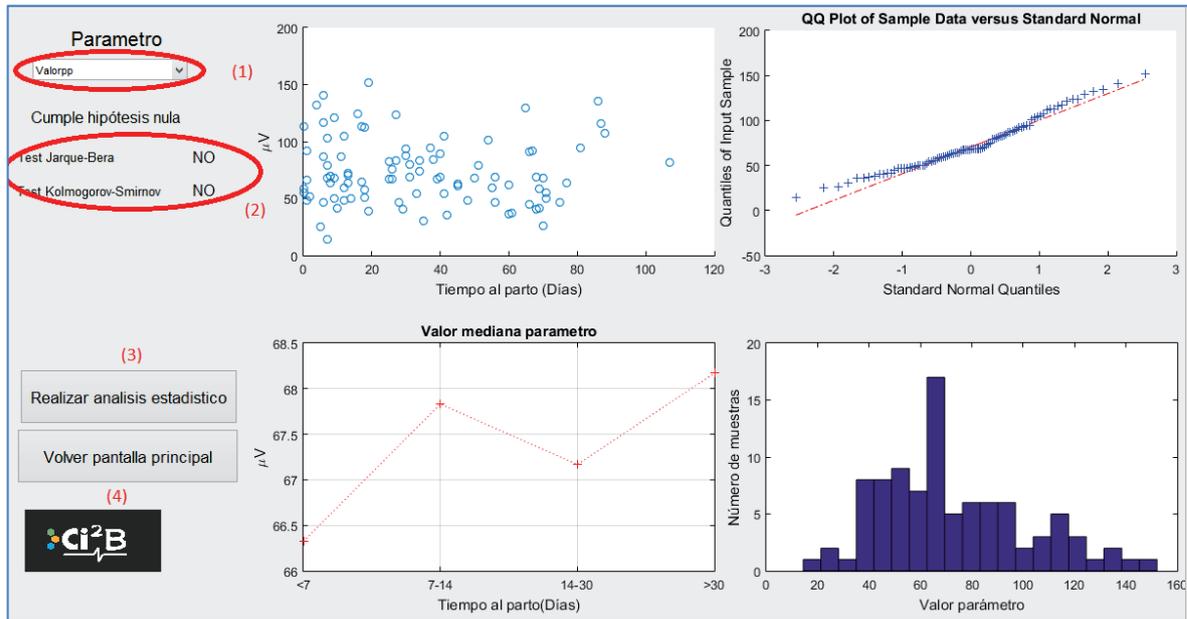


Figura 36. Ventana visualización de distribución de las variables.

La ventana que se muestra en la Figura 36 proporciona información sobre la distribución de cada parámetro. Para caracterizar dicha distribución se emplea un diagrama de dispersión, un gráfico Q-Q, una representación de las medianas de cada parámetro en función del tiempo al parto en días y un histograma.

Para seleccionar cada parámetro del cual se desea visualizar la distribución se dispone del desplegable mostrado en la marca 1 de la Figura 36. Dicho desplegable muestra únicamente los parámetros que se encuentran calculados en el análisis de la base de datos completa.

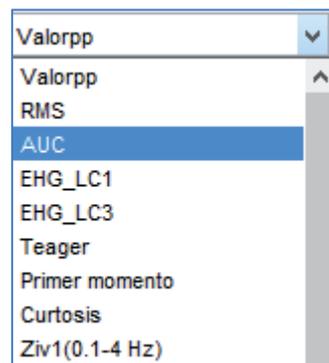


Figura 37. Desplegable con los parámetros disponibles para su visualización.

Mediante el botón de marca 4 de la Figura 36 se permite al usuario volver a la pantalla principal. Si se desea continuar a la elaboración del análisis estadístico de los datos, existe para ello el botón mostrado mediante la marca 3 de la Figura 36.

3.6 Análisis estadístico

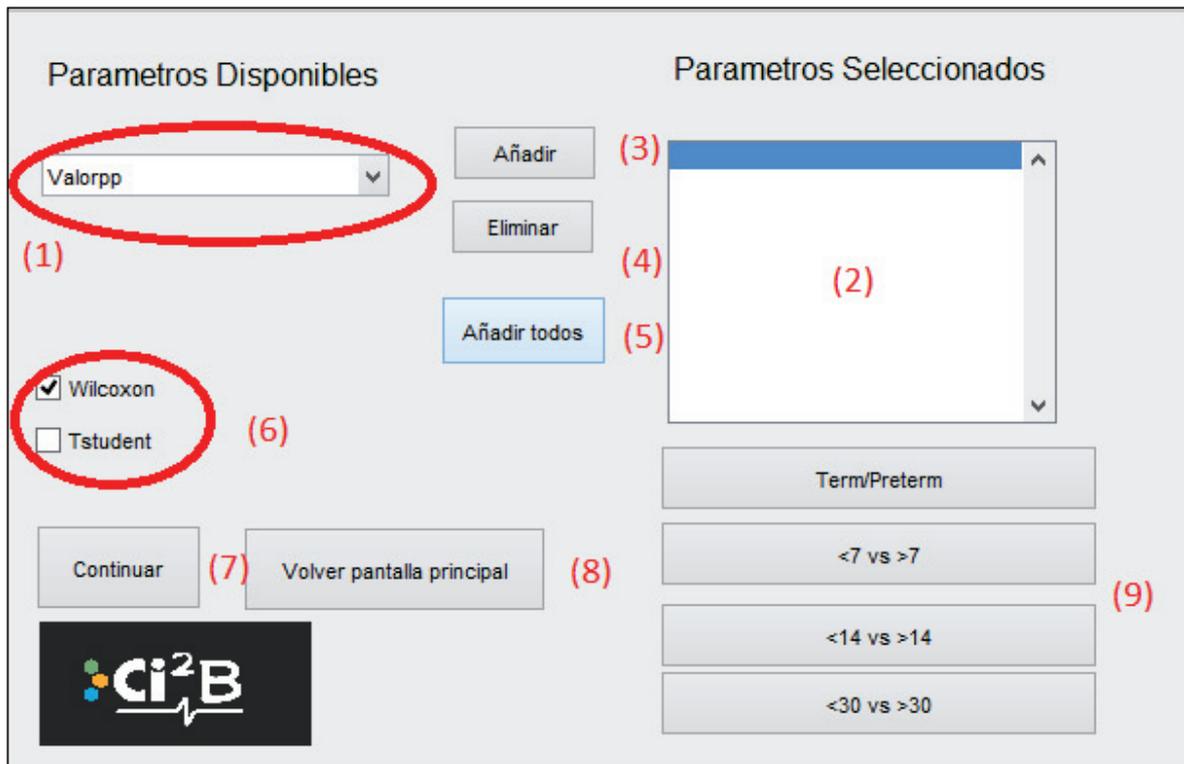


Figura 38. Ventana análisis estadístico.

La ventana que se muestra en la Figura 38 permite realizar un análisis estadístico de los datos seleccionando el grupo de clasificación. Dicho análisis genera diagramas de caja bigotes y un informe estadístico de los parámetros seleccionados. A continuación se explica cómo lleva a cabo dicho análisis.

3.6.1 Selección de parámetros a analizar

Mediante el desplegable mostrado en la Figura 38 con la marca 1, se muestran los parámetros que se encuentran disponibles en el análisis de parámetros. Por defecto, no se selecciona ningún parámetro para el análisis estadístico.

El listado mostrado mediante la marca 2 de la Figura 38 muestra los parámetros seleccionados para llevar a cabo el análisis estadístico. Para añadir un parámetro basta con seleccionarlo en el desplegable y pulsar el botón “Añadir” (Ver marca 3 Figura 38). A su vez, si se desea eliminar un parámetro del listado de marca 2, se debe seleccionar con el desplegable y pulsar el botón “Eliminar” (Ver marca 4 Figura 38). Por último, si se desea añadir todos los parámetros disponibles, basta con pulsar sobre “Añadir todos” (Ver marca 5 Figura 38).

3.6.2 Realizar análisis estadístico

Los cuatro botones visibles mediante la marca 9 de la Figura 38 muestran los cuatro grupos de clasificación disponibles para la realizar el análisis estadístico. Por defecto, para realizar el análisis solamente se mostrará el pvalor de Wilcoxon, si se desea calcular y visualizar el pvalor de Tstudent será necesario marcar el checkbox correspondiente en la marca 6 de la Figura 38. Cabe mencionar que solamente se calculará el p-valor de Tstudent

En caso de que se apriete uno de los cuatro botones de los grupos de clasificación para realizar el análisis estadístico y no se haya seleccionado ningún parámetro, aparecerá el mensaje de advertencia mostrado en la Figura 39.

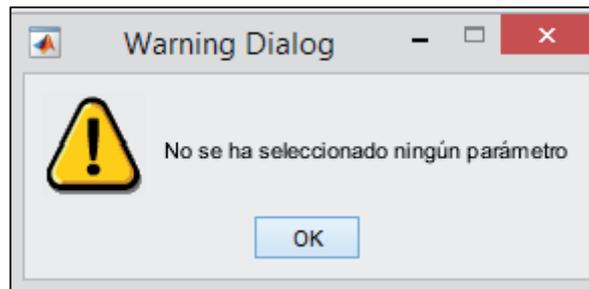


Figura 39. Mensaje de advertencia en caso de no haber seleccionado ningún parámetro.

Si se ha realizado correctamente la selección de parámetros, al pulsar sobre uno de los cuatro botones de grupos de clasificación, aparecerán los diagramas de caja bigotes de los parámetros seleccionados y se generará un informe en la ventana siguiente.

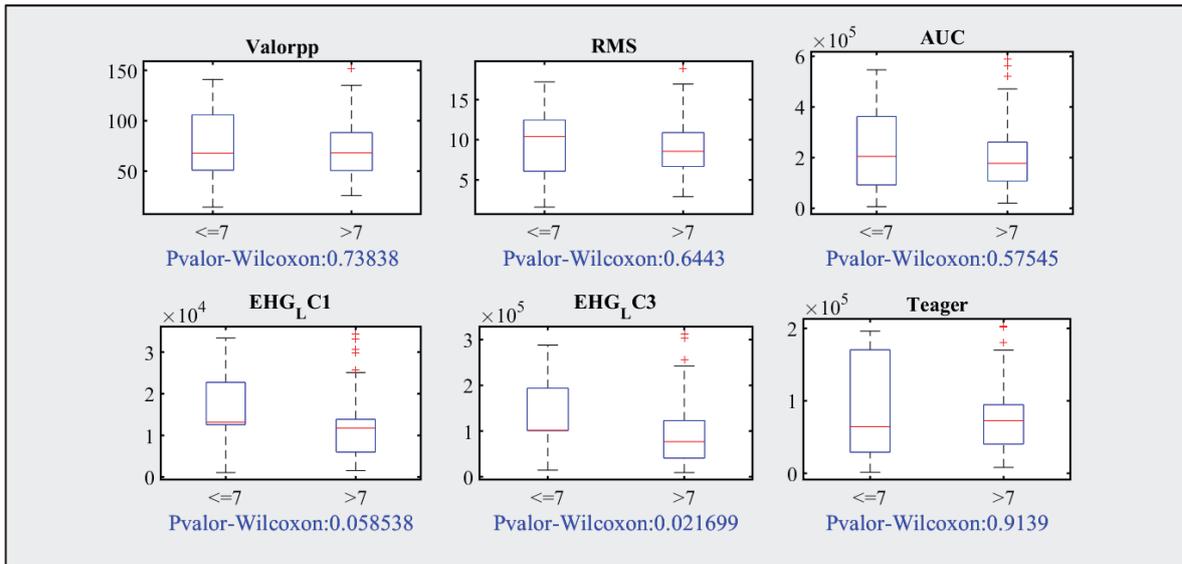


Figura 40. Ejemplo diagrama caja-bigotes generado.

Informe estadístico

[Exportar a Excel](#)

	A término		Pretérmino		
	Media	Rango Intercuartílico	Media	Rango Intercuartílico	P-valor
Time Reversibility(0,1-4 Hz)	2.8223	1.8188	3.653	1.713	0.0023026
EHG _{LC3}	85596.0677	70148.245	117394.9681	74716.8918	0.0094206
EHG _{LC1}	11502.3307	8729.2705	13917.3221	9482.4688	0.093015
Ziv1(0,1-4 Hz)	0.44291	0.11559	0.41893	0.11504	0.098401
DF2	0.36082	0.021973	0.36441	0.019531	0.12779
AUC	182682.9714	166188.6304	230919.8569	212287.9418	0.16052
RMS	8.6802	4.9289	9.4551	5.6078	0.22621
PON3	17.3271	11.7833	19.2553	10.7453	0.23803
PON2	24.2477	16.496	26.9875	15.2179	0.25028
Valorpp	70.2529	40.034	76.243	36.6855	0.28964
NE2	0.28765	0.07609	0.30283	0.081534	0.30721
Dec6	0.33132	0.026855	0.33463	0.026855	0.32276
PON4	0.13672	0.056531	0.13153	0.050887	0.32547
PON1	3.0559	1.8357	3.4062	2.4245	0.33675
Fuzzy entropy(0,1-4 Hz)	0.31257	0.11005	0.30078	0.069111	0.39904
Spectral Moment Ratio	3.4788	0.56155	3.3499	0.86525	0.40969
Dec2	0.23432	0.013428	0.2366	0.01709	0.43682
Dec5	0.30416	0.029297	0.30592	0.029297	0.44654
Primer momento	6499569.3723	6497275.1438	7404461.1622	5084605.9658	0.45371
Dec8	0.429	0.054932	0.42465	0.061035	0.48399

Figura 41. Ventana informe estadístico generado.

La ventana que se muestra en la Figura 41 se genera automáticamente al realizar el análisis estadístico que se explica en el apartado anterior. La ventana muestra la media y el rango intercuartílico para el grupo de clasificación seleccionado anteriormente (marca 2 Figura 41). Los parámetros mostrados son los seleccionados en la ventana de “Análisis estadístico” y se ordenan según su pvalor de menor a mayor.

3.6.3 Exportación informe estadístico y p-valor calculado

Mediante el botón habilitado con marca 1 de la Figura 41, el usuario puede guardar el informe generado en un fichero Excel pulsando el botón correspondiente. El directorio de almacenamiento depende del grupo de clasificación escogido, se almacena en él el fichero Excel junto con el fichero .mat que contiene la información de los pvalor calculados para los parámetros seleccionados.

Este equipo > Documentos > TFG > Analisis_estadistico > TermPreterm

Nombre	Fecha de modifica...	Tipo	Tamaño
Informe_estadistico	21/06/2018 10:52	Hoja de cálculo d...	14 KB
Pvalor_almacenados	24/06/2018 11:57	Microsoft Access ...	1 KB

Figura 42. Ejemplo almacenamiento informe y pvalor calculados para grupo de clasificación término y pretérmino.

3.7 Generación de modelos predictivos

Grupo de clasificación (1)

Preterm vs Term

<7 vs >7

<14 vs >14

<30 vs >30

Reducción de dimensión (2)

PCA

Algoritmo genético

PSO

SFFS

Método de clasificación (3)

Regresión logística

Redes neuronales

SVM

Árbol de decisión

Análisis ROC (4)

Figura 43. Ventana para la generación de modelos predictivos.

Para generar los modelos predictivos deseados, es imprescindible realizar los tres pasos siguientes:

- **Seleccionar grupo de clasificación:** es necesario seleccionar uno de los grupos disponibles para generar el modelo predictivo. Los grupos disponibles son cuatro, uno diferenciando entre término y pretérmino y el resto en función del tiempo al parto en días como se observa en la marca 1 de la Figura 43.
- **Reducción de dimensión:** actualmente, el programa cuenta únicamente con la reducción de dimensión PCA para llevar a cabo la generación de modelos predictivos. Es necesaria su selección para poder continuar, en caso de seleccionar una opción de las que no se encuentran implementadas, aparece el mensaje de advertencia que se muestra en la Figura 44.

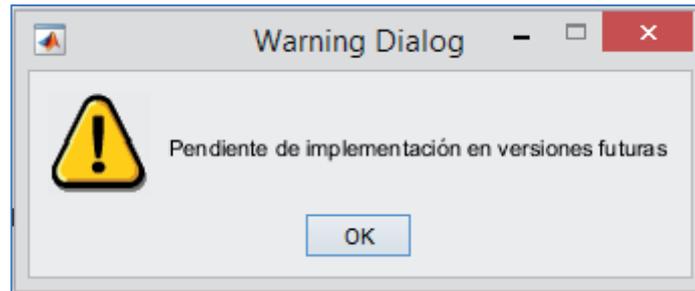


Figura 44. Mensaje de advertencia para los métodos no implementados.

- Seleccionar el método de clasificación: el programa tiene implementadas dos opciones, regresión logística y redes neuronales. A continuación se explica la selección para cada caso. Si se selecciona un método de los que no están implementados, aparecerá el mensaje de advertencia de la Figura 44.

3.7.1 Regresión logística

Una vez seleccionado el *checkbox* correspondiente a la regresión logística en el panel de marca 3 de la Figura 43, aparece el panel mostrado en la Figura 45 que solicita al usuario la introducción de los porcentajes de datos destinados al grupo de entrenamiento y al grupo de validación.

Método de clasificación	Especificar porcentaje	
	Entrenamiento	Validación
<input checked="" type="checkbox"/> Regresión logística		
<input type="checkbox"/> Redes neuronales		
<input type="checkbox"/> SVM		
<input type="checkbox"/> Árbol de decisión		

Figura 45. Panel mostrado para la introducción de los porcentajes de datos de entrenamiento y validación en regresión logística.

Es evidente que la suma de las proporciones debe resultar el 100%, en caso de introducir proporciones que no lo cumplan, el programa mostrará la siguiente advertencia (ver Figura 46) cuando se decida proceder al análisis ROC.

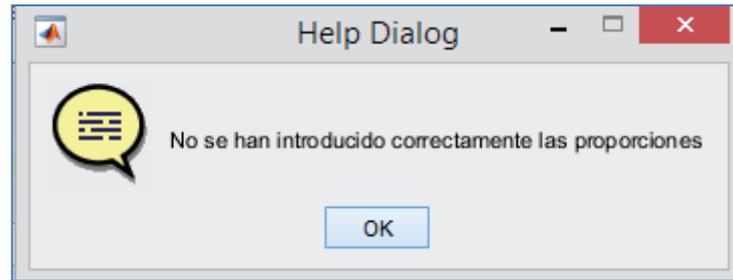


Figura 46. Mensaje de advertencia por introducción de proporciones incorrectas.

3.7.2 Redes neuronales

Una vez seleccionado el *checkbox* correspondiente a redes neuronales en el panel de marca 3 de la Figura 43, aparece el panel mostrado en la figura X que solicita al usuario la introducción de los porcentajes de datos destinados al grupo de entrenamiento, de validación y test.

Método de clasificación	Especificar porcentaje		
	Entrenamiento	Validacion	test
<input type="checkbox"/> Regresión logística			
<input checked="" type="checkbox"/> Redes neuronales	<input type="text" value="30"/>	<input type="text" value="30"/>	<input type="text"/>
<input type="checkbox"/> SVM			
<input type="checkbox"/> Árbol de decisión			
	Topología neuronas		
	<input type="text" value="1"/>		

Figura 47. Panel mostrado para la introducción de los porcentajes de datos de entrenamiento, validación y test en redes neuronales.

La suma de los porcentajes de los tres grupos debe ser el 100%, en caso contrario aparecerá el mensaje de advertencia de la figura X. A demás, debe especificarse la topología de las neuronas mediante el desplegable que se observa en la Figura 47.

Las redes generadas por esta ventana se almacenan en el siguiente directorio: Modelos→Grupo de clasificación→PCA→Redes neuronales→Redes generadas.

3.7.3 Continuar análisis ROC

Una vez realizados los pasos anteriores, para continuar a la ventana de análisis ROC se tiene habilitado el botón de marca 4 de la Figura 43.

3.8 Análisis ROC

Esta ventana se abre automáticamente al pulsar sobre el botón de análisis ROC de la ventana de generación de modelos predictivos. Dependiendo del método de clasificación escogido, la ventana mostrará unos resultados u otros, como se aprecia en la Figura 48 y Figura 49.

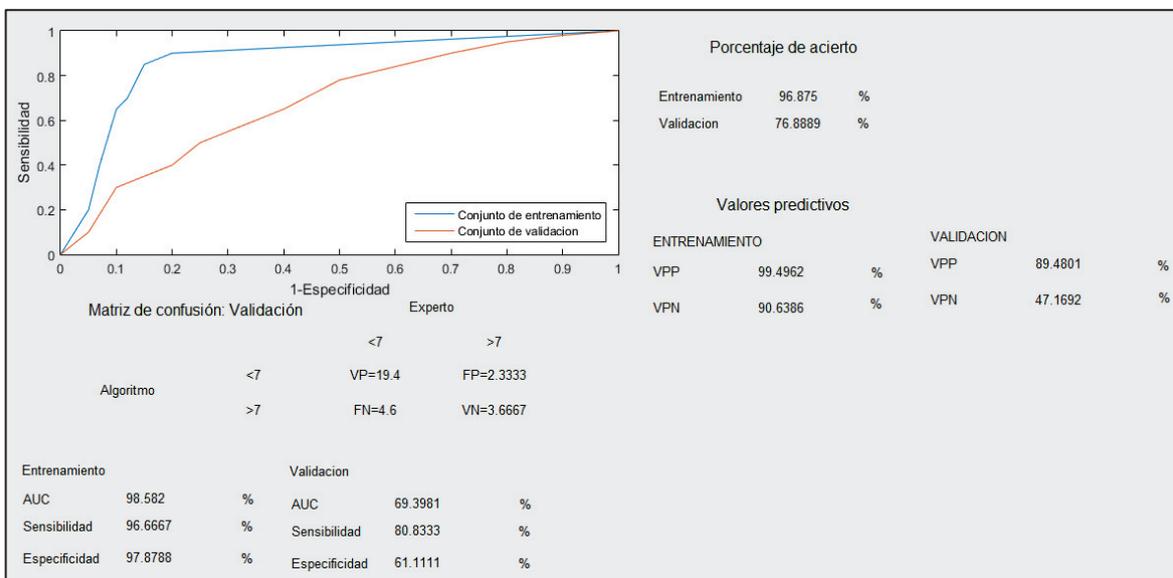


Figura 48. Ejemplo neuronales.ventana análisis ROC para redes

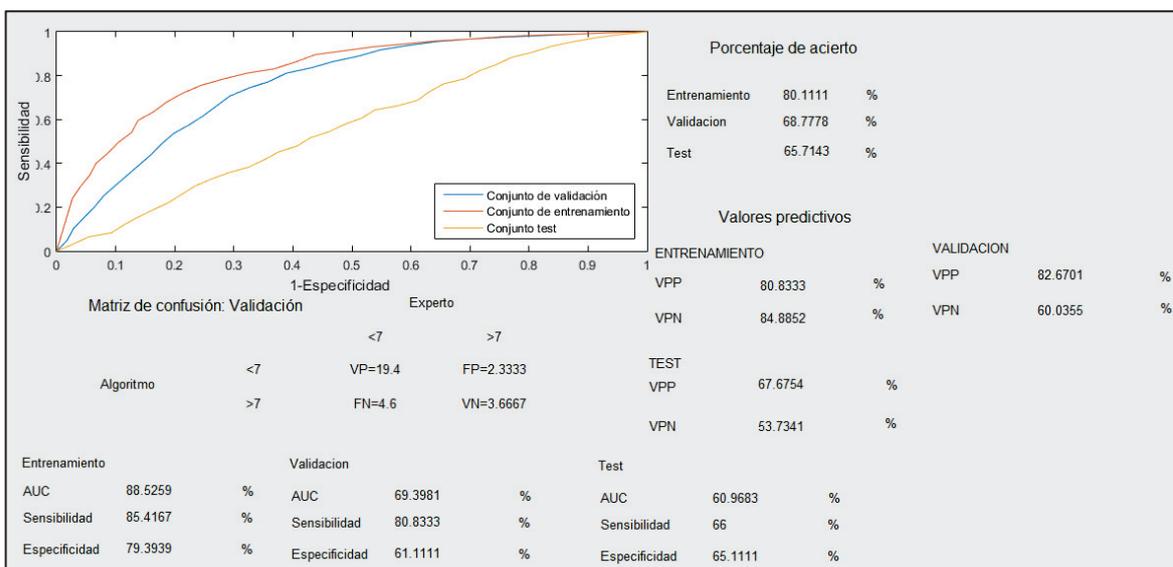


Figura 49. Ejemplo ventana análisis ROC para redes neuronales.



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Desarrollo e implementación de una aplicación para la
caracterización, el procesado de y análisis de señales
electrohisterográficas en entorno MATLAB



ESCUELA TÉCNICA
SUPERIOR INGENIEROS
INDUSTRIALES VALENCIA

PORTADA DEL PRESUPUESTO



Índice del presupuesto

1. DESARROLLO DEL TRABAJO.....	2
2. MATERIALES Y RECURSOS EMPLEADOS	3
3. PRESUPUESTO TOTAL.....	5

El presupuesto de este trabajo de fin de grado se ha dividido en dos partes, el desarrollo del trabajo, y los recursos y materiales empleados.

1. DESARROLLO DEL TRABAJO

Desarrollo del trabajo				
Descripción	Unidad	Medición	Precio unitario	Importe
Formación en el entorno gráfico GUIDE de Matlab	h	20	10€/h	200€
Desarrollo de la interfaz de usuario	h	220	30€/h	6600€
Documentación y elaboración de la memoria	h	60	15€/h	900€
Costes directos				7700€
Costes directos complementarios (2%)				154€
Total				7854€

1. Formación en el entorno gráfico GUIDE de Matlab.

Para el desarrollo del presente trabajo se contaba con los conocimientos necesarios de programación en Matlab pero ha sido necesario un tiempo de formación por parte del desarrollador del trabajo para adquirir los conocimientos necesarios que permiten implementar dicha programación en un entorno gráfico. (vamos a estimar 20h y precio 10).

2. Desarrollo de la interfaz de usuario

Por tratarse de la sección más técnica y requerir conocimientos avanzados en el entorno matemático Matlab, se estima un precio unitario de 30€/h. Se trata de la actividad principal del trabajo y ello se ve reflejado en la cantidad de tiempo requerido.

3. Documentación y elaboración de la memoria

Se estima un precio unitario de 15€/h puesto que se trata de un trabajo de redacción y elaboración de memoria básico.

2. MATERIALES Y RECURSOS EMPLEADOS

Materiales y recursos empleados				
Descripción	Unidad	Medición	Precio unitario	Importe
Ordenador portátil Asus X555LD	h	300	0.05€/h	15€
Matlab r2016b: licencia	h	220	0.16€/h	36.6€
Microsoft Office 2016 versión estudiantes	h	60	0.012€/h	0.72€
Módulos de amplificación y adquisición de diseño propio	año	3	80€/año	240€
Electrodos 3M Red Dot 2560	ud	352	1€/ud	352€
Pack Tubos Nuprep	ud	88	10€/ud	880€
Costes directos				1524,32€
Costes directos complementarios (2%)				30,49€
Total				1554,8€

1. Ordenador portátil Asus X555LD

El precio de compra de este equipo es de 600€, con una vida útil estimada para esta tecnología de 6 años. Puesto que un año tiene aproximadamente 250 días laborables, ello supone que la vida útil de este equipo es 1500 días laborables. Suponiendo la jornada laboral de 8 horas, el total de horas laborables durante la vida útil del ordenador será de 12.000 horas. Luego, el precio unitario se calcula mediante el cociente del precio de adquisición del ordenador y las horas laborables para las cuales funcionará en su vida útil. Esto es 0.05€/h.

2. Licencia Matlab r2016b



La licencia perpetua tiene un coste de 2000€. Dicha licencia se extiende a la vida útil del ordenador empleado, luego el precio unitario se calcula mediante el cociente del coste de la licencia y las 12000 horas laborables de uso. Esto resulta 0.16€/h.

3. Microsoft Office 2016 para estudiantes

El precio de la licencia es de 149€ para un ordenador, luego su amortización se extiende a la vida útil del dispositivo. El precio unitario se calcula mediante el cociente del coste de la licencia y las 12000 horas laborables de uso. Esto es 0.012€/h.

4. Módulos de amplificación y adquisición de diseño propio

El precio de este equipo es de 400€ y su amortización es de 5 años, luego el coste por año del equipo es 80€/año. Para la obtención de los registros de este trabajo han sido necesarios 3 años.

5. Electrodo 3M Red Dot 2560

El precio unitario de un electrodo es de aproximadamente 1€. Para la obtención de cada registro se necesitan 4 electrodos y en este caso se han tomado un total de 88 registros. Por tanto, se necesitarán en total de 352 unidades de electrodos

6. Tubos Nuprep

Para la toma de cada registro se necesitan cuatro tubos de Nuprep, los cuales se venden en packs de 10€. En este caso son necesarios 88 packs puesto que 88 son el número de registros tomados.

3. PRESUPUESTO TOTAL

Finalmente, el coste total del proyecto es el siguiente:

PRESUPUESTO TOTAL	
Descripción	Importe
Desarrollo del trabajo	7854€
Materiales y recursos empleados	1554.8€
Presupuesto de ejecución material	9408.8€
Beneficio industrial (6%)	564.528€
Gastos generales (13%)	1223.14€
Presupuesto de ejecución por contrata	11196.47€
IVA (21%)	2351.25€
Coste total del proyecto	13547.72€

El coste total del proyecto asciende a la suma de TRECE MIL QUINIENTOS CUARENTA Y SIETE CON SETENTA Y DOS CÉNTIMOS.