



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



ESCUELA TÉCNICA
SUPERIOR INGENIEROS
INDUSTRIALES VALENCIA

TRABAJO FIN DE GRADO EN INGENIERÍA BIOMÉDICA

DESARROLLO DE UN SISTEMA DE PREDICCIÓN DEL PARTO PREMATURO BASADO EN ELECTROHISTEROGRAFÍA Y ALGORITMO GENÉTICO PARA LA SELECCIÓN DE CARACTERÍSTICAS ÓPTIMAS

AUTOR: David García Borillo

TUTORA: Yiyao Ye Lin

COTUTORA: Gema Prats Boluda

Curso Académico: 2019-20

AGRADECIMIENTOS

“Me gustaría aprovechar estas líneas para agradecer y dedicar este trabajo...

A mis padres, por el apoyo incondicional recibido por ellos desde el día en el que nací. Gracias.

A Marta, mi mujer, sin su ayuda hubiera sido totalmente imposible conseguir este sueño que tanto ansiaba. Solo puedo decirte que gracias por estar ahí cuando el miedo me invadía, gracias por tenderme la mano cuando lo necesitaba, gracias por haber hecho esto un poco más fácil.

A Lola, mi mejor amiga, que me ha enseñado tanto con tan pocas palabras. Gracias por aparecer en mi vida.

A mis tutores, Yiyao, Gema y Javier por su paciencia conmigo y su gran trabajo diario. Sin su esfuerzo no hubiera sido posible realizar este trabajo. Gracias de corazón.

A cada una de las personas que han aparecido en este periodo tan importante de mi vida, y que de forma altruista, me han guiado en el camino.”

RESUMEN

El parto prematuro es una de las complicaciones más comunes y la primera causa de muerte en niños menores de 5 años, una de las causas que mayores complicaciones de salud provocará a lo largo de la vida del sujeto, así como un gran coste económico para la sociedad.

Los mecanismos que inducen este suceso todavía no están claros. Debido a que las contracciones uterinas son dadas por la actividad eléctrica del miometrio, su medición de manera externa, pueden ayudar a entender los mecanismos de inducción al parto y el origen del parto prematuro. Para ello, se introduce la técnica conocida como electromiograma uterino (EMG) o electrohisterograma (EHG), la cual se basa en el registro no invasivo de la actividad eléctrica relacionada con la contracción de las células miometriales del útero. Son diversos los estudios que han demostrado que mediante las características obtenidas por EHG es posible discernir entre contracciones “verdaderas” o “falsas” en los embarazos a término y prematuros. El objetivo del presente estudio es desarrollar un sistema de predicción del parto prematuro en mujeres de control rutinario en base a las características de EHG utilizando algoritmos computacionalmente eficientes. Para ello, se han comparado distintas técnicas de reducción de dimensionalidad para determinar las características relevantes: Análisis de Componentes Principales (PCA) y Algoritmo Genético (AG). Además, se han implementado y comparado los siguientes métodos de clasificación: Análisis Discriminante Lineal (LDA), Regresión Logística (RL), K vecinos más próximos (KNN) y *Extreme Learning Machine* (ELM). Los resultados muestran que el algoritmo genético permite obtener la combinación óptima de características con información complementaria para LDA y RL, obteniendo un F1-Score promedio de 89,03% y 85,78% en el grupo de test respectivamente.

En cambio, PCA ha obtenido mejores resultados que el algoritmo genético cuando se emplea KNN como método de clasificación, obteniendo un F1-Score promedio de 90,19%. En base a estos clasificadores individuales, se ha implementado un clasificador combinado con la técnica de votación por mayoría, obteniendo un F1-Score promedio del 93% en el grupo de test. Estos resultados sugieren la posibilidad de desarrollar sistemas de predicción del parto prematuro precisos y generalizables en mujeres de control rutinario en base a las características de EHG. Además, estos sistemas podrían tener una gran aceptabilidad por el personal clínico al ser métodos simples y de uso clínico, mejorando así la transferibilidad de la técnica de EHG a la praxis clínica.

Palabras Clave: electrohisterograma, pretérmino, Algoritmo Genético, *Extreme Machine Learning*, predicción del parto prematuro.

RESUM

El part prematur és una de les complicacions més habituals i la principal causa de mort en xiquets menors de cinc anys, un dels motius que permetran complicacions importants per a la salut al llarg de la vida del subjecte, així com un gran cost econòmic per a la societat.

Els mecanismes que indueixen aquest esdeveniment encara no estan clars. Com que l'activitat elèctrica del miometri dona contracció uterina, la seua mesura externa pot ajudar a comprendre els mecanismes d'inducció del treball i l'origen del treball prematur. Per a això, s'introdueix la tècnica que es coneix com a electromiograma uterí (EMG), o l'electrohisterograma (EHG), que es basa en el registre no invasiu d'activitat elèctrica relacionada amb la contracció de cèl·lules miometrials a l'úter. Diversos estudis han demostrat que utilitzant les característiques obtingudes per EHG és possible distingir entre contraccions "verdaderes" o "falses" en embarassos prematurs i a terme. L'objectiu del present estudi és desenvolupar un sistema de predicció del part prematur en dones sota control rutinari basat en les característiques d'EHG mitjançant algorismes eficients computacionalment. Per a això, s'han diferenciat tècniques de reducció de la dimensionalitat per determinar les característiques rellevants: Anàlisi de Components Principals (PCA) i Algoritme Genètic (AG). A més, s'han implementat i comparats els següents mètodes de classificació: Anàlisi Discriminant Lineal (LDA), Regressió Logística (RL), K veïns més pròxims (KNN) i *Extreme Learning Machine* (ELM). Els resultats mostren que l'algoritme genètic permet obtenir l'òptima combinació de característiques amb informació complementària per a LDA i RL, i s'ha aconseguit una puntuació F1 del 89,03% i el 85,78% en el grup de prova, respectivament.

En canvi, PCA ha obtingut millors resultats que l'algoritme genètic quan s'utilitza KNN com a mètode de classificació, ja que ha obtingut una puntuació F1 del 90,19%. Sobre la base d'aquests classificadors individuals, un classificador combinat amb la tècnica de la votació per majoria ha sigut executat, i ha aconseguit una F1-Score de 93% en el grup de prova. Aquests resultats suggereixen la possibilitat de desenvolupar un sistema exacte i generalitzable de predicció del treball prematur en dones controlades de forma rutinària basat en les característiques d'EHG. A més, aquests sistemes podrien tindre una excel·lent acceptabilitat per part del personal clínic, ja que són mètodes senzills i d'ús clínic, que milloraran així la transferibilitat de la tècnica EHG a la praxi clínic.

Paraules clau: electrohisterograma, preterme, algoritme genètic, aprenentatge extrem a màquina, predicció del part prematur.

ABSTRACT

Preterm labour is one of the most common complications and the leading cause of death in children under five years old, one of the reasons that will let significant health complications throughout the life of the subject, as well as a tremendous economic cost for the society.

The mechanisms that induce this event are still unclear because the electrical activity of the myometrium gives uterine contraction; its measurement externally can help to understand the mechanisms of labour induction and the origin of premature labour. For this, the technique is known as uterine electromyogram (EMG), or electrohysterogram (EHG) is introduced, which is based on the non-invasive recording of electrical activity related to the contraction of myometrial cells in the uterus. Several studies have shown that using the characteristics obtained by EHG. It is possible to discern between "true" or "false" contractions in term and premature pregnancies. The objective of the present study is to develop a system for predicting preterm birth in women under routine control based on the characteristics of EHG using computationally efficient algorithms. For this, different dimensionality reduction techniques have been differentiated to determine the relevant characteristics: Principal Component Analysis (PCA) and Genetic Algorithm (GA). Besides, the following classification methods have been implemented and compared: Linear Discriminant Analysis (LDA), Logistic Regression (RL), K-Nearest Neighbors (KNN), Extreme Learning Machine (ELM). The results show that the genetic algorithm allows obtaining the optimal combination of characteristics with complementary information for LDA and RL, reaching an F1-Score of 89.03% and 85.78% in the test group, respectively.

In contrast, PCA has achieved better results than the genetic algorithm when KNN is used as the classification method, obtaining an F1-Score of 90.19%. Based on these individual classifiers, a classifier combined with the majority voting technique has been executed, getting an F1-Score of 93% in the test group. These results suggest the possibility of developing an accurate and generalizable system of prediction of premature labour in routinely controlled women based on the characteristics of EHG. Furthermore, these systems could have excellent acceptability by clinical personnel as they are simple methods. As well, and for clinical use, thus improving the transferability of the EHG technique to clinical praxis.

Key Words: electrohysterogram, preterm, Genetic Algorithm, Extreme Machine Learning, prediction of preterm birth.

ÍNDICE

DOCUMENTOS CONTENIDOS

- Memoria
- Presupuesto

ÍNDICE DE LA MEMORIA

1. Introducción	1
1.1. Parto prematuro	1
1.2. Causas del inicio del parto prematuro	2
1.3. Métodos existentes para la predicción de parto prematuro	3
1.4. Electrohisterografía (EHG)	5
1.4.1. Evolución del electrohisterograma durante el embarazo	6
1.5. Estado del arte	7
2. Justificación	9
2.1. Objetivos	9
3. Materiales	11
3.1. Base de datos	11
3.2. Extracción de características	12
4. Métodos.....	15
4.1. <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	15
4.2. Partición del <i>dataset</i>	16
4.3. Selección de características	16
4.3.1. Análisis de Componentes Principales (PCA).....	17
4.3.2. Algoritmo Genético (AG).....	19
4.4. Clasificadores	23
4.4.1. Análisis Discriminante Lineal (LDA).....	23
4.4.2. Regresión Logística (RL)	24

4.4.3. K-vecinos más próximos (KNN)	25
4.4.4. Extreme Machine Learning (ELM)	26
4.4.5. <i>Bootship Aggregating (Bagging)</i>	27
4.5. Métricas utilizadas	28
4.6. Procedimiento seguido	25
5. Resultados	35
5.1. PCA y AG aplicado a LDA	35
5.2. PCA y AG aplicado a RL	37
5.3. PCA y AG aplicado a KNN	39
5.4. PCA y AG aplicado a ELM	42
5.5. Comparativa de los distintos clasificadores con diferente población	43
5.6. <i>Bootship Aggregating (Bagging)</i>	45
6. Conclusiones	48
7. Anexos	50
8. Bibliografía	53

ÍNDICE DEL PRESUPUESTO

1. Introducción	1
1.1. Desglose del presupuesto	1

INDICE DE TABLAS

Tabla 1. Clasificación de los tipos de parto según tiempo de gestación.....	1
Tabla 2. Tipo de parámetros que forman el análisis	13
Tabla 3. Recopilación del número de componentes principales generadas para cada canal.....	18
Tabla 4. Comparativa genética humana y algoritmo genético	19
Tabla 5. Matriz de confusión AG con clasificador LDA.....	45
Tabla 6. Matriz de confusión AG con clasificador RL.....	45
Tabla 7. Matriz de confusión PCA98 con clasificador KNN	45
Tabla 8. Matriz de confusión clasificador Bagging	45

Tabla 9. Distintas métricas para la comparación de los distintos clasificadores en el grupo de test ..	46
Tabla 10. Mejor cromosoma obtenido con el AG y LDA	50
Tabla 11. Mejor cromosoma obtenido con el AG y RL	51

ÍNDICE DE FIGURAS

Figura 1. Diagrama de las causas globales de muerte en menores de 5 años.	1
Figura 2. Patologías que pueden dar lugar a un parto prematuro.....	3
Figura 3. Registro mediante TOCO de actividad no relacionada con parto y registro de parto activo ..	4
Figura 4. Registro mediante EMG de actividad no relacionada con parto y registro de parto activo ...	5
Figura 5. Descripción teórica de la electromiografía uterina. IUP, siglas que hacen referencia a la presión intrauterina	6
Figura 6. Características del contenido en frecuencia de un electromiograma	6
Figura 7. Distribución de los registros en la base de datos empleada en el presente estudio	11
Figura 8. Distribución de los electrodos en el vientre materno.....	12
Figura 9. Explicación gráfica de la técnica SMOTE.	15
Figura 10. Partición del <i>dataset</i> en tres partes iguales	16
Figura 11. Matriz de observación	17
Figura 12. Representación del número de componentes principales dependiendo del canal, así como del nivel de varianza empleado.....	18
Figura 13. Explicación gráfica del concepto de mutación y cruce.....	20
Figura 14. Gráficas que representan la convergencia de la función de ajuste en el AG para LDA, RL y KNN.....	21
Figura 15. Flujograma del funcionamiento del algoritmo genético.....	22
Figura 16. Representación gráfica del funcionamiento de LDA	23
Figura 17. Curva sigmoidea que realiza la clasificación en este problema representativo	25
Figura 18. Influencia en la selección de la variable k	26
Figura 19. Esquema explicativo del funcionamiento del ELM.....	27
Figura 20. Flujograma de la técnica Bagging.....	27
Figura 21. Flujograma PCA con clasificador	30
Figura 22. Flujograma AG con clasificador	31

Figura 23. Flujograma genérico implementado en el algoritmo.....	32
Figura 24. Algoritmo que representa el flujo seguido para emplear ELM	33
Figura 25. Resultados obtenidos utilizando con el clasificador LDA	35
Figura 26. Recopilación de valores de los resultados en test del clasificador LDA	36
Figura 27. Resultados obtenidos utilizando el clasificador RL	37
Figura 28. Recopilación de valores de los resultados en test del clasificador RL.....	36
Figura 29. Resultados obtenidos utilizando el clasificador KNN	39
Figura 30. Recopilación de valores de los resultados en test del clasificador KNN	40
Figura 31. Resultados obtenidos utilizando el clasificador ELM	41
Figura 32. Recopilación de valores de los resultados en test del clasificador KNN	42
Figura 33. Población y tiempo computacional con AG - LDA	43
Figura 34. Población y tiempo computacional con AG - RL.....	44
Figura 35. Población y tiempo computacional con AG - KNN	44
Figura 36. Resultados de los distintos clasificadores comparado con los resultados obtenidos por el método Bagging	46

Documento I

Memoria

Siendo el parto prematuro la primera causa de muerte en menores de cinco años, *Figura 1*, es evidente que se hace una necesidad reducir estas cifras. Aunque se ha logrado un avance significativo en el cuidado de los bebés prematuros lo cual reduce significativamente la tasa de mortalidad perinatal, la prevalencia del parto prematuro sigue en aumento (Robert L Goldenberg *et al.*, 2008).

Diversos estudios publicados demuestran las consecuencias que puede padecer un bebé no solo en edad temprana de su vida, sino también durante su etapa de adulto al nacer pretérmino. En este caso los datos fueron extraídos de un estudio realizado por Dag Moster y su equipo en *The New England Journal of Medicine*. El estudio incluyó a 903.402 bebés nacidos vivos y sin anomalías congénitas, algunos de ellos pretérmino. 1.822 nacieron entre las 23 a 27 semanas de gestación (prematuros extremos), 2.805 a las 28 a 30 semanas (muy prematuros), 7.424 a las 31 a 33 semanas (muy prematuro a prematuro moderado), 32.945 a las 34 a 36 semanas (prematuros tardíos) y 858.406 a las 37 semanas o después (a término). Las proporciones de bebés que sobrevivieron y fueron seguidos hasta la vida adulta fueron 17.8% de los prematuros extremos, 57.3% de los muy prematuros, 85.7% de los prematuros moderados, 94.6% de los prematuros tardíos y 96.5%, de los bebés nacidos a término. Entre los sujetos sobrevivientes, la prevalencia de tener parálisis cerebral fue de 0.1% para los nacidos a término, en cambio, fue del 9.1% para los nacidos entre las 23 y 27 semanas de gestación; la prevalencia de sufrir un retraso mental, 0.4% versus 4.4%; y la prevalencia de recibir una pensión de invalidez, 1.7% versus 10.6%. Entre aquellos que no tenían discapacidades médicas, la edad gestacional al nacer se asoció con el nivel educativo alcanzado, los ingresos, la recepción de los beneficios del Seguro Social y el establecimiento de una familia, pero no con las tasas de desempleo o actividad criminal (Dag Moster *et al.*, 2016).

El parto prematuro no solo presenta un gran impacto en la salud de los bebés pretérmino, sino también supone un coste económico adicional importante. El total de gastos acumulados durante los primeros 10 años de vida, incluido el ingreso inicial al nacimiento, promedió en £ 17,819.94 para niños nacidos con <28 semanas de gestación, £ 17,751.00 para niños nacidos entre 28 y 31 semanas de gestación, £ 5,376.39 para niños nacidos entre las 32 y 36 semanas de gestación, y £ 1,658.63 para niños nacidos a las 37 semanas de gestación o superior. El número de ingresos hospitalarios, días de hospitalización y costos, respectivamente, durante los primeros 10 años de vida fue del 130%, 77% y 443% mayor para los niños nacidos con <28 semanas de gestación que para los niños nacidos a término (Stavros Petrou, 2005).

1.2. Causas del inicio del parto prematuro.

Antes que nada, es importante distinguir los partos prematuros que se dan debido a causas iatrogénicas y a los partos pretérmino dados de manera espontánea. El primero tipo de parto son debido a causas como la preclamsia, retardo de crecimiento intrauterino o las hemorragias anteparto. Por otro lado, existen otro tipo de partos prematuros que son debidos a aspectos relacionados con la distensión uterina, infección, trauma o hemorragia subplacentaria. Por último, están los partos pretérmino idiopáticos que son de causa desconocida, ya que hasta la fecha no hubo complicaciones.

Tanto el parto a término como el parto pretérmino involucran eventos clínicos similares, como pueden ser aumento de la contractilidad uterina, ruptura de las membranas corioamnióticas y la dilatación cervical.

En la *Figura 2* podemos ver algunos de los mecanismos propuestos de enfermedades implicadas en el parto pretérmino. A todas estas patologías les afectan por separado los factores tanto genéticos como ambientales. (Romero *et al.*, 2014).

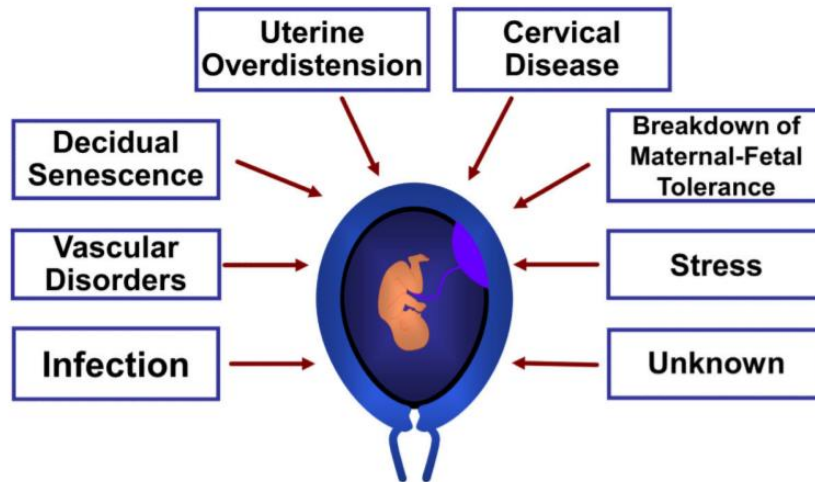


Figura 2. Patologías que pueden dar lugar a un parto prematuro. Fuente: Romero et al., 2014.

1.3. Métodos existentes para la predicción del parto prematuro.

Han sido muchos los estudios realizados cuya finalidad ha sido encontrar un método de predicción del parto pretérmino lo suficientemente fiable. El simple hecho de detectar si un parto va a ser a pretérmino o a término puede llegar a ser realmente complicado aun teniendo al alcance gran cantidad de herramientas como lo son la imagen digital, los biomarcadores y las mediciones de frecuencia de las contracciones. A continuación, se desarrollan brevemente los principales métodos actuales basados, ante todo, en los dos mecanismos principales involucrados en el parto: el acortamiento y ablandamiento del cuello uterino y la dinámica uterina.

La medida de la longitud del cuello uterino ha sido demostrada como una herramienta eficaz para la predicción del parto pretérmino. Un cuello uterino corto está directamente relacionado con el parto prematuro (Romero *et al.*, 2012). Actualmente, la medida del cuello uterino con la ayuda de la imagen mediante ecografía es considerada la técnica más factible para la predicción del parto pretérmino llegando a una sensibilidad del 98% (Honest *et al.*, 2009).

Los marcadores bioquímicos son otro de los métodos utilizados en la predicción del parto pretérmino. Algunos fluidos biológicos como el fluido amniótico, la orina y secreciones vaginales han sido utilizados con resultados satisfactorios. Estudios como el de Poletteni *et al.*, han concluido que ni un marcador ni una combinación de los mismos han podido predecir el parto pretérmino de una manera fiable (Poletteni *et al.*, 2017).

El último método aquí expuesto se trata de la predicción del parto prematuro por medio de la medida de la frecuencia de contracción uterina, esto hace referencia a la dinámica uterina. Para dicha medida es necesario la medida de la presión uterina. Pero antes, es importante diferenciar las dos maneras distintas que existen para realizar la medida de la presión interna del útero.

A continuación, se define de forma no exhaustiva tanto la tocografía externa como interna, más conocida como presión intrauterina.

- **Tocografía externa.** Se trata de un transductor que es colocado sobre el abdomen materno en la zona del fondo uterino y sujetado mediante unas cintas elásticas, habitualmente conocidas como correas. La presión del útero aumenta con la contracción y acciona el transductor, transformando la contracción en una señal eléctrica que se registra de manera digital. Es importante que el transductor se coloque en el lugar adecuado, puesto que, si no es así, el registro obtenido será diferente según su ubicación. Este método es útil para controlar la duración y la frecuencia de las contracciones (Francés *et al.*, 2014). La eficiencia de dicho método depende de la subjetividad del profesional médico y puede presentar fallos en pacientes con sobrepeso (Schlembach *et al.*, 2009). Por otro lado, pueden aparecer contracciones reflejas debido a la presión entre el transductor y el abdomen para asegurar el correcto contacto.
- **Presión intrauterina (vía transcervical).** Esta otra variante de la técnica permite medir la presión del líquido amniótico o de la cavidad uterina. Al ser una medida directa, es más fiable que el sistema externo visto anteriormente, aportando una mayor precisión en la medida de la intensidad, tono y frecuencia de la dinámica uterina. Podemos decir que se trata de una técnica invasiva, ya que, consiste en la introducción de un catéter en la cavidad uterina que, gracias al transductor, medirá los cambios de presión y serán transformado en señales eléctricas que serán posteriormente tratadas (Francés *et al.*, 2014). Esta técnica no se puede emplear para la predicción del parto prematuro ya que requiere la ruptura de la membrana amniótica.

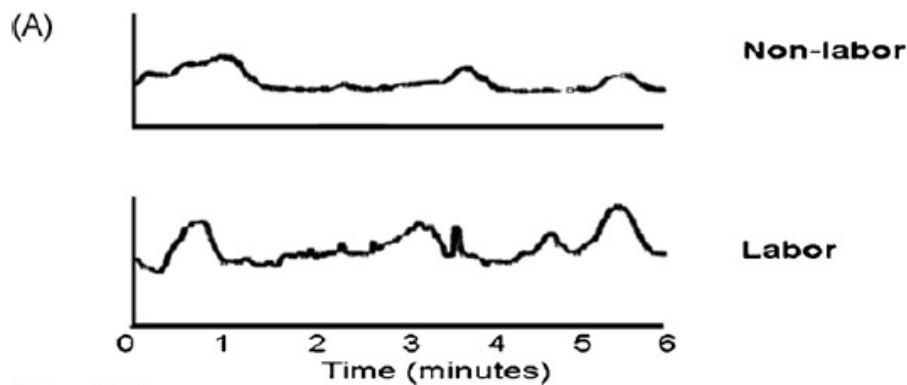


Figura 3. Registro mediante TOCO de una actividad no relacionada con parto y una actividad de parto activo. Fuente: Garfield y Maner, 2007.

En la *Figura 3* podemos observar dos registros de TOCO. En la parte superior vemos el registro de una serie de contracciones que corresponde a una paciente que no está de parto. Por otro lado, en la parte inferior observamos un registro que corresponde a otra paciente que sí está de parto. Esto es uno de los inconvenientes de los cuales hemos hablado anteriormente, y es que esta técnica puede llegar a ser muy subjetiva, ya que, como podemos observar, existen semejanzas entre ambos registros en cuanto amplitud y tasa de contracciones (Garfield y Maner, 2007).

1.4. Electrohisterografía.

Como bien es sabido, toda contracción muscular va asociada a una actividad eléctrica. Las contracciones del útero son fruto de la actividad eléctrica producida en el miometrio (capa muscular intermedia que constituye el grueso del espesor del útero). La electrohisterografía (EHG) es una técnica no invasiva basada en el registro de la actividad eléctrica en el abdomen materno que está directamente relacionada con la actividad uterina de la mujer. Además de no ser invasiva, se ha demostrado que dicha técnica es más sensible a la hora de detectar las contracciones que la técnica TOCO (Alberola-Rubio *et al.*, 2013).

Existen una gran cantidad de estudios que aseguran que mediante la técnica de electrohisterografía es posible detectar la existencia de contracciones uterinas relacionadas con el parto, así como contracciones que no están relacionadas con el mismo (Fele-Zorz *et al.*, 2008; Schlembach *et al.*, 2009; Sikora *et al.*, 2011).

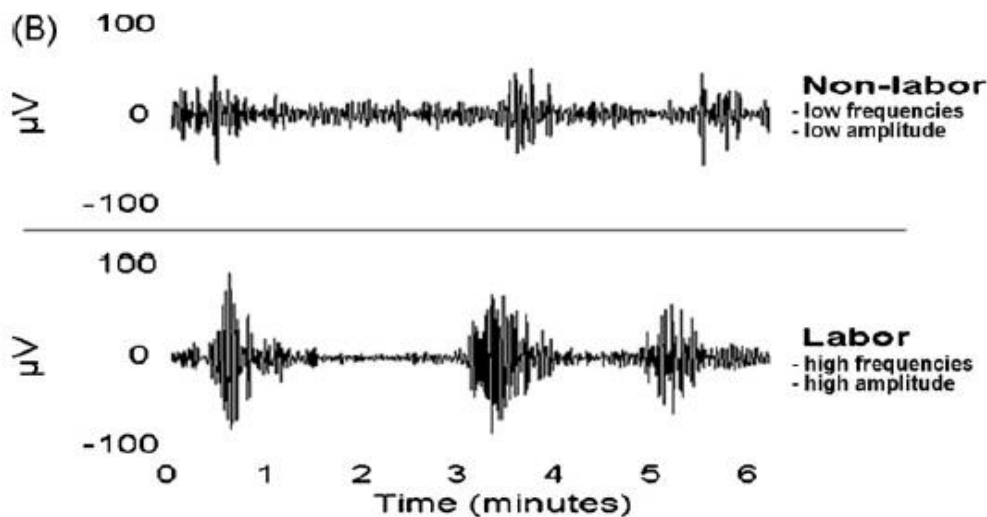


Figura 4. Registro mediante EHG de actividad no relacionada con parto y registro de parto activo.
Fuente: Garfield y Maner, 2007.

En la *Figura 4* observamos un registro EHG. La parte superior corresponde a una paciente a término con una serie de contracciones que no están relacionadas con un estado de parto. De lo contrario, en la parte inferior podemos observar un registro que corresponde a una serie de contracciones que sí están relacionadas con una paciente a término de parto. De estos registros se puede extraer que las pacientes de parto presentan registro de altas frecuencias y altas amplitudes, mientras que las pacientes que no están de parto presentan frecuencias y amplitudes más bajas (Garfield y Maner, 2007). Vemos que gracias al EHG no existe la subjetividad que podía dar lugar el TOCO.

Comparando ambas técnicas se llega a la conclusión de que el TOCO únicamente permite tener una idea de la tasa de contracción, así como una inexacta amplitud de la misma. Por otro lado, el EHG contiene información relevante sobre la electrofisiología uterina, permite ser analizado por métodos matemáticos y así analizar de manera más profunda la actividad eléctrica del miometrio.

1.4.1. Evolución del electrohisterograma durante el embarazo.

Existen suficientes evidencias de que hay diferencias en la actividad eléctrica en el miometrio durante el embarazo y el parto. Pero antes de ahondar en dichas diferencias, resulta de interés hacer una pequeña introducción a las distintas ondas que se pueden encontrar dentro de un análisis de esta índole.

Inicialmente existen dos tipos de onda: una onda lenta, *Figura 5b*, cuyo periodo será igual a la duración de la contracción, y una onda rápida, *Figura 5c*, que se haya superpuesta sobre la onda lenta (Devedeux *et al.*, 1993).

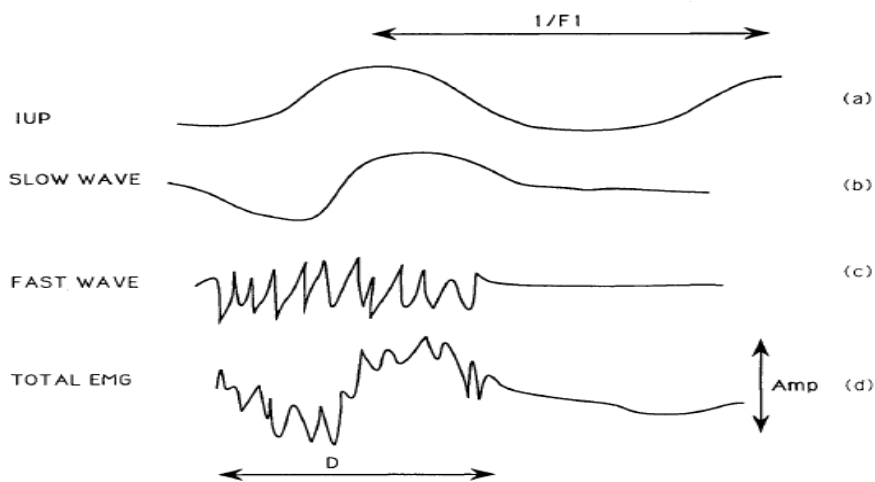


Figura 5. Descripción teórica de la electromiografía uterina. A) IUP, siglas que hacen referencia a la presión intrauterina. Fuente: Garfield y Maner, 2007.

Al mismo tiempo, la onda rápida, *Figura 5c*, puede ser dividida en una onda de baja frecuencia (FWL) y en otra de alta frecuencia (FWH), *Figura 6*.

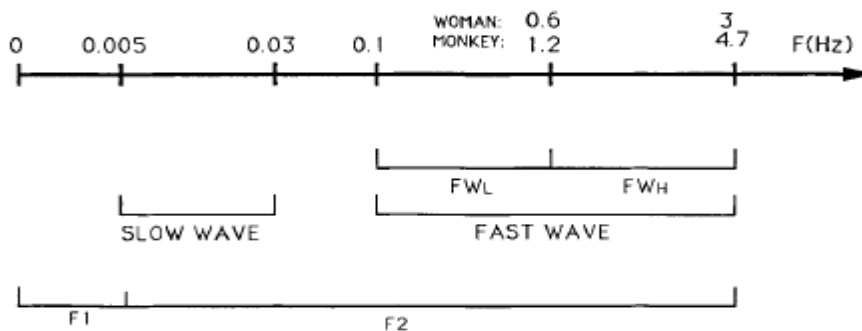


Figura 6. Características del contenido en frecuencia de un electromiograma. Fuente: Garfield y Maner, 2007.

Una gran cantidad de estudios apuntan que las características del EHG son dinámicas y varían a lo largo del embarazo (Devedeux *et al.*, 1993). En resumen, de este gran número de estudios, se concluye que en etapas tempranas de gestación la actividad eléctrica uterina es poco coordinada y conforme la edad gestacional avanza, se produce una evolución de las características del EHG (Devedeux *et al.*, 1993; Garfield and Maner, 2007). A continuación, se hace una breve introducción de las mismas:

- Aumento de la amplitud: incremento de la amplitud en los parámetros del EHG tanto en partos “verdaderos” como en partos “falsos”. Este aumento de amplitud puede llegar a ser registrado hasta dos semanas antes del parto tanto en mujeres a término como a pretérmino.
- Desplazamiento del contenido espectral hacia altas frecuencias: el total del porcentaje de tiempo que el útero presenta una actividad eléctrica de alta frecuencia aumenta de un 10% - 20% a un 80% - 90% cuando el parto dará lugar dentro de las 24 horas en mujeres a término (Garfield *et al.*, 2005).
- Reducción de la complejidad de la señal, así como aumento de su predictibilidad que se ve reflejada en los parámetros no lineales.
- Aumento de la sincronización: en edades tempranas de gestación, la actividad eléctrica uterina no presenta apenas sincronización, aumentando conforme avanza la edad gestacional.

1.5. Estado del arte.

Actualmente, la predicción de parto prematuro mediante la extracción de características de los registros obtenidos mediante EHG se considera el estado de arte de dicha temática, ya que, técnicas nombradas con anterioridad como la tocodinamometría, medida de presión intrauterina, fibronectina fetal o medida de la longitud del cérvix presentan contraindicaciones importantes. Además, estos métodos no son capaces de detectar cambios celulares característicos de un parto “verdadero”. Diversos estudios han empleado esta técnica de extracción de características con anterioridad demostrando que dichas características extraídas de EMG del útero pueden llegar a distinguir entre contracciones verdaderas y falsas (Fele-Zorz *et al.*, 2008; Schlembach *et al.*, 2009; Sikora *et al.*, 2011), así como otros estudios indican que mediante EHG es posible identificar un parto verdadero mejor que ningún otro método actualmente presente en clínica (Lucovnik *et al.*, 2011), consiguiéndose predicciones, siete días antes del parto, con una AUC del 96% utilizando características combinadas como la propagación de la propagación de velocidad y la potencia del espectro, valor elevado si es comparado con el AUC del 72%, 67% y 54% correspondientes al *Bishop Score*, contracciones y longitud cervical respectivamente.

Aun siendo una técnica que combinada con ciertos algoritmos esté dando resultados prometedores, AUC del 91.1% (Mas-Cabo J. *et al.* 2019), estos no están teniendo una repercusión clínica importante ya que existe la problemática de que la mayoría de ellos están basados en algoritmos como *Artificial Neuronal Networks* (ANN) o *Support Vector Machines* (SVM) que presentan una gran complejidad de entendimiento por parte del sector sanitario haciendo que se desconfíe de dichos resultados. Por ello, en el presente trabajo se intenta alcanzar e incluso mejorar estos resultados ya obtenidos mediante técnicas más conocidos y de gran reputación como la Regresión Logística o el Análisis Discriminante Lineal.

En cuanto a la técnica de Análisis de Componentes Principales (PCA), se trata del método convencional por excelencia para la extracción de información relevante y reducción de dimensionalidad, pero nuevos métodos están apareciendo, como el aquí utilizado algoritmo genético. Si bien es verdad que este último necesita de una mayor computación diversos estudios apuntan a que se trata de un método aun por explotar en muchos sectores, entre ellos el sector de la medicina.

CAPÍTULO 2. JUSTIFICACIÓN Y OBJETIVOS

2. Justificación.

En cuanto a la temática del parto prematuro es importante saber que este tipo de parto afecta a más de un 13% de todos los partos a nivel mundial, con la alta tasa de mortalidad que conlleva y los problemas de salud que pueden aparecer a lo largo de la vida del individuo es necesario desarrollar sistemas que ayuden a predecir que dicho parto va a producirse y así poder tomar medidas con la suficiente antelación. Por otro lado, diversos estudios y profesionales de la salud indican que en la actualidad no existe ningún método que permita predecir con exactitud este suceso. Técnicas como el EHG son capaces de realizar el registro de la actividad eléctrica del miometrio uterino mediante una aplicación no invasiva. De este registro se extraen una serie de parámetros temporales, espectrales y no lineales que permiten la caracterización de la señal EHG. Dichos parámetros podrán presentar a su vez información tanto mutua como complementaria.

Por otro lado, es conocida la importancia que presenta una selección de características óptimas para el buen funcionamiento del clasificador, ya que la escasez de las mismas puede ser perjudicial, pero un exceso de ellas también puede provocar un funcionamiento no correcto del clasificador. El Algoritmo Genético es conocido por su buen rendimiento en otras aplicaciones como en la industria, por lo que se hace necesario investigar y conocer dicho funcionamiento en el sector biomédico. Dicho rendimiento será comparado con técnicas anteriormente utilizadas como el Análisis de Componentes Principales.

2.1. Objetivos.

El objetivo del presente trabajo se basa en el desarrollo de un sistema de predicción del parto prematuro en mujeres de control rutinario basado en las características extraídas de la electrohisterografía utilizando algoritmos computacionalmente eficientes. Asimismo, es objetivo del presente trabajo la comparativa del rendimiento del sistema predictor entre el algoritmo genético para la selección de características óptimas y el método de Análisis de Componentes Principales. Una vez las características seleccionadas se utilizarán distintos clasificadores y se comparará el rendimiento entre ellos.

CAPÍTULO 3. MATERIALES

3. Materiales.

3.1. Base de datos.

Para el presente estudio se han utilizado una serie de registros de electrohisterograma recopilados. La base de datos finalmente utilizada se basa en la unión de dos subconjuntos. En primer lugar, la base de datos nombrada como *Term-Preterm ElectroHyteroGram DataBase* (TPEHGDB) disponible con licencia abierta en la plataforma Physionet. En segundo lugar, se ha incluido la base de datos TOCO, la cual aporta una suma de 26 registros.

Estos datos fueron recogidos entre 1997 y 2005 en el departamento de Obstetricia y Ginecología de la Universidad Médica Centro Liubliana. Los registros se realizaron durante los controles médicos rutinarios, algunos de ellos pertenecen a la semana 22 y otros de ellos a la semana 32 de gestación.

La unión de ambas bases de datos conforma un total de 326 registros, cada uno de ellos perteneciente a una mujer embarazada distinta. En el siguiente diagrama de árbol, *Figura 7*, se pretende exponer de manera sencilla la estructura de la base de datos trabajada.

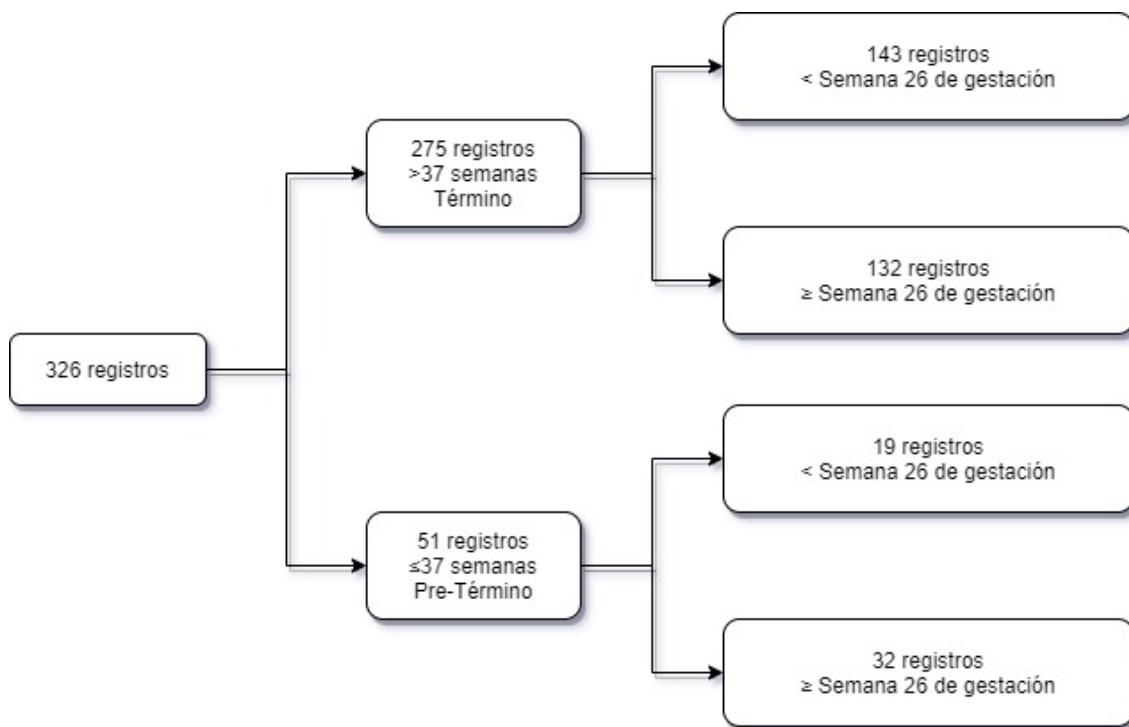


Figura 7. Distribución de los registros en la base de datos empleada en el presente estudio. Fuente: elaboración propia.

Cada registro está compuesto por tres canales nombrados como S1, S2 y S3. Cada uno de estos registros se encuentran formados por la combinación a pares de un total de cuatro electrodos. A continuación, se detalla, en la *Figura 8*, la disposición de cada uno de los electrodos.

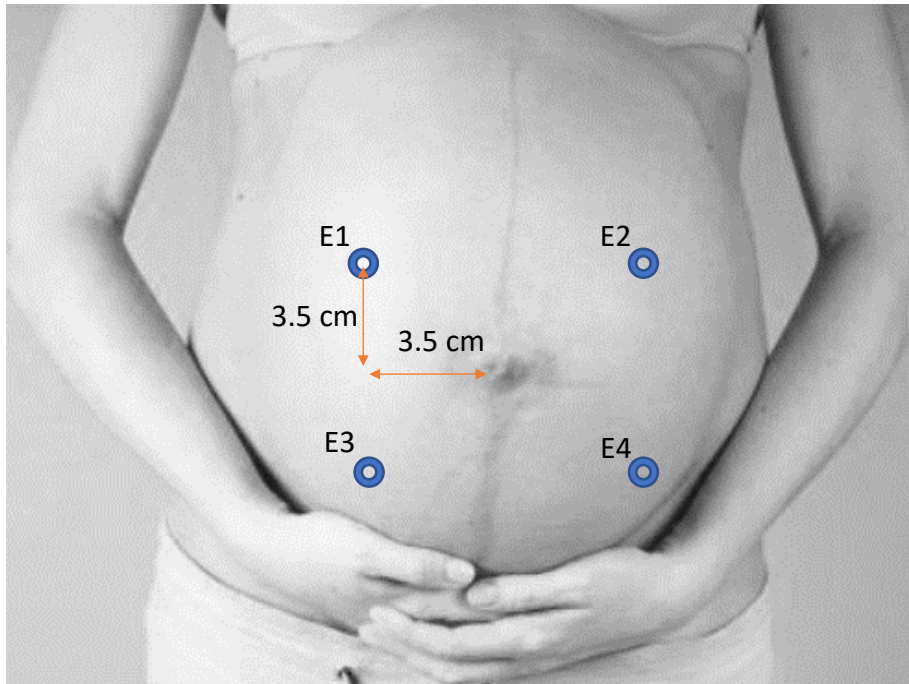


Figura 8. Distribución de los electrodos en el vientre materno. Fuente: elaboración propia.

Los tres canales S1, S2 y S3 fueron conseguidos mediante la diferencia de potencial existente entre dos electrodos. A continuación, se expone qué electrodos fueron empleados para generar cada canal.

- Canal S1 = E2 – E1
- Canal S2 = E2 – E3
- Canal S3 = E4 – E3

3.2. Extracción de características

Los miembros del *Centro de Investigación e Innovación en Bioingeniería (CI2B)* de la Universitat Politècnica de València han analizados previamente la señal EHG. Primero se identificaron visualmente los segmentos de señal fisiológicos que no correspondían ni a artefactos de movimiento ni a interferencias respiratorias. De los 326 registros, los registros válidos por canal son: 140 registros para el canal S1, 130 registros válidos para al canal S2 y 144 registros válidos para el canal S3. Respecto a la composición de parto pretérmino y término, de los 326 registros, 275 fueron a término y 51 pretérmino, *Figura 7*.

A continuación, se calculó una serie de parámetros temporales, espectrales y no lineales, *Tabla 2*, para la caracterización de la señal EHG sobre una ventana móvil de 2 minutos y solapamiento de 50%. Subsecuentemente, se obtuvo el valor mediana de las distintas ventanas de análisis para obtener un único valor representativo por parámetro y por sesión.

Desarrollo de un sistema de predicción del parto prematuro basado en electrohisterografía y algoritmo genético para la selección de características óptimas

Parámetros temporales	Parámetros espectrales	Parámetros no lineales
Amplitud pico a pico	Frecuencia media	Dimensión fractal Higuchi
	Promedio de las frecuencias pico	Entropía muestral
	Frecuencia dominante	Entropía Fuzzy
	Ratio H/L	LZ-binario (n=2)
	Decil de frecuencia	LZ-multi (n=6)
		Tiempo de reversibilidad
		SD1 SD2 SDs SDRR (Poincaré)

Tabla 2. Tipo de parámetros que forman el análisis. Fuente: elaboración propia.

Además de los datos obtenidos por electrohisterografía, se han incluido datos obstétricos, como la edad maternal, paridad, abortos y peso que podrían resultar relevantes para contextualizar la señal EHG. El conjunto de características de EHG constituye el punto de partida de este trabajo fin de grado. Para el canal S1, S2 y S3 se dispone de un total de 71 características para cada uno de ellos. Por otro lado, mediante la combinación de los tres canales obtenido el nuevo S1+S2+S3 se dispone de un total de 205 características.

En general los métodos de clasificación no obtienen un buen rendimiento con bases de datos desbalanceadas. Una consecuencia de esta desigualdad es que dichos algoritmos tiendan a favorecer a la clase con mayor proporción de observaciones, también conocida como clase mayoritaria, siendo en este caso la clase parto a término. Por lo tanto, estos algoritmos podrían llegar a fallar al hacer la predicción de la clase minoritaria si no tienen la cantidad de información necesaria (Amsantac, 2016).

La medida adoptada, en este caso, para balancear el *dataset*, se basó en la técnica conocida como *Synthetic Minority Oversampling Technique* (SMOTE) la cual trata en sintetizar nuevos ejemplos a partir de los ya existentes (Jason Brownee, 2020).

CAPÍTULO 4. MÉTODOS

4. Métodos.

4.1. Synthetic Minority Oversampling Technique (SMOTE).

Debido a la necesidad de balancear la base de datos, se aplicó la técnica SMOTE (*Synthetic Minority Oversampling Technique*) (Nitesh Chawla, *et al.*, 2002).

La técnica SMOTE se basa en primero seleccionar al azar una de las instancias pertenecientes a la clase minoritaria y a continuación se encuentran sus K vecinos de esa clase minoritaria más cercanos, en este caso en concreto el número de vecinos seleccionado fue de 5. La nueva instancia se crea uniendo la primera instancia 'a' con uno de los k vecinos más cercano 'b' mediante una línea sobre la cual se creará la nueva instancia. (Haibo H. *et al.*, 2013). Con dicha técnica se consiguió tener una base de datos balanceada con el mismo número de instancias tanto de la clase a término como de la clase pretérmino, *Figura 9*.

El número de registros que hace referencia a partos que se han dado dentro de término son 275, por otro lado, el número de partos que se han dado lugar en pretérmino son 51. Teniendo en cuenta estos números podemos decir que la base de datos se encuentra desbalanceada. El término de desbalance en un *dataset* hace referencia a que el número de observaciones no es el mismo para todas las clases, en este caso a término y pretérmino.

Synthetic Minority Oversampling Technique

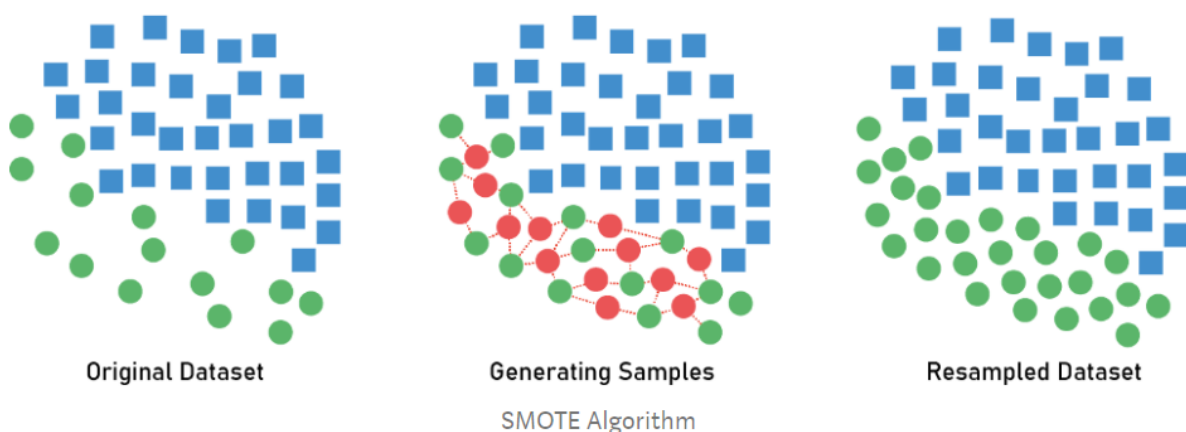


Figura 9. Explicación gráfica de la técnica SMOTE . Fuente: Charfaoui Y. (2019).

4.2. Partición del *dataset*.

Cuando se diseña el clasificador, es sumamente importante testear la capacidad de generalización del clasificador para los datos desconocidos, para ello se utilizó la técnica de partir la base de datos en tres conjuntos de datos, *Figura 10*:

- Conjunto de datos de entrenamiento: estos son los datos que van a servir para entrenar los modelos. 1/3 de los datos forman el conjunto de entrenamiento.
- Conjunto de datos de validación: este conjunto permite seleccionar el mejor de los modelos entrenados y será utilizado durante las iteraciones que se realizarán con el conjunto de entrenamiento. 1/3 de los datos forman el conjunto de validación.
- Conjunto de datos de test: este conjunto va a permitir obtener el error real producido con el modelo seleccionado. Dicho conjunto no es utilizado para el entrenamiento del modelo, por lo tanto, las observaciones aquí existentes son totalmente desconocidas para la predicción. Podemos decir que este conjunto es el que realmente nos indica cómo funciona el modelo. 1/3 de los datos forman el conjunto de datos de test.

Para evitar el sesgo que induce la partición aleatoria de la base de datos en el rendimiento del clasificador, se ha utilizado la técnica de *hold out* de 30 carpetas. Es decir, se realiza la partición aleatoria 30 veces, en cada partición, se utiliza el 33% de datos para el entrenamiento del clasificador, el 33% de datos para su validación y otro 33% para determinar su capacidad de generalización.

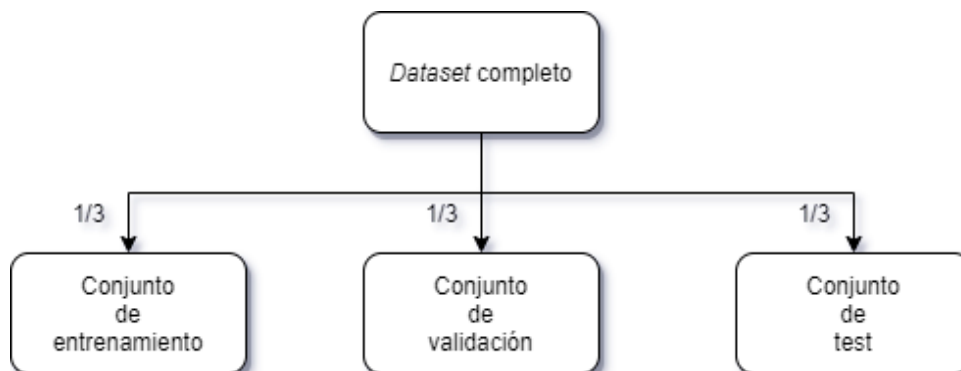


Figura 10. Partición del dataset en tres partes iguales. Fuente: elaboración propia.

4.3. Selección de características.

Mediante dicho proceso, se pretende realizar una selección de características que ayude a reducir la dimensionalidad de las mismas, ya que únicamente se selecciona un subconjunto de dichas características para crear el modelo (*Ver anexo*). Son numerosos los estudios que han demostrado que una selección de características permite mejorar el rendimiento de la predicción. Por otro lado, el uso de demasiadas características, puede llegar a empeorar el rendimiento de la predicción (MathWorks, 2020).

Existen diferentes tipos de características, entre ellas podemos encontrar las características relevantes, las irrelevantes y las redundantes. A continuación, se define cada una de ellas:

- Característica relevante: tras su eliminación, la precisión de modelo decae y el conocimiento aportado por dicha característica no puede ser suplida por ninguna.
- Característica irrelevante: el valor de dicha característica no aporta a la hora de predecir.
- Característica redundante: son aquellas características que no aportan demasiada información y pueden ser eliminadas sin afectar a la predicción.

En este trabajo, se han utilizado y comparado dos algoritmos para la selección de características: el de Análisis de Componentes Principales (PCA) y Algoritmo Genético (AG).

4.3.1. Análisis de Componentes Principales (PCA).

El principal objetivo de este análisis es reducir la dimensión del espacio de observación en el cual se estudian los objetos (Jackson y Hearne, 1973). Esta reducción de la dimensión se consigue mediante la creación de una nueva combinación lineal de las variables que caracterizan los objetos sometidos a estudio. Esta nueva combinación deberá satisfacer una serie de condiciones matemáticas y estadísticas (Mackiewicz A. y Ratajczak W., 1993).

El punto inicial en dicho método se trata de una matriz de observación X , *Figura 11*, en la cual, cada columna se trata de un vector que lista observaciones que caracterizarán un objeto a partir de unas variables.

$$X = \begin{bmatrix} x_{11} & x_{22} & \dots & x_{1n} \\ x_{21} & x_{12} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{p1} & x_{p2} & \dots & x_{pn} \end{bmatrix}$$

Figura 11. Matriz de observación.

Como se mencionó anteriormente, la función fundamental del método de componentes principales es determinar una combinación lineal con la máxima varianza posible. Estas combinaciones lineales reemplazarán el conjunto inicial de variables y formarán las nuevas variables con propiedades especiales (Mackiewicz A. y Ratajczak W., 1993).

$$V = A'X \quad (\text{Eq. 1})$$

Donde:

- V , es la matriz de las nuevas variables.
- A , matriz de vectores propios ortonormales.
- X , matriz inicial de observación.

En este apartado se decidió realizar un estudio más profundo utilizando distintos niveles de varianza. En la *Tabla 3* se hace una recopilación del distinto número de características que están disponibles para cada uno de los canales, así como para el canal combinación de los tres individuales.

Esta recopilación se realiza para cada uno de los distintos niveles de varianza estudiados, siendo estos 98, 95 y 90, *Figura 12*.

Nivel de varianza	S1	S2	S3	S1+S2+S3
98	31	30	30	51
95	27	26	26	44
90	17	16	17	28

Tabla 3. Recopilación del número de componentes principales generadas para cada canal. Fuente: elaboración propia

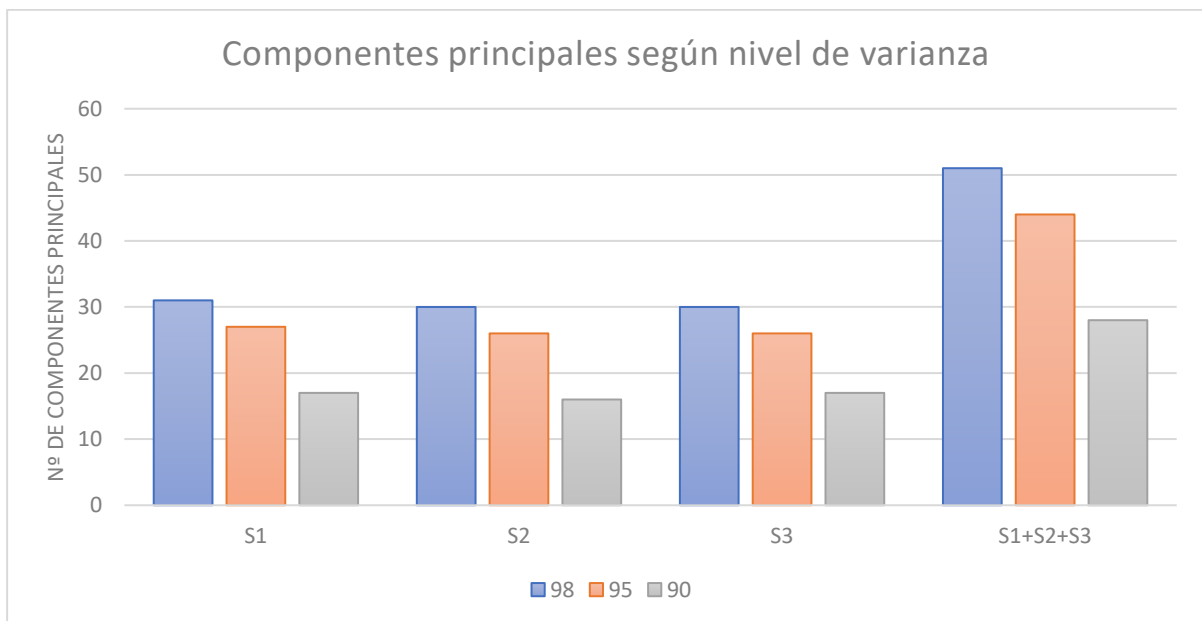


Figura 12. Representación del número de componentes principales dependiendo del canal, así como del nivel de varianza empleado. Fuente: elaboración propia.

4.3.2. Algoritmo Genético (AG).

El algoritmo genético es una técnica de optimización, una búsqueda heurística basada en la población y métodos algorítmicos que imitan el proceso de evolución natural del hombre (Melanie M., 1999).

Este método es un método iterativo, el cual manipula una población de cromosomas para producir una nueva población a través de funciones genéticas como el *crossover* y mutación. Se dice que imita la evolución natural del hombre ya que tiene cierta semejanza con el principio de evolución de Charles Darwin en cuanto a la reproducción, genética de recombinación y la supervivencia del más apto (Babatunde *et al.*, 2014).

En la *Tabla 4*, se hace un breve resumen de las semejanzas existentes entre la genética humana y el Algoritmo Genético.

Genética humana	Algoritmo genético
Cromosoma	String de bits
Genes	Características
Alelo	Valor de la característica
Locus	Posición
Genotipo	String codificada
Fenotipo	Genotipo decodificado

Tabla 4. Comparativa de la genética humana y algoritmo genético. Fuente: Sivanandam S. et al. (2008).

El ajuste de los candidatos a la solución (cromosomas), se realiza utilizando una función referida a una función de ajuste. Esta función de ajuste dará un valor numérico el cual será utilizado para dar preferencia a unos u otros cromosomas dentro de la misma población (Sivanandam S *et al.*, 2008).

Los cromosomas en el AG, son realmente cadenas binarias que sufren manipulaciones por dicho algoritmo como la evolución natural manipula los cromosomas del hombre. Este método comienza con una población generada de manera aleatoria, existe una función encargada de generar esta aleatoriedad. A lo largo de la cadena binaria que conforma los cromosomas, los valores '1' indicarán que la característica es seleccionada en dicho cromosoma, mientras que '0' indicará que la característica no fue seleccionada. A continuación, mediante las características seleccionadas y la función de ajuste, estos cromosomas son ordenados según los resultados obtenidos y los situados en las primeras 'n' posiciones son seleccionados para sobrevivir a la siguiente generación. Una vez seleccionados los cromosomas que sobreviven, los demás entran en proceso de cruce o mutación formando una nueva población. Es oportuno definir en este punto el concepto de cruce y mutación, *Figura 13*. Cruce hace referencia a la combinación de dos cromosomas para así tener una descendencia cruzada, mientras que mutación hacer referencia a la modificación de los genes en un mismo cromosoma que depende de un parámetro denominado factor de mutación. Estos cruces y mutaciones generarán hijos cruzados y mutados que formarán la nueva población para la siguiente iteración junto a los hijos élite, *Figura 15*.

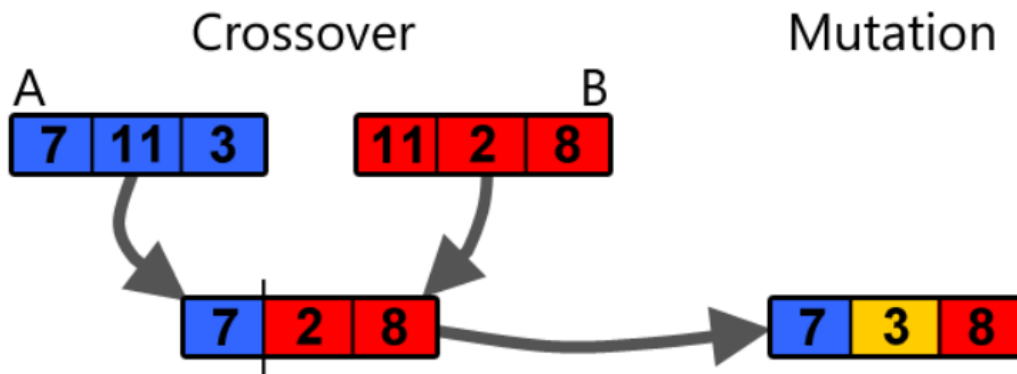


Figura 13. Explicación gráfica del concepto de mutación y cruce. Fuente: Fernández A. (2017).

Tanto el número de población, como el número de generación y el criterio de parada pueden influir en la combinación óptima de las características. Respecto al criterio de parada, en este caso se utilizó un máximo de mil generaciones sin que presentaran una mejora en F1-Score. Por último, en cuanto al número de población, se realizó un barrido paramétrico de dicho número comparándose los resultados obtenidos con el tiempo de convergencia del algoritmo.

Además, el algoritmo genético es un solo algoritmo para la selección óptima de las características y es necesario implementarlo junto con un clasificador. La mejor combinación de características evidentemente dependerá de la función de ajuste, en otras palabras, del método de clasificación que se esté empleando el cual se extraerá la métrica que se utiliza para medir el rendimiento del clasificador. Tal como se mencionará posteriormente, en el presente trabajo se han comparado los distintos clasificadores con los que se han realizado el estudio, siendo estos; Análisis Discriminante Lineal (LDA), Regresión Logística (RL) y K vecinos más próximos (KNN). Respecto a la función del ajuste, en este caso se pretende maximizar el F1-Score promedio (o minimizar 1-F1-Score), *Figura 14*, de las 30 carpetas del clasificador que se detallará más adelante. En este punto se presentan las gráficas en las cuales se puede observar la convergencia del método y ajuste del parámetro.

Una vez las funciones ajustadas mediante el AG, se obtiene un cromosoma el cual contiene los genes que han conseguido un valor más bajo de 1-F1-Score. Estos cromosomas son los encargados de seleccionar las características que serán empleadas para de lanzar de nuevo los modelos, es decir, entrenarlos, validarlos y realizar el test. A parte de los clasificadores LDA, RL y KNN se decidió lanzar con cada cromosoma obtenido el clasificador tipo ELM.

Desarrollo de un sistema de predicción del parto prematuro basado en electrohisterografía y algoritmo genético para la selección de características óptimas

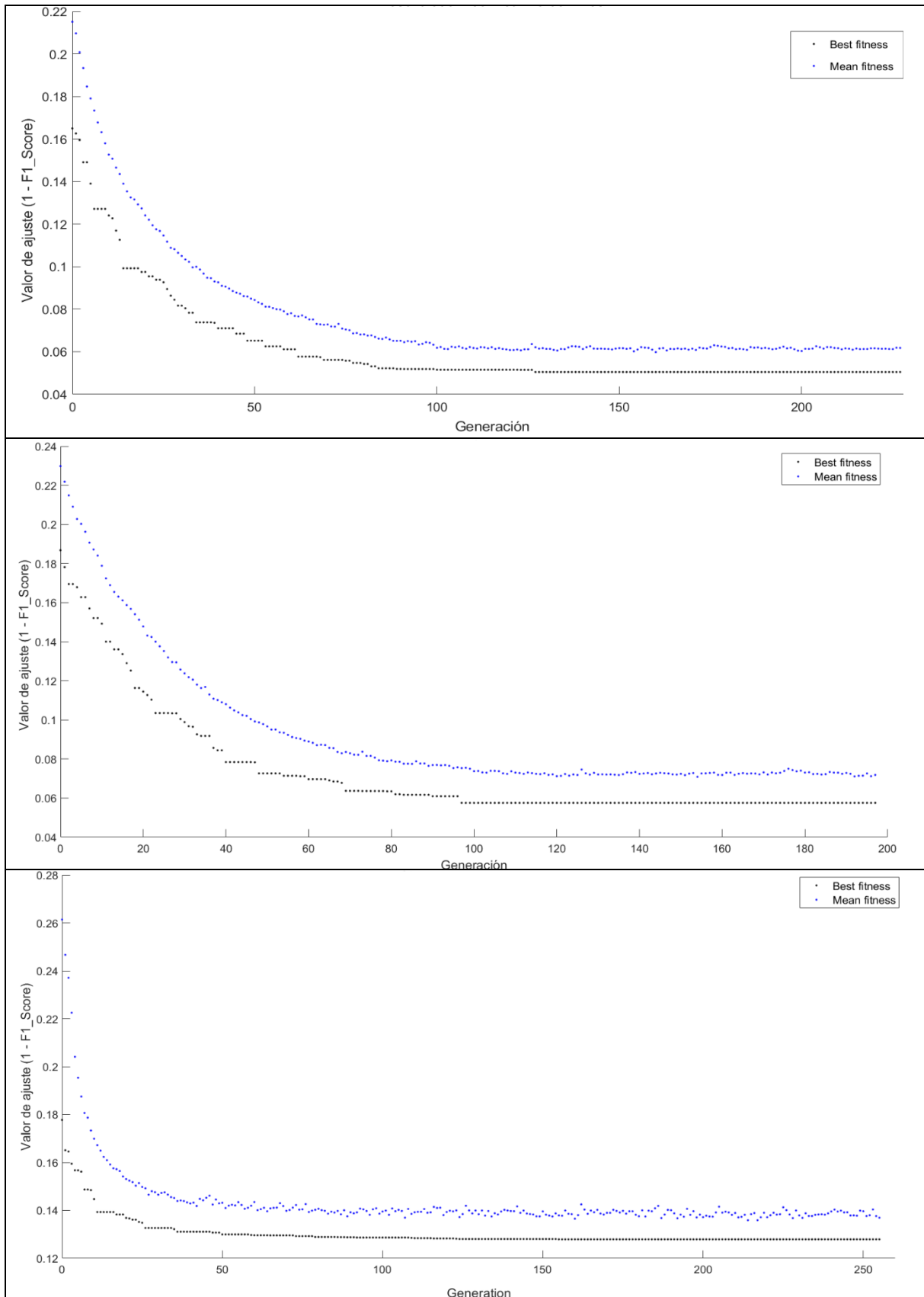


Figura 14. Gráficas que representan la convergencia de la función de ajuste en el AG para LDA, RL y KNN. Fuente: elaboración propia.

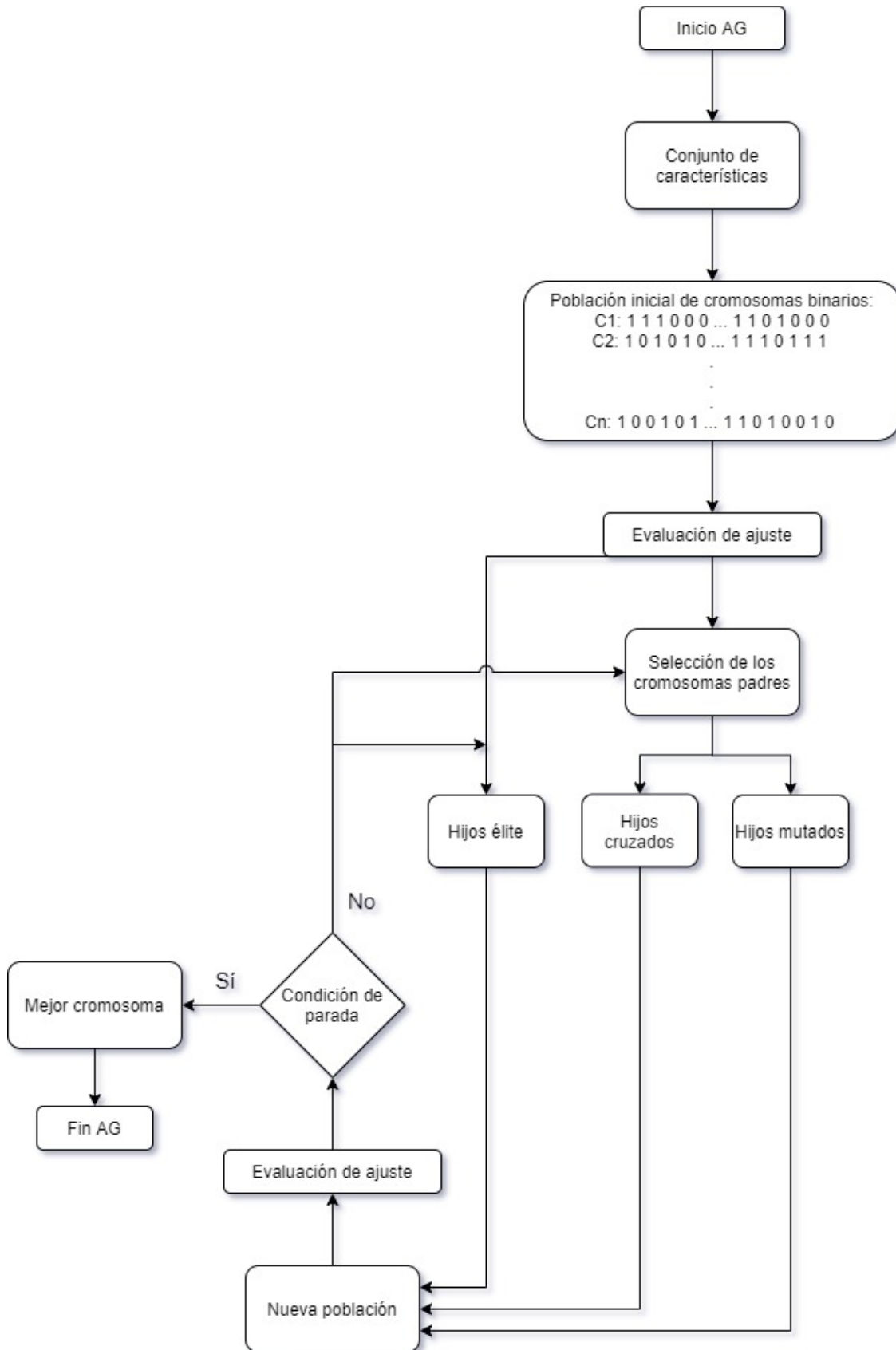


Figura 15. Flujograma del funcionamiento del algoritmo genético. Fuente: Babatunde et al., 2014.

4.4. Clasificadores.

Los clasificadores pueden ser definidos como una subcategoría del aprendizaje supervisado cuyo objetivo es predecir las etiquetas de la clase a la cual pertenecen las nuevas instancias introducidas a partir de unas observaciones que ya se conocían.

En el presente estudio se trata una clasificación binaria ya que tan solo es posible asignar dos clases distintas, pretérmino o término.

Diversos estudios recomiendan el uso de distintos algoritmos de clasificación para poder escoger entre ellos el comportamiento más adecuado al trabajo a realizar. Este comportamiento está sujeto al número de datos disponibles, a las diferentes clases existentes, si los datos son o no linealmente separables, así como el número de características.

En conclusión, se trata de conseguir una regla de decisión que sea capaz de asignar un objeto nuevo, desconocido para el clasificador, a una clase fijada con anterioridad.

4.4.1. Análisis Discriminante Lineal (LDA).

Linear Discriminant Analysis (LDA) es un método de clasificación supervisado el cual realizará una clasificación de variables cualitativas de unas nuevas observaciones a partir de unas ya conocidas basándose en sus características. Basado en el teorema de Bayes y permite estimar la probabilidad de que una observación se encuentre en una determinada clase a partir de uno determinados predictores (Amat J., 2016).

Este método trata pues en asignar la nueva observación X a la clase cuyo centro de gravedad proyectado está más cercano de la nueva observación, *Figura 16*.

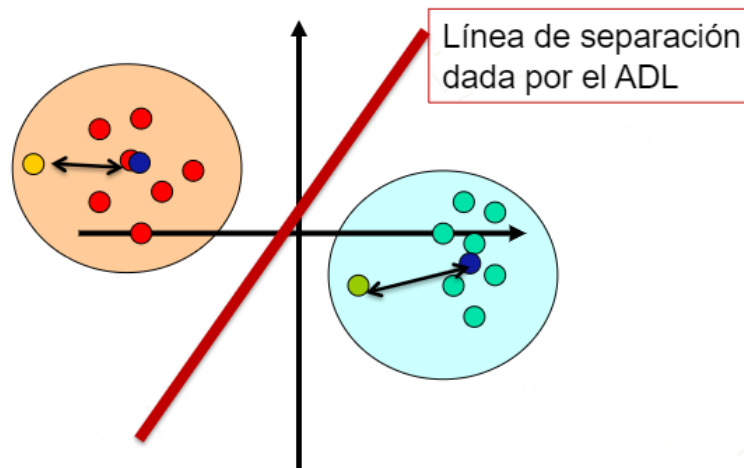


Figura 16. Representación gráfica del funcionamiento de LDA. Fuente: Oldemar, 2016.

En la *Figura 16* se pueden observar dos conjuntos de observaciones agrupados en distintos sectores. En cada uno de dichos sectores se calcula su centroide, punto en azul. Posteriormente, se calcula la distancia entre el centroide calculado y la nueva observación (punto amarillo y verde), asignando a dicha nueva observación el conjunto cuyo centroide quede más cercano. En esta misma

figura, se puede observar la recta dada por el método que divide los distintos conjuntos. Esta recta podría llegar a ser más ajustada en caso de trabajar con el análisis discriminante cuadrático.

4.4.2. Regresión Logística (RL).

El método de regresión logística lleva implícito el término regresión, que como es sabido, dicho término hace referencia a la ley experimental que traduce la relación entre variables correlacionada. Como es el caso del presente estudio, la variable a explicar únicamente puede tomar dos valores, 0 y 1 haciendo referencia al parto pretérmino o a término. Si utilizáramos la regresión lineal, al evaluar la función se obtendrían valores distintos a 0 o 1, lo cual carecería de sentido. Por lo tanto, para solucionar este inconveniente se utiliza la Regresión Logística (RL), la cual consigue pasar las entradas a través de las funciones logísticas tratando los resultados como unas probabilidades.

Por tanto, dicho método se basa en la obtención de una función logística de las variables independientes que sea capaz de clasificar a las observaciones en una de las dos subpoblaciones existentes. La función logística es la encargada de calcular, para cada observación, y según los valores de una serie de características (X_i), la probabilidad (p) de que dicha observación presente el efecto estudiado, en este caso, parto pretérmino o término (Dolores *et al.*, 2000).

Dicho esto, la regresión logística es una curva especial en forma de S (curva sigmoidea), *Figura 17*, creada a partir de la regresión lineal y transformando la estimación numérica en una probabilidad mediante la siguiente expresión.

$$h_{\theta}(x) = \sigma(\theta^T X) = \frac{e(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots)}{1 + e(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots)} \quad (Eq. 2)$$

Donde;

- $\theta^T X$, es la suma de las variables ponderadas por los coeficientes.
- e , función exponencial.
- $\sigma(\theta^T X)$, curva logística.

Siendo su finalidad localizar los parámetros θ de manera que el modelo clasifique en base a la mejor predicción de la clase para cada una de las observaciones.

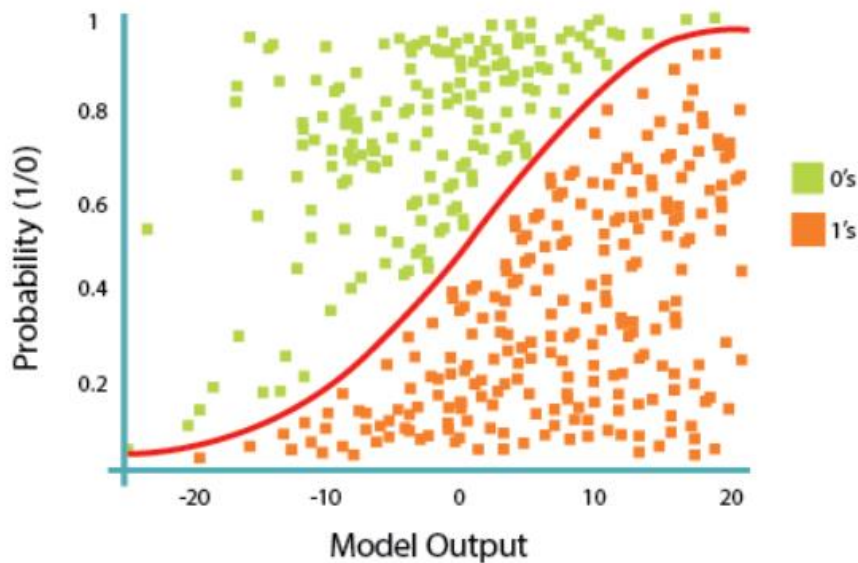


Figura 17. Curva sigmoidea que realiza la clasificación en este problema representativo.
Fuente: Start Developer, 2020.

4.4.3. K vecinos más próximos (KNN).

El método de análisis de K vecinos más próximos, perteneciente a la categoría de aprendizaje supervisado, clasifica en base al parecido de unas observaciones con otras. Creado como una herramienta capaz de encontrar patrones entre los datos sin que sea necesario una coincidencia exacta. La base de este método es que los casos parecidos se encuentran más próximos entre ellos, siendo los no parecidos los que se encuentran más alejados.

El nombre del método viene dado ya que a los casos que están situados más cercanos se les denomina como vecinos. Una vez es introducida una nueva observación en el algoritmo, este calcula la distancia con respecto a las observaciones que incluye el modelo. A continuación, se seleccionan los K elementos más cercanos y se realiza una clasificación. La nueva observación se incluye en la clase que contiene un mayor número de vecinos más próximos (IBM, 2020).

El cálculo de distancia puede hacerse mediante métodos conocidos como la distancia euclídea o el método similitud coseno.

- Distancia euclídea, es aquella distancia que existe entre dos puntos de un espacio euclídeo. Esta distancia se calcula mediante el teorema de Pitágoras.

$$d_E(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (Eq. 3)$$

- Similitud Coseno, medida de similitud que existe entre dos vectores en un espacio. Esta función trigonométrica responde con un 1 si el ángulo comprendido es cero.

$$\text{CosSim} = \frac{xy}{\sqrt{xx}\sqrt{yy}} \quad (\text{Eq. 4})$$

Donde, x e y son dos vectores.

Por último, en el presente apartado, es necesario ver la influencia del valor de la variable k , que es el valor de vecinos que son escogidos. Dicho valor es seleccionado normalmente tras una sucesión de pruebas para encontrar el valor óptimo.

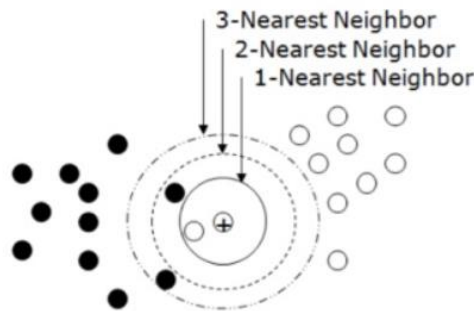


Figura 18. Influencia en la selección de la variable k . Fuente: Bing Liu, 2011.

- Si $k = 1$ el algoritmo clasificaría la bola + como blanca.
- Si $k = 2$ el algoritmo carecería de criterio para realizar la selección.
- Si $k = 3$ el algoritmo clasificaría la bola + como negra.

En este breve ejemplo se puede comprobar la gran importancia que tiene realizar una correcta selección del valor de la variable k . En el presente trabajo, lo mejores resultados fueron obtenidos con $k = 3$ para AG, excepto con PCA donde el valor óptimo del k fue $k=2$.

4.4.4. Extreme Learning Machine (ELM).

El *Extreme Learning Machine* (ELM), es un algoritmo formado por redes de neuronas artificiales con una única capa oculta, con alimentación *feedforward* usado para problemas de clasificación y regresión, *Figura 19*. Dicha red se encuentra conectada con una serie de pesos aleatorios a la capa de entrada y en la que solo serán entrenados dichos pesos que a su vez estarán conectados con la capa de salida mediante un algoritmo de mínimos cuadrados. La popularidad de dicho método es debida a que si es comparada con los algoritmos de aprendizaje convencionales de redes neuronales este supera la velocidad de entrenamiento, así como los problemas de sobreajuste existentes (Ding S. *et al.*, 2015).

En la *Figura 24*, se puede observar el flujograma que hace referencia al algoritmo programado para la implementación de ELM. Durante las iteraciones se intenta mejorar el valor de F1-Score promedio, cuando esto no es posible, la información relacionada con dicha métrica es eliminada.

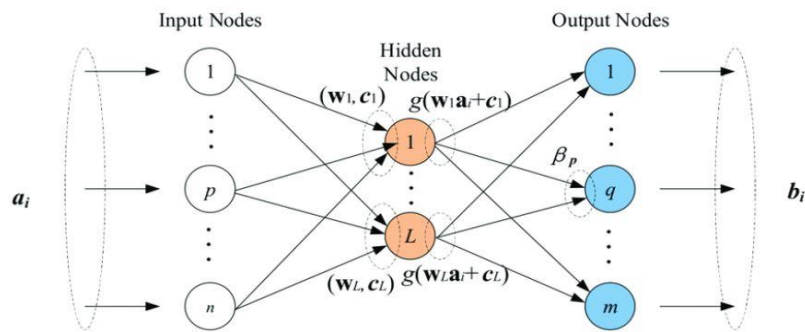


Figura 19. Esquema explicativo del funcionamiento del ELM. Fuente: Zhang Y. et al., 2018.

4.4.5. Bootstrap Aggregating (Bagging).

Bootstrap Aggregating o también conocida como técnica *Bagging*, es un método que trata de que una vez se han obtenido las predicciones de distintos modelos, estas se combinan para obtener un nuevo resultado que debería mejorar la estabilidad y precisión de los algoritmos utilizados en las clasificaciones anteriores. Esto se consigue reduciendo la varianza de los métodos de aprendizaje empleados. Otra ventaja que presenta dicha técnica es que ayuda a su vez a evitar el sobreajuste.

El fundamento de este método se basa en que mediante el promedio de un conjunto de observaciones se consigue reducir la varianza ya que cada una de estas observaciones presenta una varianza σ^2 , mientras que la media de las observaciones \bar{Z} presenta una varianza σ^2/n . Por lo tanto, los pasos que se siguieron fue obtener múltiples muestras de la población y a cada una de ellas se le aplicó un modelo distinto, *Figura 20*. Una vez obtenidos los resultados se hizo la media de las predicciones obtenidas (Amat J., 2017).

En el presente estudio se combinaron los resultados obtenidos de los clasificadores utilizando la técnica *majority voting*. Dicha técnica se basa en un sistema de votación en la que el resultado ganador es aquel que recibió más de la mitad de los votos emitidos. Este proceso se aplicó sobre las predicciones obtenidas por los tres clasificadores que dieron mejor resultado en la predicción entre partos prematuros y a término en el grupo de test, que en este caso fueron los clasificadores KNN, LDA y RL.

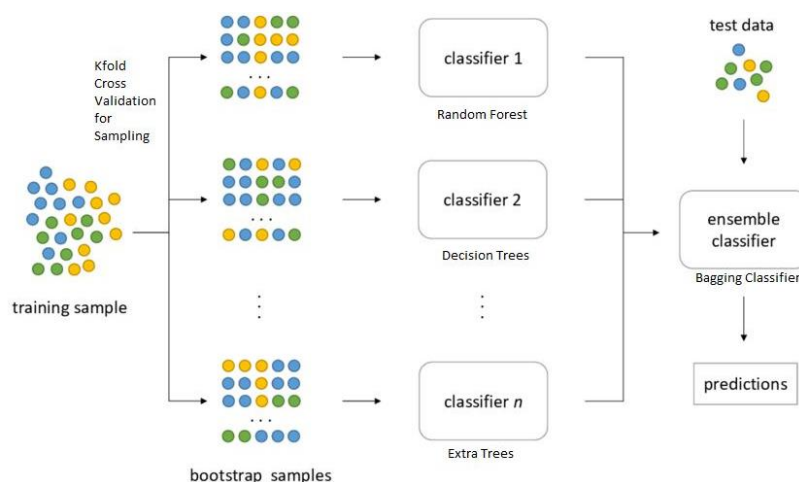


Figura 20. Flujograma de la técnica Bagging. Fuente: Patel A. (2019).

4.5. Métricas utilizadas.

A la hora de evaluar el desempeño de los clasificadores es necesario la utilización de ciertas métricas que ayudarán a valorar la capacidad de predicción de dichos métodos. Estas métricas están basadas en las matrices de confusión que se obtuvieron a partir del resultado de cada uno de los métodos empleados. Para calcular dichos valores ha sido necesario calcular con anterioridad otros parámetros que se explican a continuación:

- Verdaderos positivos VP: número de resultados que el modelo predice correctamente en la clase positiva.
- Verdaderos negativos VN: número de resultados en que el modelo predice correctamente en la clase negativa.
- Falso positivo FP: número de resultados en que el modelo predice incorrectamente la clase positiva.
- Falso negativo FN: número de resultados en que el modelo predice incorrectamente la clase negativa.

Una vez conocidos estos parámetros se procede al cálculo de las métricas. Estas son:

- Aciertos: porcentaje total de los aciertos del modelo.

$$Aciertos = \frac{VP + VN}{VP + VN + FP + FN} \quad (Eq. 5)$$

- Sensibilidad: probabilidad de que ante un caso realmente positivo el modelo lo clasifique como tal.

$$Sensibilidad = \frac{VP}{VP + FN} \quad (Eq. 6)$$

- Especificidad: probabilidad de que ante un caso realmente negativo el modelo lo clasifique como tal.

$$Especificidad = \frac{VN}{VN + FP} \quad (Eq. 7)$$

- VPP: valor predictivo positivo es el porcentaje de casos clasificados como positivos por el modelo respecto al total de casos positivos.

$$VPP = \frac{VP}{VP + FP} \quad (Eq. 8)$$

- VPN: valor predictivo negativo es el porcentaje de casos clasificados como negativos por el modelo respecto al total de casos negativos.

$$VPN = \frac{VN}{FN + VN} \quad (Eq. 9)$$

- AUC: área bajo la curva ROC. AUC mide el área existente por debajo de la curva ROC. Esta curva se construye representando la tasa de VP frente FP.
- F1-Score: métrica utilizada en el análisis estadístico de una clasificación binaria, hace referencia a la medida de la precisión que presenta un test. Por otro lado, el F1-Score es empleado como métrica en bases de datos desbalanceadas ya que es capaz de mantener el equilibrio entre precisión y exhaustividad. La diferencia entre precisión y exhaustividad reside en que mientras la primera hace referencia a qué proporción de predicciones positivas (tanto verdaderos como falsos positivos) fueron correctas la segunda hace referencia a qué proporción de positivos reales (verdaderos positivos) se identificó correctamente.

En el presente trabajo, aunque fue resuelto el problema de desbalanceo, se consideró oportuno utilizar como métrica de evaluación el clasificador F1-Score, ya que de partida la predicción del parto prematuro se trata de un problema conocido como *imbalanced data learning*.

Una vez obtenidos estos valores con los resultados de los clasificadores se puede proceder al cálculo de la métrica F1-Score. Para el cálculo de la misma se procede como se indica a continuación.

$$F1_{score} = \frac{2 * VP}{2 * VP + FP + FN} \quad (Eq. 10)$$

Esta métrica se calcula tanto para el grupo de entrenamiento como para el de validación y posteriormente se procede al cálculo de la raíz cuadrada del producto de ambas métricas. En el caso del AG, como la función busca minimizar el valor, la métrica utilizada se presenta a continuación.

$$Métrica_{AG} = 1 - F1_{score} \quad (Eq. 11)$$

Es oportuno en este punto hacer una breve comparación entre la métrica AUC y F1-Score. En general se puede decir que la curva ROC es empleada cuando existen muchos niveles diferentes de umbrales, mientras que el F1-Score se puede considerar como una métrica de aplicación en cada uno de los puntos de la curva. Por otro lado, para que el F1-Score sea alto, tanto la precisión como la exhaustividad deben de ser altos. Todo esto lleva, a que cuando se presenta un problema desbalanceado se deba utilizar la métrica F1 ya que el promedio de la curva ROC alcanza todos los umbrales posibles.

4.6. Procedimiento seguido.

Tal y como se ha comentado en el punto 3.1 se parte de la base de datos de registros *Term-Preterm ElectroHyteroGram DataBase* (TPEHGDB), disponible con licencia abierta en la plataforma Physionet, más los datos aportados por los registros TOCO. El grupo de investigación de Instrumentación y Procesado de señal del Centro de Investigación e Innovación en Bioingeniería (CI2B) de la Universitat Politècnica de València ha computado el conjunto de parámetros temporales, de no linealidad y complejidad que se detallan en la *Tabla 2*.

En dicha base de datos nos encontramos distintas carpetas para cada uno de los canales S1, S2 y S3, así como una capeta formada por la unión de los anteriores canales individuales S1+S2+S3. Dentro de cada una de estas carpetas existe una subcarpeta con el estudio que incluye las características del EHG, características de los registros TOCO y los datos obstétricos, además incluye una carpeta correspondiente a la aplicación de la técnica SMOTE.

Como se comentó en el apartado inicial, para valorar la capacidad de generalización de los resultados obtenidos con los diferentes clasificadores, la base de datos se dividió de forma aleatoria formando las treinta subcarpetas y dividiendo estos datos en tres partes iguales en cada iteración (metodología *holdout*): 1/3 para entrenamiento, 1/3 para validación y 1/3 para test.

Para llevar a cabo dicho estudio, en primer lugar, se eligieron dos maneras distintas de seleccionar las características óptimas a utilizar mediante dos técnicas diferentes: el método de Análisis de Componentes Principales (PCA) y el Algoritmo Genético (AG).

En cuanto a la parte que corresponde al PCA, se realizó el estudio para distintos niveles, entre ellos están los niveles de varianza del 98, 95 y 90%. La finalidad de esta selección es llegar a encontrar el nivel que diera un mejor resultado, *Figura 21*.

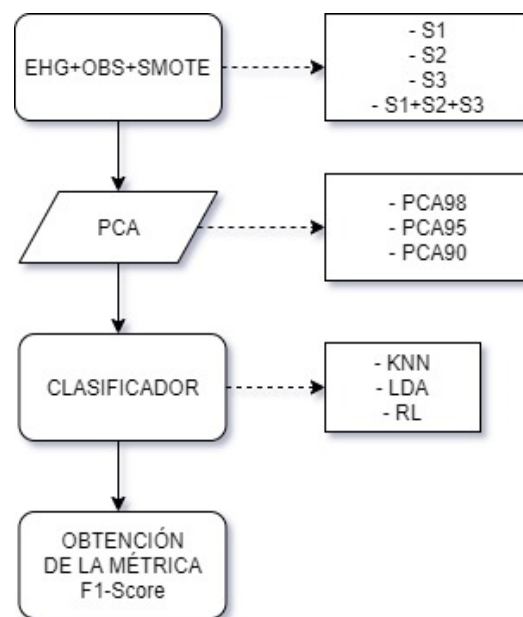


Figura 21. Flujograma PCA con clasificador. Fuente: elaboración propia.

Respecto al AG, inicialmente se realizó un primer estudio utilizando una población de 200 individuos por recomendación de la literatura existente (MathWorks, 2020). Como se ha explicado anteriormente, el AG necesita de una función de ajuste, para ello se han utilizado clasificadores LDA, KNN y RL los cuales dan un rendimiento calculado con la expresión dada en el apartado correspondiente y siguiendo el algoritmo programado, *Figura 22*. En las primeras fases del trabajo se planteó la posibilidad de incluir el método ELM como una función de ajuste más, pero tras las primeras pruebas se llegó a la conclusión de que la unión del AG y ELM necesitaba una potencia de cómputo mayor de la que ofrece un PC convencional, ya que se requiere la optimización de dos hiperparámetros como son los coeficientes C y el número de neuronas N. Por esta razón, ELM solo pudo ser utilizado como clasificador y no como función de ajuste con el AG.

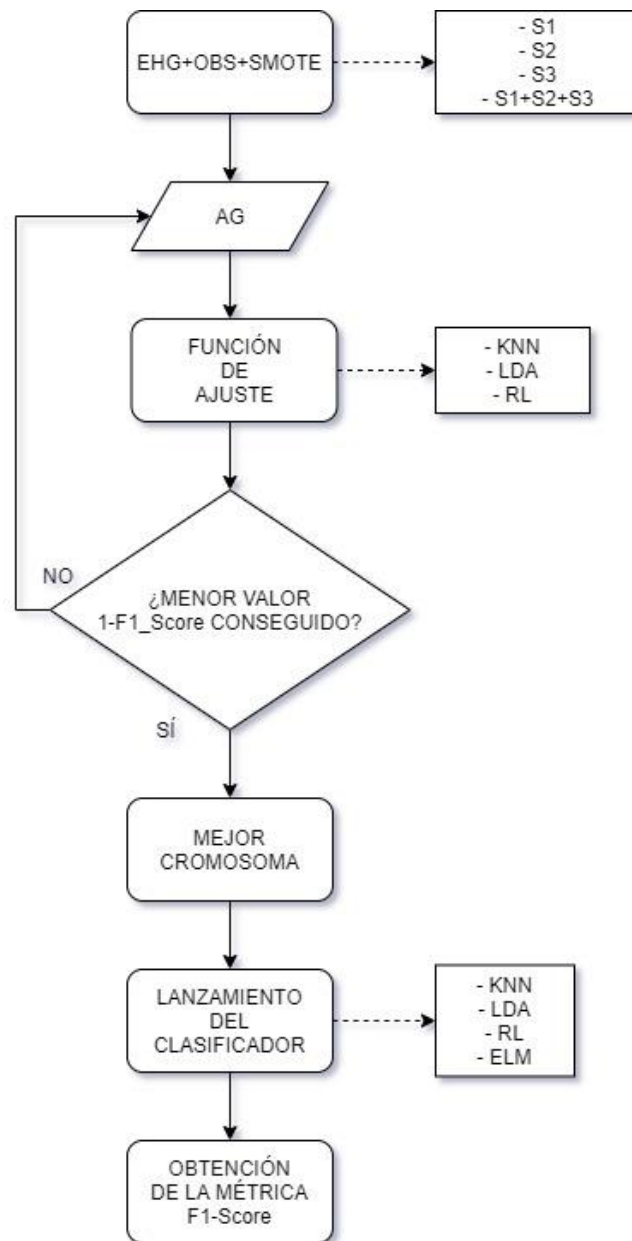


Figura 22. Flujoograma AG con clasificador. Fuente: elaboración propia.

Una vez preparados los archivos PCA y los mejores cromosomas obtenidos mediante el AG (Ver anexo) se procedió a la ejecución del algoritmo desarrollado para la clasificación, *Figura 23*.

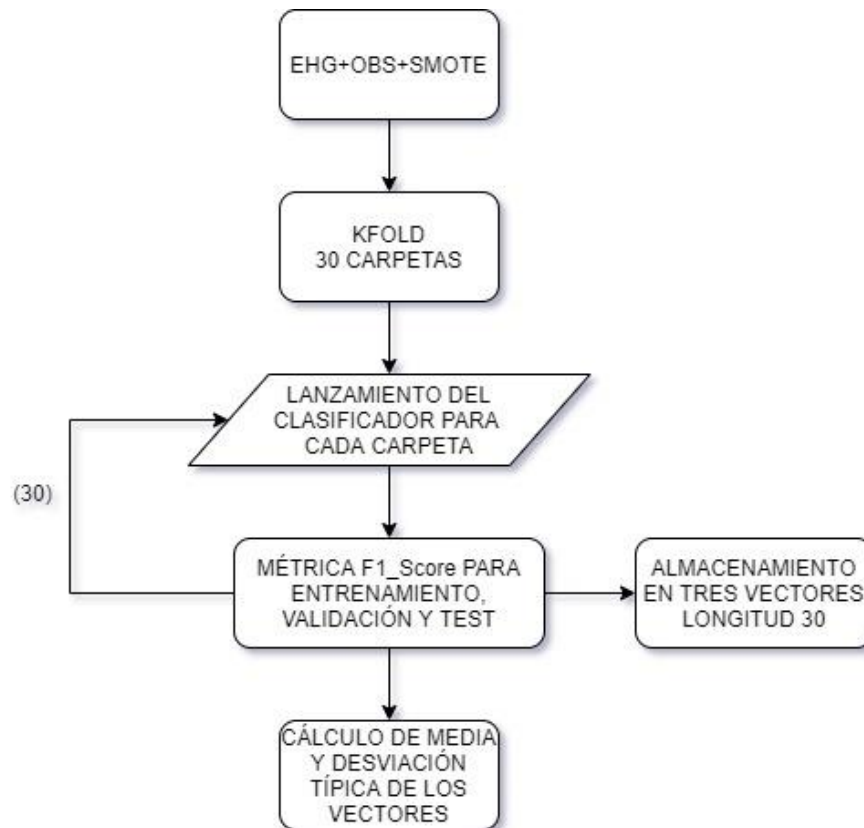


Figura 23. Flujograma genérico implementado en el algoritmo. Fuente: elaboración propia.

En la figura anterior se puede observar como el algoritmo recorre cada una de las treinta carpetas y en cada una de estas iteraciones se calcula la métrica F1-Score. Estas métricas son almacenadas en vectores de la misma longitud que veces itera el algoritmo. Una vez recorridas todas las carpetas se calcula la media y desviación típica de los vectores obteniendo así el F1-Score promedio.

El algoritmo visto en la *Figura 23*, se utilizó para los clasificadores LDA, RL, KNN y ELM. Estos dos últimos métodos presentaron la optimización de hiperparámetros como pueden ser el número de vecinos para KNN, la optimización de los coeficientes C y número de neuronas N para ELM.

En la *Figura 24*, se presenta el algoritmo desarrollado para la implementación de ELM como clasificador. El mismo está compuesto por dos iteraciones principales, las cuales se encargan de la selección de hiperparámetros como el número de neuronas y los coeficientes. Posteriormente se realiza una iteración de diez veces por cada carpeta para evitar caer en mínimos locales. A esta iteración le sigue la iteración de valor treinta que recorre cada una de las carpetas. La finalidad de este algoritmo es conocer que combinación ofrece el mejor F1-Score promedio. Una vez conseguida esta combinación se guardan los parámetros necesarios que nos permitirían llegar hasta la clasificación final.

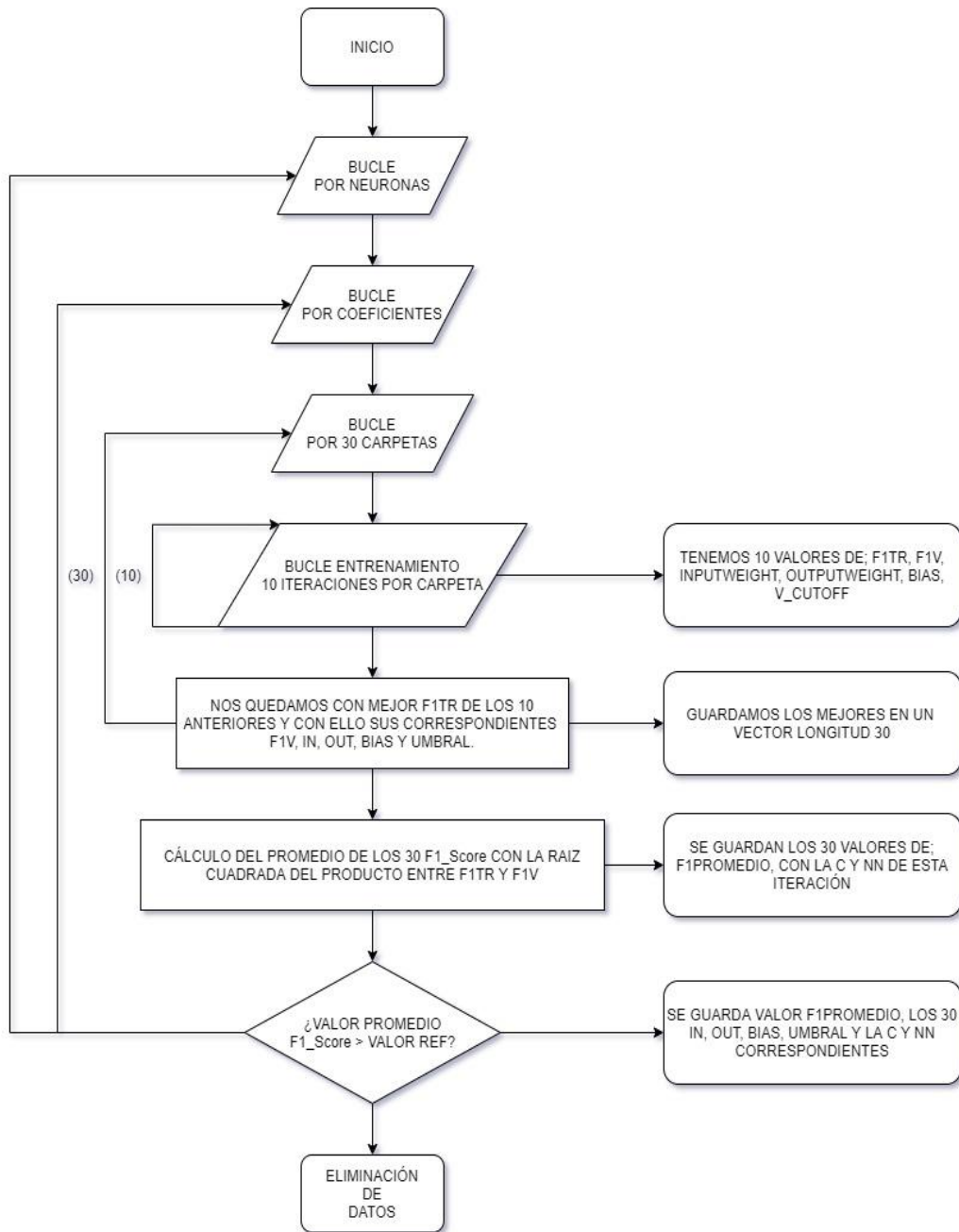


Figura 24. Algoritmo que representa el flujo seguido para emplear ELM. Fuente: elaboración propia.

Obtenidos todas las predicciones se seleccionaron los tres mejores resultados dados por el algoritmo genético, concretamente fueron las combinaciones AG+LDA, AG+RL y AG+KN. A continuación, se procedió a estudiar el efecto del tamaño poblacional en el desempeño de cada uno de ellos, así como el tiempo computacional empleado. Estos tamaños poblacionales fueron 50, 100, 200 y 400 individuos.

El último paso tras realizar el barrido poblacional fue seleccionar los tres sistemas de clasificación que arrojaron las mejores métricas (F1-Score promedio) en el grupo de test. Concretamente estos fueron PCA98+KNN, AG+LDA y AG+RL. A continuación, se procedió a realizar la técnica Bagging para combinar las predicciones obtenidas. Como se ha explicado en apartado correspondiente al Bagging, esta combinación se llevó a cabo mediante la técnica *majority voting*.

CAPÍTULO 5. RESULTADOS

5. Resultados.

5.1. PCA y AG aplicado a LDA.

En la *Figura 25* se representa los valores F1-Score promedio obtenidos usando técnicas de selección de características como PCA y AG con el clasificador de Análisis Discriminante Lineal (LDA) para los distintos canales S1, S2 y S3, así como la combinación de los tres canales en S1+S2+S3. Si se analizan los resultados arrojados al emplear dicha selección de característica se observa que los valores obtenidos para cada uno de los canales individuales son muy parecidos, por lo que no se extrae una información relevante de ninguno de ellos por separado.

Comparando los resultados obtenidos de manera individual con los obtenidos en la combinación de los tres canales, podemos observar un aumento en el porcentaje del F1-Score promedio tanto en entrenamiento, validación y test, pasando de valores del 69,10% a 89,03%. Este aumento en el porcentaje de dicha métrica, es debido a que, al unir los diferentes canales, el número de características aumenta, permitiendo de esta manera tener un proceso de selección de características más óptimo.

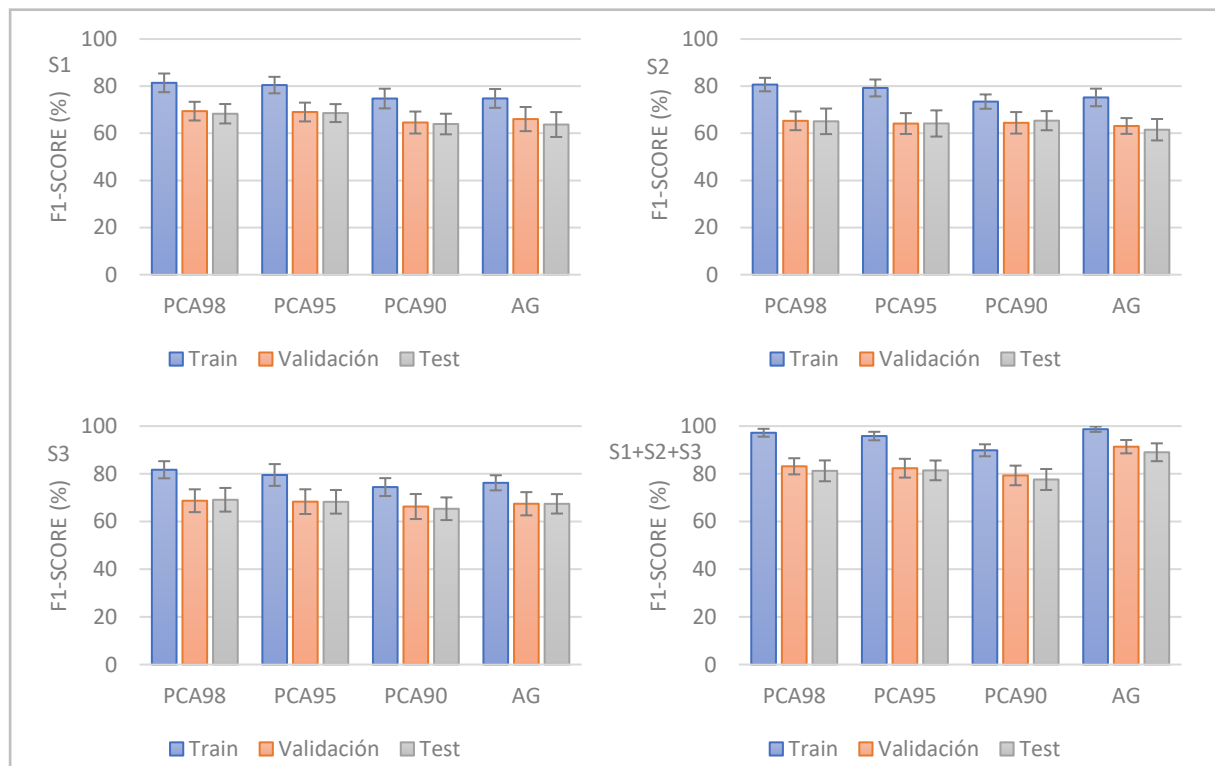


Figura 25. Resultados obtenidos utilizando con el clasificador LDA. En la fila de arriba encontramos, a la izquierda, el estudio referente al canal S1, a su derecha el canal S2. En la fila de abajo, a la izquierda, encontramos S3, a su derecha S1+S2+S3. Fuente: elaboración propia.

En la *Figura 26* se han reunido, a modo de resumen, los resultados obtenidos para el grupo de test en función de las distintas técnicas de selección de características empleadas. Como hemos dicho anteriormente, los mejores resultados corresponden al clasificador que emplea como entradas las características de todos los canales de registro (S1+S2+S3). Dentro de estos resultados destaca el F1-Score promedio del 89.03%, obtenido mediante la selección de características dada por el Algoritmo Genético, frente a los valores del 81.24%, 81.43% y 77.61% correspondiente a la selección de características generada por PCA98, PCA95 y PCA90 respectivamente.

Por lo tanto, en cuanto a la clasificación mediante LDA, se puede afirmar, que los mejores resultados se dan empleando las características seleccionadas por el Algoritmo Genético y posteriormente aplicadas al clasificador LDA. Esta combinación permite obtener un F1_Score promedio del 89.03%.

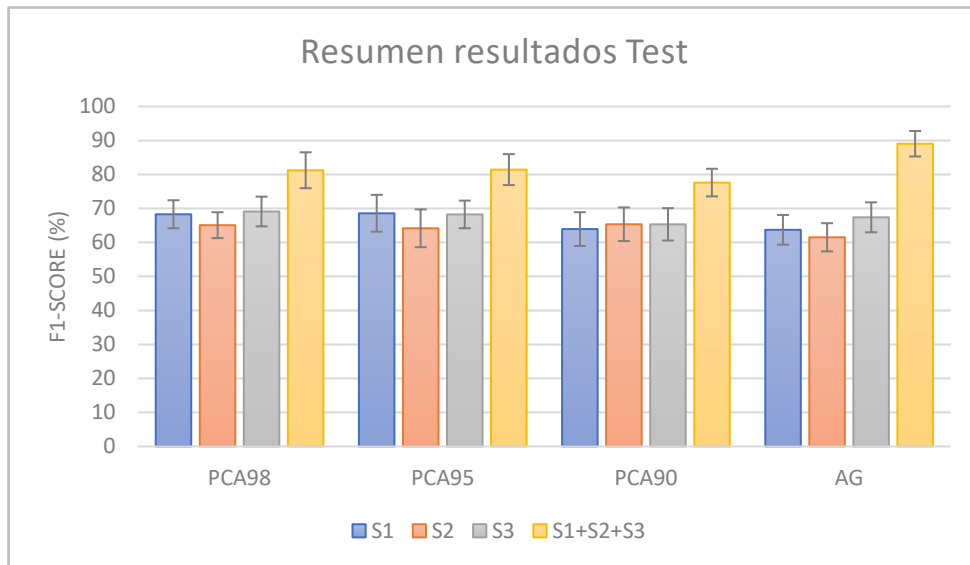


Figura 26. Recopilación de valores de los resultados en test del clasificador LDA obtenidos con distintas técnicas selección de características para los distintos canales. Fuente: elaboración propia.

5.2. PCA y AG aplicado a RL.

En la *Figura 27* se presentan los resultados de F1-Score promedio obtenidos utilizando como técnicas de selección de características el Análisis de Componentes Principales (PCA) y Algoritmo Genético (AG) cuya selección es aplicada al clasificador Regresión Logística (RL).

Como en el caso anterior, se puede observar que los valores obtenidos de F1-Score promedio obtenidos para cada canal de manera individual son más bajos que los obtenidos al emplear las características obtenidas para todos los canales. El valor máximo de dicha métrica que se obtiene en los canales individuales es del 71,9%, mientras que el máximo es del 85,78% al combinar la información de todos los canales.

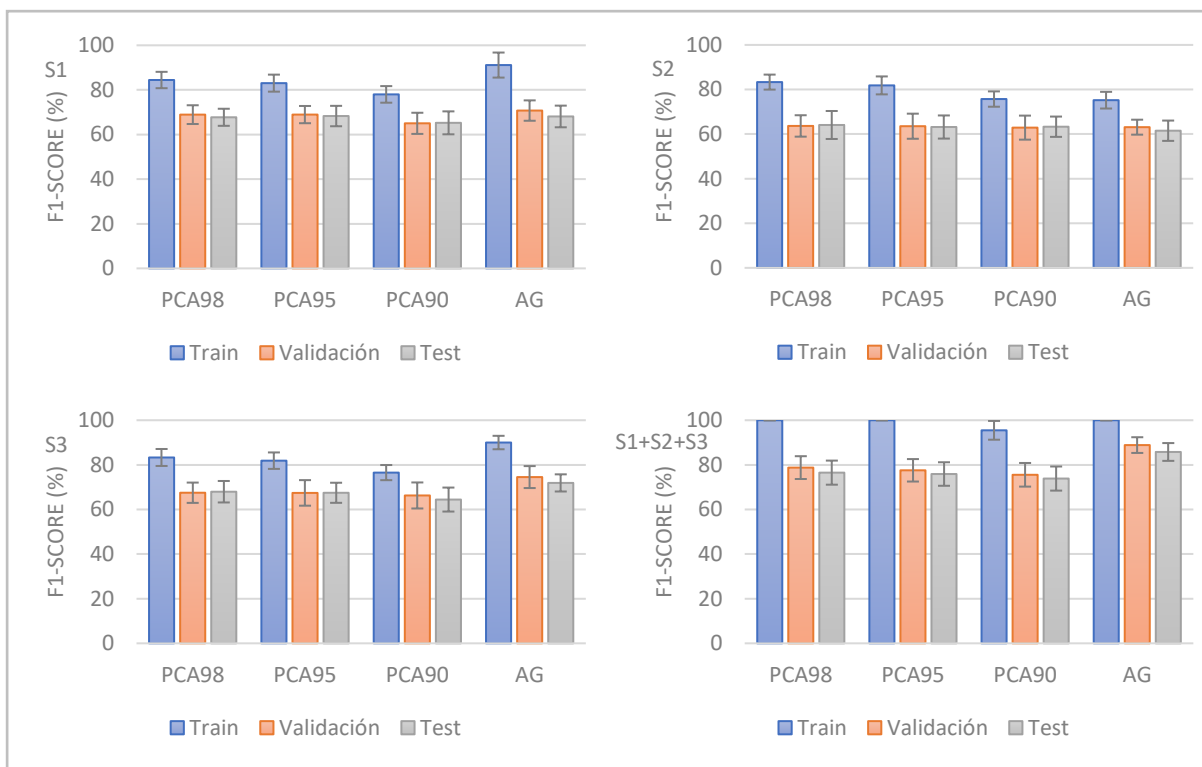


Figura 27. Resultados obtenidos utilizando el clasificador RL. En la fila de arriba encontramos, a la izquierda, el estudio referente al canal S1, a su derecha el canal S2. En la fila de abajo, a la izquierda, encontramos S3, a su derecha S1+S2+S3. Fuente: elaboración propia.

En la *Figura 28* se resumen los resultados obtenidos para el grupo de test. En dicha figura se encuentran grupos de columnas correspondientes a la misma técnica de selección de características, mientras que cada columna dentro de cada grupo hace referencia a un canal o a su combinación. Como en el caso anterior, el mejor resultado del F1-Score promedio se encuentra en la combinación de los tres canales utilizando el selector de característica del Algoritmo Genético obteniendo un F1-Score promedio del 85,78%. Este último valor es mucho mayor si es comparado con los valores 76,51%, 75,89% y 73,86% correspondientes a PCA98, PCA95 y PCA90 respectivamente.

Por lo tanto, en cuanto a la clasificación mediante RL, se puede afirmar, que los mejores resultados se dan empleando las características seleccionadas por el Algoritmo Genético y posteriormente aplicadas al clasificador RL. Esta combinación permite obtener un F1-Score promedio del 85,78%.

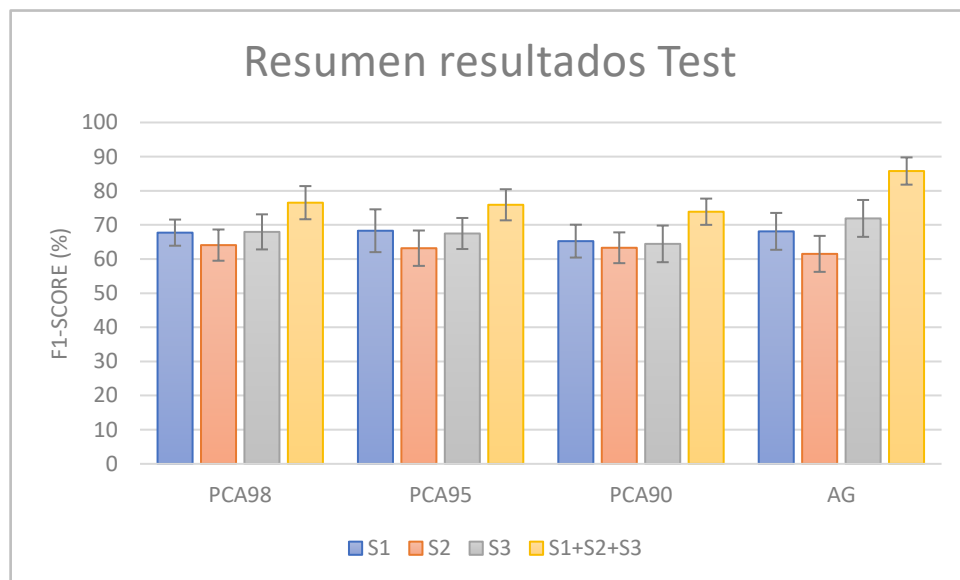


Figura 28. Recopilación de valores de los resultados en test con distintas selecciones de características para los distintos canales. Fuente: elaboración propia.

5.3. PCA y AG aplicado a KNN.

En la *Figura 29* se encuentran los resultados obtenidos para los distintos canales S1, S2 y S3, así como la combinación de los tres canales en S1+S2+S3, usando las técnicas de selección de características PCA y AG junto con el clasificador de K vecinos más próximos. En este caso se analizó el desempeño del clasificador haciendo un barrido del número de vecinos a emplear entre 2 y 5. Siendo el número de vecinos óptimo el de tres excepto para el PCA98 cuya clasificación óptima se dio con dos vecinos. Si se analiza el desempeño de los clasificadores que emplean como entradas las características de un único canal los valores obtenidos son muy similares, no existiendo de esta manera suficiente evidencia para afirmar que la información aportada por algún canal individual presenta más información que el resto para la predicción del parto prematuro.

Comparando los resultados obtenidos de manera individual con los obtenidos en la combinación de los tres canales, podemos observar un aumento significativo en el porcentaje del F1-Score promedio tanto en entrenamiento, validación y test, pasando de valores del 81,10% a 90,19%. Estos valores son acordes a los obtenidos en apartados anteriores, pero en este caso, el mayor valor alcanzado fue mediante la selección de características PCA.

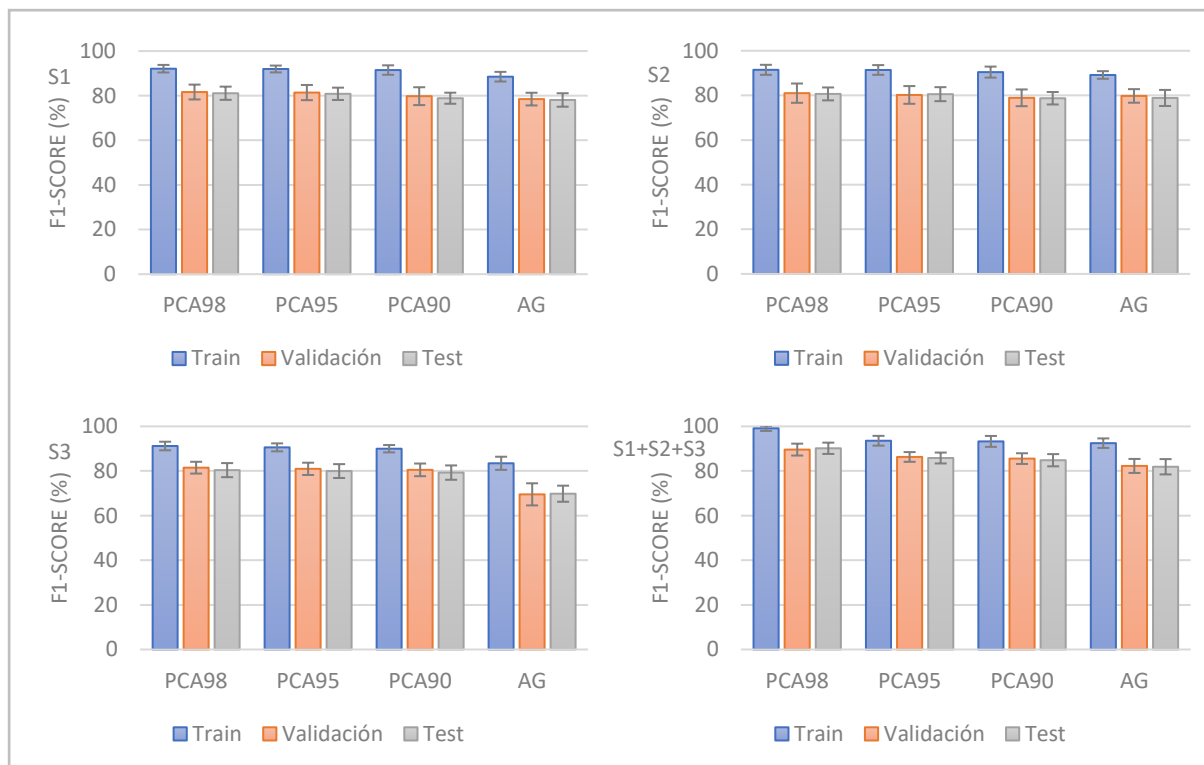


Figura 29. Resultados obtenidos utilizando el clasificador KNN. En la fila de arriba encontramos, a la izquierda, el estudio referente al canal S1, a su derecha el canal S2. En la fila de abajo, a la izquierda, encontramos S3, a su derecha S1+S2+S3. Fuente: elaboración propia.

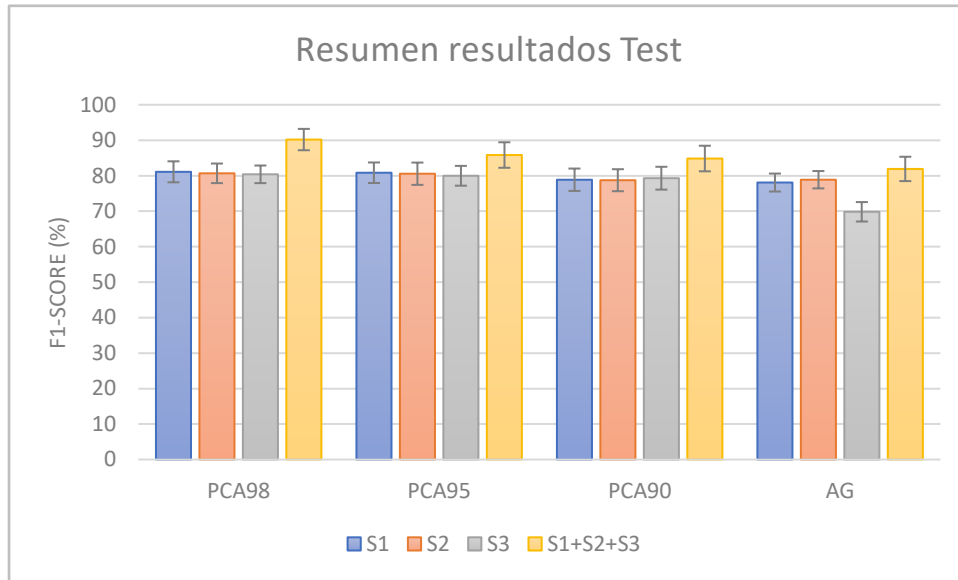


Figura 30. Recopilación de valores de los resultados en test con distintas selecciones de características para los distintos canales. Fuente: elaboración propia.

En la *Figura 30* se resumen los resultados del F1-Score promedio obtenidos para el grupo de test de cada una de las distintas técnicas de selección de características empleadas. A diferencia de los casos anteriores, el mejor resultado corresponde para el selector de características dado por el Análisis de Componentes Principales con nivel 98 con el cual se obtiene un F1-Score del 90,19%.

Por lo tanto, en cuanto a la clasificación mediante KNN, se puede afirmar, que los mejores resultados, utilizando la métrica F1-Score, se dan lugar empleando las características seleccionadas por el Análisis de Componentes Principales sobre el clasificador KNN.

5.4. PCA y AG aplicado a ELM.

En la *Figura 31*, como con las figuras anteriores, han sido recogidos los resultados obtenidos para los distintos canales individuales S1, S2 y S3, así como la combinación de los tres canales en S1+S2+S3, usando distintas técnicas de selección de características nombradas con anterioridad y utilizando *Extreme Machine Learning* (ELM) para realizar la clasificación. En el análisis de cada uno de los canales por separado podemos observar que los valores obtenidos son muy similares no existiendo diferencias significativas entre ellos.

Como se explicó en apartados anteriores, en este método es necesario una optimización de los hiperparámetros, C y NN, coeficientes y número de neuronas respectivamente. Los mejores resultados, se obtuvieron, como en los casos anteriores, con la combinación de S1+S2+S3 y la selección de los hiperparámetros C=0.1 y NN=140.

Comparando los resultados obtenidos con los canales individuales con los obtenidos en la combinación de los tres canales, pasamos de un 77,07% a un 79,08%. Por lo tanto, podemos observar un aumento en el porcentaje del F1-Score tanto en entrenamiento, validación y test en cada una de las diferentes selecciones de características.

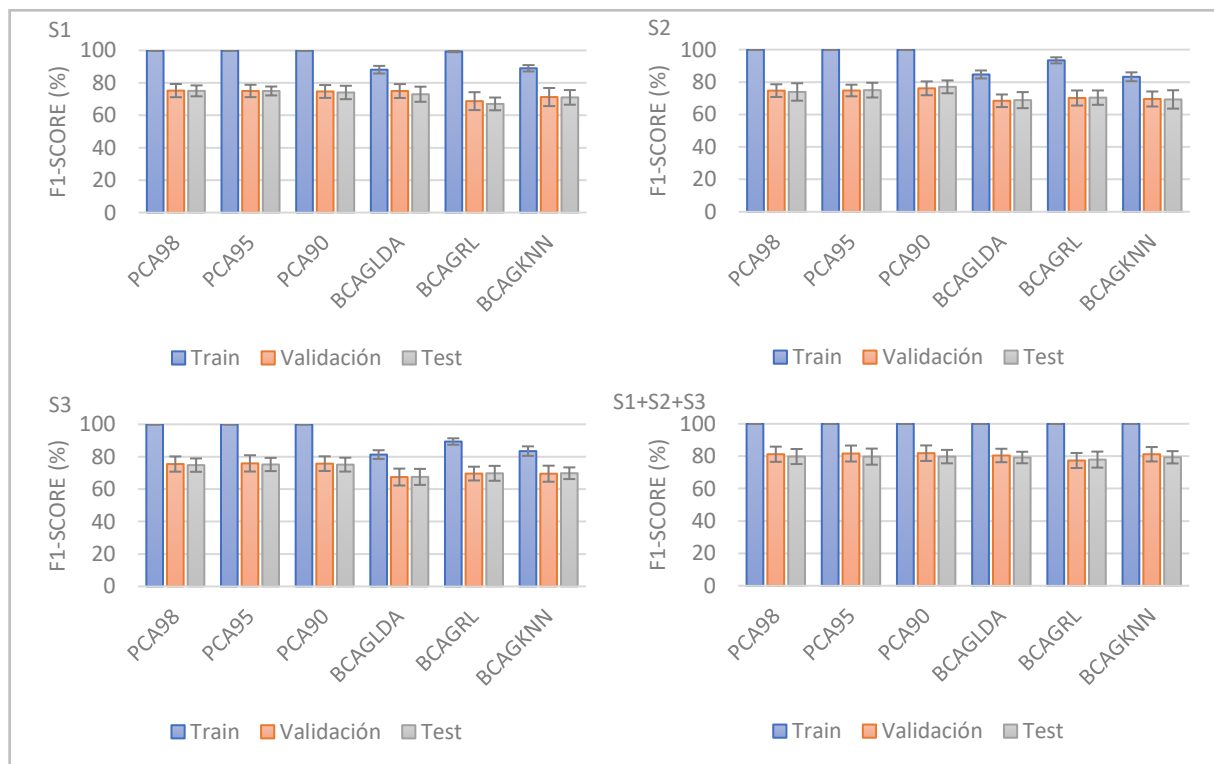


Figura 31. Resultados obtenidos utilizando el clasificador ELM. En la fila de arriba encontramos, a la izquierda, el estudio referente al canal S1, a su derecha el canal S2. En la fila de abajo, a la izquierda, encontramos S3, a su derecha S1+S2+S3. Fuente: elaboración propia.

En la *Figura 32*, se han reunido los resultados obtenidos para el grupo de test de cada una de las distintas técnicas de selección de características empleadas en cada una de sus variantes. Los mejores resultados corresponden al canal compuesto por la unión los canales individuales, concretamente para el selector de características dado por el Análisis de Componentes Principales con nivel 98, PCA98, con el cual se obtiene un F1-Score máximo del 79,78%.

Por lo tanto, en cuanto a la clasificación mediante ELM, se puede afirmar, que los mejores resultados, utilizando la métrica F1-Score, se dan lugar empleando las características seleccionadas por el Análisis de Componentes Principales, PCA98, y posteriormente aplicadas al clasificador ELM. Obteniendo de esta manera un resultado del 79,78%.

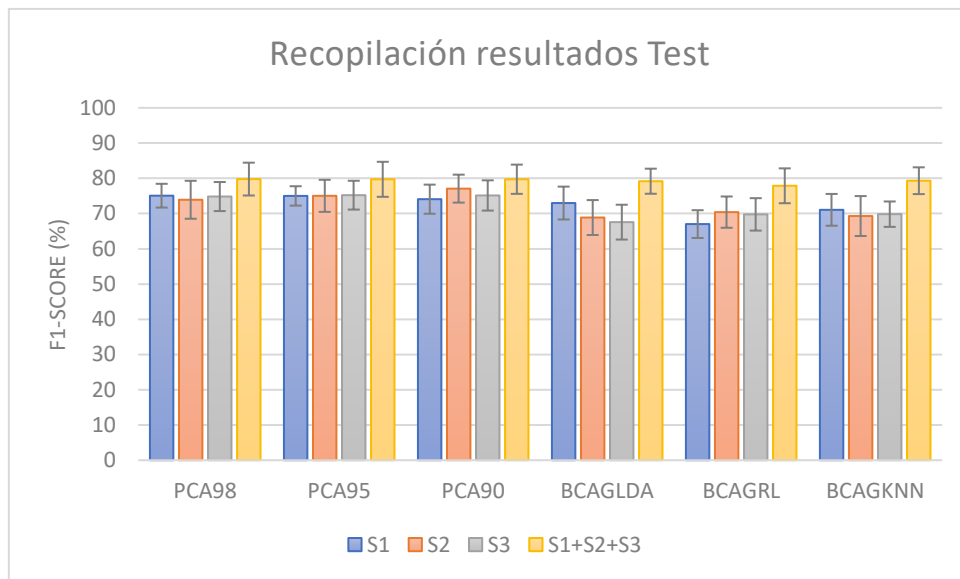


Figura 32. Recopilación de valores de los resultados en test con distintas selecciones de características para los distintos canales. Fuente: elaboración propia.

5.5. Comparativa de los distintos clasificadores con diferente número de población en AG.

Como se explicó en el apartado de métodos, una vez identificado que la unión de las características de entrada de todos los canales de registro (S1+S2+S3) resulta en un mejor desempeño de los clasificadores respecto de utilizar las características de los canales individuales cuando se aplica AG como algoritmo de selección de características, se procede al estudio a analizar cómo afecta el tamaño poblacional en el desempeño de los mismos al variar el tamaño poblacional en la técnica de selección de características.

En la *Figura 33* se observan los resultados que se obtuvieron con AG y LDA para tamaños poblacionales de 50, 100, 200, 300 y 400 individuos. Si nos fijamos en los resultados de test, el mejor valor fue del 89,13% para una población de 300, seguido por un 89,03% obtenido por una población de 200. Debido a la semejanza de los resultados se decidió añadir la métrica del tiempo computacional para ver, si de esta manera, se obtenía algo más de información que ayudara a elegir una población óptima. El tiempo computacional empleado para llegar al resultado de 89,03% con una población de 200 fue de 18955 segundos (5,3 horas), mientras que para conseguir el resultado de 89,13% con una población de 300 el tiempo empleado fue de 33238 segundos (9,2 horas). Viendo la pequeña diferencia entre los resultados de F1-Score promedio y la gran diferencia entre tiempo computacional, en este caso el tamaño poblacional óptimo fue de 200 individuos.

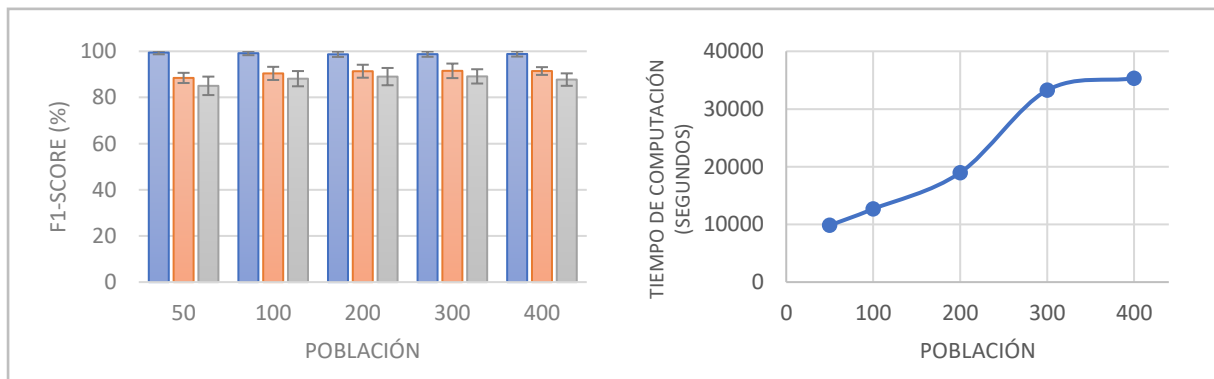


Figura 33. La figura de la izquierda hace referencia a los resultados obtenidos para la unión de los canales S1+S2+S3 con AG y el clasificador LDA, siendo los colores entrenamiento, validación y test respectivamente. A su derecha, se representa el tiempo computacional que da lugar al cómputo de distintas poblaciones. Fuente: elaboración propia.

En la *Figura 34* se resumen los resultados obtenidos para el estudio poblacional empleando el clasificador RL. Viendo los resultados de test, se puede observar que el mejor F1-Score promedio es 85,78%, obteniéndose este con una población de 200 individuos. El próximo valor más cercano es el resultado de 85,54% obtenido con una población de 100 individuos. Si se observa el tiempo computacional, tenemos un valor de 41720 segundos (11,6 horas) para la población de 100 mientras que el tiempo computacional fue de 70241 segundos (19,5 horas). Viendo la gran diferencia en tiempo computacional y la pequeña diferencia entre los resultados obtenidos, en este caso la población óptima para dicho estudio sería de una población de 100 individuos.

Desarrollo de un sistema de predicción del parto prematuro basado en electrohisterografía y algoritmo genético para la selección de características óptimas

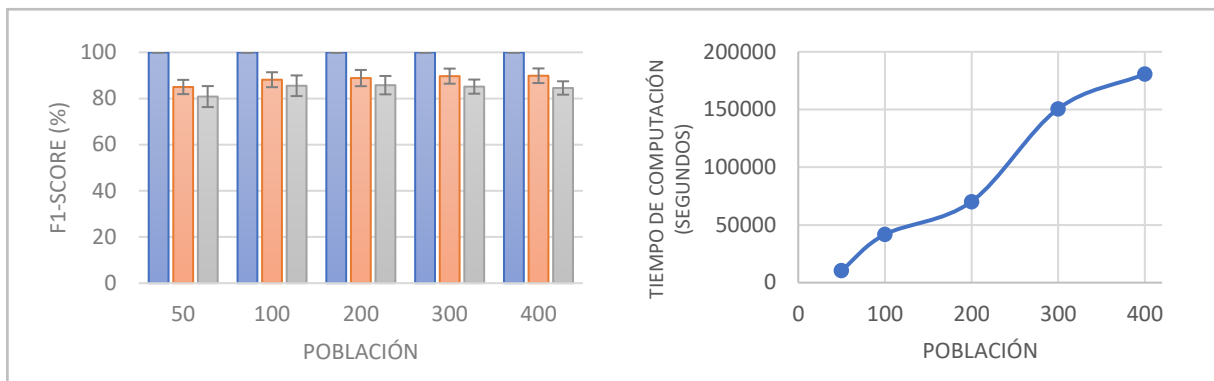


Figura 34. La figura de la izquierda hace referencia a los resultados obtenidos para la unión de los canales S1+S2+S3 con AG y el clasificador RL, siendo los colores entrenamiento, validación y test respectivamente. A su derecha, se representa el tiempo computacional que da lugar el cómputo de distintas poblaciones. Fuente: elaboración propia.

Por último, en la Figura 35, observamos los datos necesarios para realizar la elección de la población óptima utilizando el clasificador KNN con número de vecinos igual a tres. El mejor resultado fue obtenido con una población de 200, llegando al valor de F1-Score promedio del 81,91%, seguido por un 81,58% con una población de 400. Si nos centramos en el tiempo computacional, la población de 200 empleó 20397 segundos (5,7 horas) mientras que la población de 400 empleó 42683 segundos (11,9 horas). Viendo la diferencia de resultados y el tiempo computacional, la población óptima para este clasificador es de 200 individuos.

Con todos estos resultados se llegó a la conclusión de trabajar con los resultados obtenidos previamente con los valores de población 200 ya que, aun siendo verdad que en algunos casos el tiempo computacional era mayor, en la mayoría de las ocasiones fue la mejor opción para obtener resultados óptimos.

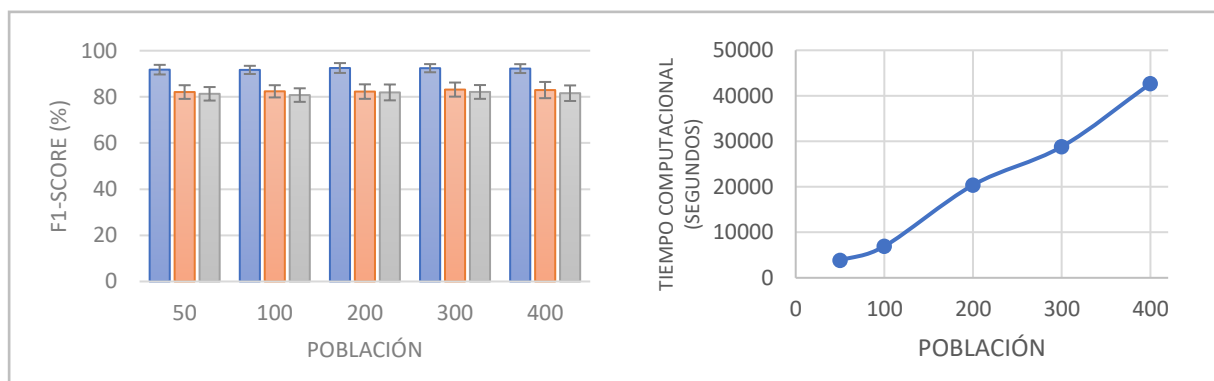


Figura 35. La figura de la izquierda hace referencia a los resultados obtenidos para la unión de los canales S1+S2+S3 con AG y el clasificador KNN, siendo los colores entrenamiento, validación y test respectivamente. A su derecha, se representa el tiempo computacional que da lugar el cómputo de distintas poblaciones. Fuente: elaboración propia.

5.6 Bootship Aggregating (Bagging).

Como fue explicado en el apartado de métodos, con la metodología *Bagging* se pretende mejorar los resultados obtenidos mediante la combinación de las salidas predichas por los tres mejores clasificadores aplicando la técnica de *majority voting*. En la Tablas 5, 6, 7 y 8 se han calculado las matrices de confusión correspondientes a los resultados obtenidos con los clasificadores LDA, RL, KNN y el obtenido tras combinar las salidas de los mismos (Bagging). Estos resultados pertenecen al grupo de test.

En la Tabla 9, se recogen las distintas métricas que permiten evaluar cada uno de los clasificadores que mejor resultado han proporcionado. En este apartado es de interés estudiar la nueva clasificación dada por esta técnica. Como se puede observar el número de aciertos es notablemente mayor llegando a conseguir un valor de 98,57%. En cuanto a la sensibilidad, que es la probabilidad de que ante un caso positivo el modelo lo clasifique como tal, la probabilidad asciende a un 95%, dos puntos superiores al más próximo por detrás. Si nos fijamos en la especificidad, probabilidad de que ante un caso negativo el método lo clasifique como tal, llega a un valor del 91%, varios puntos por encima de los otros métodos. Por último, en cuanto al F1-Score, que medía la precisión de un test, se llega a valores del 93% mientras que los demás clasificadores llegan a un 90%.

		CLASIFICADOR LDA – S1+S2+S3 – AG (200)	
		1	0
VALOR REAL	1	48.9±2.63 VP	4.1±2.63 FN
	0	7.93±2.62 FP	45.07±2.62VN

Tabla 5. Matriz de confusión AG con clasificador LDA.

		CLASIFICADOR RL – S1+S2+S3 – AG (200)	
		1	0
VALOR REAL	1	44.17±2.94 VP	8.83±2.94 FN
	0	5.77±2.75 FP	47.23±2.75 VN

Tabla 6. Matriz de confusión AG con clasificador RL.

		CLASIFICADOR KNN – S1+S2+S3 – PCA98	
		1	0
VALOR REAL	1	49.33±2.38 VP	3.67±2.38 FN
	0	7.07±2.84 FP	45.93±2.84 VN

Tabla 7. Matriz de confusión PCA98 con clasificador KNN.

		CLASIFICADOR Bagging	
		1	0
VALOR REAL	1	50.17±2.08 VP	2.83±2.08 FN
	0	4.63±2.025 FP	48.37±2.02 VN

Tabla 8. Matriz de confusión clasificador Bagging.

Desarrollo de un sistema de predicción del parto prematuro basado en electrohisterografía y algoritmo genético para la selección de características óptimas

	PCA98-KNN	AG-LDA	AG-RL	Bagging
ACIERTOS (%)	95.26	93.97	91.4	98.57
SENSIBILIDAD (%)	93	92	83	95
ESPECIFICIDAD (%)	87	85	89	91
VPP (%)	87	86	88	91
VPN (%)	93	92	84	94
AUC	0.99	0.98	1	0.93
F1-Score (%)	90	90	86	93

Tabla 9. Distintas métricas para la comparación de los distintos clasificadores en el grupo de test. Fuente: elaboración propia.

En la *Figura 36* se recogen distintas métricas que ayudan a ver la calidad de cada uno de los clasificadores, así como el nuevo obtenido mediante la técnica de Bagging. Como se puede observar, esta técnica ha proporcionado una clasificación que ha obtenido un 98,57% de aciertos, con un F1-Score promedio del 93%.

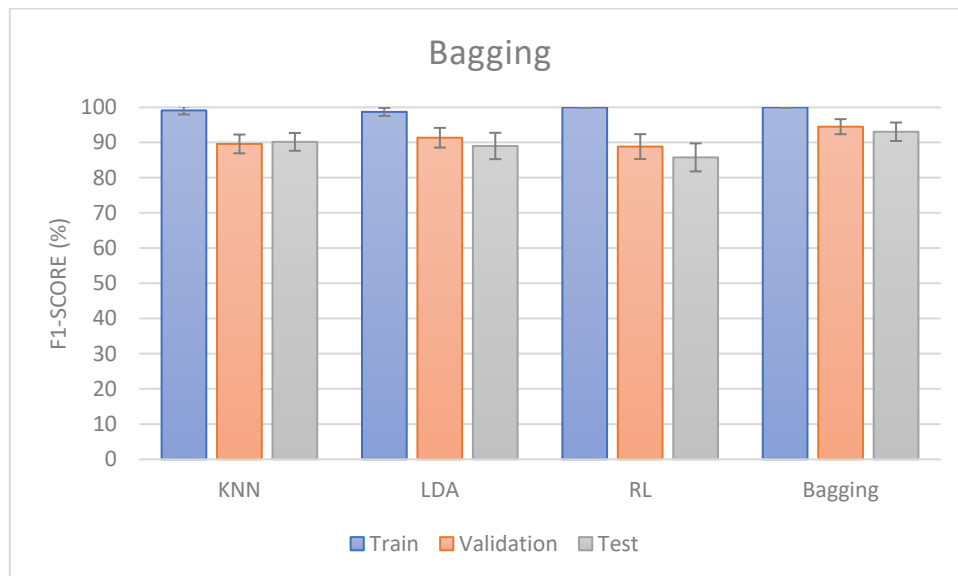


Figura 36. Resultados de los distintos clasificadores comparado con los resultados obtenidos por el método Bagging. Fuente: elaboración propia.

CAPÍTULO 6. CONCLUSIONES

6. Conclusiones

En cuanto al estudio inicial sobre información contenida en los canales de registro individuales S1, S2 y S3 para la predicción del parto prematuro, los resultados de los distintos clasificadores son muy similares para todos ellos, sin que destaque ninguno sobre los demás. Excepcionalmente, solo en el caso del canal S1 parece destacar en uno de los clasificadores desarrollados (PCA98+KNN), dichos resultados fueron computados para un número de vecinos de tres. No obstante, al combinar la información aportada por los tres canales S1+S2+S3, en general, se mejora la capacidad predictora de todos los clasificadores individuales desarrollados, independientemente de la técnica de selección de características empleada (Análisis de Componentes Principales o algoritmo genético).

Respecto a la selección de características, tal y como se ha comentado, mediante la técnica de Análisis de Componentes Principales (PCA), concretamente con PCA98 sobre las características del canal S1+S2+S3, se ha llegado a obtener, mediante un clasificador K vecinos más próximos (KNN) con parámetro óptimo de dos vecinos, un F1-Score promedio del 90,19%. Este resultado es realmente bueno si es comparado con los demás resultados obtenidos con AG, además el tiempo de computación es de tan solo 955ms. El buen resultado obtenido, así como el tiempo de computación extremadamente pequeño hace que sea posible una línea futura de investigación en torno a este resultado.

El Algoritmo Genético (AG) ha resultado ser un selector de características funcional y de altas prestaciones, pero a la vez complejo de configurar. Diversos parámetros como número de generaciones, criterios de parada, porcentaje de mutación y crossover, población, función de ajuste y algunos parámetros más, hacen de dicho método la necesidad de que sea conocido en profundidad tanto el propio método, así como el problema a estudio. Empleando AG como técnica de selección de características (*Ver anexo*) el mejor resultado se ha obtenido utilizando el clasificador LDA alcanzando un valor de F1-Score promedio del 89% con tan solo 46 características de las 205 iniciales. El único inconveniente encontrado en el uso de este algoritmo para la selección de características es el alto tiempo de cómputo, ya que para conseguir el valor de F1-Score promedio del 89% el tiempo total fue de 5,3 horas. Claro está, este tiempo podría verse reducido con un sistema de mayores prestaciones.

Los mejores resultados han sido obtenidos mediante la técnica Bagging, aplicando un *majority voting* a las salidas proporcionadas por los tres clasificadores que han dado mejores valores de F1-Score para el grupo de test empleando como características de entrada las asociadas a la combinación de las características de entrada de los tres canales de registro (S1+S2+S3) de EHG: PCA98-KNN, AG-LDA y AG-RL. Con dicho método se ha conseguido obtener un F1-Score promedio del 93% lo cual es superior a cada uno de los valores obtenidos por separado.

Por último, se exponen tres líneas futuras que podrían partir de este trabajo. La primera de ellas sería explorar el rendimiento del sistema predictor al incorporar otras características de EHG, por ejemplo, mediante el cómputo de versiones multivariantes de parámetros ya empleados en el presente TFG (como la entropía muestral multivariante o el Lempel-Ziv multivariante) o incluyendo parámetros relacionados con la sincronización y acoplamiento entre los diferentes canales de registro. La segunda línea sería la posibilidad de ahondar en técnicas de Bagging de manera que se realizara una

comparación de los resultados obtenidos mediante *majority voting* con otras técnicas de combinación de las salidas de los clasificadores. La tercera línea que podría darse lugar es extender este estudio a otros escenarios obstétricos en los que se emplea el EHG como predictor, como pudiera ser la predicción temprana del éxito de la inducción.

CAPÍTULO 7. ANEXOS

Característica	Descripción	Característica	Descripción
Fmediana_S1	Mediana frecuencia S1	FuzzyEn5_S2	Fuzzy Entropy en el ancho 1 a 4 Hz
SpecMR_S1	Ratio momento espectral S1	Tim Rev1_S2	Time Reversibility en el ancho 0,1 a 4 Hz
AUC_S1	Área bajo la curva ROC S1	Tim Rev4_S2	Time Reversibility en el ancho 0,34 a 1 Hz
EHG/LC3_S1	-	Varwt1_S2	Varianza de los coeficientes wavelet nivel 1
LZ_Bin3_S1	Lempel-Ziv Binario en el ancho 0,34 a 4 Hz S1	Varwt2_S2	Varianza de los coeficientes wavelet nivel 2
LZ_Multi2_S1	Lempel-Ziv Multiestado (7 estados) en el ancho 0,2 a 0,34 Hz S1	Varwt4_S2	Varianza de los coeficientes wavelet nivel 4
LZ_Multi3_S1	Lempel-Ziv Multiestado (7 estados) en el ancho 0,34 a 4 Hz	En1_S3	Energía normalizada en el ancho 0,1 a 0,34
SampEn2_S1	Sample Entropy en el ancho 0,2 a 0,34 Hz	En2_S3	Energía normalizada en el ancho 0,34 a 0,6
SampEn3_S1	Sample Entropy en el ancho 0,34 a 4 Hz	Dec8_S3	Deciles del Espectro
Tim Rev1_S1	Time Reversibility en el ancho 0,1 a 4 Hz	SpecMR_S3	Ratio momento espectral S1
Tim Rev3_S1	Time Reversibility en el ancho 0,34 a 4 Hz	Kurtosis_S3	Curtosis de la envolvente
SpEntr2_S1	Spectral Entropy en el ancho 0,2 a 0,34 Hz	LZ_Bin3_S3	Lempel-Ziv Binario en el ancho 0,34 a 4 Hz
SpEntr5_S1	Spectral Entropy en el ancho 1 a 4 Hz	LZ_Multi1_S3	Lempel-Ziv Multiestado (7 estados) en el ancho 0,1 a 4 Hz
Varwt1_S1	Varianza de los coeficientes wavelet nivel 1	LZ_Multi2_S3	Lempel-Ziv Multiestado (7 estados) en el ancho 0,2 a 0,34 Hz
PonC2_S1	Desviación en la dimensión 2 de la elipse de Poincare	LZ_Multi5_S3	Lempel-Ziv Multiestado (7 estados) en el ancho 1 a 4 Hz
DF1_S2	Frecuencia dominante en el rango: 0,1 a 0,2 Hz	SampleEn2_S3	Sample Entropy en el ancho 0,2 a 0,34 Hz
Dec1_S2	Deciles del Espectro	FuzzyEn2_S3	Fuzzy Entropy en el ancho 0,2 a 0,34 Hz
Dec7_S2	Deciles del Espectro	FuzzyEn5_S3	Fuzzy Entropy en el ancho 1 a 4 Hz
LZ_Bin4_S2	Lempel-Ziv Binario en el ancho 0,34 a 1 Hz	Time Rev1_S3	Time Reversibility en el ancho 0,1 a 4 Hz
LZ_Bin5_S2	Lempel-Ziv Binario en el ancho 1 a 4 Hz Lempel-Ziv	Time Rev2_S3	Time Reversibility en el ancho 0,2 a 0,34 Hz
LZ_Multi3_S2	Multiestado (7 estados) en el ancho 0,34 a 4 Hz	Time Rev3_S3	Time Reversibility en el ancho 0,34 a 4 Hz

Desarrollo de un sistema de predicción del parto prematuro basado en electrohisterografía y algoritmo genético para la selección de características óptimas

LZ_Multi5_S2	Lempel-Ziv Multiestado (7 estados) en el ancho 1 a 4 Hz	SpEntr3_S3	Spectral Entropy en el ancho 0,34 a 4 Hz
SampEn5_S2	Sample Entropy en el ancho 1 a 4 Hz	Varwt3_S3	Varianza de los coeficientes wavelet nivel 3

Tabla 10. Características seleccionadas en el mejor cromosoma obtenido con el AG y LDA.

En la *Tabla 10* se observan las características que forman el mejor cromosoma, formado por las características seleccionadas como óptimas mediante el algoritmo genético con la función de ajuste dada por LDA. Se aprecia que existen características de los tres distintos canales y una combinación de todas ellas es la que da el mejor resultado. Además, se concluye que el selector de características no selecciona, por lo general, la misma característica para distintos canales, ya que, como pudimos ver anteriormente todos los canales nos daban resultados parecidos, por lo que seleccionar las mismas características para distintos canales no aportaría información relevante.

Característica	Descripción	Característica	Descripción
E3_S1	Energía normalizada en el ancho 0,6 a 1	SampEn5_S2	Sample Entropy en el ancho 1 a 4 Hz
Dec4_S1	Deciles del Espectro	FuzzyEn4_S2	Fuzzy Entropy en el ancho 0,34 a 1 Hz
LZ_Bin4_S1	Lempel-Ziv Binario en el ancho 0,34 a 1 Hz	Time Rev3_S2	Time Reversibility en el ancho 0,34 a 4 Hz
LZ_Multi3_S1	Lempel-Ziv Multiestado (7 estados) en el ancho 0,34 a 4 Hz	Varwt2_S2	Varianza de los coeficientes wavelet nivel 2
SampEn2_S1	Sample Entropy en el ancho 0,2 a 0,34 Hz	Vart4_S2	Varianza de los coeficientes wavelet nivel 4
SampEn3_S1	Sample Entropy en el ancho 0,34 a 4 Hz	Dec4_S3	Deciles del Espectro
Time Rev1_S1	Time Reversibility en el ancho 0,1 a 4 Hz	Dec7_S3	Deciles del Espectro
SpEntr2_S1	Spectral Entropy en el ancho 0,2 a 0,34 Hz	LZ_Bin3_S3	Lempel-Ziv Binario en el ancho 0,34 a 4 Hz
SpEntr3_S1	Spectral Entropy en el ancho 0,34 a 4 Hz	LZ_Bin5_S3	Lempel-Ziv Binario en el ancho 1 a 4 Hz
SpEntr4_S1	Spectral Entropy en el ancho 0,34 a 1 Hz	SampEn2_S3	Sample Entropy en el ancho 0,2 a 0,34 Hz
SpEntr5_S1	Spectral Entropy en el ancho 1 a 4 Hz	Time Rev1_S3	Time Reversibility en el ancho 0,1 a 4 Hz
Varwt1_S1	Varianza de los coeficientes wavelet nivel 1	Time Rev2_S3	Time Reversibility en el ancho 0,2 a 0,34 Hz
DF1_S2	Frecuencia dominante en el rango: 0,1 a 0,2 Hz	Time Rev5_S3	Time Reversibility en el ancho 1 a 4 Hz
Dec1_S2	Deciles del Espectro	SpEntr1_S3	Spectral Entropy en el ancho 0,1 a 4 Hz
Dec2_S2	Deciles del Espectro	SpEntr5_S3	Spectral Entropy en el ancho 1 a 4 Hz

Desarrollo de un sistema de predicción del parto prematuro basado en electrohisterografía y algoritmo genético para la selección de características óptimas

Dec7_S2	Deciles del Espectro	Varwt1_S3	Varianza de los coeficientes wavelet nivel 1
SampEn1_S2	Sample Entropy en el ancho 0,1 a 4 Hz	Varwt2_S3	Varianza de los coeficientes wavelet nivel 2

Tabla 11. Características seleccionadas en el mejor cromosoma obtenido con el AG y RL.

En la *Tabla 11* se observan las características que forman el mejor cromosoma, formado por las características seleccionadas como óptimas mediante el Algoritmo genético con la función de ajuste dada por RL. Como en el caso anterior, se aprecia que el cromosoma está formado por características de los tres distintos canales y una combinación de todas ellas es la que da el mejor resultado. Igualmente, se concluye que el selector de características no selecciona, por lo general, la misma característica para distintos canales, ya que, como pudimos ver anteriormente todos los canales nos daban resultados parecidos, por lo seleccionar las mismas características para distintos canales no aportaría información relevante.

CAPÍTULO 8. BIBLIOGRAFÍA

8. Referencias

- ACOG (2016) 'Practice Bulletin No. 171', *Obstetrics & Gynecology*, 128(4), pp. e155–e164. doi: 10.1097/AOG.000000001711.
- Alberola-Rubio, J. *et al.* (2013) 'Comparison of non-invasive electrohysterographic recording techniques for monitoring uterine dynamics', *Med Eng Phys.* 2013, 35(12), pp. 1736–1743. doi: 10.1016/j.medengphy.2013.07.008.
- Amat J. (2017). Árboles de predicción: Bagging, random forest, boosting y C5.0. Cienciadedatos. <https://www.cienciadedatos.net/documentos/33_arboles_de_prediccion_bagging_random_forest_boosting#bagging>.
- Amat J. (2019). *Análisis de Componentes Principales (Principal Component Analysis, PCA) y t-SNE*. Cienciadedatos. <https://www.cienciadedatos.net/documentos/35_principal_component_analysis> [25 mayo 2020].
- Amat J. (2020). *Análisis discriminante lineal (LDA) y análisis discriminante cuadrático (QDA)*. <https://www.cienciadedatos.net/documentos28_linear_discriminant_analysis_lda_y_quadratic_discriminant_analysis_qda> [22 mayo 2020].
- Amat J. (2020). *Análisis discriminante lineal (LDA) y análisis discriminante cuadrático (QDA)*. <https://www.cienciadedatos.net/documentos28_linear_discriminant_analysis_lda_y_quadratic_discriminant_analysis_qda> [22 mayo 2020].
- Amsantac.co (2011). *Soluciones eficientes para el procesamiento de datos geoespaciales*. <<http://amsantac.co/blog/es/2016/09/20/balanced-image-classification-r-es.html>> [20 septiembre 2016].
- Babatunde *et al.* (1973). 'A Genetic Algorithm Based Feature Selection', *International Journal of Electronics Communication and Computer Engineering*. 2014, 5(4), pp. 889-905 (Online): 2249–071X, ISSN (Print): 2278–4209.
- Bing Liu (2011). 'Web Data Mining. Exploring Hyperlinks, Contents and Usage Data'. Ed. Springer. doi: 10.107/978-3-642-19460-3.
- Charfaoui Y. (2019). *Resampling to Properly Handle Imbalanced Datasets in Machine Learning*. Medium. <<https://heartbeat.fritz.ai/resampling-to-properly-handle-imbalanced-datasets-in-machine-learning-64d82c16ceaa>> [25 mayo 2020].
- Dag Moster, *et al.* (2008) 'Long-Term Medical and Social Consequences of Preterm Birth', *The New England Journal of Medicine*, 359(3), pp. 262-273. doi: 10.1056/NEJMoa0706475.
- Devedeux, D. *et al.* (1993) 'Uterine electromyography: a critical review', *Am J Obstet Gynecol.* 1993/12/01, 169(6), pp. 1636–1653. Doi: 10.1016/0002-9378(93)90456-S.
- Di Renzo, G. C. *et al.* (2017) 'Preterm Labor and Birth Management: Recommendations from the European Association of Perinatal Medicine', *The Journal of Maternal-Fetal & Neonatal Medicine*, 30(17), pp. 2011–2030. doi: 10.1080/14767058.2017.1323860.

- Di Renzo, G. C. *et al.* (2017) 'Preterm Labor and Birth Management: Recommendations from the European Association of Perinatal Medicine', *The Journal of Maternal-Fetal & Neonatal Medicine*, 30(17), pp. 2011–2030. doi: 10.1080/14767058.2017.1323860.
- Ding S. *et al.* (2015). 'Extreme learning machine: algorithm, theory and applications', *Artificial Intelligence*. 2015, 44(1), pp. 103-115 doi: 10.1007/s10462-013-9405-z
- Dolores *et al.* (2000). 'La regresión logística: una herramienta versátil', *Nefrología*. 2000, 20 (6), pp. 477-565 ISSN: 0211-6995, e-ISSN: 1989-2284
- Fele-Zorz, G. *et al.* (2008) '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*. 2008/04/26, 46(9), pp. 911–922. doi: 10.1007/s11517-008-0350-y.
- Fernández A. (2017). *Entendiendo los algoritmos genéticos. Un caso de uso en el entorno organizacional*. <<https://planetachatbot.com/entendiendo-los-algoritmos-gen%C3%A9ticos-un-caso-de-uso-en-el-entorno-organizacional-a745c157fa8c>> [02 junio 2020].
- Francés Ribera L. *et al.* (2014) 'Control del bienestar fetal. Monitorización biofísica anteparto', *Revista Rol de Enfermería*, 37(12), pp. 817-822. ID Medes: 96373.
- Frey, H. A. and Klebanoff, M. A. (2016) 'The epidemiology, etiology, and costs of preterm birth', *Seminars in Fetal and Neonatal Medicine*, 21(2), pp. 68–73. doi: 10.1016/j.siny.2015.12.011.
- Garfield, R. E. and Maner, W. L. (2007) 'Physiology and electrical activity of uterine contractions', *Semin Cell Dev Biol*. 2007/07/31, 18(3), pp. 289–295. doi: 10.1016/j.semcdb.2007.05.004.
- Haibo H. Yunquian M., (2013). *Imbalanced Learning*. Ed. Wiley, p.47.
- Honest, H. *et al.* (2009) 'Screening to prevent spontaneous preterm birth: systematic reviews of accuracy and effectiveness literature with economic modelling.', *Health technology assessment (Winchester, England)*, 13(43), pp. 1–627. doi: 10.3310/hta13430.
- IBM (2020). *Análisis del vecino más cercano*. IBM knowledge Center. <https://www.ibm.com/support/knowledgecenter/es/SSLVMB_sub/statistics_mainhelp_ddita/spss/base/idh_idd_knn_variables.html> [22 mayo 2020].
- Jackson J. Hearne F. (1973). 'Relationships Among Coefficients of Vectors Used in Principal Components', *Technometrics*. 1973, 15(3), pp. 601-610 doi: 10.1080/00401706.1973.10489087
- Jason Brownee (2020). *Machine Learning Mastery, Making Developers Awesome at Machine Learning*. <<https://machinelearningmastery.com/smote-oversampling-for-imbalanced-classification/>> [7 abril 2020].
- Li Liu, *et al.* (2016) 'Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals', *Lancet*, 388(1), pp. 3027-3035. doi: 10.1016/S0140-6736(16)31593-8.
- Lucovnik., M. *et al.* (2011) 'Noninvasive uterine electromyography for prediction of preterm delivery', in *American Journal of Obstetrics and Gynecology*, p. 228.e1-228.e10. doi: 10.1016/j.ajog.2010.09.024.
- Mackiewicz A. Ratajczak W. (1993). 'Principal Component Analysis (PCA)', *Computers & Geosciences*. 1993, 19(3), pp. 303-342 doi: 10.1016/0098-3004(93)90090-R

- Mas-Cabo J. *et al.* (2019). 'Design and Assessment of a Robust and Generalizable ANN-Based Classifier for the Prediction of Premature Birth by means of Multichannel Electrohysterographic Records', *Journal of Sensors*. 2019, 2019(1), pp. 1687 – 1700 doi: 10.1155/2019/5373810
- MathWorks (2020). *Algoritmo Genético — Determinación de los mínimos globales para problemas altamente no lineales*. MathWorks. <<https://es.mathworks.com/discovery/genetic-algorithm.html>> [24 mayo 2020].
- MathWorks (2020). *Introduction to Feature Selection*. <<https://es.mathworks.com/help/stats/feature-selection.html>> [14 mayo 2020].
- Melanie M., (1999). *An Introduction to Genetic Algorithms*. Ed. A Bradford Book, The MIT Press.
- Nitesh Chawla, *et al.*, 2002. 'SMOTE: Synthetic Minority Over-sampling Technique', *Journal of Artificial Intelligence Research*, 16(1), pp. 321–357.
- Oldemar (2016). *Aprendizaje Supervisado – Análisis Discriminante (Lineal y Cuadrático)*. Consulto en minería de datos. <<https://docplayer.es/1516332-Aprendizaje-supervisado-analisis-discriminante-lineal-y-cuadratico.html>>. [05 mayo 2020].
- Patel A. (2019). *Bagging — Ensemble meta Algorithm for Reducing variance*. Medium. <<https://medium.com/ml-research-lab/bagging-ensemble-meta-algorithm-for-reducing-variance-c98fffa5489f>> [23 mayo 2020].
- Polettini, J. *et al.* (2017) 'Biomarkers of spontaneous preterm birth: a systematic review of studies using multiplex analysis.', *Journal of perinatal medicine*, 45(1), pp. 71–84. doi: 10.1515/jpm-2016-0097.
- Robert L. Goldenbert, *et al.* (2008) 'Epidemiology and causes of preterm birth', *Lancet*, 371(1), pp. 75-84. doi: 10.1016/S0140-6736(08)60074-4.
- Romero, J. *et al.* (2012) 'The prediction of recurrent preterm birth in patients on 17-alpha-hydroxyprogesterone caproate using serial fetal fibronectin and cervical length.', *American journal of obstetrics and gynecology*, 207(1), p. 51.e1-5. doi: 10.1016/j.ajog.2012.04.024.
- Romero, R., Dey, S. K. and Fisher, S. J. (2014) 'Preterm labor: one syndrome, many causes', *Science*. 2014/08/16, 345(6198), pp. 760–765. doi: 10.1126/science.1251816.
- Schlembach, D. *et al.* (2009) 'Monitoring the progress of pregnancy and labor using electromyography', *Eur J Obstet Gynecol Reprod Biol.*, 144 Suppl, pp. S33-9. doi: 10.1016/j.ejogrb.2009.02.016.
- Sikora, J. *et al.* (2011) 'Recognition of premature threatening labour symptoms from bioelectrical uterine activity signals', *Archives of Perinatal Medicine*, 17(2), pp. 97–103.
- Sivanandam S. N. Deepa S. N. (2008). *Introduction to Genetic Algorithms*. Ed. Springer-Verlag, Berlin, Heidelberg, 2008.
- Stat Developer (2020). *Regresión logística en Python*. Stat Developer. <<http://www.statdeveloper.com/regresion-logistica-en-python/>> [08 junio 2020]
- Stavros Petrou, (2005) 'The economic consequences of preterm birth during th.e first 10 years of life', *BJOG: An International Journal of Obstetrics and Gynaecology*, 112(1), pp. 10-15. doi: 10.1111/j.1471-0528.2005.00577.x.

Zhang Y. *et al.*, (2018). 'Classification of Incomplete Data Based on Evidence Theory and an Extreme Learning Machine in Wireless Sensor Networks', *Sensors*. 2018, 18(4), pp. 1046 doi: 10.3390/s18041046

Documento II

Presupuesto

PRESUPUESTO GENERAL

1. Introducción.

Como todo trabajo de ingeniería es necesario presupuestar dicho proyecto y estudio. En el presente capítulo se desarrollará un presupuesto aproximado que abarque los materiales, manos de obra y unidad de obra.

Debido a la obsolescencia de los equipos informáticos en una media de 3 años, se decide que el periodo de amortización del equipo portátil sea de 3 años. Considerando que un año presenta 250 días hábiles y cada día está compuesto por 8 horas.

La licencia anual de Matlab se decide amortizar en 1 año, considerando que el año presenta 250 días hábiles y cada día está compuesto por 8 horas.

Por otro lado, la licencia, también anual, de Office 365 se decide amortizar en 1 año, como en los casos anteriores, un año presenta 250 días hábiles y cada día estará compuesto por 8 horas.

1.1. Desglose del presupuesto.

Mano de obra

REF.	Descripción	Precio	Comentario
MO.ING.BM	Ing. Biomédico	30€/h	-

Materiales

REF.	Descripción	Ud.	Cantidad	Precio en €	Importe €
MT.PC	Asus i5	U	1	1.200,00	1.200,00
MT.HDD	Toshiba 1TB	U	1	40,99	40,99
MT.MAT	Lic. Matlab anual	U	1	800	800
MT.OFFICE	Lic. Office anual	U	1	203	203
MT.KW	KW luz	U	1	0,19	0,19

Unidades de obra

CAPÍTULO 1: PROPUESTA DEL PROBLEMA A RESOLVER

Código 1.01 Estudio y propuesta del proyecto

<i>REF.</i>	<i>Descripción</i>	<i>Ud.</i>	<i>Cantidad</i>	<i>Precio en €</i>	<i>Importe €</i>
MO.ING.BM	Ing.Biomédico	h	10,00	30,00	300,00
MT.PC	Asus i5	h	10,00	0,21	21,00
	Costes directos complementarios	%	2		6,42

Subtotal	327,42 €
-----------------	-----------------

Código 1.02 Pruebas preliminares algoritmos

<i>REF.</i>	<i>Descripción</i>	<i>Ud.</i>	<i>Cantidad</i>	<i>Precio en €</i>	<i>Importe €</i>
MO.ING.BM	Ing.Biomédico	h	20,00	30,00	600,00
MT.PC	Asus i5	h	10,00	0,21	21,00
MT.MAT	Lic. Matlab	h	20	0,40	8
	Costes directos complementarios	%	2		12,58

Subtotal	641,58 €
-----------------	-----------------

Total Capítulo 1	969,00 €
-------------------------	-----------------

CAPÍTULO 2: Configuración Algoritmo Genético

Código 2.01 Investigación parámetros AG

<i>REF.</i>	<i>Descripción</i>	<i>Ud.</i>	<i>Cantidad</i>	<i>Precio en €</i>	<i>Importe €</i>
MO.ING.BM	Ing.Biomédico	h	5,00	30,00	150,00
MT.PC	Asus i5	h	5,00	0,21	1,05
	Costes directos complementarios	%	2		3,02

Subtotal	154,07 €
-----------------	-----------------

Total Capítulo 2	154,07 €
-------------------------	-----------------

CAPÍTULO 3: Cómputo de algoritmos y resultados

Código 3.01 Cómputo de algoritmos

<i>REF.</i>	<i>Descripción</i>	<i>Ud.</i>	<i>Cantidad</i>	<i>Precio en €</i>	<i>Importe €</i>
MO.ING.BM	Ing.Biomédico	h	20,00	30,00	600,00
MT.PC	Asus i5	h	200,00	0,21	42,00
MT.MAT	Lic. Matlab	h	200,00	0,40	80,00
MT.KW	KW luz	h	200,00	0,19	38,00
	Costes directos complementarios	%	2		15,20

Subtotal 775,20 €

Código 1.02 Presentación resultados

<i>REF.</i>	<i>Descripción</i>	<i>Ud.</i>	<i>Cantidad</i>	<i>Precio en €</i>	<i>Importe €</i>
MO.ING.BM	Ing.Biomédico	h	32,00	30,00	960,00
MT.PC	Asus i5	h	32,00	0,21	6,72
MT.OFFICE	Matlab	h	32,00	0,11	3,52
	Costes directos complementarios	%	2		19,41

Subtotal 989,65 €

Total Capítulo 3 1.764,85 €

Presupuesto total

Propuesta	969,00 €
Investigación	154,07 €
Cómputo	1.764,85 €

PEM	2.887,92 €
Gastos generales (13%)	375,43 €
Beneficio industrial (6%)	173,25 €

	3.436,6 €
IVA (21%)	721,69 €

Total presupuesto 4.158,29 €

El presupuesto total para dicho proyecto asciende a 4.158,29 €.

