MODELO BIO-INSPIRADO PARA LA APROXIMACIÓN DE NACIMIENTO POR CESÁREA

Autor

Jennifer Vivian Pacheco Fandiño

Directora

PhD, Luz Helena Camargo Casallas

Co-Director

MsC, Andrés Eduardo Gaona Barrera

Universidad Distrital Francisco José De Caldas

Maestría en Ingeniería

Énfasis en Ingeniería Electrónica – Línea de Investigación en Bioingeniería

Bogotá, Colombia

2023

TABLA DE CONTENIDO

Resumen 7
1. Introducción 8
1.1. Objetivos 12
1.2. Productividad académica 13
2. Marco de referencia 14
2.1. Generalidades del parto 14
2.2. Técnicas de monitoreo fetal 18
2.2.1. Electrohisterograma (EHG) 19
2.3. Modelos bio-inspirados 23
2.3.1. Clasificador K vecinos más cercanos (KNN) 24
2.3.2. Máquinas de vectores de soporte (SVM) 26
2.3.3. Redes neuronales (MLP) 27
2.3.4. Aprendizaje profundo (DP) 29
2.3.5. Desempeño de modelos bio-inspirados 30
2.4. Antecedentes 32
3. Metodología 38
3.1. Base de datos 39
3.1.1. Características demográficas 39
3.2. Caracterización señal EHG 43
3.3. Implementación y evaluación de clasificadores 48
3.3.1. Clasificador: K vecinos más cercanos (KNN) 48
3.3.2. Clasificador: perceptrón multicapa (MLP) 49
3.3.3. Clasificador: máquinas de vector soporte (SVM) 50
3.3.4. Clasificador: aprendizaje profundo (DP) 51

3.4. Implementación e interfaz gráfica 52

4. Resultados y discusión 55

4.1. Base de datos 55

4.1.1. Características demográficas 56

4.2. Caracterización señal EHG 64

4.3. Implementación y evaluación de clasificadores 68

4.3.1. Clasificador: K vecinos más cercanos (KNN) 70

4.3.2. Clasificador: perceptrón multicapa (MLP) 72

4.3.3. Clasificador: máquinas de vector soporte (SVM) 73

4.3.4. Clasificador: aprendizaje profundo (DP) 74

4.4. Implementación e interfaz gráfica 77

5. Conclusiones 80

6.Trabajo futuro 81

7. Bibliografía 83

8. ANEXOS 97

TABLA DE FIGURAS

Figura 1. Posición placentaria [33].....17
Figura 2. Monitoreo de presión intrauterina (IUC). (a) presión intrauterina [37], (b) gráfica de monitoreo [36].....18
Figura 3. Tocodinamometría [38].....19
Figura 4. Representación de la señal de IUP (a), onda lenta (b), onda rápida (c) y la suma de ambas (d) [39]. 20
Figura 5. Ubicación de electrodos para adquisición de señal EHG. (Elaboración propia).....21
Figura 6. Señal de contracción materna mediante IUC, tocodinamometría y EHG monopolar y EHG bipolar [41]. 22
Figura 7. Algoritmo Evolutivo. (Elaboración propia) 24
Figura 8. Representación gráfica del clasificador KNN. (Elaboración propia) 24
Figura 9. Kernel lineal. (Elaboración propia).....26
Figura 10. Kernel polinomial. (Elaboración propia) 27
Figura 11. Kernel gausiano. (Elaboración propia).....27
Figura 12. Diagrama MLP. (Elaboración propia).....28
Figura 13. Diferencias machine learning y DP 30
Figura 14. Modelo bio-inspirado. (Elaboración propia).....38
Figura 15. Distribución de la base de datos sin procesar. (Extraída de Weka). 41
Figura 16. Interpretación del Diagrama de Dispersión. Elaboración propia.....42
Figura 17. Electrodo Ag/AgCl [82]. 43
Figura 18. Distribución de los16 electrodos. 44
Figura 19. Diagrama de técnica de ventaneo EHG. (Extraída de Matlab).....45
Figura 20. Ubicación de electrodos (grupo 3 y 5) [73] (Elaboración propia).....48
Figura 21. Diseño de interfaz gráfica. (Elaboración propia) 54

Figura 22. Estructura de archivo con extensión .csv para interfaz gráfica. (Elaboración propia) 54
Figura 23. Fotografía escaneada [82]. 55
Figura 24. Ejemplo estructura archivo encabezado - información demográfica [82]. 56
Figura 25. Grafica de dispersión de los datos de entrada. Elaboración propia. 58
Figura 26. Gráfica del Coeficiente de correlación de Pearson r, para los partos por cesárea. Elaboración propia. 60
Figura 27. Distribución de la base de datos. (Extraída de Weka). 63
Figura 28. Estructura del dataset de información demográfica. (Extraida Dataset Physionet) 63
Figura 29. Interferencia de la Bio-señal EHG [63].....65
Figura 30. Bio-señal EHG, (a) señal total, (b) ventaneo. (Extraída de Matlab ®). ...65
Figura 31. Dataset EHG ejemplo (Extraída de Matlab ®). 66
Figura 32. Distribución de datos para el clasificador KNN. (Extraída de Weka).....70
Figura 33. MLP resultado red (Extraída de Weka). 73
Figura 34. Gráfico de pérdidas. (a) grupo A=1, (b) Grupo B=2. (Elaboración propia) 75
Figura 35. Interfaz de visualización gráfica. (Elaboración propia) 78

LISTA DE TABLAS

Tabla 1. Conjunto de entrada, clasificador KNN [47]. 25
Tabla 2. Resultados de modelo predictor de complicaciones perinatales *[64]*. 34
Tabla 3. Resultados de modelo con propagación hacia atrás [66]. 35
Tabla 4. Resultado de validación de IRM [68]. 35
Tabla 5. Resultados de predicción de parto prematuro [71]. 37
Tabla 6. Características demográficas. 39
Tabla 7. Distribución de técnica de ventaneo. (a) Preprocesamiento de EHG, (b) SM: características EHG 47
Tabla 8. Matriz de confusión. 52
Tabla 9. Requerimientos funcionales para interfaz gráfica. 52
Tabla 10. Valores de correlación [85] 59
Tabla 11. Información por grupos. 69
Tabla 12. Resultados del clasificador KNN. 71
Tabla 13. Resultados de clasificador MLP. 72
Tabla 14. Resultados de clasificador SVM. 73
Tabla 15. Resultados del clasificador DP. 74

RESUMEN

El parto por cesárea es una de las causas de mortalidad materna con mayor porcentaje, en Colombia para el 2021 fue del 46,4 %. Para mitigar esto los profesionales de la salud realizan seguimiento a la gestante, sin embargo, la densidad de información y el volumen de pacientes hacen complejo considerar su sintomatología en detalle.

Teniendo en cuenta una base de datos pública que contiene información demográfica de 45 gestantes y el registro de la bioseñal electrohistereograma (EHG), la cual es una técnica no invasiva que representa la actividad eléctrica de las células del útero. Se desarrolla un modelo bio-inspirado implementando los clasificadores: K vecinos cercanos (KNN), perceptrón multicapa (MLP), máquinas de vectores de soporte (SVM) y aprendizaje profundo (DP), para procesar grupos de datos con información demográfica, características de la bioseñal EHG, y su combinación. Finalmente, se evalúa el modelo bio-inspirado calculando los índices de desempeño: sensibilidad, especificidad y precisión.

El mejor rendimiento en el análisis de datos demográficos se obtiene con: KNN donde la sensibilidad es del 100 % y la especificidad es superior al 80 % junto a SVM con una sensibilidad del 75 % y una especificidad del 83,3 %. Los clasificadores de mayor desempeño en el análisis de la bio-señal EHG son: MLP con una sensibilidad del 82,3 % y una especificidad del 85,7 %, le sigue el DP con una especificidad de 72,8 %.

Los porcentajes de mayor acierto se encuentran cuando se consideran solo los datos demográficos o la bioseñal EHG, en contraste con el obtenido con los datos combinados. El modelo bio-inspirado permite una precisión del 90 % con información demográfica y del 90,9 % con la información de la EHG.

Palabras Claves: Embarazo, cesárea, inteligencia artificial, mortalidad materna, aproximación, predicción.

1. INTRODUCCIÓN

Uno de los objetivos de Desarrollo del Milenio de la Organización Mundial de la Salud (OMS) es mejorar la salud materna, una de las metas de este es reducir la mortalidad materna en tres cuartas partes entre 1990 y 2015 a nivel mundial [1]. Sin embargo, en el 2005 la OMS y el Fondo de las Naciones Unidas para la Infancia (Unicef) presentan cifras globales de 400 muertes maternas por 100.000 nacidos vivos en el informe de la mortalidad materna [2] [3].

En este mismo contexto la Organización Panamericana de la Salud (OPS) indica en el informe de abril de 2019 que “En el mundo diariamente fallecen 830 mujeres gestantes por 100.000 nacidos vivos”, en consecuencia, se presenta como objetivo a 2030 que ningún país debería superar la tasa de mortalidad materna de 70 muertes por cada 100.000 niños nacidos vivos [4]. En particular, la tasa de mortalidad materna en Colombia aumento de 75 a 83.2 por 100.000 nacidos vivos en el periodo 2010-2021 [5].

Respecto a las causas que generan esta tasa de mortalidad materna, se ha establecido que obedece a complicaciones y afecciones significativas, tales como: la eclampsia, diabetes gestacional, hipertensión [6] [7], hemorragias, cesárea, abortos y desprendimiento de placenta [8].

Ante este panorama mundial, en Colombia la evidencia muestra que la cesárea representa una de las problemáticas a resolver. Por ello, en el año 2015 se determinó que la tasa ideal de mortalidad por esta causa debía ser menor al 15 % [9] [10]. Pero, a pesar de los esfuerzos por reducir estas cifras, los resultados no han sido satisfactorios, dado que las cesáreas realizadas a 2014 correspondieron al 46,6 %, mientras que en el 2021 fue del 46,4 % [11].

El nacimiento por cesárea, es una operación o intervención quirúrgica, la cual tiene la finalidad de extraer el feto por vía abdominal [12]. Esto obedece a dos factores básicos; los maternos relacionados con hemorragias, cáncer cérvico, hipertensión arterial en el embarazo, también conocida como preeclampsia [13] [14] [15] [16], insuficiencia renal, enfermedad hepática o feto placentarias. Y los factores fetales como anomalía en la posición, parto prematuro y embarazo gemelar [17]. Efectuar la cesárea en lugares que carecen de medidas de bioseguridad para realizar cirugías de forma biosegura suele provocar complicaciones y discapacidades significativas [8].

Se han realizado estudios para poder determinar si la labor de parto requiere de esta práctica, por ejemplo, se usan métodos no lineales para analizar el electrohisterograma (EHG) de mujeres gestantes [18]. En otros trabajos se desarrolla una herramienta de clasificación de mujeres que presentan una morbilidad hipertensiva y su relación con la cesárea [19].

Sin embargo, la información necesaria para realizar una cesárea se toma durante la emergencia. Por lo tanto, el enfoque principal de este trabajo es desarrollar un modelo bio-inspirado utilizando técnicas de inteligencia computacional como KNN, MLP, SVM y DP para respaldar el diagnóstico durante el embarazo de abordar un parto por cesárea, basado en información demográfica y la bioseñal EHG de la mujer gestante. Los resultados muestran la efectividad del modelo porque se logra una sensibilidad de hasta el 90 %.

El documento está organizado de la siguiente manera: en el capítulo 1 se plantea el contexto del problema, la pregunta problema y los objetivos del proyecto; en el capítulo 2 se exponen los conceptos a tener en cuenta para la comprensión del documento: generalidades del parto, técnicas de monitoreo fetal, modelos bio-inspirados y antecedentes encontrados en la literatura sobre la solución de la pregunta problema; en el capítulo 3 se ilustra el modelo bio-inspirado y se describe

la metodología; en el capítulo 4 se muestran los resultados obtenidos y su análisis. Finalizando con las conclusiones en el capítulo 5 y el trabajo futuro.

Planteamiento del problema

Durante la labor de parto es posible que se presenten complicaciones que concluyan en una intervención quirúrgica, por ejemplo: eclampsia, diabetes gestacional, hemorragias, abortos y desprendimiento de placenta. Para la detección de estas distocias se han usado análisis de diferentes parámetros como lo son [6]:

- La triada de la hipertensión conformada por los valores de tensión sistólica, diastólica y arterial media.
- La proteinuria.
- La diabetes.
- El edema.
- El código rojo que se da por hipotonía del útero.
- Las distocias cefalopélvicas, es decir, la pelvis de la gestante no es apta para la labor de parto.

Desde la academia se encuentran estudios de exámenes como el Holter (presión arterial) o Doppler (comportamiento sanguíneo), los cuales permiten evidenciar problemas en la salud materna, entre ellos: la hipertensión o la vasoconstricción de las arterias del útero [7]. Se han planteado diferentes alternativas para procesar esta información, por ejemplo, se emplean las redes neuronales para el tratamiento de una base de datos que contiene 17 parámetros, con el fin de clasificar las maternas que puedan padecer morbilidades hipertensivas, obteniendo un acierto del 50 % [20], es decir, la mitad de las gestantes clasificadas concluyeron su embarazo con el pronóstico dado. Por otra parte, se utilizan redes bayesianas en el tratamiento de 11 parámetros donde se incluye información demográfica e historial clínico, determinando si la paciente sufre de preeclampsia y en consecuencia parto por cesárea, esto con un porcentaje de acierto del 66.7 % [21].

Otro método encontrado es el agrupamiento, con el cual se procesa una base de datos que contiene información demográfica como edad, número de embarazos, último período menstrual, fecha estimada del parto, altura y grupo sanguíneo, proceso que presenta un 78 % de aciertos frente a una base de datos ya conocida, no obstante, no se realizan pruebas con una base de datos diferente [22] [23].

Por lo anterior, es necesario realizar un estudio sobre los parámetros y métodos computacionales que pueden ser tenidos en cuenta a la hora de implementar una herramienta tecnológica que sea un apoyo para el profesional de salud al momento del diagnóstico de labor de parto mediante cesárea.

Formulación del problema

¿Es eficiente detectar el riesgo de cesárea mediante un modelo bio-inspirado considerando los clasificadores KNN, SVM, MLP y DP, a partir de parámetros demográficos y la bioseñal de una electrohisterografía (EHG) de mujeres gestantes?

Hipótesis

Los modelos bio-inspirados que consideran los clasificadores KNN, SVM, MLP y DP son eficaces en el análisis y procesamiento de parámetros demográficos y de bioseñal EHG para la detección de riesgo de cesárea en mujeres gestantes.

1.1. OBJETIVOS

Objetivo General

Desarrollar un modelo bio-inspirado para el apoyo en el diagnóstico de nacimiento por cesárea basado en información demográfica de la madre y señales biológicas del binomio madre-hijo.

Objetivos Específicos

- Establecer las características demográficas más relevantes para determinar el riesgo de cesárea.
- Caracterizar la señal biológica electrohisterografía (EHG) para implementar un clasificador de madres gestantes con riesgo de cesárea.
- Implementar y evaluar el modelo propuesto de la clasificación de madres gestantes que puedan presentar riesgo de intervención por cesárea.

1.2. PRODUCTIVIDAD ACADÉMICA

Anexo 1.

J. Pacheco Fandiño, E. Camargo Casallas, A. Gaona Barrera y V. Fandiño Osorio, “Approach to the diagnostic of cesarean birth using bio-inspired models”, “Aproximación al diagnóstico del parto por cesárea mediante modelos bioinspirados”, XVII Congreso Internacional De Electrónica, Control Y Telecomunicaciones, noviembre 2022, ISBN: 978-958-44-5254-2 Vol. 13.

Anexo 2.

Certificado de PONENTE en XVII Congreso internacional de Electrónica, Control y Telecomunicaciones, realizado el 24 y 25 de noviembre de 2022 de manera virtual. Código: 0X7484844B708.

2. MARCO DE REFERENCIA

El presente capitulo contiene los conceptos básicos necesarios para el entendimiento de este proyecto desde la perspectiva médica e ingeniería.

En primer lugar, se definen conceptos correspondientes a la finalización de la gestación y distocias que se pueden presentar en el desarrollo del embarazo. Posteriormente, se describen conceptos básicos del monitoreo fetal incluyendo la bioseñal EHG, información de los modelos bio-inspirados, clasificadores de datos y parámetros de desempeño para evaluar modelos. Finalmente, se presentan investigaciones relacionadas con la clasificación de gestantes y diferentes alteraciones durante el embarazo.

2.1. GENERALIDADES DEL PARTO

Nacimiento por cesárea: el parto es un proceso natural que se da al final de la gestación, este puede ser de inicio espontáneo o por inducción, dependiendo de la evolución del trabajo de parto puede generar salida por vía vaginal (parto eutócico) o por vía abdominal (cesárea). La cesárea es una intervención quirúrgica, la cual tiene la finalidad de extraer el feto por vía abdominal [12] [24]. Existen varios factores que llevan a este:

1. Factores maternos: hemorragias internas, cáncer cérvico, hipertensión o falta de progreso en el embarazo.
2. Factores fetales: anomalía en la posición del feto idealmente debe estar en posición transversa donde los hombros ingresan primero al cuello del útero, sin embargo, la posición podálica (pies y glúteos ingresan primero al canal del parto) resulta en cesárea; anomalía en la ubicación de la placenta ya que esta puede estar obstruyendo el canal de parto, parto prematuro, embarazo gemelar, sufrimiento fetal es decir cambios en la frecuencia cardiaca del feto.

En el momento en que la gestante acude al centro médico se realizan una serie de exámenes diagnósticos para determinar el estado del feto y de la madre, y así poder determinar si es necesario intervenir [17].

Proteinuria gestacional: la proteinuria es la presencia de proteínas en la orina de la gestante, en específico se trata de 150 mg o más en lo que es la orina de 24 horas. Esto se puede deber a un fallo durante o previa al embarazo en los riñones u otros órganos relacionados, a una infección en el tracto urinario, diabetes, tensión alta, intoxicación o preeclampsia [25] [15]. Existen tres tipos de proteinuria:

- Ortostática, se presenta aumento cuando se pasa mucho tiempo de pie.
- Transitoria, que es provocada por situaciones tales como fiebre, estrés, exposición al frío o al calor.
- Persistente, tiene lugar cuando los niveles de proteína son reflejo de una enfermedad.

Hipertensión en el embarazo: la hipertensión en el embarazo se diagnostica cuando la gestante tiene presión arterial alta pero la orina no contiene proteínas, teniendo en cuenta que la presión previa al embarazo estaba en los rangos normales y aumenta después de la semana 20 de gestación y postparto regresa a la normalidad [26] [16]. Existen 3 distocias posibles de diagnosticar:

- Preeclampsia leve: es posible diagnosticarla si la presión arterial sistólica es mayor igual a 140 mmHg, o presión arterial diastólica mayor o igual a 90 mmHg.
- Preeclampsia grave: se diagnostica si la presión arterial sistólica es de 160 mmHg o mayor, o presión arterial diastólica es igual o mayor a 110 mmHg, teniendo en cuenta que esto sucede en al menos dos ocasiones con una

diferencia de 6 horas. También se puede diagnosticar si en una muestra de orina de 24 horas contiene 5 o más gramos de proteína.

- Eclampsia: se diagnostica cuando la gestante presenta preeclampsia y además desarrolla convulsiones [23] [13].

Diabetes gestacional: la diabetes gestacional es una dolencia caracterizada por intolerancia a los carbohidratos, lo cual puede conllevar a hiperglucemia variable.

La mujer gestante con esta alteración es propensa de padecer trastornos hipertensivos, nacimientos por vía abdominal, entre otras. Además, en 50 % de las pacientes con diabetes gestacional tienen riesgo de llegar a padecer diabetes tipo 2 en los siguientes 10 años [27].

Algunas de las complicaciones que se pueden presentar asociadas con las concentraciones maternas de glucosa y con la duración son: macrosomía fetal, hipoglucemia neonatal, hiperbilirubinemia, posible nacimiento por cesárea o parto instrumentado, distocia de hombro, trauma fetal durante el nacimiento y muerte fetal-neonatal [28].

Índice de masa corporal: es la relación que se calcula entre el peso y la talla con el fin de determinar el sobrepeso y la obesidad en los adultos. Se obtiene a partir de la división del peso de una persona en kilos en el cuadrado de su altura en metros (kg/m2) [29].

$$IMC = \frac{Peso(Kg)}{Altura(m)^2} \qquad (1)$$

Gravidez: termino usado para referirse a la cantidad de veces que una mujer ha estado en gestación, es decir, el período que transcurre desde la implantación en el útero del óvulo fecundado al momento del parto, se incluyen los abortos [30].

Paridad: en el contexto de este trabajo es el número total de partos que ha tenido una mujer [31].

Posición placentaria: la placenta se desarrolla en el útero durante el embarazo, y según sea su ubicación, se describe como placenta de cara anterior, posterior o fúndica [32], se observan en la figura 1:

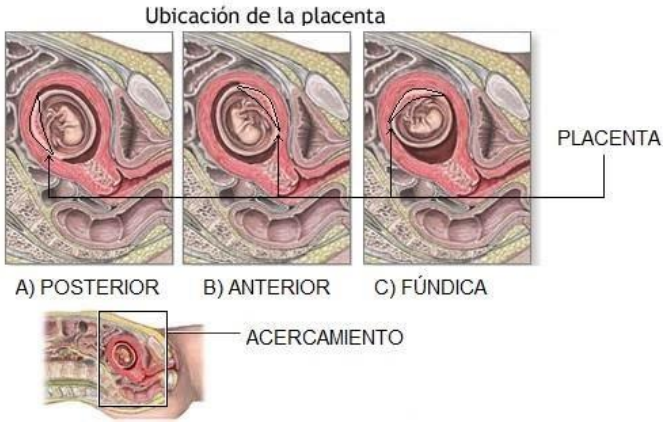


Figura 1. Posición placentaria [33].

La placenta anterior: se encuentra implantada en la cara uterina más próxima al ombligo.

La placenta posterior: se implanta en la cara uterina más próxima a la espalda.

La placenta fúndica: se implanta en la parte más alta del útero.

Edad gestacional: es el término que se emplea para describir el tiempo de desarrollo que tiene el feto. La unidad de medida es en semanas, un embarazo normal tiene una duración desde 38 hasta 42 semanas. Los bebes nacidos antes de la semana 37 se denominan prematuros y los nacidos después de 42 semanas son posmaduros [34]. Se puede determinar de las siguientes maneras:
Previo al alumbramiento: se realiza una ecografía desde la cual es posible medir el tamaño de la cabeza, fémur y abdomen.

Posterior al alumbramiento: es posible determinar la edad gestacional mediante el peso, talla, tamaño de la cabeza, entre otros.

Oxitocina sintética: su aplicación por vía intravenosa es usada para inducir el parto y estimular las contracciones uterinas una vez que se ha iniciado la labor de parto [35].

2.2. TÉCNICAS DE MONITOREO FETAL

Monitoreo de presión intrauterina (IUC): Esta técnica es invasiva, se lleva a cabo haciendo uso de un catéter el cual realiza la medición a la presión intrauterina presente durante una contracción, en la figura 2b. se observa cómo es introducido el catéter hasta el feto. Sin embargo, esta medición se realiza después de que las membranas se han roto y ha dado inicio el trabajo de parto. El rango de la medición va entre 0 y 100 mmHg como se ve en la figura 2a. [36].

(a) (b)

Figura 2. Monitoreo de presión intrauterina (IUC). (a) presión intrauterina [37], (b) gráfica de monitoreo [36].

Tocodinamometría: esta técnica no invasiva hace uso de un cinturón que se ubica alrededor del abdomen sobre el cual se conecta un Doppler para detectar la frecuencia cardiaca fetal, ubicado entre el ombligo de la gestante y la fosa iliaca anterior (según posición fetal) y un segundo Doppler que mide las contracciones y su intervalo ubicado en el fondo uterino, esto se puede observar en la figura 3 [38].

La ubicación de los transductores se debe a que se fija la medición lo más próximo al corazón del feto.

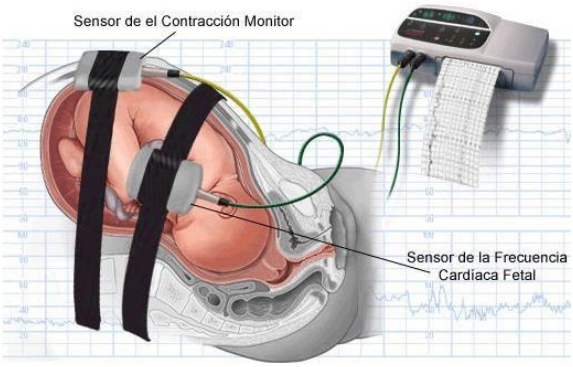


Figura 3. Tocodinamometría [38].

Otra de las técnicas es la electrohisterografía descrita a continuación.

2.2.1. Electrohisterograma (EHG)

La electrohisterografía es una técnica en la que se realiza el monitoreo de las contracciones uterinas de forma no invasiva, mediante el registro de la actividad eléctrica del miometrio, obtenida mediante una serie de electrodos de superficie ubicados en el abdomen de la gestante [39] [40], este es comúnmente usado para el monitoreo de la actividad eléctrica uterina, predicción del parto prematuro y detección de distocias en las contracciones. Se puede implementar después de la semana 18, antes de esta la actividad presente en el miometrio no es perceptible por los electrodos.

La señal del EHG está compuesta por:

Onda lenta (SW). Tiene un periodo igual a la duración de las contracciones y su amplitud se relaciona con su intensidad. Con frecuencias desde 0.014 a 0.033 Hz, su forma es normalmente bifásica, es decir, comienza con una deflexión negativa y se encuentra sincronizada con la presión intrauterina.

Onda rápida (FW). Esta superpuesta a SW y se produce antes del máximo que alcanza la presión intrauterina al elevarse. Se distinguen dos componentes con diferentes bandas de frecuencia: Onda rápida de baja frecuencia (FWL), está relacionada con la propagación de la actividad uterina. El rango de frecuencias es [0.13, 0.26] Hz y según se acerca el parto se va desplazando hacia mayores frecuencias.

Y la onda rápida de alta frecuencia (FWH), está relacionada con la excitabilidad del útero es decir las contracciones efectivas durante el parto. Su frecuencia es [0.36, 0.88] Hz y su densidad espectral de potencia aumenta conforme se va acercando el parto. Todo esto se observa en la figura 4.

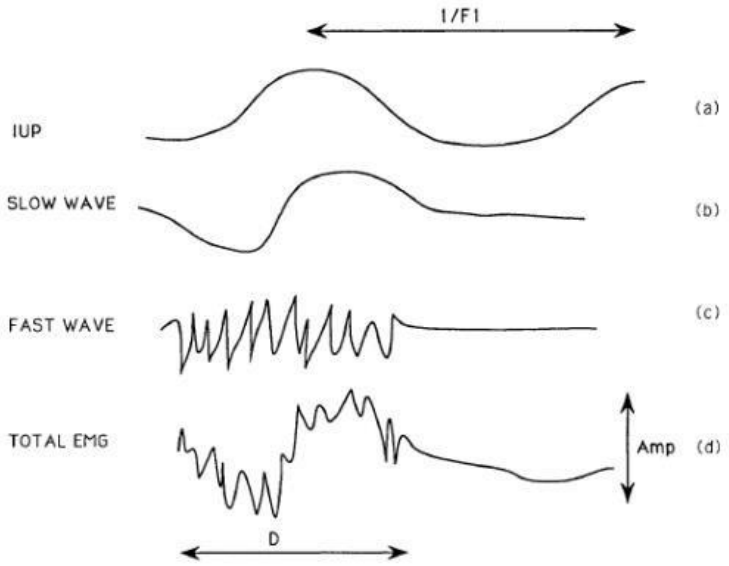
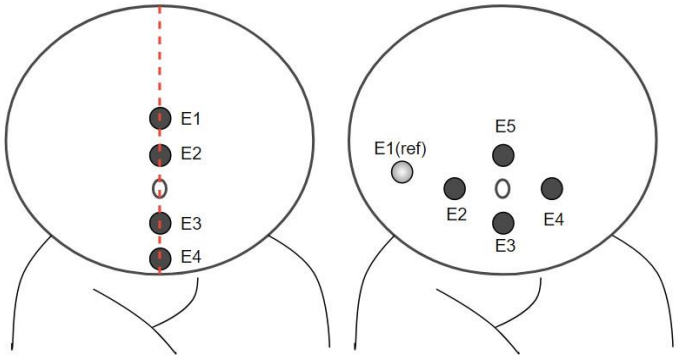


Figura 4. Representación de la señal de IUP (a), onda lenta (b), onda rápida (c) y la suma de ambas (d) [39].

En la actualidad no existe un estándar sobre la ubicación de los electrodos en el abdomen de la gestante, la cual permita una correcta medición o adquisición de la señal. Hasta el momento se conocen configuraciones monopolares y bipolares, la ubicación de los electrodos depende directamente al objetivo del estudio a realizar, por ejemplo, si se ubican de forma vertical en medio del abdomen materno como se observa en la figura 5a. se puede adquirir la señal que indica la propagación de las

contracciones, al ubicarse en rombo con un electrodo de referencia figura 5b. se obtiene otra perspectiva de las contracciones durante la labor de parto.



(a) (b)

Figura 5. Ubicación de electrodos para adquisición de señal EHG. (Elaboración propia)

En comparación con las otras técnicas de monitoreo, esta permite ser tomada sin poner en riesgo la vida del feto y la vida materna, tampoco presenta riesgo de infección o alguna intervención al desarrollo del feto ni de la gestación, en la figura 6 se muestra la toma en forma simultánea de la señal de una contracción con las técnicas IUC, tocodinamometría, EHG monopolar y EHG bipolar. En la figura 6 se observa que las señales presentan un bajo nivel cuando el útero se encuentra en reposo y empieza a aumentarse con forme la contracción aparece.

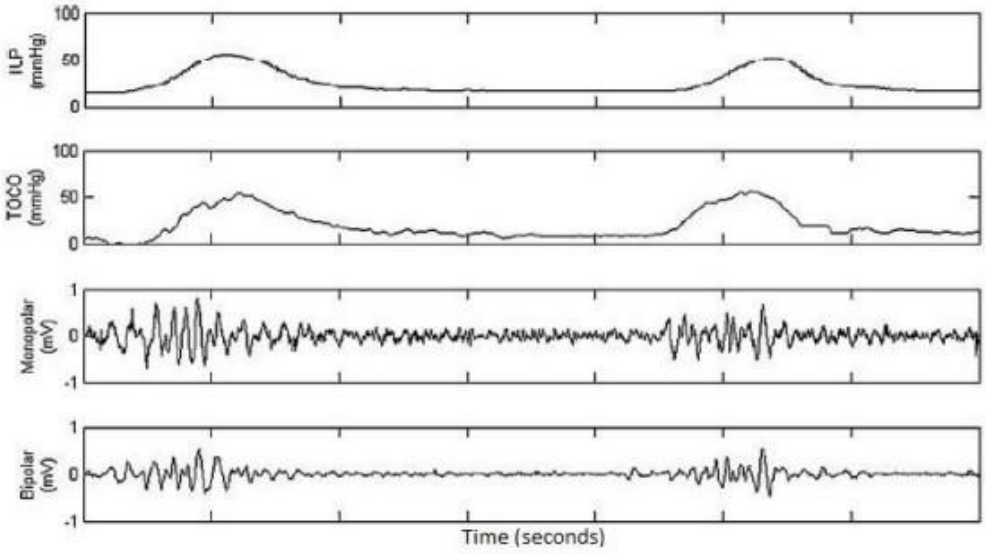


Figura 6. Señal de contracción materna mediante IUC, tocodinamometría y EHG monopolar y EHG bipolar [41].

Métodos de análisis: existen 3 formas de analizar la señal EHG, métodos espectrales, temporales y no lineales [42], a continuación, se describen cada uno de ellos:

Método espectral: se basa en parámetros en el dominio de la frecuencia entre ellos la frecuencia máxima, media y potencia espectral. Permite detectar características de la contracción en la gestante, por ejemplo, la fatiga en el musculo uterino.

Método temporal: se basa en el cálculo de parámetros en el dominio del tiempo, por ejemplo, la amplitud de la señal la cual permite identificar la fuerza que realiza el musculo en cada contracción.

Método no lineal: debido a que los sistemas del cuerpo humano presentan cambios con el tiempo es necesario desarrollar métodos no lineales o dinámicos, por ejemplo, la entropía muestral, la cual proporciona información adicional de los evidenciados en los exámenes físicos es decir en los cambios fisiológicos de la gestante.

2.3. MODELOS BIO-INSPIRADOS

Un problema de optimización es aquel que tiene como objetivo encontrar la mejor solución, cumpliendo con ciertas restricciones, realizando un análisis detallado de todas las posibles soluciones, sin embargo, en algunos casos este tipo de problemas requieren métodos diferentes a los convencionales, por lo cual se aplican modelos inspirados en el funcionamiento de la naturaleza o bio-inspirados [43].

Estos modelos emulan la evolución natural de las especias frente a los desafíos que se le presentan al interactuar socialmente, debido a esto presentan una estructura paralela y tienen capacidad de adaptarse a nueva información [44], a continuación, se mencionan los componentes de los modelos bio-inspirados:
Población: conjunto de datos.

Mecanismo de Evolución: métodos por los cuales se modifica la población (preprocesamiento de los datos).

Mecanismo de Calificación: métodos predictivos o de clasificación, por los cuales se transforma cada elemento de la población. Dentro de este mecanismo se pueden encontrar algoritmos evolutivos, la inteligencia de enjambre, los algoritmos inmunológicos y las redes neuronales.

Condiciones de Parada: procedimiento que determina si se ha encontrado una solución o se ha superado un número de iteraciones o repeticiones determinado.

Uno de los mecanismos de calificación más implementados son los algoritmos evolutivos, los cuales replican el proceso de evolución de las especies, el proceso se muestra en la figura 7, partiendo de un conjunto inicial de soluciones es sometido a ciclos de transformaciones también llamadas generaciones, tras cierto número de estas se espera que la mejor solución sea encontrada [45].

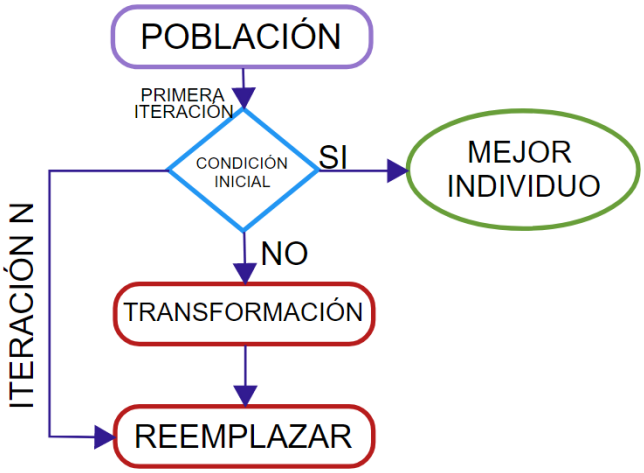


Figura 7. Algoritmo Evolutivo. (Elaboración propia)

Los clasificadores se basan en la simulación del comportamiento de sistemas presentes en la naturaleza para el diseño de algoritmos o métodos heurísticos de búsqueda o aprendizaje. Dentro de esos se encuentran los algoritmos supervisados y no supervisados, a continuación, se explican los clasificadores implementados en este proyecto:

2.3.1. Clasificador K vecinos más cercanos (KNN)

Este paradigma se basa en que la clasificación de un nuevo individuo está definida por la clase más frecuente a la que pertenecen sus vecinos más cercanos. Clasifica los datos calculando la distancia entre cada dato no etiquetado y todos los demás puntos [46] cómo se observa en la figura 8.

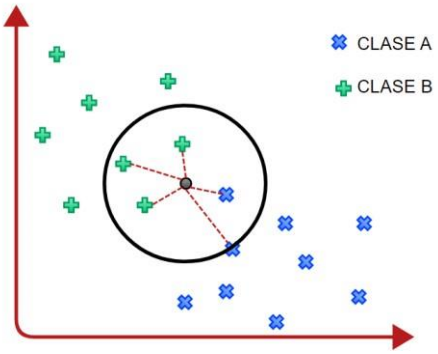


Figura 8. Representación gráfica del clasificador KNN. (Elaboración propia)

Teniendo en cuenta el conjunto de datos:

Tabla 1. Conjunto de entrada, clasificador KNN [47].

Individuos		X_1	$\cdots$	X_j	$\cdots$	X_n	C
(x_1, c_1)	1	x_{11}	$\cdots$	x_{1j}	$\cdots$	x_{1n}	c_1
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
(x_i, c_i)	i	x_{i1}	$\cdots$	x_{ij}	$\cdots$	x_{in}	c_i
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
(x_N, c_N)	N	x_{N1}	$\cdots$	x_{Nj}	$\cdots$	x_{Nn}	c_N

Donde n son el número de variables de cada individuo, N es el número de individuos de la población, entonces $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ con clasificación c_i.

La distancia para un conjunto de datos que contiene variables continuas puede ser calculada de las siguientes formas [48]:

Distancia Euclidiana

$$D(x_i, x_j) = \sqrt{\sum_{k=1}^{n} (x_{ik} - x_{jk})^2} \qquad (2)$$

$$Donde\ x_j\ es\ el\ nuevo\ individuo\ y\ i \neq j$$

La distancia Euclidiana se basa en medir la distancia según el teorema de Pitágoras.

Distancia Manhattan

$$D(x_i, x_m) = \sum_{j=1}^{n} |x_{ij} - x_{mj}|, \qquad (3)$$

$$x_i = (x_{i1}, x_{i2}, \dots, x_{in})\ y\ x_m\ es\ el\ nuevo\ individuo$$

La distancia Manhattan mide la distancia entre dos puntos como la suma de los catetos.

La clasificación se obtiene con el siguiente procedimiento:

1. Seleccionar el número de K vecinos.
2. Seleccionar nuevo elemento.

3. Tomar los K vecinos más cercanos al nuevo elemento posterior al cálculo de la distancia.
4. Contar los K vecinos que pertenecen a cada categoría.
5. Asignar la categoría de la que se contó más vecinos al nuevo elemento

2.3.2. Máquinas de vectores de soporte (SVM)

Es un algoritmo de aprendizaje supervisado, este clasificador se basa en la teoría del clasificador de margen máximo, el cual encuentra un hiperplano que separa efectivamente dos clases diferentes de conjuntos de datos. Los kernel son funciones que transforman un espacio de baja dimensión en un espacio de mayor dimensión, algunos de estos son [50]:

Kernel lineal: Es mayormente implementado cuando el problema es linealmente separable y los datos son representados en 2 dimensiones (figura 9). Su transformación está dada por la ecuación (4)

$$K_L(x_i, x_{i'}) = \langle x_i \cdot x_{i'} \rangle \quad (4)$$

donde $x_i, x_{i'} \in X$ conjunto de entrada. $\langle \cdot \rangle$ hace referencia al producto punto.

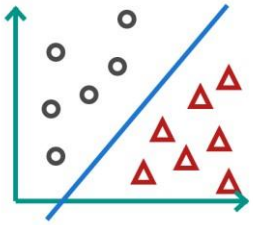


Figura 9. Kernel lineal. (Elaboración propia)

Kernel polinomial: el grado del polinomio esta dado por la variable d (ecuación (5)), permite más flexibilidad en la separación de los datos, como se observa en la figura 10, además de que es usado cuando la dimensión de los datos de entrada es mayor a 2.

$$K_P(x_i, x_{i'}) = (\langle x_i \cdot x_{i'} \rangle + c)^d \quad (5)$$

donde $x_i, x_{i'} \in X$ conjunto de entrada, c es una constante de ajuste y d son los grados de la función polinomial.

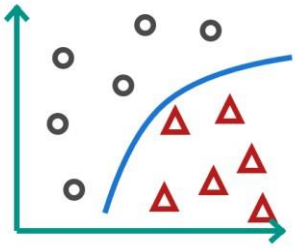


Figura 10. Kernel polinomial. (Elaboración propia)

Kernel Gausiano: también conocido como función de base radial, es mayormente usado para procesar datos no lineales (figura 11), también ayuda a hacer una separación adecuada cuando no hay un conocimiento previo de los datos. En la ecuación (6) se muestra cómo se realiza esta transformación:

$$K_G(x_i, x_{i'}) = exp\left(-\frac{\|x_i - x_{i'}\|^2}{2\sigma^2}\right) \quad (6)$$

donde $x_i, x_{i'} \in X$ conjunto de entrada y σ es la anchura del kernel.

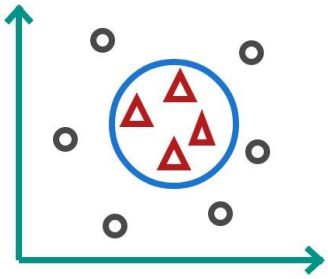


Figura 11. Kernel gausiano. (Elaboración propia)

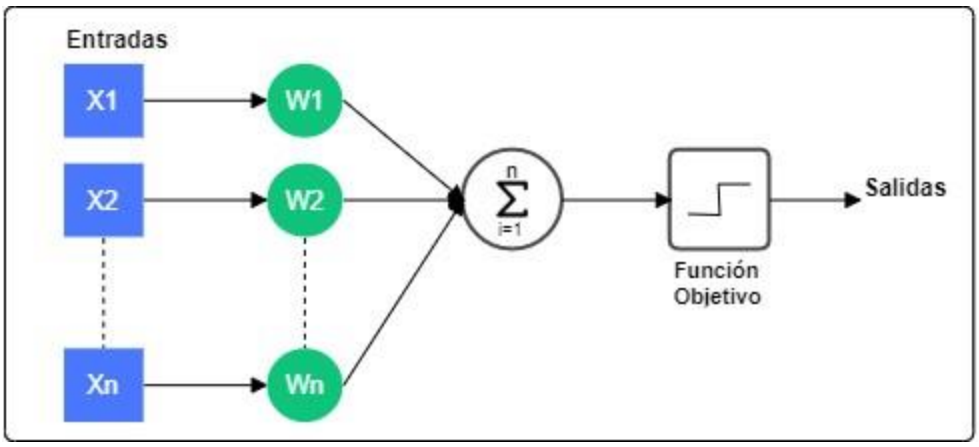
2.3.3. Redes neuronales (MLP)

Las redes neuronales se basan en el procesamiento de información bajo ciertas condiciones o funciones, las cuales le asignan un peso o relevancia a cada una de los parámetros de la base de datos, los procesos que este lleva a cabo se basan en

la generalización de modelos matemáticos asemejando el cerebro humano [51], basado en cuatro definiciones:

- El procesamiento de la información ocurre en unidades simples llamadas neuronas.
- La información es transmitida entre neuronas, a esto se le conoce como enlace.
- Cada enlace tiene un peso asociado.
- Las unidades simples aplican una función de activación (generalmente no lineal).

Estas definiciones son implementadas en la red, las cuales indican el patrón de conexiones entre las neuronas, es decir, la arquitectura o el método con el cual se determina el peso de los enlaces y por último la función de activación para obtener un valor en la salida, como se muestra en la figura 12, lo anterior define el proceso de entrenamiento o aprendizaje de la arquitectura [52] [53].



Wi:peso de entrada i , Xi:entrada i, F:función objetivo.

Figura 12. Diagrama MLP. (Elaboración propia)

MLP es una red neuronal con una capa de entrada, una capa de salida y N capas ocultas [54]. Las matrices de peso calculan las conexiones de capa, la ecuación de un nodo en las capas ocultas es la ecuación (7).

$$h_{ij} = f\left(\sum_{j=1}^{n} w_{ij}x_j + b_i\right) \tag{7}$$

donde h_{ij} es el j-ésimo nodo de hi y w_{ij} es el peso asociado. b_i como el peso w_{i0} con la convención de $x_0 = -1$.

2.3.4. Aprendizaje profundo (DP)

El método de aprendizaje profundo o Deep learning utiliza varias capas de unidades de procesamiento no lineales que extraen y transforman los datos de entrada. El proceso se itera el número de veces necesario para obtener la precisión determinada y pasar a la siguiente capa [55], este clasificador requiere grandes cantidades de potencia informática e información para resolver problemas complicados. El aprendizaje puede ser supervisado, semi-supervisado o no supervisado. Existen varios tipos de algoritmos de aprendizaje profundo: redes convolucionales, redes de memoria a largo plazo, redes neuronales recurrentes (RNN), redes adversarias generativas (GAN), Mapas auto-organizados, entre otros [56].

En la figura 13 se ilustra la diferencia entre el aprendizaje automático (ML) y el DP, en el ML se debe realizar una extracción de variables de la base de datos para que el clasificador pueda ejecutarse lo cual concluye en más tiempo de procesamiento y costo computacional, por otra parte, en el DP realiza la extracción de variables y la clasificación en un solo paso.

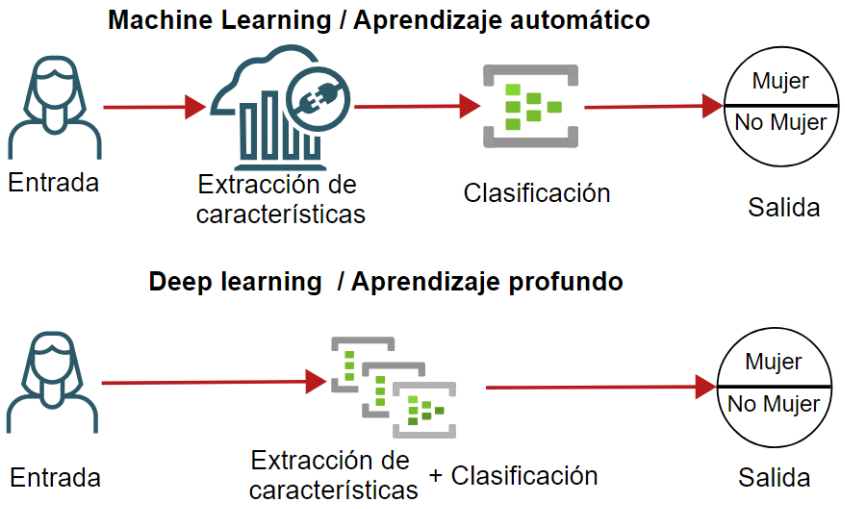


Figura 13. Diferencias machine learning y DP

2.3.5. Desempeño de modelos bio-inspirados

En un problema de clasificación binaria, se emplean las siguientes definiciones [44] [57] [58]:

- Verdaderos positivos, en la ecuación se mencionan como VP, Número de gestantes que son diagnosticadas por el modelo como parto vaginal y en la base de datos también lo son.

- Falsos positivos, en la ecuación de mencionan como FP, Número de gestantes que son diagnosticadas por el modelo como parto vaginal y corresponden con parto por cesárea.

- Falsos negativos, en la ecuación de mencionan como FN, Número de gestantes que son diagnosticadas por el modelo como parto por cesárea y corresponden a parto vaginal.

- Verdaderos negativos, en la ecuación de mencionan como VN, Número de gestantes que son diagnosticadas por el modelo como parto por cesárea.

Basados en estas definiciones existen tres índices de desempeño para un clasificador.

Sensibilidad:

Representa la capacidad del modelo para discriminar los casos positivos de los negativos. La sensibilidad es el fragmento de verdaderos positivos.

$$Sensibilidad(S) = \frac{VP}{VP+FN} \quad (8)$$

Especificidad:

Representan la capacidad del modelo para discriminar los casos positivos de los negativos. La especificidad, es el fragmento de verdaderos negativos

$$Especificidad(E) = \frac{VN}{FP+VN} \quad (9)$$

Exactitud:

Representa la capacidad del modelo de dar el resultado verdadero de una medición.

$$Exactitud(Ex) = \frac{VP-VN}{VP+VN+FP+FN} \quad (10)$$

2.4. ANTECEDENTES

Los hallazgos en la literatura acerca de la clasificación de mujeres gestantes que puedan presentar alteraciones durante el embarazo hacen uso de diferentes grupos de datos, a continuación, se ilustran dos de estos y los desarrollos encontrados:

Basado en parámetros demográficos

Realizando la revisión de las investigaciones en torno al tema se encuentran desarrollos donde se plantea el uso de los árboles de decisión para la clasificación de gestantes que pueden presentar alto riesgo, se aplican 3 tipos de clasificadores para el tratamiento de una base de datos que contiene 11 parámetros demográficos, de este trabajo se obtienen un acierto del 75,8 % [59].

Se emplea el método Back-Propagatión o propagación hacia atrás, para la predicción de parto prematuro en Estados Unidos de América, se tienen en cuenta 13 parámetros demográficos. La precisión del modelo para predecir si la paciente tuvo un parto prematuro fue del 87.3 %, en el modelo entrenado se evidencia que los factores con los mayores pesos relativos para la predicción fueron los resultados de la prueba física realizada a las 24 semanas, que contiene la dilatación cervical, información demográfica y su clasificación de semanas a la hora de labor de parto [60].

También se han realizado investigaciones basadas en métodos diferentes a la redes neuronales, por ejemplo: la calculadora de parto vaginal después de una cesárea (VBAC), desarrollada por la Red de Unidades de Medicina Materno-Fetal de Indonesia este plantea el uso de métodos estadísticos para ayudar a las mujeres a tomar la decisión del modo de parto después de un historial de cesárea, esto teniendo en cuenta que los parámetros analizados son el historial de parto de la madre y la posición de la placenta en el embarazo actual, esta calculadora presenta un 64,9 % de precisión [61].

La combinación entre análisis multivariante y áreas bajo las curvas ROC (Receiver Operating Characteristics) se usa para determinar la capacidad predictiva teniendo en cuenta 28 potenciales parámetros, tras la validación los datos que se mantienen en el modelo final son: talla materna, índice de masa corporal, nuliparidad (pacientes sin partos precedentes), test de Bishop, edad gestacional, macrosomía (Tamaño de bebe mayor a 4 Kgr), sexo fetal y tasa global de cesárea del ginecólogo. La capacidad predictiva del modelo indica que el 76 % de las mujeres evaluadas concluyeron su gestación según la predicción realizada [62].

Otra de las alteraciones que se puede presentar durante el embarazo es la preeclampsia en el trabajo que se llevó a cabo en la Universidad del Norte de Barranquilla, Colombia se diseña un modelo que permite brindar la asistencia médica y predice la preeclampsia en las gestantes, esto mediante técnicas de minerías de datos, por ejemplo: arboles de decisión, algoritmo C5.0, PCA (Principal Component Analysis) y NMDS (Non Metric Dimensional Scaling).
La propuesta analiza los datos demográficos de la gestante tomados al inicio de la gestación los cuales incluyen: hipertensión, proteinuria, presencia de hormono lactato deshidrogenasa (alimenta la placenta), creatinina, plaquetas, edad, entre otros; realizando el entrenamiento con el 75 % de los datos y la validación con el 25 %, obteniendo una tasa de detección de un 83.6 % [63].

El modelo realizado en la Universidad de Valparaíso de Chile usa el aprendizaje automático para la predicción de complicaciones perinatales en gestantes con diabetes, lo datos con los que se desarrolla el modelo son datos demográficos de la gestante: edad, estado civil, nivel escolar, partos previos, obesidad e información del neonato: preeclamsia, metrorragia en el embarazo (sangrado abundante), problemas con la placenta como placenta previa, dopler alterado, entre otras. Los clasificadores comparados incluyen SVM, LDA, bosque aleatorio y MLP, dando como resultado la tabla 2 [64].

Tabla 2. Resultados de modelo predictor de complicaciones perinatales *[64]*.

Índice	SVM	LDA	Arbol desc	MLP
Exactitud	73 %	61 %	72 %	61%
Precisión	72 %	63 %	69 %	62 %
Sensibilidad	60 %	55 %	62 %	56 %

Basado en parámetros demográficos y señales biológicas

Haciendo uso de redes neuronales convolucionales multimodales (MCNN) se detecta el compromiso fetal durante el trabajo de parto y el parto, específicamente la hipoxia fetal (el feto se queda sin oxígeno aun estando en el vientre materno), para ello se procede con la interpretación de la cardiotocografía CTG, incluyendo información sobre la pérdida de señal y su impacto en el rendimiento del parto. Se desarrolla MCNN apilado para analizar por separado y vincular secuencialmente el CTG, antes y después del inicio de la segunda etapa dinámica del trabajo de parto, se valida el rendimiento del modelo utilizando conjuntos de datos externos multicéntricos, no obstante, la tasa no presenta variación ambos fueron del 65 % entre los dos métodos. Sin embargo, no en todas las labores de parto es posible realizar la toma de la señal CTG, por lo cual este modelo es aplicable para las maternas de bajo riesgo [65].

El algoritmo propagación hacia atrás es implementado para la predicción de labor de parto eutócico (vaginal) en mujeres que han realizado labor de parto mediante cesárea en embarazos previos, se realiza con una base de datos que contiene 400 registros con información obstétrica demográfica, así como información del índice de embarazo, se realizan dos modelos donde se tiene una variación de la cantidad de reglas y la cantidad de variables, en la siguiente tabla se presentan los resultados obtenidos [66].

Tabla 3. Resultados de modelo con propagación hacia atrás [66].

Número de Reglas	8 variables	4 variables
Sensibilidad	63 %	59 %
Especificidad	42 %	44 %

En el siguiente desarrollo se contempla en análisis de electrohisterograma con el objetivo de predecir partos prematuros, para esto se utilizan siete diferentes redes neuronales artificiales para identificar los registros. Después de la validación de modelos se encuentra que la combinación de la red neuronal Feed-Forward Neural, la red neuronal de función de base radial y la red neuronal aleatoria entrenados por Levenberg – Marquardt obtuvieron los mejores resultados, con un 84 % de especificidad [67].

Al uso de señales biológicas se une el análisis de resonancias magnéticas (IRM), el objetivo es evaluar si el uso de aprendizaje de máquina para analizar si las texturas derivadas de las IRM pueden proporcionar un nuevo punto de vista metodológico útil para evaluar la presencia o ausencia de espectro de placentaria accreta (placenta que se adhiere al útero) en pacientes con placenta previa (placenta ubicada en el cuello del utero), dando como resultado la tabla 4; concluyendo que los modelos con porcentajes mayores a 90 % presentan mayor sensibilidad [68].

Tabla 4. Resultado de validación de IRM [68].

Método	Exactitud
Random Forest	97.4 %
K Nearest	98 %
Naives Bayes	77.3 %
Perceptrón multicapa	84.9 %

Por otro lado, se encuentra un modelo de predicción desarrollado con base a la Herramienta de Evaluación de Riesgo de Sesgo del Modelo de Predicción (PROBAST) que contiene pautas recientes para el desarrollo del modelo de predicción para evitar riesgos de sesgo. Se realiza un comparativo con modelos de predicción similares de estudios publicados en registros de PUBMED, EMBASE y SCOPUS desde 2015 al 2020. El modelo desarrollado aplica un algoritmo de aprendizaje automático que utiliza parámetros demográficas y diagnósticos. Se utilizan 17 predictores extraídos por un método de bosques aleatorio, los resultados de las validaciones realizadas tienen una precisión del 77 %. El estudio concluye que el resultado es similar a los encontrados en la literatura [69].

La señal EHG ha sido ampliamente usada para el monitoreo del binomio madre-hijo por ejemplo en el trabajo realizado en el “Hospital Universitari i Politècnic La Fe de València” por la Universidad Politécnica de Valencia, España busca predecir el éxito de un parto inducido farmacológicamente haciendo uso de la señal EHG. La práctica de inducir el parto por parte del profesional médico que atiende la gestante es llevada a cabo si el trabajo de parto tiene determinadas situaciones o condiciones de riesgo ya que se debe extraer el neonato lo más rápido posible sin embargo esto no siempre funciona y debe acudirse a la cesárea como método de parto.

El trabajo consistió en la toma de 94 registros de EHG tomados antes y después de la inducción del parto y posterior a ello se calcularon 25 características sobre distinto tramos de la señal, y se desarrolló un predictor basado en redes neuronales, dando como resultado un 87 % de acierto es decir 87 de cada 100 mujeres analizadas culminaron su trabajo de parto exitosamente después de la inducción [70].

Por otro lado, el estudio realizado por la Escuela Técnica Superior Ingeniería industrial de Valencia consiste en el desarrollo de un sistema predictor de parto prematuro tomando como entradas la señal EHG y datos obstétricos. Se implementa el aprendizaje automático mediante algoritmos de clasificación

“Extreme Learning Machine (ELM)” y “Weigthed Extreme Learning Machine (WELM)” las cuales son redes feedforward basados en modelos lineales.

Tabla 5. Resultados de predicción de parto prematuro [71].

Método	**Exactitud**
ELM	51.01 %
WELM	66.70 %

En la tabla 5 se observan los resultados obtenidos en el desarrollo del trabajo con los dos clasificadores, se tiene hasta un poco más del 50 %, posterior a esto se usó la técnica de SMOTE con la cual se obtuvieron resultados del 95.39 %, demostrando una mejora en la predicción de partos prematuros [71].

3. METODOLOGÍA

El método que se emplea en el desarrollo de este trabajo será el método cuasi experimental el cual es útil para estudiar problemas en los cuales no se posee el control absoluto de los datos, por lo cual se incluyen "grupos intactos", es decir, grupos ya constituidos [72].

El modelo bio-inspirado de predicción de parto por cesárea se muestra en la figura 14, en primer lugar, el conjunto de datos se compone de características demográfica y características de la bio-señal de EHG, el cual es preprocesado con el objetivo de seleccionar las entradas al bloque clasificador. A continuación, se diseñan y desarrollan los clasificadores teniendo en cuenta la evaluación del desempeño. Finalmente, el modelo presenta el resultado la aproximación.

Este modelo se pretende implementar en los controles realizados durante los primeros meses por el obstetra y, de acuerdo a lo observado por los profesionales médicos y al resultado de este modelo bio-inspirado, es posible proponer un plan de embarazo y plan de parto acorde, con el fin de preservar la vida del binomio madre-hijo.

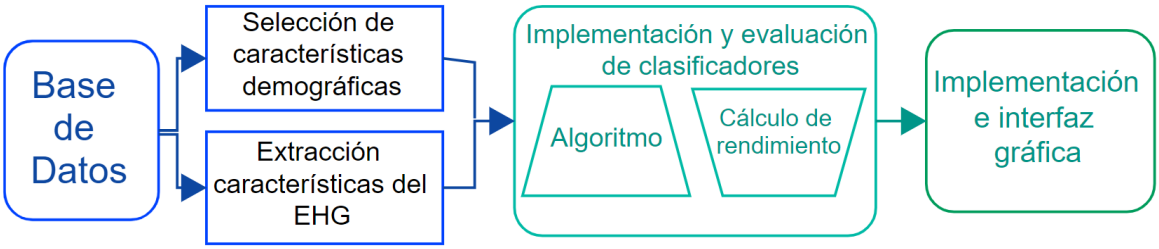


Figura 14. Modelo bio-inspirado. (Elaboración propia)

A continuación, se describe cada uno de estos procesos:

3.1. BASE DE DATOS

La información para llevar a cabo este proyecto es extraída de la página Physionet, la cual es un repositorio de datos de investigación médica de libre acceso, administrado por el Laboratorio de Fisiología Computacional del MIT [73], la base de datos titulada "Icelandic 16-electrode Electrohysterogram Database", consta de 122 electrohisterogramas (EHG) de mujeres embarazadas, con edades entre 19 y 39 años, los registros y los datos se toman antes y durante el trabajo de parto.

Las grabaciones se realizaron en Islandia entre el 2008 y 2010, esta investigación se llevó a cabo entre el Centro de Atención Primaria de Salud de Akureyri, el Hospital de Akureyri y el Hospital Universitario Landspitali. Los registros son mediciones realizadas en pacientes con sospecha de trabajo de parto y presentes en las salas de parto [74]. Cada registro contiene un archivo de cabecera con la información demográfica de la gestante.

3.1.1. Características demográficas

La selección de la información demográfica es guiada por la estructura, el tamaño de la base de datos [75] y la correlación de cada una de estas variables con el objetivo de la presente investigación. Las características demográficas contendidas en el encabezado de cada grabación se muestran en la siguiente tabla:

Tabla 6. Características demográficas.

Información demográfica	Tipo de Datos
ID de participante	Numérico
Número de registro	Numérico
Tipo de registro	Parto, embarazo

Edad	Numérico (años)
Edad gestacional al momento del registro	Numérico (semanas)
Edad gestacional al momento del parto	Numérico (semanas)
IMC antes del parto	Numérico
IMC durante la grabación	Numérico
Gravidez	Numérico
Paridad	Numérico
Cesárea previa	Si, No
Posición placentaria	Fúndico, posterior, anterior
Uso de oxitocina sintético	Si, No
Epidural durante el trabajo de parto	Si, No
Comentarios para el registro	Alfanumérico
Comentarios para el parto	Alfanumérico
Clasificación	Vaginal, Cesárea

Los criterios de exclusión e inclusión son los siguientes:

Criterios de inclusión: edad gestacional entre 32 a 38 semanas y un máximo de dos registros por paciente, esto se debe a que este modelo trata de apoyar el diagnóstico preventivo, por lo que, si la edad gestacional es mayor a 38 semanas el parto está próximo a suceder, por el contrario, si la edad gestacional es menor a 32 semanas, el feto es inmaduro. En consecuencia, el registro no puede mostrar el comportamiento del binomio madre-hijo [76] [77].

Criterios de exclusión: se omiten los datos que forman parte de la identificación del paciente y los comentarios que no contienen información médica o demográfica.

La distribución inicial de los datos se muestra en la figura 15:

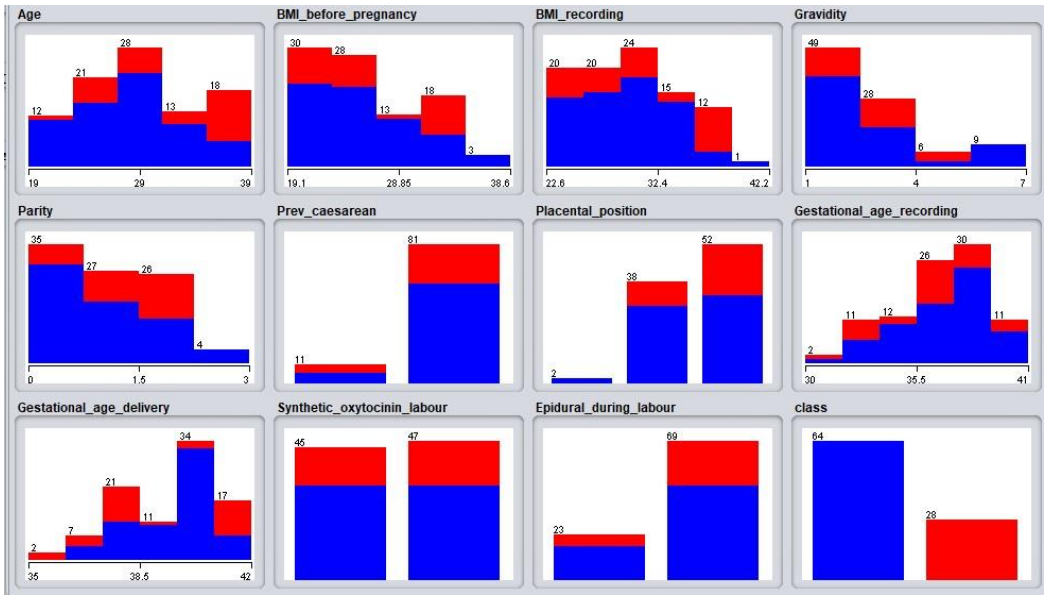


Figura 15. Distribución de la base de datos sin procesar. (Extraída de Weka).

Correlación de los datos

Adicional a los criterios de exclusión e inclusión previamente presentados se contempla el cálculo de la correlación presente entre las variables de entrada compuestas por características demográficas y la salida del modelo, por lo cual se tiene en cuenta la correlación, para ello se calcula la matriz de correlación [78].

El gráfico de dispersión [79] permite realizar una inspección visual de la distribución de los datos, este grafico es apoyo para la determinar la correlación y permite identificar el rango de flujo de datos, es decir, el valor máximo y mínimo de las variables. En la Figura 16 se presentan los tipos de relación que se pueden encontrar en el gráfico.

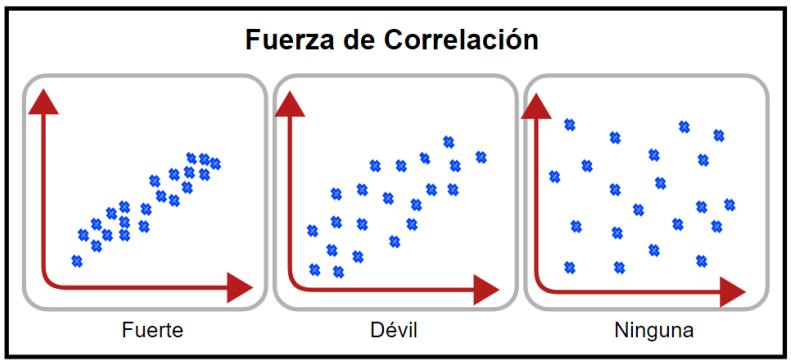


Figura 16. Interpretación del Diagrama de Dispersión. Elaboración propia.

La correlación es un método estadístico que permite conocer la relación entre dos variables, el coeficiente de correlación se denota r y se define como el cociente entre la covarianza de las variables y la raíz cuadrada del producto de la varianza de cada variable, esto se muestra en la ecuación (11) [80], este coeficiente puede tomar valores entre -1 y 1:

$$r_{xy} = \frac{cov(x,y)}{\sqrt{Var(x) * Var(y)}} \quad (11)$$

Dadas las variables de entrada $X = (x_1, x_2, \dots, x_i)$ la matriz de correlación [81] se representa de la siguiente forma:

$$R = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_i \end{array} \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & \cdots & x_i \end{array} \\ \begin{pmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1i} \\ r_{21} & 1 & r_{23} & \cdots & r_{2i} \\ r_{31} & x_{32} & 1 & \cdots & r_{3i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{i1} & x_{i2} & x_{i3} & \cdots & 1 \end{pmatrix} \end{array}$$

La propiedad conmutativa del coeficiente de correlación describe que el orden de las variables no afecta el cálculo, es decir, que el coeficiente de correlación x_{12} es igual al coeficiente x_{21}, lo cual concluye que la matriz es simétrica, a continuación, se ilustra el resultado:

$$R = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_i \end{array} \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & \cdots & x_i \end{array} \\ \begin{pmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1i} \\ r_{12} & 1 & r_{23} & \cdots & r_{2i} \\ r_{13} & x_{23} & 1 & \cdots & r_{3i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{1i} & x_{2i} & x_{3i} & \cdots & 1 \end{pmatrix} \end{array}$$

3.2. CARACTERIZACIÓN SEÑAL EHG

Para realizar la medición de la señal se utilizó un registrador de mediciones fisiológicas multipropósito de 16 canales “Embla A10” el cual es usado para los estudios de sueño realizado en hospitales con una frecuencia de muestreo de 200 Hz, este equipo usa el software “Somnológica” para digitalizar las señales a 16 bits en un archivo EDF (formato europeo). Para disminuir el ruido se implementa un filtro antisoplamiento con frecuencia de 100 Hz pasa bajos.

El registro se tomó mediante un protocolo estandarizado de EHG de 4 por 4 electrodos para lo cual se usan electrodos Ag/AgCl (plata / cloruro de plata) como se ve en la figura 16, los cuales son reutilizables con diámetro exterior D = 13 mm e interior d = 8 mm.

Figura 17. Electrodo Ag/AgCl [82].

La señal contiene los datos de 16 electrodos monopolares, la figura 18 muestra la distribución de los electrodos en el abdomen de la gestante, la distancia entre los centros de los electrodos es de 17.5 mm. Para ubicar los electrodos se tuvo en cuenta que los electrodos 9 al 12 en el medio del útero teniendo en presente que el electrodo 10 y 11 están entre el fondo uterino y la sínfisis púbica. El ombligo se evita desplazando la matriz de electrodos hacia arriba o hacia abajo según la gestante analizada. Se ubico un electrodo de tierra y uno de referencia en las crestas ilíacas.

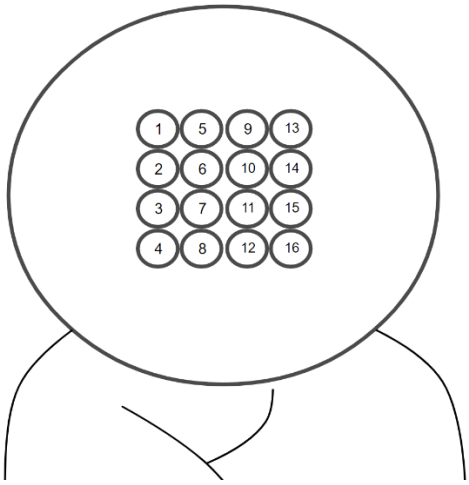


Figura 18. Distribución de los16 electrodos.

Para incorporar la bio-señal al modelo, es necesario realizar un preprocesamiento, mediante el método de ventana [83], este proceso toma un pequeño subconjunto de un conjunto de datos más grande con un filtro tipo ventana rectangular, la amplitud usada en este artículo es del 30 %. Esto implica truncar el conjunto de datos antes y después de esta ventana.

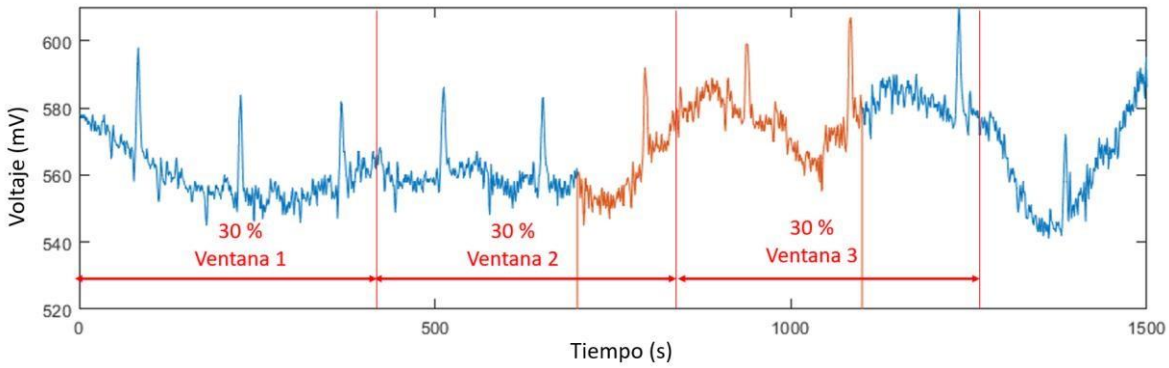


Figura 19. Diagrama de técnica de ventaneo EHG. (Extraída de Matlab)

En la figura 19 se ilustra en color naranja una alteración a la señal EHG, es decir, la gestante realizo un movimiento lo que ocasionó que la señal cambiara los niveles de voltaje. Esta misma figura muestra el proceso de ventaneo donde la ventana 1 corresponde a los primeros 450 segundos el 30% de la señal completa, seguido de la ventana 2 con la información entre el segundo 451 al 900, esto se repite hasta terminar de analizar la señal EHG.

Dada la naturaleza dinámica de los sistemas biológicos, las señales adquiridas pueden presentar en su mayoría comportamientos estocásticos y no estacionarios por lo que ciertas características cambian su comportamiento con el tiempo, en consecuencia, se deben analizar por secciones como se plantea con la técnica de ventaneo, analizando así la señal por completo. Las características en el dominio del tiempo y la frecuencia son calculadas para cada subventana, estas son [84]:

- *Raíz media cuadrática:* es la raíz cuadrada de la media de la señal EHG permite identificar la amplitud de las contracciones o movimientos uterinos presentes en el registro [82], es decir la intensidad de las mismas:

$$RMS = \sqrt{\left(\frac{1}{N}\sum_{i=1}^{N} x(i)^2\right)} \tag{12}$$

donde $x(i)$ es la ventana de EHG, $i = 1,2, \dots, N$ donde N es la longitud de $x(i)$.

- *Varianza*: representa la medida en que los datos fluctúan en torno a la media de la bioseñal [84].

$$VAR = \frac{1}{N-1}\sum_{i=1}^{N}(x(i) - \mu)^2, \text{ donde } \mu \text{ es la media aritmética} \quad (13)$$

- *Valor medio absoluto:* representa el promedio de los valores absolutos de la N muestras de la EHG [84]:

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|x(i)| \quad (14)$$

donde $x(i)$ es la ventana de EHG, $i = 1,2, \dots, N$ donde N es la longitud de $x(i)$.

- *Potencia media:* es la potencia promedio dentro de un rango espectral de la señal EHG *[*84]:

$$P = \lim_{L\to\infty}\left[\frac{1}{2L}\int_{-L}^{L}|x(t)|^2\,dt\right] \quad (15)$$

Donde $x(\mathrm{i})$ es la ventana de EHG, $i = 1,2, \dots, N$ donde N es la longitud de $x(\mathrm{i})$.

- *Frecuencia media:* es el valor medio de la distribución en términos de frecuencia del EHG [84]:.

$$MF = i_m\frac{f_s}{N}\sum_{i=1}^{i=i_m}P(i) = \sum_{i=i_m}^{i=N}P(i) \quad (16)$$

donde $x(i)$ es la ventana de EHG, $i = 1,2, \dots, N$ donde N es la longitud de $x(i)$ y $\mathrm{P(i)}$ es el espectro de potencia de los $\mathrm{x(i)}$.

A continuación, en la tabla 7a. se muestra la estructura de la matriz donde se guardan el resultado del proceso de ventaneo y una vez terminado este se realiza el cálculo de las características anteriormente descritas y son almacenadas en la matriz titulada SM (safe Machine) ilustrada en la tabla 7b.

Tabla 7. Distribución de técnica de ventaneo. (a) Preprocesamiento de EHG, (b) SM: características EHG

(a)

	Vent 1	Vent 2	$\cdots$	Vent n
Electrodo 1	$a_{1,1}$	$a_{1,2}$	$\cdots$	$a_{1,n}$
Electrodo 2	$a_{2,1}$	$a_{2,2}$	$\cdots$	$a_{2,n}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
Electrodo 16	$a_{16,1}$	$a_{16,2}$	$\cdots$	$a_{16,n}$

(b)

RMS	Var	MAV	P	MF
RMS_1	VAR_1	MAV_1	P_1	MF_1
RMS_2	VAR_2	MAV_2	P_2	MF_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
RMS_{16}	VAR_{16}	MAV_{16}	P_{16}	MF_{16}

Para definir el conjunto de datos de entrada del modelo se tienen en cuenta cinco grupos de parámetros descritos a continuación:

El grupo 1 tiene datos demográficos de la mujer embarazada como la edad, el número de partos, el número de embarazos, el IMC, las cesáreas anteriores y la ubicación de la placenta, estos son mencionados previamente en la sección 3.1. El grupo 2 contiene un promedio de las variables de la bio-señal EHG descritos en la sección 3.2 y datos demográficos.

El grupo 3 tiene variables de la bio-señal EHG descritos en la sección 3.2. donde los datos son un promedio de las variables de los electrodos en forma vertical (figura 20, electrodos con recuadro rojo). El grupo 4 tiene la raíz cuadrática media de los 16 electrodos. El grupo 5 tiene variables de la bio-señal, tomando un promedio de las variables de los electrodos de forma horizontal (figura 20, electrodos con recuadro azul).

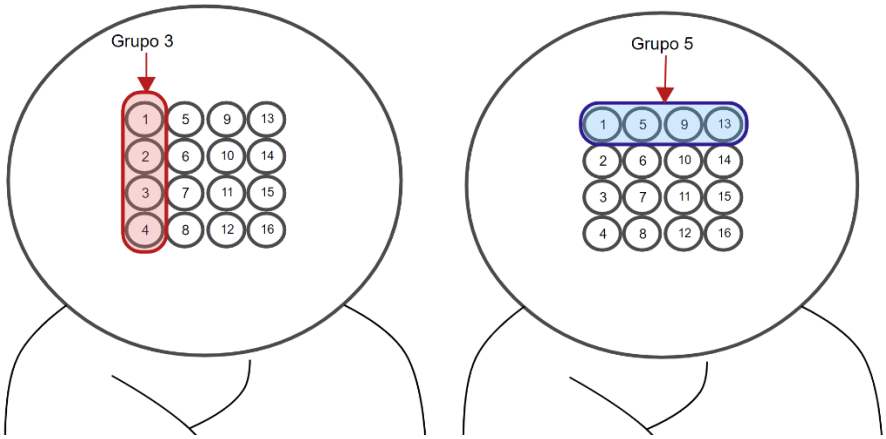


Figura 20. Ubicación de electrodos (grupo 3 y 5) [73] (Elaboración propia)

3.3. IMPLEMENTACIÓN Y EVALUACIÓN DE CLASIFICADORES

Para comparar los resultados del modelo bio-inspirado, los clasificadores se ejecutan con cinco grupos de parámetros, cada uno de ellos se divide en un 70 % para entrenamiento y un 30 % para realizar una validación cruzada. Se utilizan dos herramientas computacionales libres WEKA y Python (IDLE). Finalmente, con los resultados es posible calcular los indicadores de desempeño mostrados en la sección 2.

3.3.1. Clasificador: K vecinos más cercanos (KNN)

Para la clasificación con KNN, el número de vecinos comienza en 1 y aumenta según el rendimiento y el tiempo de procesamiento, la función de distancia implementada es la euclidiana (sección 2) y las iteraciones comienzan en 50 y aumentan si es necesario. A continuación, el pseudocódigo desarrollado.

Algoritmo 1. Pseudocódigo para clasificador KNN.

Cargar
$D = \{(x_1, c_1), \dots, (x_N, c_N)\} \leftarrow Conjunto\ de\ entrada\ y\ su\ clase$
Inicio
$x = (x_1, \dots, x_n) \leftarrow Nuevo\ conjunto\ de\ datos\ a\ clasificar$

$Para\ i = 1{:}\ 1{:}\ N$
$d_i = d(x_i, x) \leftarrow Calcula\ la\ distancia\ entre\ el\ objeto\ clasificado\ y\ el\ nuevo$
$D_x{}^K \leftarrow Organizar\ d_i\ en\ orden\ ascendente$ y guardar
$D_x{}^K \leftarrow Depurar\ para\ quedase\ con\ los\ K\ casos\ más\ cercanos\ a\ x$
$Asignar\ a\ x\ la\ clase\ más\ frecuente\ en\ D_x{}^K$
Retornar
$Sensibilidad, Especificidad\ y\ Exactitud$
Fin

3.3.2. Clasificador: perceptrón multicapa (MLP)

El MLP es una red unidireccional (feed forward), para este proyecto se comienza en 1 capa oculta y aumenta según los resultados, además el número de neuronas por capa varía según el rendimiento y el tiempo de procesamiento, la función de activación para la capa de entrada elegida es la función lineal para evitar valores negativos en el proceso, para las capas ocultas es la función sigmoidal aumentando así el rendimiento y finalmente para la capa de salida se utiliza la función sigmoide debido a que el problema de esta investigación es clasificación binaria es decir existen 2 clases mutuamente excluyentes, además para usar este clasificador se normalizan los valores de entrada. A continuación, el pseudocódigo desarrollado:

Algoritmo 2. Pseudocódigo para clasificador MLP.

Cargar
$X_{in} \leftarrow$ Datos de entrada
$Y \leftarrow$ Salida de la red neuronal
Inicializar la red (
$n_{in} \leftarrow$ número de neuronas de entrada
$k_{hi} \leftarrow$ número de capas ocultas
$n_{hi} \leftarrow$ número de neuronas ocultas
$n_{out} \leftarrow$ número de neuronas de salida
)
$W_k \leftarrow$ Inicializar los pesos de todas las neuronas a un valor aleatorio.
Mientras convergencia **hacer {**
1. $Propagar\ las\ entradas\ X_{in}\ hacia\ adelante\ y\ obtener\ las\ salidas$

2. $Propagar\ los\ errores\ de\ atrás\ hacia\ delante$
3. $Actualizar\ los\ pesos\ W_k$
4. $Actualizar\ red(W_k)$

}
Retornar
$Sensibilidad, Especificidad\ y\ Exactitud$
Fin

3.3.3. Clasificador: máquinas de vector soporte (SVM)

El clasificador SVM encuentra un hiperplano que separa los datos, este clasificador emplea un mapeo no lineal para separar los datos nuevos empleando las funciones Kernel, en este proyecto se varía los tipos de kernel entre lineales y polinomiales cuyas ecuaciones se muestran en la sección 2. A continuación, el pseudocódigo desarrollado.

Algoritmo 3. Pseudocódigo para clasificador SVM.

Cargar
X_{in} $\leftarrow D$atos de entrada
Y $\leftarrow S$alida de la red neuronal
Inicializar set de entrenamiento y validación
$x_{train} = 0,\quad y_{train} = 0$
$x_{test} = 0,\quad y_{test} = 0$
Inicializar SVM **(**
$K \leftarrow$ Kernel
$x_{train} = 70\ \%\ de\ X_{in},\quad y_{train} = 70\ \%\ de\ Y$
$x_{test} = 30\ \%\ de\ X_{in},\quad y_{test} = 30\ \%\ de\ Y$
$C \leftarrow Condición\ de\ convergencia$
)
Mientras $C\ no\ se\ cumpla$ **hacer {**

1. $Entrenar\ el\ modelo\ SVM$
2. $Calcular\ Y_{train}$
3. $Realizar\ test\ y\ calcular\ Y_{test}$

}
Retornar
$Sensibilidad, Especificidad\ y\ Exactitud$
Fin

3.3.4. Clasificador: aprendizaje profundo (DP)

El aprendizaje profundo varía el tamaño del lote de entrenamiento, es decir, hay N muestras para entrenamiento y el tamaño del lote se establece en 5 o 10. El segundo parámetro a ajustar es el número de épocas por procesamiento entre 50, 100 y 1000. A continuación, el pseudocódigo desarrollado.

Algoritmo 4. Pseudocódigo para clasificador DP.

Cargar
X_{in} ← Datos de entrada
Y ← Salida de la red neuronal

Inicializar set de entrenamiento y validación
$x_{train} = 0,$ $y_{train} = 0$
$x_{test} = 0,$ $y_{test} = 0$
Inicializar la red (
n_{in} ← número de neuronas de entrada
n_{out} ← número de neuronas de salida
N_b ← $Bachsize$
C ← $Condición\ de\ convergencia$
)
Mientras $C\ no\ se\ cumpla$ **hacer {**
$Ejecutar\ deep\ leraning$
$Calcular\ Y_{train}, Y_{test}$ **}**
Retornar
$Sensibilidad, Especificidad\ y\ Exactitud$
Fin

Dado el objetivo de este trabajo la prueba de desempeño se realiza mediante el cálculo de los índices explicados en la sección 2, teniendo en cuenta la matriz de confusión que se ilustra en la siguiente tabla:

Tabla 8. Matriz de confusión.

Matriz de confusión		
a	b	
VN	FP	a: Vaginal
FN	VP	b: Cesárea

Donde VP (Verdaderos positivos), VN (Verdaderos negativos), FP (Falsos positivos), FN (Falsos negativos).

3.4. IMPLEMENTACIÓN E INTERFAZ GRÁFICA

Para el estudio y prueba de los clasificadores anteriormente descritos se utilizan 2 herramientas de uso libre: WEKA con interfaz gráfica propia y Python (IDLE) en su interfaz gráfica en Visual Studio. Dentro de cada uno de ellos es necesario seleccionar los paquetes y algoritmo adecuado para la aplicación.

Para la implementación clínica del modelo bio-inspirado, es necesario desarrollar una interfaz gráfica de visualización personalizada para ordenador, el objetivo de la interfaz es contemplar el cargue de datos para entrenar, entrenamiento y validación de los clasificadores antes mencionados y la visualización de los indicadores de rendimiento en las etapas de entrenamiento y validación. Los requerimientos funcionales se muestran a continuación:

Tabla 9. Requerimientos funcionales para interfaz gráfica.

N° de Req	**Detalles**	
1	**Descripción**	1. Cargue de base de datos 2. Lectura de base de datos para realizar entrenamiento y validación
	Procesos relacionados	Elaborar archivo .csv que contenga la base de datos de acuerdo al grupo de parámetros seleccionados.

2	**Descripción**	3. Seleccionar clasificador.
	Procesos relacionados	Desarrollar los clasificadores KNN, SVM, MLP y DP
3	**Descripción**	4. Entrenar modelo 5. Exportar modelo entrenado 6. Guardar resultados obtenidos en entrenamiento y validación.
	Procesos relacionados	Generar grupo de datos de entrenamiento con el 70 % y de validación con el 30 % de los datos ingresados. Generar un .csv con la información de salida de cada proceso.
4	**Descripción**	7. Calculo e impresión de indicadores de desempeño
	Procesos relacionados	Lectura de resultados y desarrollo de matriz de confusión
5	**Descripción**	8. Cargue y lectura de archivo con información de una clasificación nueva.
	Procesos relacionados	Elaborar archivo .csv que contenga la información de la nueva paciente.
6	**Descripción**	9. Clasificación de la nueva paciente de acuerdo con modelo previamente entrenado.
	Procesos relacionados	Aplicar los clasificadores KNN, SVM, MLP y DP.
7	**Descripción**	10. Ver resultados de clasificación de nueva paciente
	Procesos relacionados	Lectura de archivo con extensión .csv guardado con los resultados de la clasificación. Cálculo de Indicadores de desempeño.

A continuación, se mencionan los requerimientos no funcionales para el diseño de la interfaz gráfica:

- Lenguaje de desarrollo Python, el cual permite la programación de los clasificadores de acuerdo con la metodología anteriormente descrita.

- Entorno de desarrollo Visual Studio.
- Selección de paquetes de clasificadores.

Para el diseño de la interfaz gráfica se propone la siguiente distribución:

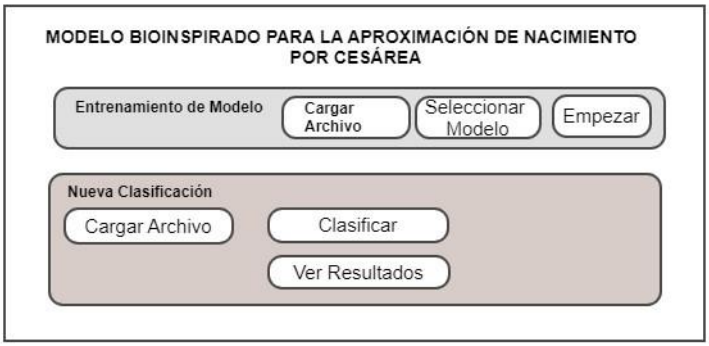


Figura 21. Diseño de interfaz gráfica. (Elaboración propia)

Para realizar el cargue correcto del archivo que contiene la base de datos para entrenar el modelo y la información de la nueva paciente, es necesario tener en cuenta la siguiente estructura:

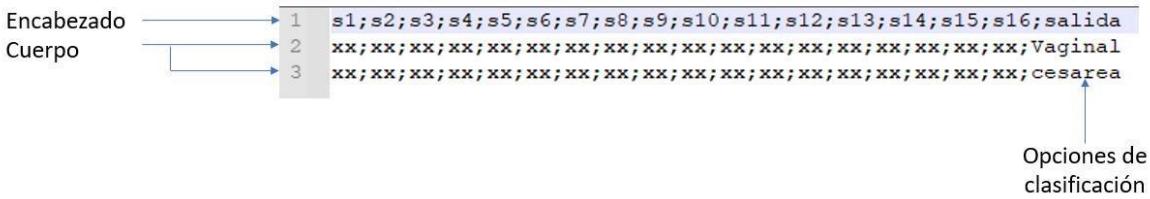


Figura 22. Estructura de archivo con extensión .csv para interfaz gráfica. (Elaboración propia)

4. RESULTADOS Y DISCUSIÓN

4.1. BASE DE DATOS

Los datos para realizar este proyecto son extraídos de la base titulada "Icelandic 16-electrode Electrohysterogram Database", la cual contiene 122 electrohisterogramas (EHG) de 45 mujeres gestantes con edades entre los 19 a los 39 años, no se realizó ningún procesamiento de la señal previamente. Está compuesta por: un archivo con las grabaciones de la señal que se convirtieron archivos de señal al ser digitalizadas (.dat), una imagen (.jpg) y encabezado (.hea).

Tocógrafo escaneado: Es una imagen con extensión .jpg mostrada en la figura 23 la cual representa la señal tomada de la mujer gestante, esta es modificada manualmente para indicar inicio y fin de la toma, cada una de las divisiones horizontales representa 30 segundos.

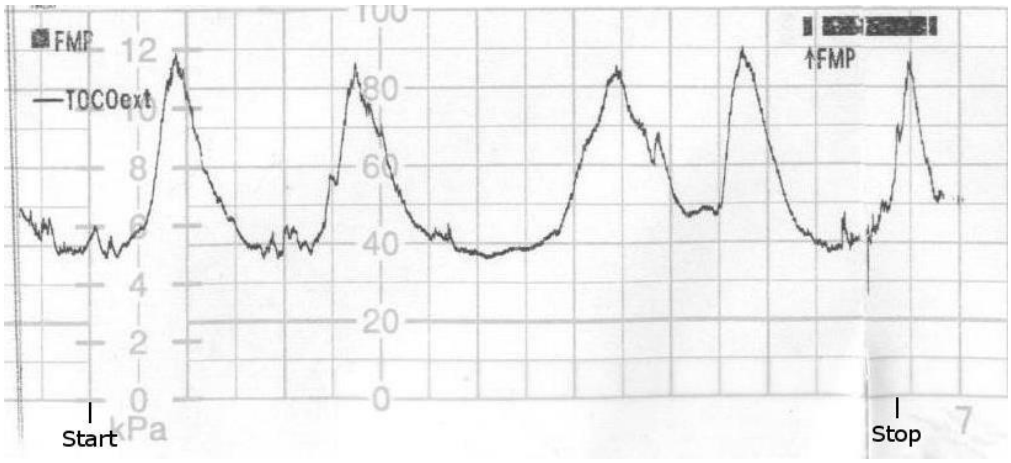


Figura 23. Fotografía escaneada [82].

El archivo que contiene la información demográfica de cada paciente se titula *Número_Paciente.hea,* en la figura *24* se muestra la estructura y posible contenido de este archivo, el cual es procesado para crear la base de datos.

```
#Info:
#ID:ice001
#Record type:labour
#Record number:1/1
#Age(years):31
#BMI before pregnancy:23.3
#BMI at recording:27.6
#Gravidity:3
#Parity:2
#Previous caesarean:No
#Placental position:Fundus
#Gestational age at recording(w/d):39/3
#Gestational age at delivery:39/3
#Mode of delivery:Vaginal
#Synthetic oxytocin use in labour:No #Epidural
during labour:No
#Comments for recording: #Electrodes placed 5-10 mins prior
to beginning of recording.
```

Figura 24. Ejemplo estructura archivo encabezado - información demográfica [82].

4.1.1. Características demográficas

En este trabajo se establecieron las características demográficas más relevantes para determinar el riesgo de cesárea, a partir de los trabajos de investigación reportados por W.L. Moreira [21], M. K. Nallakaruppan [22], E. Nodelman [60] y R. S. Solano [63]; quienes consideraron que la información demográfica que aporta capacidad de decisión al modelo bio-inspirados son : edad, gravidez, altura, grupo sanguíneo, peso, grupo sanguíneo, índice de masa corporal, genero del feto, hipertensión, examen sanguíneo, entre otros.

La selección de los datos demográficos se realiza de acuerdo con los criterios expuestos en el apartado 3, se considera la base de datos de entrada que contiene variables demográficas de la gestante, descritas a continuación: edad de la gestante, edad gestacional, índice de masa corporal (IMC) (antes del embarazo y al momento del registro), posición de la placenta, cantidad de gravidez, paridad (partos anteriores

mayores de 22 semanas de gestación) y cesáreas previas, y los datos de salida son el modo de parto (parto vaginal y cesárea).

Correlación de Pearson de los datos

Seleccionar las características demográficas permite disminuir la dimensionalidad de la base de datos de entrada para los clasificadores, para apoyar el resultado obtenido por la evaluación de los criterios de exclusión e inclusión se realizó un análisis de la dispersión de los datos de entrada y el cálculo del coeficiente de correlación de Pearson r.

La gráfica de dispersión se genera usando el lenguaje de programación Python con el siguiente código:

```
sns.pairplot(data1, height=1.5)
```

El coeficiente de correlación de Pearson r y la matriz de correlación se calcula y grafica por medio del lenguaje Python, con el siguiente código:

```
data=pd.read_csv('base.csv')
data.info()
data.describe()

ces=data.loc[data.categoria == "caesarean"]
corr_mat_ces=ces.corr()

plt.figure(figsize=(10,8))
plot=sns.heatmap(ces.corr().round(2), annot=True)
```

En la Figura 25 se observa la dispersión de los datos de entrada y los datos de salida del modelo, la dispersión de las variables de entrada como la edad, cesáreas previas, posición placentaria y edad gestacional no presentan una relación entre sí, esto de acuerdo con lo ilustrado en la Figura 16.

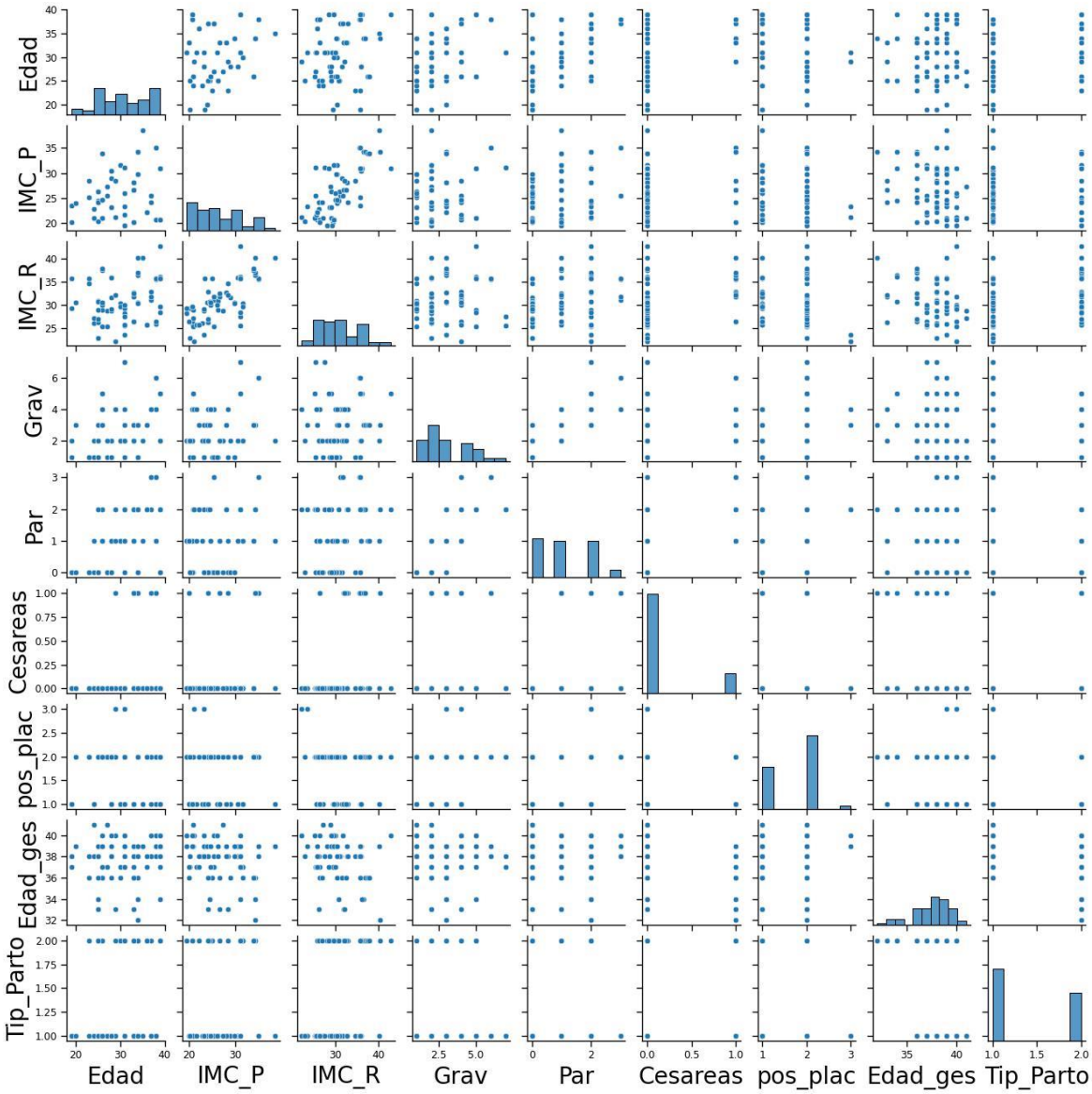


Figura 25. Grafica de dispersión de los datos de entrada. Elaboración propia.

En la Tabla 10 se presentan la descripción de la relación entre variables dependiendo del valor que resulta del cálculo del coeficiente de correlación:

Tabla 10. Valores de correlación [85]

Coeficiente de Correlación	**Descripción**
$r = \pm 1$	Correlación Perfecta
$\pm 0.8 < r < \pm 1$	Correlación Muy Alta
$\pm 0.6 < r < \pm 0.8$	Correlación Alta
$\pm 0.4 < r < \pm 0.6$	Correlación Moderada
$\pm 0.2 < r < \pm 0.4$	Correlación Baja
$0 < r < \pm 0.2$	Correlación Muy Baja
$r = 0$	Correlación Nula

En la Figura 26 se observa el resultado del cálculo del coeficiente de correlación para las características demográficas ya que son las variables de entrada del modelo para la clasificación de cesárea y de parto vaginal, los coeficientes de correlación entre las variables de entrada: edad, cesáreas previas, posición placentaria y edad gestacional, el coeficiente de correlación presenta un rango de $-0.45 < r < 0.46$ lo cual indica que tienen una correlación baja, muy baja o nula en algunos casos.

A excepción del coeficiente de correlación entre las características IMC antes de la gestación - IMC durante la gestación, paridad – gravidez, las cuales son consideradas para la detección de parto por cesárea bebido a que no poseen una correlación muy alta o perfecta, dado que el coeficiente de Pearson considerado para determinar que una variable es redundante es $r < 0.8$ [86][87] [88][89].

Por lo anterior, las características demográficas mencionadas son consideradas para la detección de riesgo de cesárea, teniendo en cuenta que los profesionales de la salud consideran que son representativo para el apoyo del ya que están relacionadas con las morbilidades en la gestación que concluyan en una cesárea para evitar el sufrimiento fetal [86][87].

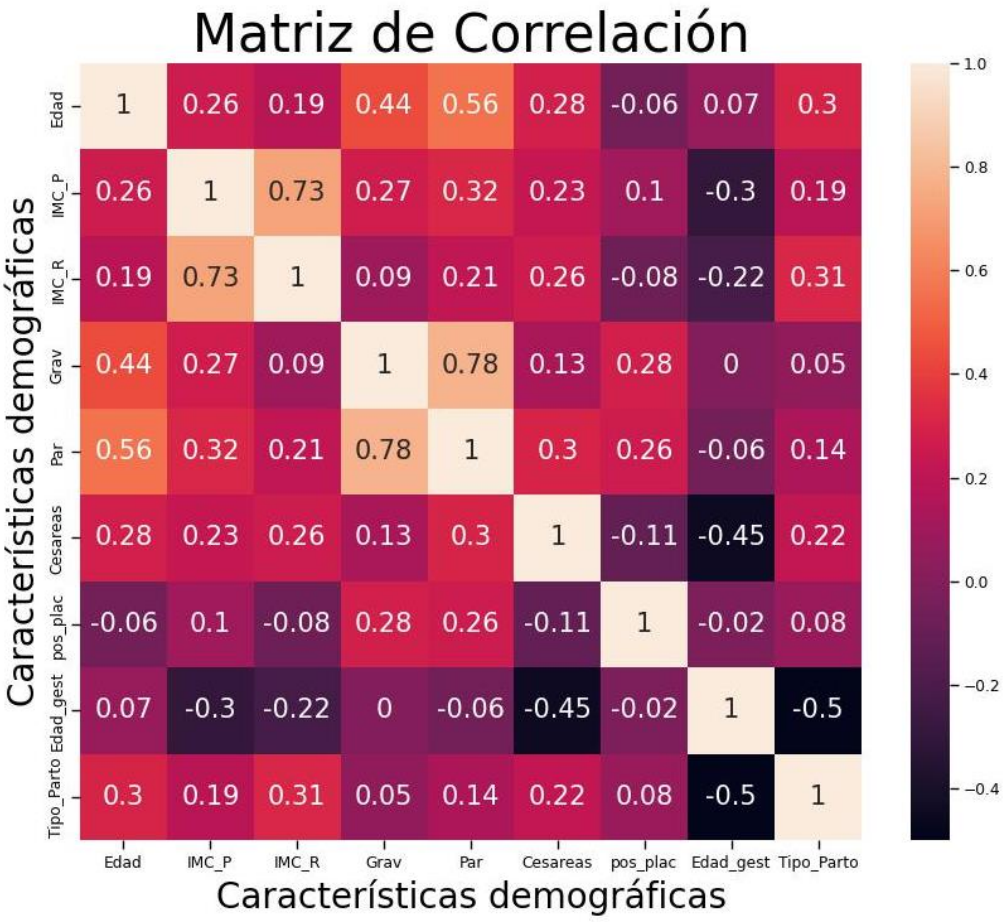


Figura 26. Gráfica del Coeficiente de correlación de Pearson r, para los partos por cesárea. Elaboración propia.

La correlación clínica de la salida del presente modelo con los datos de entrada se describe a continuación:

Información demográfica de la gestante

- La edad de la gestante: las mujeres con edad avanzada, es decir, más de 35 años son asociadas en su mayoría a morbilidades durante el embarazo y a partos por cesárea [84].

- Índice de masa corporal (IMC) previo a la gestación: este valor es calculado de la siguiente manera, es el peso de una persona en kilogramos dividido por el cuadrado de la estatura en metros [91].

$$IMC = \frac{Peso(Kg)}{Altura(m)^2} \tag{17}$$

  Este índice permite identificar si la gestante se encontraba en bajo peso, peso saludable, sobrepeso u obesidad previo al embarazo, lo cual influye en los cuidados nutricionales que se deben considerar para el desarrollo de la gestación, ya que en caso de presentar sobre peso o desnutrición se contempla un plan nutricional detallado y se requiere seguimiento por parte del nutricionista.
- Índice de masa corporal (IMC) durante la grabación: las maternas que presentan un IMC elevado, aumentan la posibilidad de padecer morbilidades y por ende complicaciones para el término del mismo, en la literatura se encuentra que el riesgo de cesárea aumenta con el peso de la gestante [92], el aumento del IMC durante el embarazo se considera normal si esta entre los 10 a 18 Kg de diferencia, por el contrario si este índice disminuye se requiere intervención médica ya que esto compromete el desarrollo del feto.

Formula obstétrica

En las instituciones médicas se implementa la fórmula para resumir los antecedentes obstétricos de la gestante se compone de gestaciones, partos a término, partos pretérmino, abortos e hijos nacidos vivos [92]. A continuación, se mencionan los componentes de esta que son usados para el desarrollo de este trabajo:

- Gestaciones previas: la cantidad de gestaciones previas es un parámetro que permite identificar si durante estas gestaciones se presentaron anomalías o morbilidades.

- Partos previos: es la cantidad de embarazos culminados, sin embargo, este parámetro se correlaciona directamente con las cesáreas previas ya que si la gestante tiene más de 3 partos previos y estos han sido por cesárea es una contraindicación absoluta para el parto vaginal por lo que directamente el trabajo de parto seria planeado mediante cesárea [94].
- Cesáreas previas: Si han existido cesáreas previas, es posible que el modo de parto se repita. Cada caso se debe valorar individualmente ya que esto depende de los resultados y secuelas de la cirugía uterina anterior [91].
- La posición de la placenta: si es previa, el parto debe ser con cesárea porque la placenta obstruye el cuello uterino [92].
- Edad gestacional: este parámetro permite identificar la evolución esperada del feto, de acuerdo a lo encontrado en la literatura las gestaciones entre 29 y 36 semanas tienen una probabilidad de culminar en cesárea con un 57 %, este porcentaje disminuye a 43.3% si la gestante es primigesta, es decir, está en su primera gestación [97].

El resultado de la selección de información demográfica se ilustra en la figura 27, donde se observa la distribución que tiene cada una de las características, teniendo en cuenta las dos clases identificadas como vaginal con color azul y cesárea con color rojo.

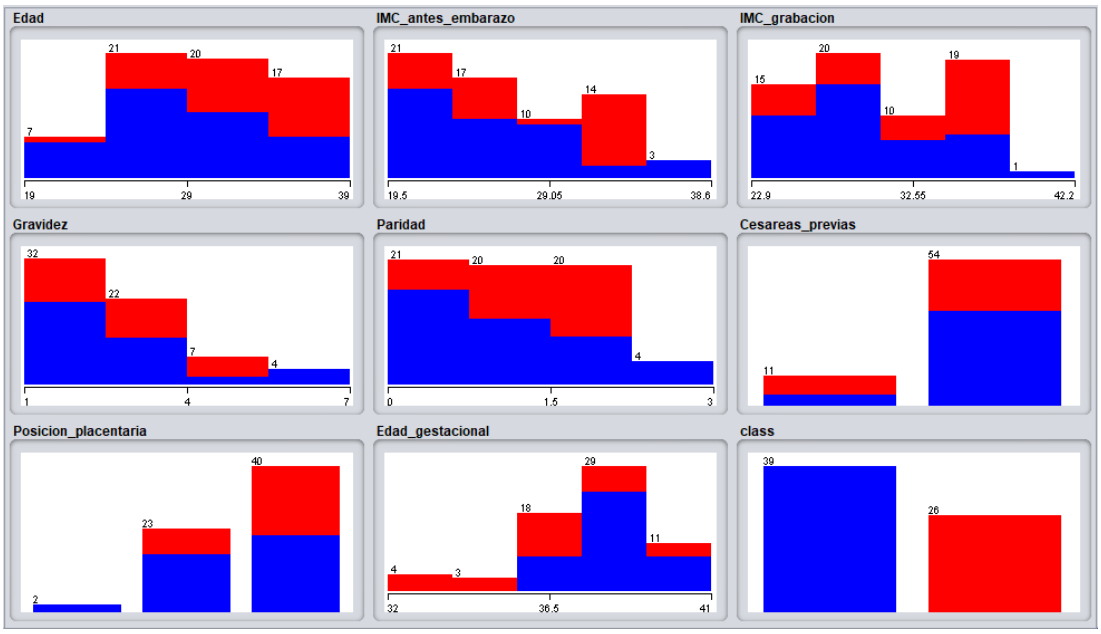


Figura 27. Distribución de la base de datos. (Extraída de Weka).

La estructura del archivo que contiene la información demográfica se puede observar en la figura 28, este se basa en dar nombre a cada atributo y cada fila tendrá la información de cada paciente y su información:

```
1   @RELATION partos
2
3   @ATTRIBUTE Age                                numeric
4   @ATTRIBUTE BMI_before_pregnancy               numeric
5   @ATTRIBUTE BMI_recording                      numeric
6   @ATTRIBUTE Gravidity                          numeric
7   @ATTRIBUTE Parity                             numeric
8   @ATTRIBUTE Prev_caesarean                     {Yes,No}
9   @ATTRIBUTE Placental_position                 {Fundus,Posterior,Anterior,}
10  @ATTRIBUTE Gestational_age_recording          numeric
11  @ATTRIBUTE Gestational_age_delivery           numeric
12  @ATTRIBUTE Synthetic_oxytocinin_labour        {Yes,No}
13  @ATTRIBUTE Epidural_during_labour             {Yes,No}
14  @ATTRIBUTE class                              {Vaginal,caesarean}
15
16  @DATA
17  31, 23.3,   27.6,   3,  2,  No,  Fundus ,   39  ,   39, No, No  ,   Vaginal
18  37, 24.1,   28.9,   4,  2,  Yes, Posterior, 38  ,   38, No, No  ,   caesarean
```

Figura 28. Estructura del dataset de información demográfica. (Extraida Dataset Physionet)

4.2. CARACTERIZACIÓN SEÑAL EHG

Como resultado de la caracterización de la señal EHG, se obtuvo una matriz SM (Safe Machine) que contiene los valores de raíz media cuadrática, varianza, valor medio absoluto, potencia media y frecuencia media de la señal EHG de cada paciente calculados mediante el método de ventaneo, con esta información es posible actualizar la base de datos de entrada que ya contenía la información demográfica, la cual puede ser procesada de forma ágil por el modelo bio-inspirado.

En el marco conceptual se observa la importancia de esta señal durante la labor de parto, sin embargo, ha sido poco explorado la incidencia del EHG cuando es registrada en las semanas previas a la fecha prevista de parto, considerando que, durante la labor de parto esta mide las características de las contracciones, pero durante la gestación permite analizar la actividad del útero, por ejemplo, el movimiento del feto, el intercambio sanguíneo de la placenta, entre otros. Para poder realizar este análisis se obtiene las características anteriormente calculadas.

El preprocesamiento de la señal requiere implementar una técnica de filtración, debido a que esto disminuye el ruido presente durante la captura de la bio-señal EHG, como se muestra en la figura *29*, este ruido está presente al inicio y al final de cada grabación y según los comentarios en la base de datos, esto se debe a los movimientos de la gestante e interferencias de electrodos mal conectados.

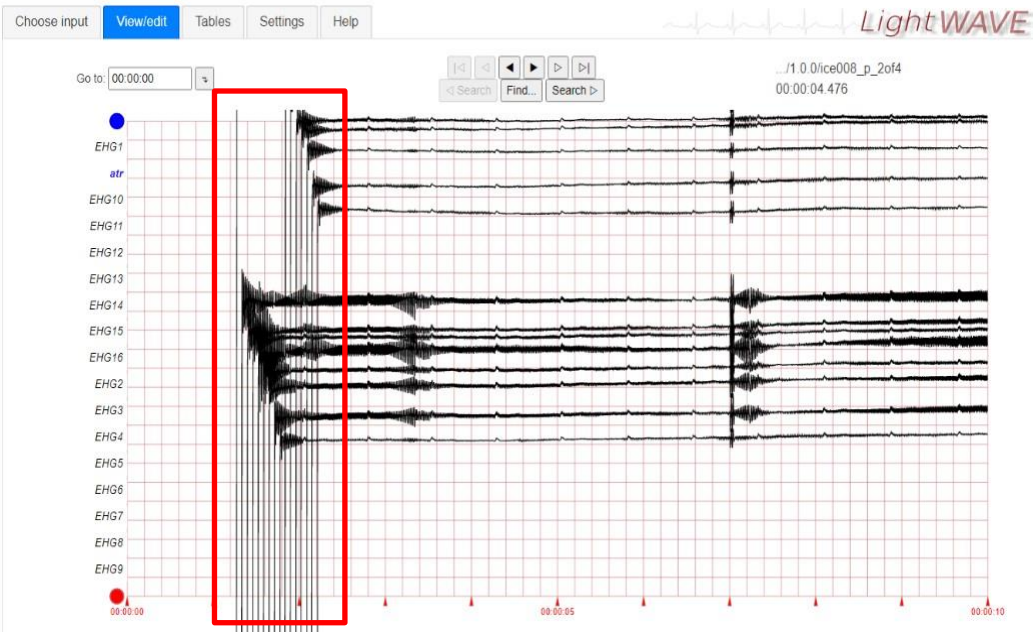
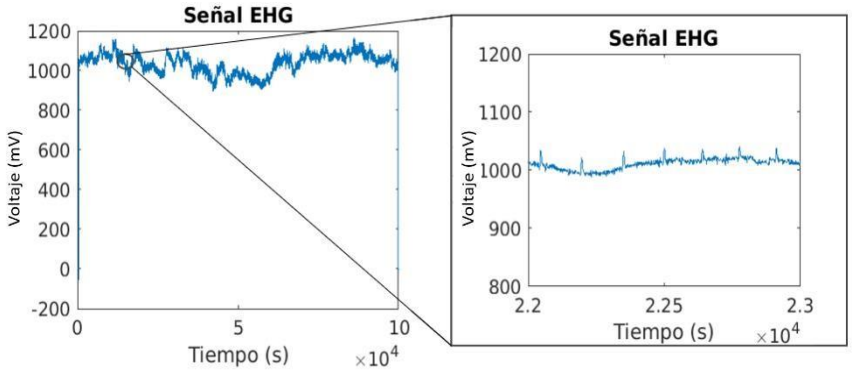


Figura 29. Interferencia de la Bio-señal EHG [63].

Seguido del paso de filtración, los registros de EHG son representados por características en el dominio de tiempo y la frecuencia mencionados en la sección 3, los cuales son calculados con el método de ventaneo, este permite el cálculo de las variables teniendo en cuenta la información en detalle. En la figura 30 se muestra la diferencia que existe al ver toda la información del EHG y ver el resultado del ventaneo.



(a) (b)

Figura 30. Bio-señal EHG, (a) señal total, (b) ventaneo. (Extraída de Matlab ®).

Para caracterizar la señal EHG se implementa la herramienta computacional MATLAB ®, mediante el siguiente flujo:

Se cargan los datos de las grabaciones de las gestantes obtenidas después de la depuración de los criterios de la sección 3.

Algoritmo 5. Código paso 1: caracterización de la señal EHG.

```
1. %PASO 1
2. clc; clear all;
3. val=importdata('ice001_l_1of1m.mat');
```

El resultado de este código se muestra en la figura 31, son cargadas 16 filas y el número de columnas varía según el tiempo de la grabación.

16x30001 double

	1	2	3	4	5	6	7	8	9
1	578	578	577	578	577	577	578	577	
2	306	307	309	309	307	306	310	309	
3	471	471	472	472	470	470	472	471	
4	1185	1182	1179	1181	1183	1180	1179	1181	
5	1041	1044	1041	1039	1042	1044	1043	1039	
6	624	624	625	626	624	624	627	625	
7	807	809	807	806	807	808	808	804	
8	1412	1412	1409	1409	1410	1409	1408	1407	
9	481	480	479	480	480	480	480	479	
10	895	897	895	895	895	897	897	894	
11	1382	1383	1384	1384	1382	1382	1385	1384	
12	595	597	596	595	596	597	598	595	
13	755	756	754	753	755	756	755	753	
14	906	906	906	906	905	906	908	906	

Figura 31. Dataset EHG ejemplo (Extraída de Matlab ®).

Después se realiza el cálculo del 30% de la señal y construcción de ventana.

Algoritmo 6. Código paso 2: caracterización de la señal EHG.

```
%PASO 2
f=size(val,1); c=size(val,2);
treinta=c.*0.30;
valu=zeros(16,treinta);
n=1;
for j=1:1:16
for i=c1:1:c2
valu(j,n)=val(j,i);
n=n+1;
end
n=1;
end
```

Con cada ventana se realiza el cálculo de las características de la señal EHG, mediante el siguiente código.

Algoritmo 7. Código paso 3: caracterización de la señal EHG.

```
%PASO 3
f=size(val1);
%RMS
n=1;
for i=1:1:f(1)
     val=val1(i,:);
     rms=sum((val).^2);
     Sm(n,i)=sqrt (rms/size(val,2));
end

%VARIANZA
xn=0; n=2;
for i=1:1:size(val1,1)
     val=val1(i,:);
     xn=sum((val).^2);
     Sm(n,i+48)=(1/(size(val1,2)-1))*xn;
end

 %MEDIA
n=3;
for i=1:1:f(1)
     val=val1(i,:);
     MAV=sum(abs(val));
     Sm(n,i+32)=MAV/f(2);
end
```

```
%POTENCIA PROMEDIO
n=4;
for i=1:1:size(val1,1)
     Sm(n,i+80)=bandpower(val1(i,:));
end

%FRECUENCIA MEDIANA
n=5;
for i=1:1:size(val1,1)
     Sm(n,i+96)=medfreq(val1(i,:));
end
```

Por último, cada resultado es guardado en la matriz llamada *SM*, una vez culminado el análisis de la señal se calcula el promedio para hallar el valor por la señal completa y se almacena, la cual es la entrada al módulo del clasificador.

4.3. IMPLEMENTACIÓN Y EVALUACIÓN DE CLASIFICADORES

Para la implementación del modelo bio-inspirado se desarrollan cuatro clasificadores, los cuales se seleccionaron de acuerdo con los resultados de precisión hallados en la literatura superiores al 70% [60][62][64][67][68], los clasificadores desarrollados en este trabajo son KNN, SVM, MLP y DP. Se realizó la evaluación de cada uno de estos mediante el cálculo de los índices de desempeño señalados en la sección 2.3.5, los valores computados permiten conocer el rendimiento del modelo para la clasificación de madres gestantes que puedan presentar riesgo de intervención por cesárea.

En el desarrollo del modelo bio-inspirado se implementaron 2 herramientas computacionales WEKA la cual es una herramienta de libre acceso y la segunda es la implementación con el lenguaje Python en el entorno gráfico de Visual Studio. Para desarrollar e implementar el modelo inspirado se tienen en cuenta cinco grupos de parámetros descritos en la tabla 10, esto con el fin de comparar los resultados al incluir las características extraídas de la señal EHG y los obtenidos sin esta información, cada una de estas es dividida en: un grupo de entrenamiento que contendrá el 70 % de los datos y un grupo de prueba o validación con el 30 % de los datos:

Tabla 11. Información por grupos.

Grupo 1 de parámetros	Grupo 2 de parámetros	Grupo 3 de parámetros
Datos demográficos: Edad Edad gestacional al registro Edad gestacional al parto IMC antes del parto IMC durante la grabación Gravidez Paridad Cesárea previa Posición placentaria Uso de oxitocina sintético Epidural durante el trabajo de parto Clasificación	**Datos demográficos:** Edad IMC antes del parto IMC durante la grabación Gravidez Paridad Cesárea previa Posición placentaria Edad gestacional al registro Edad gestacional al parto Uso de oxitocina sintético Epidural durante el trabajo de parto **Características del EHG:** RMS Varianza Valor medio absoluto Potencia media Frecuencia media Clasificación	**Características del EHG:** RMS Varianza Valor medio absoluto Potencia media Frecuencia media Clasificación
Grupo 4 de parámetros	**Grupo 5 de parámetros**	
VRMS	**Características del EHG:** RMS Varianza Valor medio absoluto Potencia media Frecuencia media Clasificación	

A continuación, se presentan los resultados obtenidos por cada clasificador implementado de acuerdo con el grupo de datos procesado:

4.3.1. Clasificador: K vecinos más cercanos (KNN)

El objetivo de este clasificador es hallar un plano que clasifique óptimamente los datos, en la figura 32 se observa la distribución de los datos, en este caso en el eje X tenemos la edad de la paciente y en los ejes Y el número de partos y embarazos.

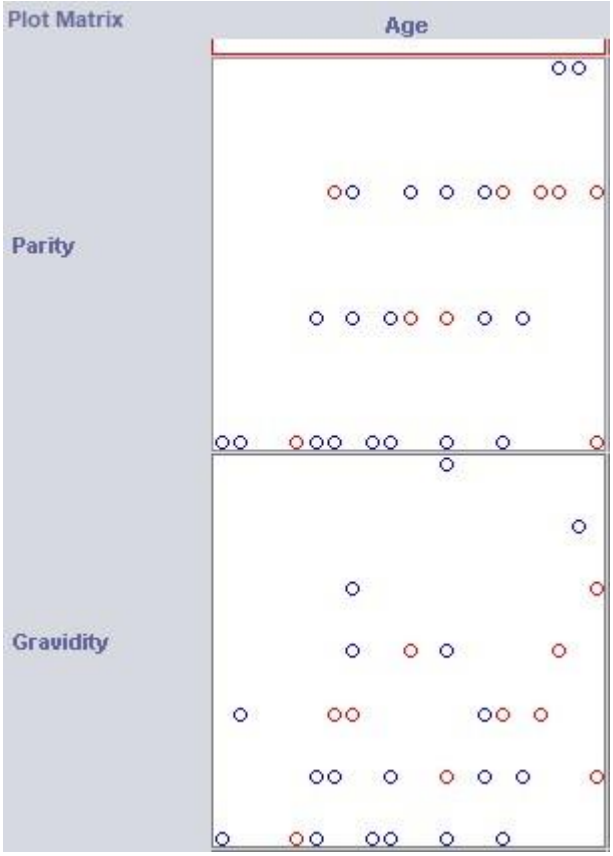


Figura 32. Distribución de datos para el clasificador KNN. (Extraída de Weka)

El número de vecinos elegido es 2, ya que, si cambia ese valor, el tiempo se incrementa sin variaciones en el resultado. Por otro lado, el porcentaje de sensibilidad disminuye del 100 % al 87,5 % entre el grupo 1 y 2 (tabla 12). En otras palabras, el modelo clasifica mejor al paciente que puede sufrir alguna complicación que a los que no tienen riesgo. La especificidad del grupo 3 es del 70 %, sin embargo, ese indicador decrece con respecto a los grupos 1 y 2. Los grupos 4 y 5 presentan los menores porcentajes de sensibilidad. Esto demuestra que este grupo

de datos no es diferenciable, es decir, sus valores no presentan suficiente información para que el modelo aprenda de las tendencias.

Tabla 12. Resultados del clasificador KNN.

Grupo de datos	Sensibilidad	Especificidad	Exactitud
1	100 %	83,3 %	90 %
2	87,5 %	83,3 %	77,7 %
3	29,4 %	72,7 %	44,4 %
4	35,2 %	72,7 %	50,5 %
5	25 %	72,5 %	40,7 %

De acuerdo con la descripción presentada en la implementación de la segunda herramienta este modelo realiza el cálculo de la distancia del elemento nuevo a cada uno de los existentes, luego ordena las distancias de menor a mayor para ir seleccionando la clase al cual pertenece.

El modelo KNN haciendo uso de la información demográfica obtiene una precisión del 90 %, y al procesar los grupos que solo contienen las características de bioseñal de EHG, disminuye al 77,7 %, lo que significa que la información demográfica representa mayor capacidad de decisión para el modelo. En la investigación desarrollada por la Universidad de Islamabad, Pakistán para clasificar los nacimientos prematuros que nacen mediante cesáreas, presenta el mayor porcentaje de sensibilidad del 80 % con el clasificador KNN [94]. La diferencia entre trabajos es del 10 % y radica en que el tipo de información demográfica es diferente y en el caso de la investigación de la Universidad de Islamabad algunos de ellos no representan capacidad de decisión.

La diferencia entre el grupo 1 y 2 se debe a que el uso de señales biológicas como EHG requiere una base de datos más robusta, a pesar de esto, los porcentajes mostrados son cercanos a las investigaciones encontradas en la literatura, por ejemplo, en la investigación de la Universidad de Ciencia y Tecnología de Polonia, obtuvieron una precisión del 96 % para predecir nacimientos prematuros teniendo en cuenta una gran base de datos con más de 400 registros [95].

4.3.2. Clasificador: perceptrón multicapa (MLP)

Tabla 13. Resultados de clasificador MLP.

Grupo de datos	Sensibilidad	Especificidad	Exactitud
1	75 %	100 %	90 %
2	75,5 %	75 %	75 %
3	58,8 %	54,4 %	57,1 %
4	82,3 %	90,6 %	90,9 %
5	17,6 %	100 %	50 %

El clasificador es implementado con 1 y 2 capas ocultas pero los resultados del rendimiento son inferiores al 50 % como se observa en la tabla 13, y si el número de capas ocultas es superior a 3, el rendimiento es igual pero este proceso sobrecarga (no es óptimo) la máquina, por estas razones, el MLP se implementa con 3 capas ocultas. Los porcentajes de acierto con el grupo 1 son superiores al 70 %. Al incluir la señal biológica del grupo 2, el porcentaje de precisión disminuye, lo que indica que estos valores no aportan valor de decisión al modelo. mientras que para el grupo 3, 4 y 5, modelo compuesto por 3 capas los porcentajes son superiores al 60 %.

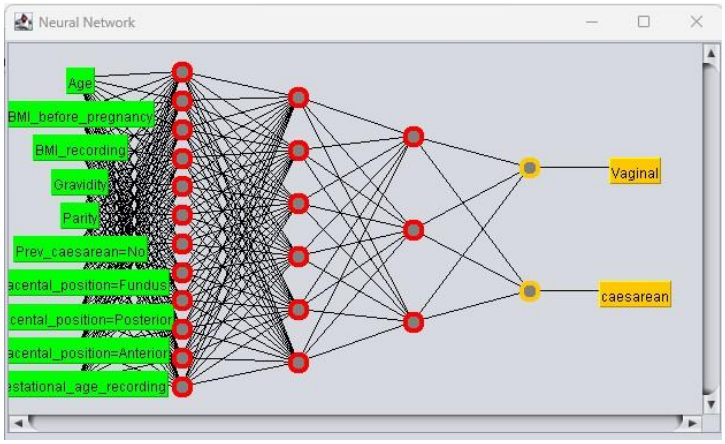


Figura 33. MLP resultado red (Extraída de Weka).

El MLP y el grupo 2 tienen una sensibilidad del 75,5 % y el grupo 4 es del 82,3 %. En comparación a estos resultados, la investigación realizada por la Universidad Politécnica de Valencia [96], hace uso de la señal biológica EHG y redes neuronales para predecir el éxito al inducir el parto, la sensibilidad es del 68,9 %. Estos resultados concuerdan con la presente investigación, si bien la herramienta no es un diagnóstico médico, permite soportarlo.

4.3.3. Clasificador: máquinas de vector soporte (SVM)

Tabla 14. Resultados de clasificador SVM.

Grupo de datos	Sensibilidad	Especificidad	Exactitud
1	75 %	83,3 %	80 %
2	28 %	50 %	25 %
3	100 %	12,5 %	39,2 %
4	35,2 %	54,5 %	42,8 %
5	11,7 %	54,4 %	20,2 %

Los resultados del presente proyecto con el modelo SVM se observan en la tabla 14 y presenta valores inferiores al 50 % de precisión con los grupos 2,3,4,5 pero con el grupo 1 los valores son superiores al 70 %. Esto coincide con los métodos encontrados en la literatura, porque los resultados están entre el 50 % y el 98 %, a veces sin especificar si la validación se realiza con datos que no son conocidos por el modelo.

Sin embargo, en el trabajo presentado por la Universidad Tecnológica de Bolívar, donde se predice la morbilidad materna, se afirma que la precisión del modelo varía de acuerdo a la edad gestacional cuando se está registrando el EHG, teniendo en cuenta el primer y segundo trimestre, la precisión es del 61 % y el último trimestre el porcentaje de precisión es superior al 70 % [97].

4.3.4. Clasificador: aprendizaje profundo (DP)

Tabla 15. Resultados del clasificador DP.

Grupo de datos	Sensibilidad	Especificidad	Exactitud
1	77,7 %	81,81 %	77,7 %
2	62,5 %	66,67 %	55,56 %
3	58,8 %	45.5 %	54 %
4	23,5 %	72,8 %	52 %
5	18,2 %	81,8 %	46 %

Para el grupo 1, la sensibilidad del modelo aumenta si se amplía el número de épocas, considerando un bach size de 10 y 1000 épocas. Esto es típico en este modelo, en vista de que cuantas más veces se entrena cada capa, más precisa es. La sensibilidad máxima es del 77,7 %. El grupo 2 presenta una sensibilidad del 62,5

% al configurar con un bach size de 5 y 1000 épocas, lo anterior de acuerdo con la tabla 15.

Para el grupo 4 el porcentaje de precisión es bajo, la sensibilidad es de 23.5 % y la especificidad de 72.8 %, esta última aumenta al procesar el grupo 5 con 81.8 %, sin embargo, la sensibilidad se mantiene con valores menores al 30 %.

El gráfico de las pérdidas de entrenamiento, indica la capacidad del modelo para adaptarse a los datos de entrenamiento, mientras que las pérdidas durante la validación indica la capacidad del modelo para adaptarse a los nuevos datos dado que el método de validación cruzada presenta datos que el modelo desconoce. La figura 34a se refiere al grupo 1, la capacidad de ajustarse a los datos de entrenamiento es más rápida que el ajuste a los nuevos datos.

Además, las pérdidas de entrenamiento final son 0,46 y las pérdidas de validación son 0,5, lo que significa que hay un sobreajuste parcial. El objetivo de estos modelos es que la pérdida de validación sea la menor posible. En el resultado es aceptable cierto sobreajuste, teniendo en cuenta la cantidad de datos con los que se entrenó y validó el modelo. En la figura 34b, el error de entrenamiento de validación muestra un sobreajuste parcial, al igual que el grupo 1.

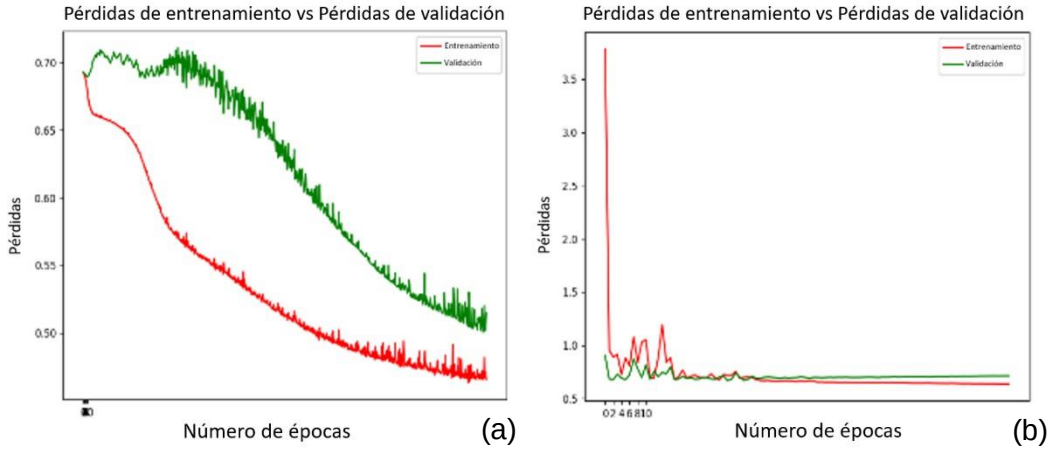


Figura 34. Gráfico de pérdidas. (a) grupo A=1, (b) Grupo B=2. (Elaboración propia)

En la literatura se encuentran investigaciones que predicen morbilidades en el embarazo desarrolladas con herramientas diferentes a las propuestas en esta investigación, por ejemplo, un modelo predictivo de parto por cesárea realizado por la doctora Sonia de Miguel Manso [97], es capaz de predecir si la labor de parto se realiza mediante cesárea con una precisión del 70.74 % haciendo uso de métodos estadísticos y basada en información demográfica. En el presente trabajo el modelo bio-inspirado se obtuvo una precisión del 90.2 %, lo que permite inferir que el uso de modelos bio-inspirados incrementa la capacidad del modelo de predecir si una gestante realizará el trabajo de parto mediante cesárea.

Por otro lado, la investigación realizada por el departamento de Obstetricia y Ginecología de la Universidad de Columbia realiza un modelo de predicción del parto vaginal después de una cesárea en el período prematuro, con una muestra de 1295 pacientes, este es procesado por un modelo basado en la regresión logística multivariable, las pruebas realizadas dan como resultado un 80 % de precisión en la predicción [99]. De acuerdo con los resultados del modelo bio-inspirado y los modelos encontrados en la literatura se evidencia que la capacidad adaptativa de los modelos basados en la biología incrementa la precisión de la predicción.

Herramientas tecnológicas como este modelo bio-inspirado que permiten el uso de información demográfica o bioseñales como el EHG, permiten que sea posible implementarlas en los centros médicos que cuentan con tecnología para tomar la bioseñal, pero también permite que en los centros de salud que no tienen estos equipos se pueda hacer uso de esta herramienta con la información de la gestante, para realizar la clasificación y brindar apoyo al diagnóstico preventivo y así disminuir la morbilidad materna.

4.4. IMPLEMENTACIÓN E INTERFAZ GRÁFICA

De acuerdo con la descripción de la metodología en la sección 3.4. es necesario seleccionar los paquetes o módulos que son usados en las dos herramientas tecnológicas usadas para implementar los clasificadores y finalmente validar el modelo.

Weka

Class IBk: es el clasificador *KNN, con este se puede seleccionar* el valor apropiado de K basado en la validación cruzada. Y calcula de distancia de cada dato. *Class LibSVM:* es el clasificador SVM*,* permite realizar experimentos con One-class SVM, Regressing SVM y nu-SVM. Esta librería también tiene como salida estadísticas útiles sobre el clasificador entrenado, por ejemplo, la matriz de confusión, precisión, recuperación, puntaje ROC, entre otros,

Class MultilayerPerceptron: es el clasificador MLP. Para realizar la clasificación usa Backpropagation.

Python

Math: este módulo permite hacer uso de funciones. Algunas de estas son: funciones trigonométricas, logarítmicas, hiperbólicas, entre otras.

Numpy: modulo usado para el uso y manipulación de matrices N-dimensional y con. Permite la integración de bases de datos.

Matplotlib.pyplot: modulo implementado para realizar figuras y gráficos programáticos.

Sklearn: modulo que contiene una amplia biblioteca utilizada para el aprendizaje automático. En se encuentran los algoritmos de clasificación mencionados en la sección 3.3.

MLPClassifier: clasificador perceptrón multicapa. *Kmeans*: Clasificador KNN *keras.models.Sequentiall*: Clasificador basado en el aprendizaje profundo.

La interfaz desarrollada para este proyecto se ilustra en la figura 32, para implementar los clasificadores en la interfaz es necesario crear un módulo de carga de información en el cual se ingresa una base de datos de entrenamiento, que es dividida aleatoriamente, luego se selecciona el clasificador. Finalmente, se desarrolla un módulo para realizar la clasificación de un nuevo paciente. Adicionalmente esta aplicación guarda el clasificador, los resultados obtenidos en el entrenamiento y en la clasificación del nuevo paciente.

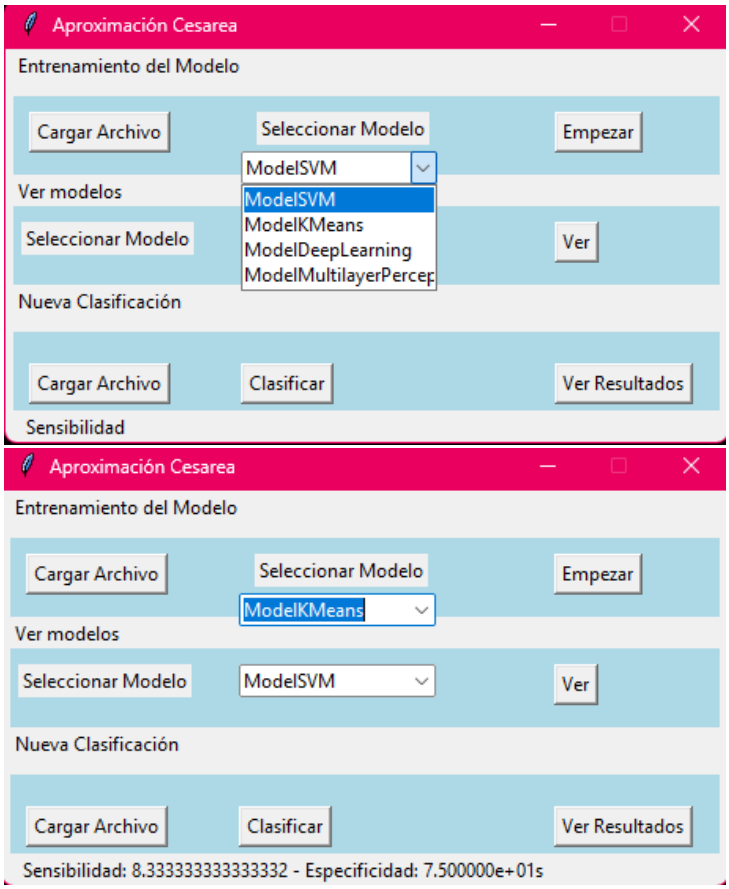


Figura 35. Interfaz de visualización gráfica. (Elaboración propia)

A continuación, se describe el desarrollo de los clasificadores:

El clasificador KNN se implementa mediante la librería sklearn.cluster para hacer uso de kmeans, se configura para que realice 10 ejecuciones con diferentes semillas de centroide, debido a que son 2 tipos de clases y con los resultados previamente

obtenidos se asignan 2 clusters, con un máximo de 300 iteraciones. Esta librería entrena bajo el algoritmo (EM) xpectation–maximization o esperanza-maximización.

```
self.model = KMeans(n_clusters = 2, n_init=100,  verbose=1, algorithm="full")
self.model.fit(X_train.values)
```

El clasificador SVM se implementa haciendo uso de la librería sklearn.svm.SVC, teniendo en cuenta que tipo de kernel que se utilizará en el algoritmo es lineal.

El modelo basado de en redes neuronales se implementa mediante la librería sklearn.neural_network.MLPClassifier, se realiza el código teniendo en cuenta 200 neuronas en las capa oculta con la tasa de aprendizaje adaptativa es decir mantiene constante la tasa la cual es de 0.001, y se evalúa cada dos épocas con el fin de aumentar la validación y finalmente 200 iteraciones.

```
self.model = MLPClassifier(hidden_layer_sizes=200,learning_rate="adaptive")
self.model.fit(X_train, Y_train)
```

Para el desarrollo del clasificador de aprendizaje profundo se implementa usando la librería keras.models.Sequentiall, la cual se configura de la siguiente manera:

El argumento de la primera capa indica que recibe solo datos positivos (None) y son 16 datos de entrada:

```
tf.keras.layers.Input((None,16)),
```

Compuesta por 3 capas con diferente cantidad de neuronas:

```
tf.keras.layers.Dense(100,activation='relu'),
tf.keras.layers.Dropout(0.5), tf.keras.layers.Dense(128,
activation='relu'), tf.keras.layers.Dense(1,
activation='sigmoid')
```

Finalmente se compila el modelo para ser ejecutado, teniendo en cuenta que la división de la base de datos entrenamiento y validación es de manera aleatoria:

```
self.model.compile(optimizer='adam',
loss=tf.keras.losses.binary_crossentropy,
metrics=[tf.keras.metrics.CategoricalAccuracy()])
```

5. CONCLUSIONES

Se establecen las características demográficas más relevantes para realizar la predicción de culminar el embarazo mediante la cesárea, estas son: edad de la gestante, edad gestacional al momento de la grabación, índice de masa corporal (IMC) (antes del embarazo y al momento del registro), posición de la placenta, cantidad de gravidez, paridad (partos anteriores mayores de 22 semanas de gestación), cesáreas previas, y el tipo de parto (parto vaginal y cesárea), estas características demográficas fueron analizadas mediante el cálculo de la correlación de Pearson y la gráfica de dispersión, dando como resultado que la correlación con la salida del modelo es baja.

Los datos como identificación de la paciente, edad gestacional al momento del parto, uso de oxitocina sintética y uso de epidural, no se consideraron en este trabajo por que hacen parte del momento del parto y el modelo tiene como objetivo una predicción previa al nacimiento.

Se caracterizó la señal biológica EHG para la ejecución del modelo bio-inspirado para apoyo al diagnóstico de nacimiento por cesárea, para lo cual se realizó la extracción de las variables: RMS, VAR, MAV, P y MF, utilizando herramientas computacionales, mediante el método de ventaneo, lo que permitió obtener la información que describe el comportamiento eléctrico uterino para la clasificación de gestantes con riesgo de culminar su embarazo por medio de la cesárea.

En esta investigación se implementaron, evaluaron y compararon los clasificadores KNN, MLP, SVM y DP para realizar la clasificación entre parto vaginal y cesárea. Con estos modelos bio-inspirados propuestos se encontró la mayor predicción de partos por cesárea, con el clasificador KNN se obtuvo un 90 % de exactitud con el

procesamiento de las características demográficas y el clasificador MLP obtuvo un 90,9 % de exactitud al procesar la información de la bio-señal EHG.
Los modelos bio-inspirados basados en KNN y MLP son eficientes para la predicción de partos por cesárea en mujeres gestantes en comparación con los modelos bio-inspirados basados en los clasificadores SVM y DP, puesto que a partir del diseño experimental efectuado muestran indicadores de desempeño $S, E, EX < 80\%$.

6.TRABAJO FUTURO

De acuerdo con los resultados obtenidos del trabajo expuesto anteriormente, es posible realizar la aproximación de nacimiento por cesárea a través de un modelo bio-inspirado, a partir de datos demográficos y de las características extraídas de un EHG, esto ofrece una oportunidad de ampliar los conocimientos e investigaciones acerca del uso de la inteligencia artificial como una estrategia para resolver problemas que tienen como base un conjunto de datos no balanceado.

Como trabajo prospectivo, se considera la implementación en centros médicos para poder alimentar la base de datos ya establecida y ampliar el espacio muestral, brindando así a la aplicación y al modelo una mayor capacidad de aprendizaje; en consecuencia, aumenta su adaptabilidad a nueva información ya que se incrementa la diversidad de pacientes y permitiendo así que el conjunto de datos sea balanceado. Como se evidencia en el modelo desarrollado por el Hospital de Mujeres e Infantes de Rhode Island y la facultad de Medicina Alpert de la Universidad de Brown, el cual tiene como objetivo predecir parto por cesárea primaria en pacientes con diabetes mellitus gestacional haciendo uso de inteligencia artificial, los resultados arrojan un 95 % de precisión, el espacio muestral es compuesto de 3.631 pacientes, indicando así que aumenta la capacidad de predicción si se aumenta la base de datos con la que se entrena el modelo [90].

Existen diversas áreas de investigación que pueden ampliar el desarrollo de modelos predictores basados en uso de bio-señales, en especial para preservar la vida del binomio madre-hijo, tal como se indicó al inicio de este documento la cesárea es una de las causas de mortalidad materna, con modelos autónomos es posible detectar otras distocias que se pueden presentar en la gestación.

7. BIBLIOGRAFÍA

[1] OMS, «Objetivo de Desarrollo del Milenio 5: Mejorar la salud materna,» de *ttps://www.who.int/maternal{_}child{_}adolescent/topics/maternal/mdg/es/*, 2008.

[2] OMS, «Mortalidad materna en 2005: estimaciones elaboradas por la OMS,el UNICEF, el UNFPA y el Banco Mundial,» Reporte Anual 2006, OMS, 2008.

[3] U. N. C. Fund, «Maternal, Newborn & Child Survival. Tracking Progress in Maternal, Newborn & Child Survival,» *Countdown to 2015, ISSN 9789280642841,* pp. 1-111, 2008.

[4] G. Tambini, «La mortalidad materna es inaceptable,» 24 abril 2019. [En línea]. Available: https://www3.paho.org/col/index.php?option=com_content&view=article&id=3165:la-mortalidad-materna-es-inaceptable-ginatambini&Itemid=551#:~:text=En%20el%20mundo%20diariamente%20fallece n,el%20parto%20y%20el%20postparto.. [Último acceso: 20 Junio 2021].

[5] D. d. E. y. D. Ministerio de Salud y Protección Social, «MORTALIDAD EN COLOMBIA PERIODO 2020-2021,» MINSALUD, Bogotá, 2021.

[6] O. Carrasco, «Protocolo diagnóstico terapéutico de la preeclampsia grave y eclampsia,» *Revista-Cuadernos (Perú),* vol. 53, nº 1, pp. 71-75, 2008.

[7] C. M. R. Walter Ventura Laverano, «Significado clínico del doppler patológico en la arteria cerebral media en fetos del tercer trimestre,» *Rev Chil Obstet Ginecol,* vol. 57, nº 6, pp. 405-410, 2010.

[8] T. Konechne, «Cesarean,» *Performance Research, doi 10.1080/13528165.2001.10871770,* vol. 6, nº 1, pp. 90-93, 2001.

[9] OPS, «OPS, Declaración de la OMS,» 10 Abril 2015. [En línea]. Available: https://www3.paho.org/hq/index.php?option=com_content&view=article&id=10646:2015-la-cesarea-solo-deberia-realizarse-cuando-es-medicamentenecesaria&Itemid=1926&lang=es. [Último acceso: 20 Junio 2021].

[10] M. d. S. y. P. Social, «Atlas de Variaciones Geográficas en Salud de Colombia 2015- Estudio,» MinSalud, Bogotá, 2015.

[11] DANE, «DIRECCIÓN DE CENSOS Y DEMOGRAFÍA - ESTADÍSTICAS VITALES EEVV,» DANE, Bogotá, 2021.

[12] S.YASUI, «Cesarean section,» *Japanese journal of medical science and biology,* vol. 49, nº 6, pp. 1036-1040.

[13] E. M. A. Verónica Chamy P., «Perfil clínico de embarazadas con preeclampsia y embarazos no complicados,» *Revista chilena de obstetricia y ginecología, ISSN 0717-7526,* vol. 69, nº 5, p. 1, 2004.

[14] M. U. Gustavo Francisco Petro, «GuÍa de trastornos hipertensivos del embarazo Periodo de investigación y desarrollo de la GuÍa (2013),» *Guia Maternidad-Trastornos,* vol. 1, nº 1, p. 1, 2014.

[15] J. Á.-J. Marcos Castillo-Zamora, «Massive proteinuria as prognostic factor for maternal-fetal morbimortality in patients suffering severe preeclampsia: a case report and literature review,» *Revista Colombiana de Obstetricia y Ginecología,* vol. 63, nº 3, pp. 280-285, 2012.

[16] E. M. A. Verónica Chamy P., «Management of hypertension in pregnancy,» *Anales Sis San Navarra, ISSN 1137-6627,* vol. 32, nº 1, p. 1, 2009.

[17] D. N. M.-. M. P. Padrón, «Consideraciones actuales sobre la operación cesárea,» *MEDISAN, ISSN 1029-3019,* vol. 16, nº 10, 2012.

[18] D. T. a. B. M. a. S. M. Far, «Prediction of preterm labor from EHG signals using statistical and non-linear features,» *BMEiCON 2015 - 8th Biomedical Engineering International Conference, ISSN 09123814,* vol. 1, nº 1, p. 75, 2016.

[19] M. W. a. R. J. J. a. O. A. M. a. S. K. a. N. A. J. Moreira, «Predicting hypertensive disorders in high-risk pregnancy using the random forest approach,» *IEEE International Conference on Communications, ISSN 15503607,* vol. 1, nº 1, p. 1, 2017.

[20] M. T. y. T. B. e. I. Syarif, «Algoritmo de redes neuronales para investigar factores previos de preeclampsia en mujeres con hipertensión crónica durante el embarazo en el proceso de parto,» *2018 Simposio internacional de electrónica sobre creación de conocimiento e informática inteligente (IESKCIC),* vol. 1, nº 1, pp. 55-55, 2018.

[21] M. W. a. R. J. J. a. O. A. M. a. R. R. F. a. S. K. Moreira, «A preeclampsia diagnosis approach using Bayesian networks,» *2016 IEEE International Conference on Communications, ICC 2016, ISSN 97814799666461,* vol. 1, nº 1, pp. 1-5, 2016.

[22] D. C. R. a. S. R. S. a. M. K. Nallakaruppan, «Prediction of Women Survival Rate during Pregnancy Using Classification and Clustering,» *2017 World Congress on Computing and Communication Technologies (WCCCT),* vol. 1, nº 1, pp. 231-235, 2017.

[23] C. a. P. M. a. P. H. J. a. N. I. Poncelet, «FYS-3012 PATTERN RECOGNITION Exercises for Week 40,» *FYS-3012 PATTERN RECOGNITION Exercises for Week 40, ISSN 0301-2115,* vol. 86, nº 1, p. 3012, 2012.

[24] C. SCHNAPP , E. SEPÚLVEDA S y J. A. ROBERT S, «Operación cesárea,» *Revista Médica de Clínica Las Condes (RMCLC) ,* vol. 25, nº 6, pp. 987-992, 2014.

[25] P. O. C. Encalada Chérrez Wendy, «Proteinuria y presión arterial de pacientes preeclámpticas severas y eclámpticas: relación con el resultante neonatal y la presencia de complicaciones maternas en pacientes de terapia intensiva del hospital gineco-obstétrico “Enrique C. Sotomayor” de ener,» *Rev. Med. FCMUCSG, ISSN 1390-0218,* vol. 16, nº 4, pp. 273-279, 2010.

[26] A. B. Moncloa, E. Alegría Valdivia y G. Valezuela Rodríguez, «Hypertension in pregnancy,» *Revista Peruana de Ginecología y Obstetricia,* vol. 64, nº 2, 2018.

[27] E. Medina-Pérez, «Diabetes gestacional. Diagnóstico y tratamiento en el primer nivel de atención,» *Medicina interna de México, ISSN 0186-4866,* vol. 33, nº 1, p. 1, 2017.

[28] P. V.-D. Gracia, «Diabetes gestacional: conceptos actuales,» *Ginecología y obstetricia de México, ISSN 0300-9041,* vol. 85, nº 6, p. 1, 2017.

[29] OMS, «Obesidad y sobrepeso,» 09 Junio 2021. [En línea]. Available: https://www.who.int/es/news-room/fact-sheets/detail/obesity-and-overweight. [Último acceso: 20 Junio 2021].

[30] I. B. D. M. A. M. V. Anna Sandra Hernandez Aguado, «¿Qué es el Embarazo?,» 20 Febrero 2018. [En línea]. Available: https://www.clinicbarcelona.org/asistencia/cuida-tu-salud/embarazo-yparto/embarazo. [Último acceso: 20 Junio 2021].

[31] A. R. J. M. Agüero Oscar, «EDAD, PARIDAD, EMBARAZO Y PARTO,» *Revista de Obstetricia y Ginecología de Venezuela, ISSN 0048-7732,* vol. 6, nº 3, 2001.

[32] M. d. S. P. d. Ecuador, «Anomalías de inserción placentaria y vasos sanguíneos fetales,» Octubre 2017. [En línea]. Available: https://www.salud.gob.ec/wpcontent/uploads/2019/02/GPC_anomalias_de_insercion_placentaria_2017.p df. [Último acceso: 20 Junio 2021].

[33] A.D.A.M., «Ubicación de placenta,» 2018. [En línea]. Available: https://www.adam.com/. [Último acceso: 20 Junio 2021].

[34] F. L. Carlos Grandi, «ESTIMACIÓN DE LA EDAD GESTACIONAL,» *Revista del Hospital Materno Infantil Ramón Sardá,* vol. 23, nº 3, pp. 138-143, 2004.

[35] Y. C. G. M. M. M. H. A. C. C. M. Vidala, «Exposición a oxitocina sintetica en el periodo perinatal y resultados en la lactancia materna: un estudio piloto,» *Perinatologia y Reproducción Humana,* vol. 1, nº 30, 2016.

[36] J. C. M. Cevallos Ochoa Andrea Patricia, «"MONITORIZACIÓN ELECTRÓNICA FETAL INTRAPARTO",» Universidad Católica de Loja, Ecuador, 2013.

[37] D. U. O. Ama-haurren Ospitalea, «GUÍA DE MONITORIZACIÓN ELECTRÓNICA FETAL INTRAPARTO,» Hospital Universitario Donostia, 2013.

[38] an Donald Interuniversity School of MedicalUltrasound, «Intrapartum fetal monitoring,» *Progresos de Obstetricia y Ginecología,* vol. 48, nº 4, pp. 207216, 2005.

[39] G. E.-A. Escalante-Gaytán J, «Utilidad de la electrohisterografía como técnica de monitorización uterina en el ámbito clínico: revisión bibliográfica,» *Ginecologia y Obstetricia de Mexico,* vol. 87, nº 1, pp. 280-285, 2019.

[40] J. M. Marshall, «Regulation of activity of uterine smooth muscle.,» *Physiol,* vol. 42, nº 1, pp. 213-235, 1962.

[41] G. P.-B. Y. Y.-L. J. V. A. P. J. G.-C. J. Alberola-Rubio, «Comparison of noninvasive electrohysterographic recording techniques for monitoring uterine dynamics,» *Medical Engineering & Physics,* vol. 35, nº 12, pp. 1736-1743, 2013.

[42] Y. Y.-L. G. P.-B. J. M.-C. J. A.-R. a. A. P. J Garcia-Casado, «Electrohysterography in the diagnosis of preterm birth: a review,» *Physiological Measurement,* vol. 39, nº 2, p. 23, 2018.

[43] A. Cardenas, «Inteligencia Artificial, Métodos Bio-Inspirados: Un Enfoque Funcional Para Las Ciencias De Computación,» *Universidad Tecnólogica de Pereira,* 2012.

[44] D. A. L. S. E. G. V. Diego F. Rocha., «BIO-INSPIRED SYSTEMS AND THEIR APPROACH TO ENGI-NEERING PROBLEM-SOLVING,» *Redes de Ingeniería,* vol. 1, nº 2, pp. 22-29, 2012.

[45] W. S. A. Michael Winkler, «Genetic algorithms and genetic programming: Modern concepts and practical aplications,» *CRC Press Tylor \& Francis Group.,* vol. 1, 2009.

[46] I. G. M. Cristina García Cambronero, «ALGORITMOS DE APRENDIZAJE: KNN & KMEANS,» Universidad Carlos III de Madrid, Madrid.

[47] I. I. y. P. L. Abdelmalik Moujahid, «Tema 5. Clasificadores K-NN,» Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad del País Vasco–Euskal Herriko Unibertsitatea.

[48] M. G. I. Mikel Xabier Uriz Martin, «Aprendizaje de distancias basadas en disimilitudes para el algoritmo de clasificación kNN,» Universidad pública de Navarra, Pamplona, 2015.

[49] P. H. T. M. COVER, «Nearest neighbor pattern classification,» *IEEE Transactions on Information Theory,* vol. 13, pp. 21-27, 1967.

[50] A. Kowalczyk, «Support Vector Machines Succinctly,» *ISSN 978-1-64200150-1, https://www.syncfusion.com/succinctly-free-ebooks/support-vectormachines-succinctly,* nº 1, 2017.

[51] M. {. a. M. {Srivastava}, «A view of Artificial Neural Network,» *2014 International Conference on Advances in Engineering Technology Research (ICAETR - 2014),* pp. 1-3, 2014.

[52] M. B. A. {. a. T. K. A. {Rahman}, «Short Term Load Forecasting Using an Artificial Neural Network Trained by Artificial Immune System Learning Algorithm,» *2010 12th International Conference on Computer Modelling and Simulation,* pp. 408-413, 2010.

[53] G. P. J. {. a. C. {. a. F. S. {Gouws}, «ANN-DT: an algorithm for extraction of decision trees from artificial neural networks,» *IEEE Transactions on Neural Networks,* vol. 10, nº 6, pp. 1392-1401.

[54] M. B. Daniel I. MORARIU, «THE WEKA MULTILAYER PERCEPTRON CLASSIFIER,» vol. 1, nº 1, p. 1, 2017.

[55] . J.-H. Wang y S.-J. Lou, «Predicting the success rate of natural spontaneous delivery through deep learning,» *2019 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW),* 2019.

[56] J. Z. A. J. H. Laith Alzubaidi, «Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,» *Journal of Big Data volume,* vol. 8, nº 53, 2021.

[57] D. L. Guzmán, M. Carreno y L. H. Camargo, «Identifying Sleep Apnoea and Hypopnoea Episodes From Respiratory Polygraphy Signals,» *Revista Colombiana de Psiquiatría,* vol. 46, nº 2, pp. 88-94, 2017.

[58] P. C. C. C. A. Marcano cedeno, «predecir la evolución de pacientes con Daño Cerebral Adquirido,» vol. 6, pp. 145--148.

[59] M. W. a. R. J. J. a. K. N. a. N. J. a. W. I. Moreira, «Performance Assessment of Decision Tree-Based Predictive Classifiers for Risk Pregnancy Care,» *2017 IEEE Global Communications Conference, GLOBECOM 2017 - Proceedings,* pp. 1-6, 2018.

[60] E. a. M. J. a. H. M. Nodelman, «Using artificial intelligence to predict spontaneous preterm delivery,» *American Journal of Obstetrics and Gynecology, ISSN 00029378,* vol. 222, nº 1, p. 1, 2020.

[61] J. E. Fonseca, «Validation of a predictive model for successful vaginal birth after cesarean section,» *Colombia Medica, ISSN 16579534,* vol. 50, nº 1, pp. 13-21, 2019.

[62] P. Cient, «Elaboración y validación de un modelo predictivo de riesgo de cesárea en gestantes con inducción del trabajo de parto - Publicaciones Científicas,»

https://www.enfermeriadeciudadreal.com/articulo{_}imprimir.asp?idarticulo=430{\&}accion=, nº 7, pp. 1-13, 2020.

[63] R. S. P. SOLANO, «DISEÑO DE UN MODELO PARA ASISTENCIA MÉDICA Y PREDICCIÓN DE PREECLAMPSIA EN MUJERES EMBARAZADAS,» FUNDACIÓN UNIVERSIDAD DEL NORTE, BARRANQUILLA, 2015.

[64] M. A. G. Riotti, «Modelos explicables de Aprendizaje Automático para la predicción de complicaciones perinatales en mujeres embarazadas con diabetes,» Universidad de Valparaíso , Chile, 2021.

[65] A. a. R. C. W. G. a. P. A. T. a. J. I. a. G. A. Petrozziello, «Multimodal Convolutional Neural Networks to Detect Fetal Compromise During Labor and Delivery,» *IEEE Access, ISSN 2169-3536,* vol. 7, nº 1, pp. 112026--112036, 2019.

[66] G. A. a. H. N. a. E. R. a. S. D. M. a. M. S. J. Macones, «Predicting outcomes of trials of labor in women attempting vaginal birth after cesarean delivery: A comparison of multivariate methods with neural networks,» *American Journal of Obstetrics and Gynecology,* vol. 184, nº 3, pp. 409--413, 2001.

[67] P. a. I. I. a. H. A. a. D. C. Fergus, «Advanced artificial neural network classification for detecting preterm births using EHG records,» *Neurocomputing, ISSN 18728286,* vol. 188, pp. 42-49, 2016.

[68] V. a. R. C. a. C. R. a. S. A. a. V. F. a. S. L. a. I. G. a. M. P. P. a. D. M. a. B. A. a. M. S. Romeo, «Machine learning analysis of MRI-derived texture features to predict placenta accreta spectrum in patients with placenta previa,» *Magnetic Resonance Imaging,* vol. 64, nº April, pp. 71--76, 2019.

[69] H. a. W. Sufriyana, «Artificial intelligence-assisted prediction of preeclampsia: Development and external validation of a nationwide health insurance dataset

of the BPJS Kesehatan in Indonesia,» *E BioMedicine, ISSN 23523964,* vol. 54, 2020.

[70] M. A. de Felipe, B.-P. C. y G.-C. J. , «Sistema de predicción del éxito de la inducción farmacológica basado en electrohisterografía,» *Libro de Actas del XXXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica,* vol. 1, nº 1, pp. 335-338, 2018.

[71] F. G. ALIÑO, «DESAROLLO DE UN SISTEMA DE PREDICCIÓN DEL PARTO PREMATURO EN BASE A LA ELECTROHISTEROGRAFÍA MEDIANTE APRENDIZAJE DE DATOS DESBALANCEADOS,» Escuela Técnica Superor Ingeniería Industrial Valencia, Valencia, España, 2020.

[72] C. S. Maria A, «Diseños Cuasiexperimentales, Facultad de Salud Publica,» *Universidad de Antioquia,* vol. 4, 2003.

[73] M. L. f. C. Physiology, «PhysioNet,» 2022. [En línea]. Available: https://physionet.org/. [Último acceso: 30 09 2022].

[74] A. Goldberger, L. Amaral y L. Glass, «Components of a new research resource for complex physiologic signals,» *PhysioNet,* p. 101, 2000.

[75] T. M. Nair, «Statistical and artificial neural network-based analysis to understand complexity and heterogeneity in preeclampsia,» *Computational Biology and Chemistry,* vol. 75, pp. 222-230, 2018.

[76] R. E. Behrman y A. Stith Butler, «Preterm Birth: Causes, Consequences, and Prevention,» *the National Academies Press,* p. 790, 2007.

[77] C. D.-B. Manuel Gómez-Gómez y M. Aceves-Gómez, «Classification of the newborns,» *Revista Mexicana de pediatría,* vol. 79, nº 1, pp. 32-39, 2012.

[78] R. R. R. M. P. R. a. L. P. C. I. R. García, «"Correlación: no toda correlación implica causalidad"».

[79] Rivas-Ruiz, Rodolfo; Roy-García, Ivonne; Pérez-Rodríguez, Marcela; Berea, Ricardo; MorenoPalacios, Jorge; Moreno-Noguez, Moisés; Palacios-Cruz, Lino; Ureña-Wong, Kingston Rodolfo Pertinencia e impertinencia de los gráficos en la investigación clínica Revista alergia México, vol. 67, núm. 4, 2020, Octubre-Diciembre, pp. 381-396 Colegio Mexicano de Inmunología Clínica y Alergia, A.C. DOI: https://doi.org/10.29262/ram.v67i4.854

[80] Hernández Lalinde, Juan Diego; Espinosa Castro, Franklin; Rodríguez, Johel E; Chacón Rangel, José Gerardo; Toloza Sierra, Cristian Andrés; Arenas Torrado, Marlly Karina; Carrillo Sierra, Sandra Milena; Bermúdez Pirela, Valmore José Sobre el uso adecuado del coeficiente de correlación de Pearson: definición, propiedades y suposiciones Archivos Venezolanos de Farmacología y Terapéutica, vol. 37, núm. 5, 2018 Sociedad Venezolana de Farmacología Clínica y Terapéutica, Venezuela Disponible en: https://www.redalyc.org/articulo.oa?id=55963207025

[81] Crombet-Grillet, Sandra; Abalos-Rodríguez, Arelis; Zamora-Matamoros, Larisa Determinación de los parámetros ambientales de mayor incidencia en las aguas residuales de la comunidad universitaria Antonio Maceo Revista Cubana de Química, vol. 31, núm. 1, 2019, Enero-Abril, pp. 137-154 Universidad de Oriente

[82] Natus, «Natus Medical Incorporated,» 2022. [En línea]. Available: https://partners.natus.com/asset/resource/file/neuro/asset/2020-05/028620D%20Grass%20Disposable%20Electrodes%20Flyer_ES%20Final.pdf. [Último acceso: 14 11 2022].

[83] K. Kammerer, R. Pryss y B. Hoppenstedt, «Process-Driven and Flow-Based Processing of Industrial Sensor Data,» *Sensors (Basel),* vol. 20, nº 18, p. 5245, 2020.

[84] Q. Q. a. X. Z. b. Y. A. a. J. P. a. L. Y. a. D. Z. c. Dongmei Hao a, «Application of decision tree in determining the importance of surface electrohysterography,» *Biocybernetics and Biomedical Engineering,* vol. 39, nº 3, pp. 806-813, 2019.

[85] L. L. C. T. P. L. M. M. O. Lic. Rosa María Martínez Ortega, «EL COEFICIENTE DE CORRELACION DE LOS RANGOS DE SPEARMAN CARACTERIZACION,» *Revista Habanera de Ciencias Médicas,* vol. 8, nº 2, 2009.

[86] O. A. L. E. M. B. 1. M. S. N. F. M. D. F. O. T. A. G. G. María De la Calle FM.1, «SOBREPESO Y OBESIDAD REGESTACIONAL COMO FACTOR DE RIESGO DE CESÁREA Y COMPLICACIONES PERINATALES,» *Revista chilena de obstetricia y ginecología,* vol. 74, nº 4, pp. 233-238, 2009.

[87] M. d. C. P. H. M. d. C. G. M. F. C. S. L. C. C. F. Á. V. S. R. T. C. y. L. J. M. C. Juan Jesús Fernández Alba, «Maternal overweight and obesity as independent risk factors for cesarean delivery,» *Nutrición Hospitalaria,* vol. 33, nº 6, 2016.

[88] A. P. B. G. A.-B. M. Ö. T. C. B. L. K. L. A. Jonathan Bearak, «Unintended pregnancy and abortion by income, region, and the legal status of abortion: estimates from a comprehensive model for 1990–2019,» *The Lancet,* vol. 8, nº 9, 2020.

[89] S. S. L. R. a. C. V. Elena Prada, Embarazo no deseado y aborto inducido en Colombia: Causas y Consecuencias, Colombia: GuttMacher Institute, 2011.

[90] J. G. T. P. M. C. A. A. G. Blanca Heras Pérez, «Maternal age as an obstetric risk factor. Perinatal outcomes in pregnant women of advanced age,» *Progresos de Obstetricia y Ginecología,* vol. 54, nº 11, 2011.

[91] B. B. T. Poliana Cristina Ferrari Schiavetto, «Índice de masa corporal en embarazadas en la unidad de salud de la familia,» *Enfermería Global,* vol. 17, nº 52, 2018.

[92] D. R. A. G. M. M. L. C. MsC. Rosa D. Gutiérrez de Alarcón, «Índice de masa corporal como factor de riesgo en parto por cesárea.,» *REVISTA DE INVESTIGACIÓN ESTADÍSTICA, Revista Electrónica Universidad Nacional de Trujillo,* vol. 1, nº 7, pp. 69-77, 2018.

[93] C. A. R. T. Jorge A. Carvajal Cabrera, Manual Obstetricia y Ginecología., Chile: Escuela de Medicina - Pontificia Universidad Católica de Chile, 2017.

[94] Hospital Sant Joan de Déu - Universitat de Barcelona., CONTROL GESTACIONAL EN GESTANTES CON CESÁREA ANTERIOR.

[95] H. C. |. H. S. J. d. D. |. U. d. Barcelona, «CONTROL GESTACIONAL EN GESTANTES CON CESÁREA ANTERIOR,» *medicinafetalbarcelona,* p. 11, 2019.

[96] J. Panduro-Barón, E. Panduro-Moore y J. Pérez-Molina, «Changes in fetal static and their predisposition to risk factors throughout of pregnancy,» *Ginecología y obstetricia de México,* vol. 85, nº 8, 2017.

[97] S. E. y. S. A. V. P. L. D.-A. Joaquín Ruiz-Sánchez, «Cesarean: Trends and outcomes,» *Perinatología y reproducción humana,* vol. 28, nº 1, 2014.

[98] K. H. R. F. Muhammad Ehsan Ul Haq, «Use of Machine learning models for predicting and improving maternal and child health indicators,» *IEEE-SEM,* vol. 9, nº 8, pp. 1-12, 2021.

[99] M. B. M. S. a. P. A. Somayeh Mohammadi Far, «Prediction of Preterm Delivery from Unbalanced EHG Database,» *Sensors (Basel),* vol. 22, nº 4, p. 1507, 2022.

[100] Y. Y.-L. J. G.-C. ,. J. A.-R. A. P. a. G. P.-B. Carlos Benalcazar-Parra, «Prediction of Labor Induction Success from the Uterine Electrohysterogram,» *Journal of Sensors,* vol. 2019, nº 1, p. 12, 2019.

[101] E. L. A. Rodr´ıguez, «Early Prediction of Extreme Maternal Morbidity Using Machine Learning,» TECHNOLOGICAL UNIVERSITY OF BOLIVAR, Cartagena, 2017.

[102] S. de Miguel Manso, «CLINICAL VALIDATION OF AN EXPERIMENTAL MODEL FOR PREDICTING THE MODE OF LABOR AFTER INDUCTION,» Universidad de Valladolid, 2020.

[103] R. M. G. L. V. A. D. Sebastian Z., «Prediction model for primary cesarean delivery in patients with gestational diabetes mellitus,» *American Journal of Obstetrics and Gynecology,* vol. 226, nº 1, p. S44, 2022.

[104] C. V. A. A. G. C. G.-B. Anne H. Mardy, «A prediction model of vaginal birth after cesarean in the preterm period,» *American Journal of Obstetrics and Gynecology,* vol. 215, nº 4, p. 513, 2016.

8. ANEXOS

Anexo 1.

J. Pacheco Fandiño, E. Camargo Casallas, A. Gaona Barrera y V. Fandiño Osorio, “Approach to the diagnostic of cesarean birth using bio-inspired models”, “Aproximación al diagnóstico del parto por cesárea mediante modelos bioinspirados”, XVII Congreso Internacional De Electrónica, Control Y Telecomunicaciones, noviembre 2022, ISBN: 978-958-44-5254-2 Vol. 13.

XVII CONGRESO INTERNACIONAL DE ELECTRÓNICA

CONTROL Y TELECOMUNICACIONES

"Ciencia, Tecnología e Innovación avanzadas para transitar hacia un nuevo sistema sociotécnico: Transformación Social Sostenible"

Libro de memorias

ISBN: 978-958-44-5254-2 Vol. 13

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Proyectos Curriculares de Tecnología en Electrónica Industrial, Ingeniería en Control y Automatización e Ingeniería en Telecomunicaciones

Bogotá D.C. (Colombia) 24 y 25 de noviembre de 2022

UNIVERSIDAD DISTRITAL
FRANCISCO JOSE DE CALDAS

XVII CONGRESO INTERNACIONAL DE ELECTRÓNICA, CONTROL Y TELECOMUNICACIONES

CIENCIA, TECNOLOGÍA E INNOVACIÓN AVANZADAS PARA TRANSITAR HACIA UN NUEVO SISTEMA SOCIOTÉCNICO: TRANSFORMACIÓN SOCIAL SOSTENIBLE

24 y 25 de noviembre de 2022
Universidad Distrital Francisco José de Caldas
Facultad Tecnológica
Bogotá - Colombia

BIOINGENIERÍA Y SALUD PREVENTIVAS

UNIVERSIDAD DISTRITAL FRANCISCO JOSE DE CALDAS

XVII CONGRESO INTERNACIONAL DE ELECTRÓNICA, CONTROL Y TELECOMUNICACIONES

CIENCIA, TECNOLOGÍA E INNOVACIÓN AVANZADAS PARA TRANSITAR HACIA UN NUEVO SISTEMA SOCIOTÉCNICO: TRANSFORMACIÓN SOCIAL SOSTENIBLE

24 y 25 de noviembre de 2022
Universidad Distrital Francisco José de Caldas
Facultad Tecnológica
Bogotá - Colombia

Approach to the diagnostic of cesarean birth using bio-inspired models

Aproximación al diagnóstico del parto por cesárea mediante modelos bio-inspirados

Jennifer V. Pacheco Fandiño[1], Luz H. Camargo Casallas[2], Andrés E. Gaona Barrera[3], Vilma C. Fandiño Osorio[4]

Abstract

The cesarian section-related maternal mortality rate in Colombia was 46.4 % in 2021, a rate that health professionals have been attempting to reduce by monitoring women. However, the high volume of information and of patients makes it complex to take into account all the possible symptomatology. This article develops a bio-inspired model for the classification of cesarean deliveries, based on demographic information and electrohystereographic biosignals (EHG) of the mother-child binomial. The classifiers implemented are K nearest neighbors (KNN), multilayer perceptron (MLP), support vector machines (SVM), and deep learning.

[1] Electronic Engineering. Research Group Ingeniería y Nanotecnología para la Vida (INVID), Universidad Distrital Francisco José de Caldas – Facultad de Ingeniería, Bogotá, Colombia. E-mail:: jvpachecof@correo.udistrital.edu.co ORCID: https://orcid.org/0000-0003-3624-662X

[2] Doctoral in Engineering, Master in Biomedical Engineering, and Specialist in Bio-Engineering. Research Group Ingeniería y Nanotecnología para la Vida (INVID), Universidad Distrital Francisco José de Caldas – Facultad de Ingeniería, Bogotá, Colombia. Contact: lhcamargoc@udistrital.edu.co ORCID: https://orcid.org/0000-0002-3416-018X

[3] Master in Engineering. Research Group Laboratorio de Automática e Inteligencia Computacional (LAMIC), Universidad Distrital Francisco José de Caldas – Facultad de Ingeniería, Bogotá, Colombia. Contact: aegaonab@udistrital.edu.co ORCID: https://orcid.org/0000-0002-3966-5646

[4] Master in Nursing with emphasis on maternal perinatal. Universidad de Ciencias Aplicadas y Ambientales (UDCA) - Programa de Enfermería, Bogotá, Colombia. Contact: vfandino@udca.edu.co ORCID: https://orcid.org/0000-0001-9088-5866

The following models provide the best analysis of demographic data: KNN with a sensitivity (S) of 100 % and a specificity (ES) of more than 80 %, together with SVM with an S of 75 % and an ES of 83.3 %. The best performers in EHG analysis are MLP with S of 82.3 % and ES of 85.7 %, followed, in order, by deep learning with S of 72.8 %. This tool supports the early detection of cesarean births, taking into account the mother's history and the behavior of the fetus.

Keywords: Artificial Intelligence, Bio-Inspired, Birth, Cesarean section, Classification, Electrohystereography (EHG).

Resumen

En Colombia la tasa de mortalidad materna por cesárea en 2021 fue el 46,4 %, para disminuirlo, profesionales de la salud dan seguimiento a las gestantes, sin embargo, la densidad de información y el volumen de pacientes hace complejo tener en cuenta la sintomatología. Este artículo desarrolla un modelo bio-inspirado para la clasificación de parto por cesárea, basado en información demográfica y la bioseñal Electrohistereograma (EHG) del binomio madre-hijo. Los clasificadores son k vecinos más cercanos (KNN), perceptrón multicapa (MLP), máquinas de vectores de soporte (SVM) y aprendizaje profundo.

El mejor desempeño del análisis de datos demográficos se obtiene con: KNN con sensibilidad (S) del 100 % y especificidad (ES) de más del 80 % junto a SVM con S del 75 % y ES del 83,3 %. El mayor rendimiento en el análisis del EHG son MLP con S de 82.3 % y ES de 85.7 %. Esta herramienta brinda apoyo a la detección temprana de nacimientos por cesárea, teniendo en cuenta los antecedentes de la gestante y el comportamiento del feto.

Palabras clave: Inteligencia Artificial, Bio-Inspirado, Parto, Cesárea, Clasificación, Electrohistereografía (EHG).

1. Introducción

El Objetivo N°5 de Desarrollo del Milenio de las Naciones Unidas busca reducir la mortalidad materna [1], los riesgos durante la cesárea representan las principales causas de estas muertes. En particular se propone que en Colombia la tasa de mortalidad ideal por esta causa no puede superar el 15 % para el 2015 [2], a pesar de los esfuerzos, la prevalencia de muerte por cesáreas en este país es del 46,4 % en 2021 [3].

La inteligencia artificial es útil para respaldar los métodos actuales para la detección temprana de este método de parto; por ejemplo, los modelos desarrollados sobre el manejo de la información demográfica para la predicción de partos prematuros [4] o la clasificación de gestantes que pueden presentar morbilidad hipertensiva presentan una precisión superior al 80 % [5]. Otra alternativa es el análisis de bio-señales, por ejemplo: con redes neuronales convolucionales es posible detectar hipoxia fetal con una precisión del 65 % usando la bio-señal de la cardiotocografía [6], sin embargo, estos modelos requieren una gran base de datos para realizar la validación cruzada y obtener alto rendimiento en la predicción.

La información necesaria para tomar la decisión de realizar el alumbramiento mediante una cesárea se toma durante la emergencia. Por lo tanto, el enfoque principal de este artículo es desarrollar un modelo bio-inspirado utilizando técnicas de inteligencia computacional como KNN, MLP, SVM y aprendizaje profundo, para respaldar el diagnóstico previo de parto por cesárea, basado en información demográfica y bio- señal EHG de la mujer embarazada, es decir, el diagnostico en las etapas tempranas de la gestación para que de esta manera el

sistema de salud pueda desarrollar un plan de gestación de acuerdo con la conclusión del modelo. Los resultados de la clasificación entre parto vaginal y cesárea muestran la efectividad del modelo porque se logra una sensibilidad de hasta el 85 %.

El trabajo está organizado de la siguiente manera: la sección 2 presenta los conceptos para tener en cuenta: cesárea, algoritmos de clasificación e indicadores de desempeño, la sección 3 detalla el modelo bio-inspirado propuesto, la sección 4 presenta el proceso a llevar a cabo para la implementación del modelo bio-inspirado, la sección 5. se presentan los resultados obtenidos durante los ensayos y su análisis. Finalmente, en la sección 6 se describen las conclusiones del trabajo.

2. Marco Teórico

Cesárea: el parto por cesárea es una intervención quirúrgica, que tiene como finalidad la extracción del neonato a través de una incisión realizada desde el abdomen de la gestante [7] [8].

Electrohisterografía: esta herramienta de diagnóstico es una técnica no invasiva que mide la actividad eléctrica de las contracciones empleando electrodos de superficie ubicados en el abdomen de la gestante [9].

El k-vecino más cercano (KNN): clasifica los datos calculando la distancia entre cada dato sin etiquetar y todos los demás puntos etiquetados con la distancia euclidiana [10]. La clasificación se obtiene con la ecuación (1):

$$C(x_i) = \text{argmax}_k \sum_{X_j \in KNN} C\left(X_j, Y_k\right) \qquad (1)$$

donde $x_i$es el dato de prueba, X_j es uno de los k vecinos más cercanos en el conjunto de entrenamiento, $C(X_j, Y_k)$ indica si X_j pertenece a la clase Y_k.

Perceptrón multicapa: MLP es una red neuronal con una capa de entrada, una capa de salida y N capas ocultas [11]. Las conexiones que existen entre las capas tienen asignado un peso, la matriz que los contienen es calculada nodo por nodo mediante la ecuación (2).

$$h_{ij} = f(\sum_{i=1}^{n} w_{ij} x_i + b_j) \qquad (2)$$

donde h_{ij}es el j-ésimo nodo de h1 yw_{ij} es el peso asociado.

Máquinas de vectores de soporte: este clasificador se basa en la búsqueda de un hiperplano que separe efectivamente las n clases diferentes de datos. Los Kernel son funciones que transforman un espacio de baja dimensión en un espacio de mayor dimensión, algunos de estos son kernel lineal y polinómicos [12], cuyo cálculo se muestra en la ecuación (3).

$$K_L(x_i, x_{i'}) = \langle x_i \cdot x_{i'} \rangle \, , K_P(x_i, x_{i'}) = (\langle x_i \cdot x_{i'} \rangle)^d \qquad (3)$$

donde $x_i \, , x_{i'} \in X$ conjunto de entrada y d es el grado de la función polinomial.

Aprendizaje profundo: este método de aprendizaje busca imitar las funciones cerebrales empleando múltiples capas de unidades de procesamiento no lineales que extraen y transforman los datos de entrada. El proceso se itera el número de veces necesario para obtener la precisión determinada y pasar a la siguiente capa [13]. Permite la combinación de métodos de aprendizaje automáticos.

Indicadores de rendimiento: el significado de estos indicadores es calcular la capacidad de generalización de un modelo sobre datos futuros no vistos. Con base en esto, las ecuaciones (4-6) muestran tres indicadores de desempeño, sensibilidad (S), especificidad (E)

y exactitud (EX), considerando que verdaderos positivos (VP), falsos positivos (FP), falsos negativos (FN), verdaderos negativos (VN). [14]:

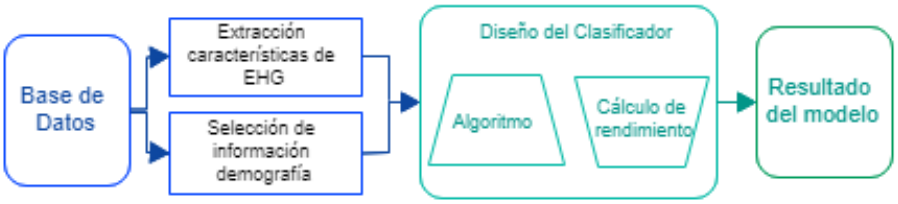
$$S = \frac{VP}{VP+FN} \quad (4) \qquad \mathrm{E} = \frac{VN}{FP+VN} \quad (5) \qquad \mathrm{EX} = \frac{VP-VN}{VP+VN+FP+FN} \quad (6)$$

3. Modelo bio-inspirado propuesto

El modelo bio-inspirado de predicción de parto por cesárea se muestra en la Figura 1, en primer lugar, el conjunto de datos se compone de información demográfica y la bio-señal de EHG, el cual es preprocesado con el objetivo de seleccionar las entradas al bloque clasificador. A continuación, se diseña y desarrolla el clasificador teniendo en cuenta la evaluación del desempeño. Finalmente, el modelo presenta el resultado de la aproximación. Este modelo se pretende implementar durante los controles realizados durante los primeros meses por el obstetra y, de acuerdo a lo observado por los profesionales y al resultado de este modelo bio-

inspirado, es posible proponer un plan de embarazo y plan de parto acorde, con el fin de preservar la vida del binomio madre-hijo.

Figura 1. Diagrama de bloques del modelo bioinspirado.



Fuente: Elaborado por los autores.

Base de datos: la información para llevar a cabo este proyecto extraída de la página Physionet, la cual es un repositorio de datos de investigación médica de libre acceso, administrado por el Laboratorio de Fisiología Computacional del MIT, la base de datos es titulada "Icelandic 16-electrode Electrohysterogram Database", consta de 122 electrohisterogramas (EHG) de mujeres embarazadas, con edades entre 19 y 39 años, los registros y los datos se toman antes y durante el trabajo de parto. Las grabaciones se realizaron en Islandia entre el 2008 y 2010, esta investigación se llevó a cabo entre el Centro de Atención Primaria de Salud de Akureyri, el Hospital de Akureyri y el Hospital Universitario Landspitali. Los registros son mediciones realizadas en pacientes con sospecha de trabajo de parto y presentes en las salas de parto [15].

Cada registro contiene un archivo de cabecera con la información demográfica de la gestante, a continuación se lista la información contenida: ID de participante, número de registro, tipo de registro (parto, embarazo), edad de la gestante (años), edad gestacional (semanas), índice de masa corporal (IMC) (antes del embarazo y al momento del registro),

posición de la placenta, cantidad de gravidez, paridad (partos anteriores mayores a 22 semanas de gestación) y cesáreas previas (Sí, No), edad gestacional al parto (semanas/días), comentarios para el registro, comentarios para el parto y por último el modo de parto (parto vaginal y cesárea). El bio-señal EHG se describe en la sección 3.3.

Preprocesamiento de datos demográficos: la selección de la información demográficos es guiada por la estructura y el tamaño de la base de datos [16]. Por lo tanto, los criterios de inclusión son los siguientes: edad gestacional entre 32 a 38 semanas y un máximo de dos registros por paciente, esto se debe a que este modelo trata de apoyar el diagnóstico preventivo, por lo que, si la edad gestacional es mayor a 38 semanas, el parto está próximo a suceder, de lo contrario si la edad gestacional es menor a 32 semanas, el feto es inmaduro. En consecuencia, el registro no puede mostrar el comportamiento del binomio madre-hijo [17] [18]. Se omiten los datos que forman parte de la identificación del paciente y los comentarios que no contienen información médica o demográfica.

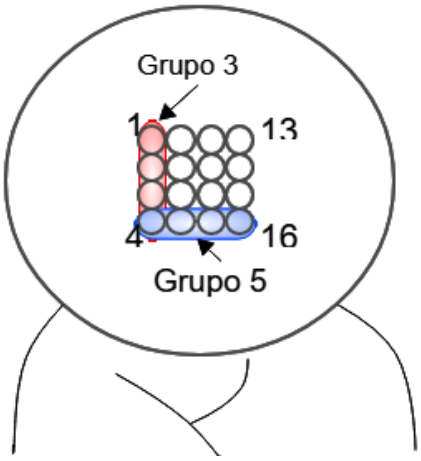
Por lo anterior, los datos de entrada se componen de: edad de la gestante, edad gestacional, índice de masa corporal (IMC) (antes del embarazo y al momento del registro), posición de la placenta, cantidad de gravidez, paridad (partos anteriores mayores de 22 semanas de gestación) y cesáreas previas, edad gestacional al momento del parto, y los datos de salida son el modo de parto (parto vaginal y cesárea). El modo de parto esta correlacionado con cada

una de las entradas, por ejemplo, la posición de la placenta, si es previa, el parto debe ser con cesárea porque la placenta obstruye el cuello uterino [19].

Preprocesamiento EHG: esta señal contiene los datos de los 16 electrodos monopolares, la figura 2 muestra la distribución de los electrodos en el abdomen de la gestante.

Para incorporar la bio-señal al modelo, es necesario realizar un pre-procesamiento; para lo cual se divide el EHG mediante el método de ventana [20], este proceso toma un pequeño subconjunto de un conjunto de datos más grande con un filtro tipo ventana rectangular, la amplitud usada en este artículo es del 30 %. esto implica truncar el conjunto de datos antes y después de esta ventana.

Figura 2. Ubicación de electrodos (grupo 3,4 and 5) [15].



Fuente: Tomado de [15].

Las características en el dominio del tiempo y la frecuencia son calculadas para cada subventana, estas se muestran en la Tabla 1 [21].

Tabla 1. Características en el dominio de tiempo y frecuencia de la Bio-señal EHG [21].

Raíz media cuadrática	**Valor Medio Absoluto**	**Varianza**
$\boldsymbol{RMS = \sqrt{\left(\frac{1}{N}\sum_{i=1}^{N} x(i)^2\right)}}$	$\boldsymbol{MAV = \frac{1}{N}\sum_{i=1}^{N}\lvert x(i)\rvert}$	$\boldsymbol{VAR = \frac{\sum_{i=1}^{N}(x(i)-\mu)^2}{N}}$
Frecuencia Media		Potencia Media
$\boldsymbol{MF = i_m \frac{f_s}{N}\sum_{i=1}^{i_m} P(i) = \sum_{i=i_m}^{i=N} P(i)}$		$P = \lim_{L\to\infty}\left[\frac{1}{2L}\int_{-L}^{L}\lvert x(t)\rvert^2\right]dt$

Donde $x(i)$ es el segmento EHG, $i = 1,2,\ldots,N$ donde N es la longitud de $x(i)$, $\mu = \sum_{i=1}^{N}\frac{x(i)}{N}$ es la media de la $x(i)$, y $P(i)$ representa la potencia espectral de $x(i)$.

Para definir el conjunto de datos de entrada del modelo se tienen en cuenta cinco grupos de parámetros: el grupo 1 tiene datos demográficos de la mujer embarazada como la edad, el número de partos, el número de embarazos, el IMC, las cesáreas anteriores y la ubicación de la placenta, estos son mencionados previamente. El grupo 2 contiene un promedio de las variables de la bio-señal EHG (Tabla 1) y datos demográficos. El grupo 3 tiene variables de la bio-señal EHG (Tabla 1) donde los datos son un promedio de las variables de los electrodos en forma vertical (Figura 2, electrodos rojos). El grupo 4 tiene la raíz cuadrada media de los 16

electrodos. El grupo 5 tiene variables de la bio-señal, tomando un promedio de las variables de los electrodos de forma horizontal (Figura 2, electrodos azules).

4. Metodología

Para comparar los resultados entre modelos bio-inspirados, los clasificadores se prueban con cinco grupos de parámetros explicados en la sección 3, que se dividen en un 70 % para entrenamiento y un 30 % para validación cruzada. Se utilizan dos herramientas computacionales WEKA y Python. Finalmente, con los resultados es posible calcular los indicadores de desempeño mostrados en la sección 2.

Para la clasificación con KNN, el número de vecinos comienza en 1 y aumenta según el rendimiento y el tiempo de procesamiento, la función de distancia implementada es la euclidiana (sección 2) y las iteraciones comienzan en 50 y aumentan si es necesario. Por otro lado, el MLP comienza en 1 capa oculta y aumenta según los resultados, además el número de neuronas por capa varía según el rendimiento y el tiempo de procesamiento, la función de activación para la capa de entrada elegida es la función lineal y para las capas ocultas es la función sigmoidal, además para usar este clasificador se normalizan los valores de entrada.

El clasificador SVM encuentra un hiperplano que separa los datos, se varía los tipos de kernel entre lineales y polinomiales. Finalmente, el aprendizaje profundo varía el tamaño del lote de entrenamiento, es decir, hay N muestras para entrenamiento y el tamaño del lote se

establece en 5 o 10. El segundo parámetro a ajustar es el número de épocas por procesamiento entre 50, 100 y 1000.

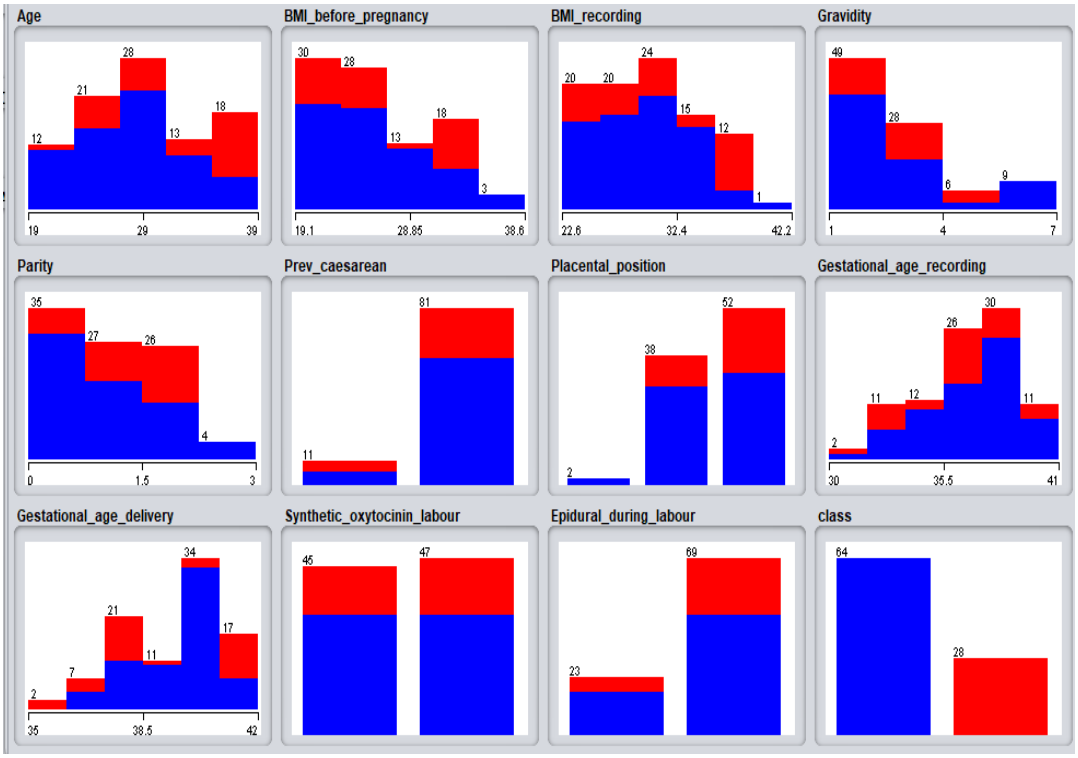
Para la implementación clínica del modelo bio-inspirado, es necesario desarrollar una interfaz gráfica de visualización personalizada para computador, esta es diseñada en el entorno de programación gráfica de Visual Studio y el lenguaje Python, la interfaz debe contemplar los clasificadores antes mencionados y la visualización de los indicadores de rendimiento en las etapas de entrenamiento y validación.

5. Resultados y discusiones

El modelo bio-inspirado construido permite clasificar si el método parto podría ser vaginal o por cesárea. La selección de los datos demográficos se realiza de acuerdo con los criterios expuestos en el apartado 3, el resultado se ilustra en la Figura 3, donde se observa la

distribución que tiene cada una de las características, teniendo en cuenta las dos clases identificadas como Vaginal con color azul y cesárea con color rojo.

Figura 3. Distribución de base de datos de acuerdo a tipo de parto. En azul parto vaginal y en rojo parto por cesárea.

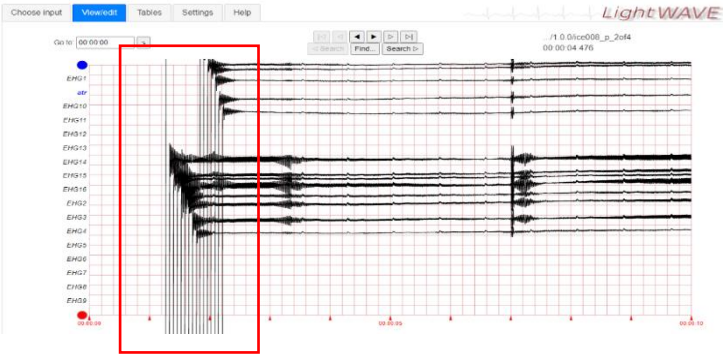


Fuente: Elaborado en el programa WEKA por los autores.

El preprocesamiento de la señal requiere implementar una técnica de filtración, debido a que esto disminuye el ruido presente durante la captura de la bio-señal EHG, como se muestra en la Figura 5, este ruido está presente al inicio y al final de cada grabación y según

los comentarios en la base de datos, esto se debe a los movimientos de la gestante e interferencias de electrodos mal conectados.

Figura 4. Ruido presente en la Bio-señal EHG, indicado con el recuadro rojo

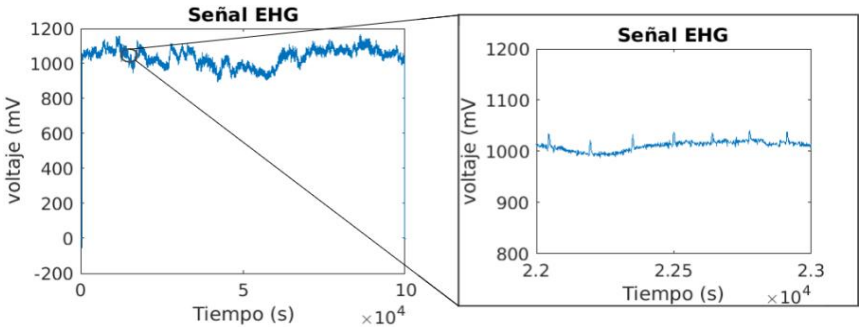


Fuente: tomado del visor de señales de [22].

Seguido del paso de filtración, los registros de EHG son representados por características en el dominio de tiempo y la frecuencia mencionados en la Tabla 1, los cuales son calculados con el método de ventaneo, este permite el cálculo de las variables teniendo en cuenta la información en detalle.

En la Figura 3 se muestra la diferencia que existe al ver toda la información del EHG y ver el resultado del ventaneo.

Figura 5. Señal EHG, a) señal completa, b) ventaneo



A B

Fuente: Elaborado en el programa Matlab por los autores.

En la tabla 2 se presentan los resultados obtenidos por cada clasificador implementado de acuerdo con el grupo de datos procesado:

Tabla 2. Resultados de clasificación. S, E y EX están en porcentaje (%). A= 1, 11, 13, B= 100, 128.

Grupo	KNN				MPL				SVM		
	K	S	E	EX	#	S	E	EX	S	E	EX
1	2	100	83,3	90	A	75	100	90	75	83,3	80
2	2	87,5	83,3	77,7	A	75,5	75	75	28	50	25
3	2	29,4	72,7	44,4	B	58,8	54,4	57,1	100	12,5	39,2
4	2	35,2	72,7	50,5	B	82,3	90,9	85,7	35,2	54,5	42,8
5	2	25	72,5	40,7	B	17,6	100	50	11,7	54,4	20,2

KNN: el número de vecinos elegido es 2, ya que, si cambia ese valor, el tiempo se incrementa sin variaciones en el resultado. Por otro lado, el porcentaje de sensibilidad disminuye del 100 % al 87,5 % entre el grupo 1 y 2 (tabla 2). En otras palabras, el modelo clasifica mejor al paciente que puede sufrir alguna complicación que a los que no tienen riesgo. La especificidad del grupo 3 es del 70 %, sin embargo, ese indicador decrece con respecto a los grupos 1 y 2. Los grupos 4 y 5 presentan los menores porcentajes de sensibilidad. Esto demuestra que este grupo de datos no es diferenciable, es decir, sus valores no presentan suficiente información para que el modelo aprenda de las tendencias.

MLP: el clasificador es implementado con 1 y 2 capas ocultas pero los resultados del rendimiento son inferiores al 50 %, y si el número de capas ocultas es superior a 3, el rendimiento es igual pero este proceso sobrecarga (no es óptimo) la máquina, por estas

razones, el MLP se implementa con 3 capas ocultas. Los porcentajes de acierto con el grupo 1 son superiores al 70 %. Al incluir la señal biológica del grupo 2, el porcentaje de precisión disminuye, lo que indica que estos valores no aportan valor de decisión al modelo. mientras que para el grupo 3, 4 y 5, modelo compuesto por 3 capas los porcentajes son superiores al 60 % (Tabla 2).

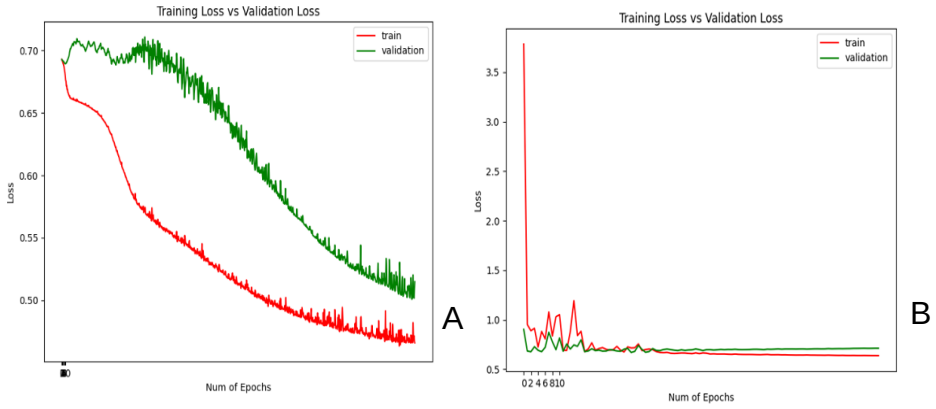
SVM: el kernel lineal presenta los valores de sensibilidad más bajos. Sin embargo, cuando configura el kernel polinomial y el grupo 1, el valor de precisión es del 75 % (Tabla 2). Aprendizaje profundo: para el grupo 1, la sensibilidad del modelo aumenta si se amplía el número de épocas, considerando dos variaciones 10 y 1000 épocas. Esto es típico en este modelo en vista de que cuantas más veces se entrena cada capa, más precisa es. La sensibilidad máxima es del 77,7 %. El grupo 2 presenta una sensibilidad del 75 % al configurar con 5 y 1000 épocas.

El gráfico de las pérdidas de entrenamiento, indica la capacidad del modelo para adaptarse a los datos de entrenamiento, mientras que las pérdidas durante la validación indica la capacidad del modelo para adaptarse a los nuevos datos dado que el método de validación cruzada presenta datos que el modelo desconoce. La figura 4.A se refiere al grupo 1, la capacidad de ajustarse a los datos de entrenamiento es más rápida que el ajuste a los nuevos datos. Además, las pérdidas de entrenamiento final son 0,46 y las pérdidas de validación son 0.5, lo que significa que hay un sobreajuste parcial. El objetivo de estos modelos es que la

pérdida de validación sea la menor posible. En el resultado es aceptable cierto sobreajuste, teniendo en cuenta la cantidad de datos con los que se entrenó y validó el modelo. En la figura 4.B, el error de entrenamiento de validación muestra un sobreajuste parcial, al igual que el grupo 1.

Para el grupo 4 el porcentaje de precisión es bajo, la sensibilidad es de 23.5 % y la especificidad de 72.8 %, esta última aumenta al procesar el grupo 5 con 81.8 %, sin embargo, la sensibilidad se mantiene con valores menores al 30 %.

Figura 6. Gráfica de pérdidas. A) Grupo 1, B) Grupo 2. En rojo pérdidas de entrenamiento y en verde pérdidas de validación.



Fuente: Elaborado en el programa Matlab por los autores.

El modelo KNN haciendo uso de la información demográfica obtiene una precisión del 90 %, y al procesar los grupos que solo contienen las características de bio-señal de EHG, disminuye al 77,7 %, lo que significa que la información demográfica representa mayor capacidad de decisión para el modelo. En la investigación desarrollada por la Universidad de Islamabad, Pakistán para clasificar los nacimientos prematuros que nacen mediante cesáreas, presenta el mayor porcentaje de sensibilidad del 80 % con el clasificador KNN [23]. La diferencia entre trabajos es del 10 % y radica en que el tipo de información demográfica es

diferente y en el caso de la investigación de la Universidad de Islamabad algunos de ellos no representan capacidad de decisión.

La diferencia entre el grupo 1 y 2 se debe a que el uso de señales biológicas como EHG requiere una base de datos más robusta, a pesar de esto, los porcentajes mostrados son cercanos a las investigaciones encontradas en la literatura, por ejemplo, en la investigación de la Universidad de Ciencia y Tecnología de Polonia, obtuvieron una precisión del 96 % para predecir nacimientos prematuros teniendo en cuenta una gran base de datos con más de 400 registros [24].

El MLP y el grupo 2 tienen una sensibilidad del 75,5 % y el grupo 4 es del 82,3 %. En comparación a estos resultados, la investigación realizada por la Universidad Politécnica de Valencia [25], hace uso de la señal biológica EHG y redes neuronales para predecir el éxito al inducir el parto, la sensibilidad es del 68,9 %. Estos resultados concuerdan con la presente investigación, si bien la herramienta no es un diagnóstico médico, permite soportarlo.

Finalmente, los resultados de la presente investigación muestran que el modelo SVM presenta valores inferiores al 50 % de precisión con los grupos 2,3,4,5 pero con el grupo 1 los valores son superiores al 70 %. Esto coincide con los métodos encontrados en la literatura, porque los resultados están entre el 50 % y el 98 %, a veces sin especificar si la validación se realiza con datos que no son conocidos por el modelo.

Sin embargo, en el trabajo presentado por la Universidad Tecnológica de Bolívar, donde se predice la morbilidad materna, se afirma que la precisión del modelo varía de acuerdo a la edad gestacional cuando se está registrando el EHG, teniendo en cuenta el primer y segundo

trimestre, la precisión es del 61 % y el último trimestre el porcentaje de precisión es superior al 70 % [26].

En lo que se refiere al resultado de un modelo predictivo de parto por cesárea en la literatura se encuentra el trabajo realizado por la doctora Sonia de Miguel Manso [27], la cual predice si la labor de parto se realiza mediante cesárea con una precisión del 70.74 % haciendo uso de métodos estadísticos basada en información demográfica. En este trabajo el modelo bio-inspirado obtuvo una precisión del 90.2 % lo que permite observar que el uso de modelos bio-inspirados incrementa la capacidad del modelo de predecir si una gestante realizará el parto mediante cesárea. Adicional a esto el modelo citado anteriormente se realiza con 231 pacientes y el presente modelo se realizó con 122, lo que quiere decir que con un tamaño muestral inferior se lograron conseguir valores de precisión más altos.

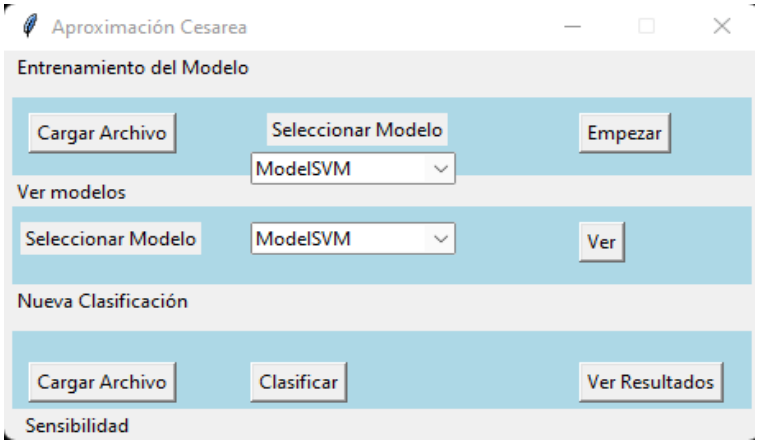
El modelo desarrollado por el Hospital de Mujeres e Infantes de Rhode Island y la facultad de Medicina Alpert de la Universidad de Brown tiene como objetivo predecir parto por cesárea primaria en pacientes con diabetes mellitus gestacional haciendo uso de inteligencia artificial, los resultados arrojan un 95 % de precisión, a diferencia del estudio anterior el espacio muestral es mayor, compuesto de 3.631 pacientes, indicando así que aumenta la capacidad de predicción si se aumenta la base de datos con la que se entrena el modelo [28].

Por otro lado, la investigación realizada por el departamento de Obstetricia y Ginecología de la Universidad de Columbia realiza un modelo de predicción del parto vaginal después de una cesárea en el período prematuro, el espacio muestral es de 1295 pacientes él es procesado por un modelo basado en la regresión logística multivariable, las pruebas realizadas dan como resultado un 80 % de precisión en la predicción [29]. De acuerdo con los

resultados del modelo bio-inspirado y los modelos encontrados en la literatura se evidencia que la capacidad adaptativa de los modelos basados en la biología incrementa la precisión de la predicción.

La interfaz desarrollada para este proyecto se ilustra en la Figura 7. Para implementar los clasificadores en la interfaz es necesario crear un módulo de carga de información en el cual se ingresa una base de datos de entrenamiento, la cual es dividida aleatoriamente, luego se selecciona el clasificador. Finalmente, se desarrolla un módulo para realizar la clasificación de un nuevo paciente. Adicionalmente esta aplicación guarda el clasificador, los resultados obtenidos en el entrenamiento y en la clasificación del nuevo paciente.

Figura 7. Interfaz de visualización gráfica desarrollada para la implementación del modelo.



Fuente: Elaborado por los autores.

6. Conclusiones

El modelo bio-inspirado propuesto es validado mediante los experimentos ilustrados y muestra su efectividad en la predicción de partos por cesárea, con el clasificador KNN que permite un 90 % de precisión con la información demográfica y el clasificador MLP para llegar

al 90,9 % de precisión al procesar la información de la bio-señal EHG. Por tanto, el siguiente paso en esta implementación es realizar ensayos clínicos para aumentar la diversidad de pacientes y dotar de mayor adaptabilidad al modelo.

Herramientas tecnológicas como este modelo bio-inspirado que permiten el uso de información demográfica o bioseñales como el EHG, permiten que sea posible implementarlas en los centros médicos que cuentan con tecnología para tomar la bioseñal, pero también permite que en los centros de salud que no tienen estos equipos se pueda hacer uso de esta herramienta con la información de la gestante, para realizar la clasificación y brindar apoyo al diagnóstico preventivo y así disminuir la morbilidad materna. Además, esta investigación compara distintos clasificadores para optimizar la detección y permite utilizar diversidad de datos en función de las instalaciones del centro médico.

Referencias

[1] U. Nations, "ECLAC - United Nations," 2015. [Online]. Available: https://www.cepal.org/es/temas/objetivos-de-desarrollo-del-milenio-odm/.

[2] E. A. Alfonso, A. D. Arcila and M. L. Latorre, "Atlas de Variaciones Geográficas en Salud de Colombia 2015- Estudio," Ministerio de Salud y Protección Social, 2015.

[3] DANE, "DIRECCIÓN DE CENSOS Y DEMOGRAFÍA - ESTADÍSTICAS VITALES EEVV," DANE, Bogotá, 2021.

[4] E. a. M. J. a. H. M. Nodelman, "Using artificial intelligence to predict spontaneous preterm delivery," *American Journal of Obstetrics and Gynecology, ISSN 00029378,* vol. 222, no. 1, p. 1, 2020.

[5] M. Tahir and T. Badriyah, "Neural Networks Algorithm to Inquire Previous Preeclampsia Factors in Women with Chronic Hypertension During Pregnancy in Childbirth Process," *International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC),* pp. pp. 51-55, 2018.

[6] A. a. R. C. W. G. a. P. A. T. a. J. I. a. G. A. Petrozziello, "Multimodal Convolutional Neural Networks to Detect Fetal Compromise During Labor and Delivery," *IEEE Access, ISSN 2169-3536,* vol. 7, no. 1, pp. 112026--112036, 2019.

[7] K. M. M. N. P. T. P. M. Mcdonnell NJ, "Analgesia after Caesarean Delivery," *Anaesthesia and Intensive Care,* pp. 539-551, 2009.

[8] S. YASUI, "Cesarean section," *Japanese journal of medical science and biology,* vol. 49, no. 6, pp. 1036-1040.

[9] G. E.-A. Escalante-Gaytán J, "Utilidad de la electrohisterografía como técnica de monitorización uterina en el ámbito clínico: revisión bibliográfica," *Ginecologia y Obstetricia de Mexico,* vol. 87, no. 1, pp. 280-285, 2019.

[10] A. Moujahid, I. Inza and P. Larranaga, "Universidad del País Vasco," 14 07 2022. [Online]. Available: http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t9knn.pdf..

[11] D. Morariu, R. G. Cretulescu and M. Breazu, "THE WEKA MULTILAYER," *International Journal of Advanced Statistics and IT&C for Economics and Life Sciences,* pp. 1-9, 2017.

[12] A. Kowalczyk, "Support Vector Machines Succinctly," *Succinctly free ebooks,* p. 114, 2017.

[13] . J.-H. Wang and S.-J. Lou, "Predicting the success rate of natural spontaneous delivery through deep learning," *2019 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW),* 2019.

[14] D. L. Guzmán, M. Carreno and L. H. Camargo, "Identifying Sleep Apnoea and Hypopnoea Episodes From Respiratory Polygraphy Signals," *Revista Colombiana de Psiquiatría,* vol. 46, no. 2, pp. 88-94, 2017.

[15] A. Goldberger, L. Amaral and L. Glass, "Components of a new research resource for complex physiologic signals," *PhysioNet,* p. 101, 2000.

[16] T. M. Nair, "Statistical and artificial neural network-based analysis to understand complexity and heterogeneity in preeclampsia," *Computational Biology and Chemistry,* vol. 75, pp. 222-230, 2018.

[17] R. E. Behrman and A. Stith Butler, "Preterm Birth: Causes, Consequences, and Prevention," *the National Academies Press,* p. 790, 2007.

[18] C. D.-B. Manuel Gómez-Gómez and M. Aceves-Gómez, "Classification of the newborns," *Revista Mexicana de pediatría,* vol. 79, no. 1, pp. 32-39, 2012.

[19] J. Panduro-Barón, E. Panduro-Moore and J. Pérez-Molina, "Changes in fetal static and their predisposition to risk factors throughout of pregnancy," *Ginecología y obstetricia de México,* vol. 85, no. 8, 2017.

[20] K. Kammerer, R. Pryss and B. Hoppenstedt, "Process-Driven and Flow-Based Processing of Industrial Sensor Data," *Sensors (Basel),* vol. 20, no. 18, p. 5245, 2020.

[21] Q. Q. a. X. Z. b. Y. A. a. J. P. a. L. Y. a. D. Z. c. Dongmei Hao a, "Application of decision tree in determining the importance of surface electrohysterography," *Biocybernetics and Biomedical Engineering,* vol. 39, no. 3, pp. 806-813, 2019.

[22] A. A. L. G. L. Goldberger, "physionet," 09 05 2015. [Online]. Available: https://physionet.org/content/ehgdb/1.0.0/. [Accessed 20 05 2022].

[23] K. H. R. F. Muhammad Ehsan Ul Haq, "Use of Machine learning models for predicting and improving maternal and child health indicators," *IEEE-SEM,* vol. 9, no. 8, pp. 1-12, 2021.

[24] M. B. M. S. a. P. A. Somayeh Mohammadi Far, "Prediction of Preterm Delivery from Unbalanced EHG Database," *Sensors (Basel),* vol. 22, no. 4, p. 1507, 2022.

[25] Y. Y.-L. J. G.-C., J. A.-R. A. P. a. G. P.-B. Carlos Benalcazar-Parra, "Prediction of Labor Induction Success from the Uterine Electrohysterogram," *Journal of Sensors,* vol. 2019, no. 1, p. 12, 2019.

[26] E. L. A. Rodríguez, "Early Prediction of Extreme Maternal Morbidity Using Machine Learning," TECHNOLOGICAL UNIVERSITY OF BOLIVAR, Cartagena, 2017.

[27] S. de Miguel Manso, "CLINICAL VALIDATION OF AN EXPERIMENTAL MODEL FOR PREDICTING THE MODE OF LABOR AFTER INDUCTION," Universidad de Valladolid, 2020.

[28] R. M. G. L. V. A. D. Sebastian Z., "Prediction model for primary cesarean delivery in patients with gestational diabetes mellitus," *American Journal of Obstetrics and Gynecology,* vol. 226, no. 1, p. S44, 2022.

[29] C. V. A. A. G. C. G.-B. Anne H. Mardy, "A prediction model of vaginal birth after cesarean in the preterm period," *American Journal of Obstetrics and Gynecology,* vol. 215, no. 4, p. 513, 2016.

Anexo 2.

Certificado de PONENTE en XVII Congreso internacional de Electrónica, Control y Telecomunicaciones, realizado el 24 y 25 de noviembre de 2022 de manera virtual. Código: 0X7484844B708.

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS
Acreditacion Institucional de Alta Calidad

CONGRESO INTERNACIONAL DE ELECTRÓNICA, CONTROL Y TELECOMUNICACIONES
XVII
24 y 25 de noviembre de 2022
Ciencia tecnología e innovación avanzadas para transitar hacia un nuevo sistema Sociotécnico: transformación social sostenible.

RITA
RED DE INVESTIGACIONES DE TECNOLOGÍA AVANZADA

UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS

CERTIFICA QUE:

Jennifer Pacheco

Con número de identificación: 1072707727

Participó en calidad de

PONENTE

XVII Congreso internacional de Electrónica, Control y Telecomunicaciones, realizado el 24 y 25 de noviembre de 2022 de manera virtual.

En constancia firman:

Firma Certificado Congreso

Harold Vacca González Coordinador General CIECT XVII	Ruth Esperanza Román C. Decana Decana Facultad Tecnológica	Lely Adriana Luengas Contreras Coordinadora Tecnología en Electrónica Industrial

En la ciudad de Bogotá D.C.,a los doce (12) días del mes de Diciembre de 2022

Código: 0X7484844B708