

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Facultad de Ingeniería, Arquitectura y Diseño
Programa de Maestría y Doctorado en Ciencias e Ingeniería



UNA TÉCNICA DE RED NEURONAL ARTIFICIAL Y UN
ALGORITMO DE AUMENTACIÓN DE DATOS PARA SISTEMATIZAR
EL DIAGNÓSTICO DE TROMBOSIS VENOSA PROFUNDA MEDIANTE
EL USO DE CRITERIOS DE WELLS

T E S I S

que para cubrir parcialmente los requisitos para obtener el grado de
DOCTORA EN CIENCIAS

presenta:

MARÍA BERENICE FONG MATA

Director de tesis:

Dr. Everardo Inzunza González

Co-director de tesis:

Dr. Enrique Efrén García Guerrero

Ensenada, B.C. México, Diciembre 2020.

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

Facultad de Ingeniería, Arquitectura y Diseño

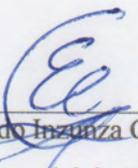
**UNA TÉCNICA DE RED NEURONAL ARTIFICIAL Y UN ALGORITMO DE
AUMENTACIÓN DE DATOS PARA SISTEMATIZAR EL DIAGNÓSTICO DE
TROMBOSIS VENOSA PROFUNDA MEDIANTE EL USO DE CRITERIOS DE
WELLS**

T E S I S

que para obtener el grado de DOCTORA EN CIENCIAS presenta:

MARÍA BERENICE FONG MATA

Y aprobada por el siguiente comité:



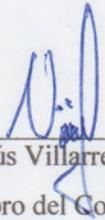
Dr. Everardo Inzunza González
Director del Comité



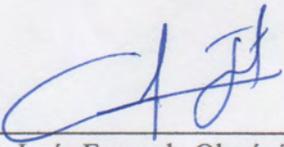
Dr. Enrique Efrén García Guerrero
Co-director del Comité



Dr. Oscar Roberto López Bonilla
Miembro del Comité



Dr. Luis Jesús Villarreal Gómez
Miembro del Comité



Dr. Jesús Everardo Olguín Tizado
Miembro del Comité

Diciembre de 2020

RESUMEN de la tesis de María Berenice Fong Mata, presentada como requisito parcial para la obtención del grado de **DOCTORA EN CIENCIAS EN ELÉCTRICA**, del programa de Maestría y Doctorado en Ciencias e Ingeniería de la UABC. Ensenada, B.C., México, Diciembre del 2020.

UNA TÉCNICA DE RED NEURONAL ARTIFICIAL Y UN ALGORITMO DE AUMENTACIÓN DE DATOS PARA SISTEMATIZAR EL DIAGNÓSTICO DE TROMBOSIS VENOSA PROFUNDA MEDIANTE EL USO DE CRITERIOS DE WELLS

Resumen aprobado por:



Dr. Everardo Inzunza González

Director de tesis



Dr. Enrique Efrén García Guerrero

Co-director de tesis

El presente trabajo de tesis doctoral propone el uso de una Red Neuronal Artificial (RNA) de retropropagación para sistematizar la confiabilidad en el diagnóstico de la Trombosis Venosa Profunda (TVP) mediante el uso de los criterios establecidos por Wells. Para ello, se plantea un nuevo modelo de RNA que mejora la precisión en el diagnóstico cuando se dispone de un conjunto de datos altamente desbalanceado. También se propone un nuevo algoritmo de aumentación de datos basado en datos estadísticos de casos reales reportados en la literatura (conocidos como prevalencia de TVP) así como de un hospital público, para generar un conjunto de 10,000 datos sintéticos utilizados en el entrenamiento de la RNA; en donde a cada instancia generada le corresponden nueve factores de riesgo de acuerdo con los criterios de Wells, además de incluir dos factores adicionales los cuales son el género y la edad. Por otro lado, el esquema del conjunto de datos elegido para el entrenamiento de la red se deriva de los resultados obtenidos de realizar entrevistas a médicos con diferentes especialidades. En el presente trabajo se encontraron dos umbrales de decisión para alcanzar una mejor Precisión y Sensibilidad (también llamada Recall) en el modelo, 0.484 y 0.138, respectivamente. La exactitud alcanzada por el modelo de RNA propuesto es de 90.99%, superando el porcentaje obtenido con otros algoritmos de aprendizaje automático bien conocidos en la literatura. La validación del modelo se realizó implementando la técnica de la validación cruzada con las 10,000 instancias sintéticas generadas. Finalmente, para prueba final se utilizaron 59 registros clínicos obtenidos de un hospital regional, logrando alcanzar una Exactitud del 98.30%.

Palabras clave: aprendizaje automático; redes neuronales; aumentación de datos; sistemas de apoyo a toma de decisiones clínicos; trombosis venosa profunda.

ABSTRACT of the thesis of María Berenice Fong Mata, presented as a partial requirement to obtain the degree of **DOCTOR in ELECTRICAL SCIENCES**, of the program of MSc and PhD in Sciences and Engineering of UABC. Ensenada, B.C., México, December 2020.

AN ARTIFICIAL NEURAL NETWORK APPROACH AND A DATA AUGMENTATION ALGORITHM TO SYSTEMATIZE THE DIAGNOSIS OF DEEP VEIN THROMBOSIS BY USING WELLS' CRITERIA.

Abstract approved by:



Dr. Everardo Anzunza González

Thesis supervisor



Dr. Enrique Efrén García Guerrero

Thesis supervisor

The use of a back-propagation artificial neural network (ANN) to systematize the reliability of a Deep Vein Thrombosis (DVT) diagnostic by using Wells' criteria is introduced herein. In this Ph.D. thesis, a new ANN model is proposed to improve the Accuracy when dealing with a highly unbalanced dataset. To create the training dataset, a new data augmentation algorithm based on statistical data known as the prevalence of DVT of real cases reported in literature and from the public hospital is proposed. The above is used to generate one dataset of 10,000 synthetic cases. Each synthetic case has nine risk factors according to Wells' criteria and also the use of two additional factors, such as gender and age, is proposed. According to interviews with medical specialists, a training scheme was established. In addition, a new algorithm is presented to improve the Accuracy and Sensitivity/Recall. According to the proposed algorithm, two thresholds of decision were found, the first one is 0.484, which is to improve Accuracy. The other one is 0.138 to improve Sensitivity/Recall. The Accuracy achieved is 90.99%, which is greater than that obtained with other well-known machine learning methods reported in literature. The proposed ANN model was validated performing the k-fold cross validation technique using a dataset with 10,000 synthetic cases. The test was performed by using 59 real cases obtained from a regional hospital, achieving an Accuracy of 98.30%.

Keywords: machine-learning; neural network; data augmentation; clinical decision support systems; CDSS; deep-vein thrombosis.

Agradecimientos

Al Dr. Everardo Inzunza González por darme la oportunidad de realizar mis estudios de doctorado bajo su dirección.

Al Dr. Enrique Efrén García Guerrero por sus observaciones y comentarios durante la presente investigación.

Al Dr. Oscar Roberto Gómez Bonilla por su apoyo en los cursos y comentarios realizados en los avances de tesis.

Al Dr. Jesús Everardo Olguín Tizado por aceptar ser miembro de mi comité de tesis.

Al Dr. Luis Jesús Villarreal Gómez por su amistad sincera, sus consejos y su disposición siempre que fue necesario.

A mi esposo David Abdel Mejía Medina, por su apoyo incondicional, su amor infinito, su paciencia y sus desveladas. Por compartirme su conocimiento y enseñarme como siempre cosas nuevas. ¡Siempre juntos en esta y en todas las vidas, hasta el infinito y más allá!

A mis hijos Iker Abdel y Keira Marie Mejía Fong, por acompañarse uno al otro durante las horas dedicadas a la presente investigación, por sacrificar fines de semana y tardes de juego, por su comprensión y amor demostrado desde el primer día del presente trabajo.

A la Facultad de Ciencias de la Ingeniería y Tecnología, en particular a los maestros Alonso Hernández Guitrón y Antonio Gómez Roa, quienes durante su respectiva gestión como directores me apoyaron con las condiciones laborales necesarias para poder llevar a cabo este trabajo de tesis.

A la Universidad Autónoma de Baja California y al Sindicato de Profesores Superación Universitaria de la UABC por contribuir con el financiamiento de las colegiaturas durante mis estudios de doctorado.

Al MYDCI por proporcionarme un espacio de preparación y superación académica.

Tabla de Contenido

Lista de Figuras	ix
Lista de Tablas.....	x
Lista de Ecuaciones	xi
Capítulo 1 Introducción	1
1.1 Introducción.....	1
1.2 Motivación.....	4
1.3 Planteamiento del problema	5
1.4 Pregunta de investigación.....	7
1.5 Objetivos.....	8
1.5.1 Objetivo General.....	8
1.5.2 Objetivos Específicos	8
1.6 Delimitación del objeto de estudio	8
1.7 Conclusiones del Capítulo 1.....	10
Capítulo 2 Antecedentes.	11
2.1 Modelos probabilísticos.....	12
2.1.1 Modelo de Wells.....	12
2.1.2 Regla de Oudega.....	12
2.1.3 Protocolos de Atención de la Trombosis Venosa Profunda con diagnóstico por algoritmos probabilísticos.	14
2.2 Dímero-D.....	17
2.3 Equipo de diagnóstico	17
2.3.1 Equipo de diagnóstico en etapa de investigación	18
2.3.2 Equipo de diagnóstico en Unidades de atención médica.....	18

2.4	Inteligencia Artificial en el diagnóstico médico.....	19
2.4.1	Máquinas de Soporte Vectorial (MSV).....	20
2.4.2	K-vecinos más cercanos (KNN).....	20
2.4.3	Árboles de Decisión (AD).....	20
2.4.4	Redes Neuronales Artificiales (RNA).....	21
2.5	Conclusiones del Capítulo 2.....	23
Capítulo 3	Metodología.....	25
3.1	Análisis del problema.....	25
3.1.1	Revisión bibliográfica.....	26
3.1.2	Análisis del proceso de diagnóstico de la Trombosis Venosa Profunda.....	27
3.2	Diseño del modelo de predicción de trombosis venosa profunda.....	28
3.2.1	Configuración de la red neuronal artificial.....	29
3.2.2	Algoritmo de aumentación de datos.....	32
3.2.3	Pre-procesamiento del conjunto de datos.....	40
3.2.4	Estrategias de Entrenamiento de la Red Neuronal.....	42
3.3	Validación del Modelo de predicción de la trombosis profunda.....	49
3.4	Diseño del modelo clasificador.....	51
Capítulo 4	Resultados.....	55
4.1	Evaluación de resultados del modelo de predicción de la trombosis venosa profunda.....	56
4.2	Los resultados desde la perspectiva del conjunto de datos.....	59
4.3	Comparación con otras técnicas de análisis de datos.....	66
4.4	Validación Externa.....	67
4.5	Propuesta de aplicación del modelo de predicción de la trombosis venosa profunda.....	70
4.5.1	Escenario de uso.....	70

4.6	Limitaciones del trabajo de tesis	73
4.7	Conclusiones del Capítulo 4.....	73
Capítulo 5	Conclusiones	74
5.1	Conclusiones generales	74
5.2	Aportaciones del trabajo de tesis.....	76
5.3	Trabajo Futuro	77
Referencias	78

Lista de Figuras

Figura 1.1 Campos del conocimiento.....	9
Figura 2.1. Diagrama de flujo del algoritmo clínico para TVP (Fox & Bertoglio, 2011)..	15
Figura 2.2 Diagrama de flujo del algoritmo clínico para pacientes de traumatología (Modi et al., 2016).	16
Figura 2.3 Estructura de una neurona artificial (Pano-Azucena et al., 2018).	22
Figura 3.1 Metodología General de Investigación.	26
Figura 3.2 Configuración de la red neuronal.....	31
Figura 3.3 Entrenamiento del modelo de RNA propuesto.	31
Figura 3.4 Algoritmo para generar el conjunto de datos sintéticos para el entrenamiento y validación de la red neuronal.....	33
Figura 3.5 Cantidad y porcentaje de ocurrencia de factores de riesgo de Wells observados en los registros médicos del hospital ISSSTE, Ensenada, B.C. (Elaboración propia).....	38
Figura 3.6 Pseudocódigo para generar el conjunto de datos sintéticos.	39
Figura 3.7 Distribución de probabilidad.....	42
Figura 3.8 Validación Cruzada del conjunto de datos sintético con k=10.	50
Figura 3.9 Representación de un umbral de clasificación. A. En un caso ideal con datos balanceados. B. En un caso real con datos desbalanceados.	52
Figura 3.10 Algoritmo de optimización de umbral de clasificación.	53
Figura 4.1 Curva PR (Precision-Recall).....	61
Figura 4.2 Comparación de las matrices de confusión de los modelos propuestos.	62
Figura 4.3 Comparación de los porcentajes de precisión, recall y exactitud obtenidos con los modelos propuestos.....	63
Figura 4.4 Resultados de la evaluación externa.	69
Figura 4.5 Uso sugerido del SADC en módulos de atención primaria.	71
Figura 4.6 Diagrama a bloques del Protocolo sugerido en unidades de atención primaria.	72

Lista de Tablas

Tabla 2.1 Incidencia del TEV.....	11
Tabla 2.2 Modelo de Wells para la predicción de TVP (Modi et al., 2016).	13
Tabla 2.3 Escala de Oudega para predicción de TVP (Oudega et al., 2005).	13
Tabla 2.4. Comparación de las herramientas de diagnóstico de TVP.	19
Tabla 2.5. Inteligencia Artificial en aplicaciones médicas.....	21
Tabla 3.1 Modelo de predicción propuesto para el diagnóstico de TVP.....	30
Tabla 3.2 Detalles de configuración del algoritmo de retropropagación en Matlab.	32
Tabla 3.3 Distribución de casos en el conjunto de datos para 10,000 instancias.....	36
Tabla 3.4 Cantidad de factores establecidos en la Escala de Wells que pueden contribuir al diagnóstico de la TVP (P. S. Wells et al., 2006).....	37
Tabla 3.5 Asignación por género.....	40
Tabla 3.6 Asignación por edad.....	41
Tabla 3.7 Asignación por factor de riesgo.....	41
Tabla 3.8 Fragmento de datos del Dataset1.....	44
Tabla 3.9 Fragmento de datos del Dataset2.....	46
Tabla 3.10 División en grupos del conjunto de datos de 10,000 instancias.....	48
Tabla 3.11 Fragmento de datos del Dataset3.....	49
Tabla 3.12 Matriz de Confusión para Modelo de Redes Neuronales.....	51
Tabla 4.1 Resultados de la validación cruzada con el modelo generado con el Dataset3..	58
Tabla 4.2 Resultados de la validación cruzada con el modelo generado con el Dataset1..	58
Tabla 4.3 Influencia que ejerce cada una de las variables de entrada en el desempeño de la red.	65
Tabla 4.4 Desempeño de los modelos de apoyo en la toma de decisiones para el diagnóstico de la TVP.....	67
Tabla 4.5 Comparación con otros algoritmos de aprendizaje automático.	68
Tabla 4.6 Matriz de confusión de los resultados de evaluar el modelo de RNA con los casos reales.....	69

Lista de Ecuaciones

Ecuación 3.1. Sistema de ecuaciones para determinar el porcentaje de casos que se presentaron de cada tipo de riesgo.....	35
Ecuación 3.2 Fórmula para calcular el porcentaje de ocurrencia de cada factor de Wells.....	37
Ecuación 3.3 Fórmula para calcular la probabilidad de TVP positiva en cada grupo.....	45
Ecuación 3.4 Función umbral para obtener máxima Exactitud.....	54
Ecuación 3.5 Función umbral para obtener la máxima exactitud con un Recall igual o menor al presentado por Wells (P. S. Wells et al., 2006).....	54
Ecuación 4.1 Sensibilidad del modelo.	56
Ecuación 4.2 Especificidad del modelo.	56
Ecuación 4.3 Precisión del modelo.	56
Ecuación 4.4 Exactitud del modelo.....	56
Ecuación 4.5 Exactitud máxima que puede alcanzar el modelo.	60

Capítulo 1 Introducción

1.1 Introducción

La Trombosis Venosa Profunda (TVP) y la Embolia Pulmonar (EP), son dos condiciones del Trombo-Embolismo Venoso (TEV); en países como los Estados Unidos de América (EE. UU.) pueden resultar afectados hasta 900,000 pacientes por año, con más de 300,000 muertes en un mismo año (Myers et al., 2020). La TVP es una afección vascular en la que un trombo dentro de una vena se desprende y viaja a través del torrente sanguíneo, el cual en caso de alcanzar los pulmones puede causar una Embolia Pulmonar con un desenlace fatal (Farhat et al., 2018; Taha et al., 2019; White, 2003). La TEV ocurre por primera vez en aproximadamente 100 personas por 100,000 habitantes por año en EE. UU. y aumenta exponencialmente de menos de cinco casos por cada 100,000 personas menores de 15 años a 500 casos en personas de 80 años en el mismo número de casos.

Las actividades propias de la vida moderna como los vuelos trans-oceánicos (Abramowitz & Gertz, 2007; Clarke et al., 2016; Kuipers et al., 2007), la exigencia de un alto desempeño en el deporte (Drobnic et al., 2015) e incluso los trastornos de sueño (Takeuchi & Kodama, 2017) han incrementado la probabilidad de sufrir alteraciones en el sistema circulatorio, propiciando el padecimiento de las ET's. De ahí que se diga que este padecimiento es un problema universal derivado del desarrollo de las sociedades en función de sus características geográficas, raciales y socioculturales (Middeldorp & Levi, 2007) citado en (Martínez-Murillo et al., 2010).

Con el diagnóstico temprano de una ET es posible iniciar oportunamente el tratamiento que evite la manifestación pulmonar de la enfermedad. Aunque la venografía es el estándar de oro para el diagnóstico de la TVP (Goldhaber & Bounameaux, 2012) es un método invasivo, costoso y presenta cierto grado de riesgo en la salud del paciente, además de ser innecesario en el 50% de los casos; lo que hace limitado su uso en aplicaciones clínicas. Otras de las técnicas más utilizadas para la detección de ET's son el Dímero D y las pruebas de imagen no invasivas, como la ecografía para la trombosis venosa profunda y la angiotomografía computadorizada (TC) para el embolismo pulmonar (Alfonso Piérola & Páramo Fernández, 2012). Las angiografías, son realizadas únicamente cuando el paciente

presenta síntomas de ET y en muchas ocasiones no es una prueba suficiente, por lo que es necesario realizar una segunda evaluación apoyándose de una ecografía. En contraste, la ecografía Doppler a color es segura, no invasiva, fácil de usar y tiene alta sensibilidad 88%-98% y especificidad 97%-100%; además de ser el método de diagnóstico inicial recomendado en pacientes con una alta probabilidad de sufrir una TVP en extremidades inferiores (Pedraza García et al., 2018). La principal desventaja de utilizar el ultrasonido Doppler es la dependencia de un operador especializado, sobre todo partiendo de que el radiólogo tiene un entrenamiento que presenta una curva lenta de aprendizaje (Paola Paolinelli, 2013). Así mismo, Bolado et al. comentan en (Bolado et al., 2011) que las exploraciones realizadas en su trabajo de investigación fueron hechas por radiólogos con diferentes formaciones por lo que no todos tenían la misma experiencia en las técnicas utilizadas. De manera que realizar un diagnóstico clínico de TVP puede ser un proceso no tan fácil de llevar a cabo. Como los signos y síntomas de la TVP a menudo son inespecíficos, hay un umbral bajo para ordenar un estudio de ultrasonido de las extremidades inferiores para descartar la TVP (PC et al., 2015). En (T et al., 2018) señalaron que debido a que el diagnóstico de las ET se sospecha con frecuencia pero se confirma en menos del 20% de los casos, no es ideal realizar imágenes en todos los casos sospechosos. Por ello, una alternativa a estos estudios son los algoritmos probabilísticos, los cuales clasifican el riesgo de que un paciente presente diagnóstico positivo a una TVP con base en un conjunto de síntomas y factores asociados a la enfermedad. Entre ellos, la Escala de Wells (P. S. Wells et al., 1997) y la regla de Oudega (Oudega et al., 2005) son los más utilizados.

Por otra parte, la Inteligencia Artificial (IA) está siendo cada vez más utilizada en el campo de la medicina. En (Tan et al., 2017) mencionan que el diagnóstico asistido por computadora juega en el futuro un papel muy importante en el cuidado de la salud. De ahí que en los últimos años el sector médico se haya visto beneficiado particularmente con los Sistemas de Apoyo a la Toma de Decisiones Clínicas (SADC), los cuales son utilizados tanto en la elección de medicamentos (Kadhim, 2018; Morales et al., 2018; Y. Shen et al., 2018; Simões et al., 2018); como en el diagnóstico de enfermedades (Bourouis et al., 2014; Christopher et al., 2015; Hsu, 2018; Kadhim, 2018; Lin et al., 2006; Osman et al., 2010; Uzoka et al., 2011); debido a su alto nivel de confiabilidad. Sin embargo, existen cuestiones

éticas en torno a compartir información médica de los pacientes, lo que representa una barrera para poder entrenar adecuadamente los algoritmos de aprendizaje automático para el diagnóstico de enfermedades. Como solución a la limitante de la baja cantidad de datos para el entrenamiento, a menudo se utilizan técnicas de aumentación de datos, las cuales consisten en generar nuevos casos a partir del conjunto de datos original sin que el patrón de los mismos se vea afectado. En (Choquenaira Florez et al., 2020) se demostró que es posible aumentar los datos clínicos para mejorar el desempeño de sistemas predictivos automáticos. Los autores utilizaron dos métodos para crear historias clínicas sintéticas (trayectorias) basadas en datos existentes. En ambientes médicos, por ejemplo, es muy común utilizar técnicas de aumentación para incrementar el conjunto de imágenes requeridas para realizar diagnósticos basados en imágenes (de Souza et al., 2020; Gao et al., 2019; Z. Liu et al., 2020; Zuluaga-Gomez et al., 2020). Sin embargo, no se han encontrado investigaciones relevantes de técnicas de aumentación de datos que no requieran el uso de imágenes para llevar a cabo el diagnóstico. Debido a ello, en este trabajo se propone una nueva técnica de aumentación de datos utilizando una aproximación de mezcla de patrones, la cual toma en cuenta como primer patrón (i) el conjunto de 59 registros médicos obtenidos de un hospital público y (ii) los porcentajes de distribución de las instancias entre las clases reportadas en (P. S. Wells et al., 2006).

A partir de la revisión del estado del arte, se encontraron algunos trabajos que utilizan algoritmos de aprendizaje automático como herramientas de apoyo para el diagnóstico de enfermedades Trombo-Embólicas (G. et al., 2019; S. Liu et al., 2019; Martins et al., 2020; Mishra & Ashraf, 2020; Nafee et al., 2020; Wang et al., 2020; P. Wells, 2020; Willan et al., 2019). Sin embargo, ninguno de ellos se enfoca en el desarrollo de algoritmos para sistematizar la detección de la TVP en las unidades de atención primaria. Dentro del grupo de los algoritmos de aprendizaje automático más utilizados para el diagnóstico de enfermedades se encuentran las Máquinas de Soporte Vectorial (MSV) (Srivastava et al., 2020), los Árboles de Decisión (AD) (Doma & Pirouz, 2020), los Bosques Aleatorios (BA) (Syed-Abdul et al., 2020) y las Redes Neuronales Convolucionales (RNC) (Z. Liu et al., 2020; Sun et al., 2020), en todos ellos se reportan excelentes resultados cuando se dispone de un conjunto de datos balanceado. Otros esfuerzos de investigación incluyen la implementación de algoritmos de aprendizaje automático en sistemas integrados y arreglos

de compuertas de campos programables (FPGA por sus siglas en inglés) que permiten desarrollar sistemas más eficientes y portátiles (Tlelo-Cuautle et al., 2016).

1.2 Motivación

La TVP es una condición que debido a lo inespecífico de sus síntomas puede confundirse con algún otro padecimiento. Se requiere de un diagnóstico diferencial para confirmar la existencia de TVP. En muchos países, entre ellos México, el médico familiar o de atención primaria no dispone del equipo ni de la experiencia para realizar pruebas de imagen como el ultrasonido o la venografía, por lo que recurre (en el mejor de los casos) a los modelos probabilísticos. El diagnóstico clínico de la TVP por sí solo es poco fiable, por lo tanto, se han desarrollado modelos de probabilidad clínica como son la Escala de Wells y la Regla de Oudega para guiar su investigación y tratamiento (Hardy & Bevis, 2019). El modelo más estudiado y validado es el sugerido por Wells (Bernardi & Camporese, 2018); el cual clasifica a los pacientes en tres grupos de acuerdo a la probabilidad que presentan de sufrir una TVP. Aunque existen estos algoritmos probabilísticos para indicar la probabilidad de que un paciente sufra una TVP, raramente son utilizados por los médicos (van der Hulle et al., 2016) y, cuando esto sucede, el cien por ciento de la decisión sobre continuar con el protocolo de detección de TVP depende de la experiencia y percepción del médico.

Por su parte, las redes neuronales artificiales (RNA) proporcionan una herramienta poderosa para ayudar a los médicos a analizar, modelar y dar sentido a datos clínicos complejos en una amplia gama de aplicaciones médicas. La mayoría de las aplicaciones de las RNA en el campo de la medicina se refieren a problemas de clasificación (Al-Shayea & Kadhim Al-Shayea, 2011). Las instituciones de atención médica están implementando técnicas de aprendizaje automático, como las RNA, para mejorar la prestación de atención a los pacientes a un costo reducido. Además, los métodos basados en RNA se han vuelto cada vez más atractivos en el ámbito médico a medida que los marcos de aprendizaje profundo maduran y se vuelven populares. Así pues, en el diagnóstico de enfermedades Trombo-Embólicas, varios autores se han mostrado a favor de las RNA, por presentar un mejor rendimiento (Martins et al., 2020; Qatawneh et al., 2019; Willan et al., 2019) incluso

en comparación con otros algoritmos de aprendizaje automático (Mishra & Ashraf, 2020). Algunas ventajas del uso de las RNA en las predicciones médicas (Tu, 1996) son: (i) requieren menos capacitación estadística formal, (ii) pueden detectar implícitamente relaciones complejas no lineales entre variables dependientes e independientes, (iii) tienen la capacidad de detectar todas las posibles interacciones entre las variables predictoras, y (iv) se pueden desarrollar utilizando múltiples algoritmos de entrenamiento diferentes. Por todas las ventajas descritas anteriormente y por el desempeño observado en la detección de enfermedades Trombo-Embólicas, es que en este trabajo se eligió desarrollar un modelo de clasificación utilizando redes neuronales artificiales y el algoritmo de aprendizaje de retropropagación; basados en el modelo probabilístico de Wells para diseñar un sistema de apoyo a la toma de decisiones clínica en la detección de la TVP que sea utilizado en unidades de atención primaria.

1.3 Planteamiento del problema

El diagnóstico oportuno de la TVP puede contribuir a reducir el impacto de la enfermedad, como lo son las secuelas o incluso la muerte. Si bien existen tecnologías que facilitan el diagnóstico de la trombosis venosa profunda, como la venografía que es considerada el estándar de oro para el diagnóstico de la TVP (Goldhaber & Bounameaux, 2012), ya se ha establecido que es un método invasivo, costoso y presenta cierto grado de riesgo en la salud del paciente.

Existen otras técnicas de diagnóstico como son el Dímero-D, el cual opera como un marcador sensible en la detección de TEV y excluye el TEV sin necesidad de pruebas adicionales entre pacientes con baja probabilidad clínica de TEP. Sin embargo, a medida que el dímero D aumenta con la edad, los pacientes mayores con mayor frecuencia tienen resultados positivos falsos, lo que disminuye la especificidad de la prueba en estos pacientes. Por otro lado, las pruebas de imagen no invasivas, como la ecografía para trombosis venosa profunda y la angiotomografía computadorizada para embolismo pulmonar requieren de un operador especializado para su correcta interpretación.

Por ello, una alternativa a estos estudios son los algoritmos probabilísticos. El problema asociado a los algoritmos probabilísticos es su baja exactitud (alrededor del

76%), y las consecuencias derivadas de ello. Por ejemplo, en el caso de la escala de Wells (Wells, 2006), más del 20% de los diagnósticos los clasifica erróneamente de manera positiva, lo que implica que se requiera de manera innecesaria la comprobación del diagnóstico a través del Dímero-D, o algún otro medio de diagnóstico, lo que eleva de sobremanera el costo asociado a la consulta.

Adicionalmente, la Inteligencia Artificial, particularmente los SADC son cada vez más utilizados en aplicaciones médicas. Sin embargo, la mayor parte de los esfuerzos realizados son orientados hacia el diagnóstico por imagen, el cual requiere de equipo y especialistas que únicamente en muy escasas ocasiones pueden estar disponibles en una unidad médica de atención primaria, por lo que no es posible la implementación de modelos inteligentes basados en imágenes en este nivel de atención. Un segundo inconveniente asociado a los sistemas inteligentes, en particular a las redes neuronales es que se requiere un volumen grande datos para realizar el entrenamiento, lo que trae consigo algunos puntos a considerar.

Primeramente, se debe mencionar que la obtención de los datos no es una actividad fácil, ya sea porque hay poca información disponible por el protocolo de acceso a la información; la información no está procesada de manera que pueda ser utilizada tal cual dentro de un esquema de entrenamiento de una red neuronal, o los datos extraídos están correlacionados entre sí, lo que afecta el resultado de la clasificación.

Por otro lado, se debe destacar el hecho de que las Redes Neuronales Artificiales al igual que otros algoritmos de aprendizaje automático, están diseñadas para trabajar con conjuntos de datos balanceados, lo que significa que la cantidad de ejemplos de cada clase representada en el conjunto de datos de aprendizaje deben de ser muy similar. En aquellas aplicaciones en las que sea necesario trabajar con conjuntos de datos desbalanceados, las salidas del modelo de clasificación tienden a orientarse hacia la clase mayoritaria. Esto es particularmente problemático en un ambiente médico, ya que la mayoría de los pacientes que se presentan con sospecha de una enfermedad, terminan con un diagnóstico negativo. De hecho, según el trabajo publicado por Wells (2006) dentro de un conjunto de 10,000 casos sospechosos de TVP, 8,098 de esos casos resultaron ser negativos, mientras que

únicamente 1,902 casos resultaron positivos, lo que implica una relación aproximada de un caso positivo por cada cuatro casos negativos.

Aunado a lo anterior, también es importante mencionar que generalmente, el rendimiento de la clasificación utilizando una red neuronal se mide a través de la exactitud; sin embargo, en hospitales resulta de mucho más interés conocer la precisión (en adelante llamada *presicion* en inglés) y la exhaustividad (en adelante llamada *Recall* en inglés) de un modelo de predicción, ya que estos conceptos están asociados al tipo de error que se presenta, los cuales pueden ser: (i) que un paciente sano sea diagnosticado con una TVP (falso positivo), lo que tiene una implicación económica para quien cubre los costos asociados a la atención médica del paciente, ya que se requieren estudios adicionales para corroborar el diagnóstico; o (ii) que un paciente enfermo sea diagnosticado como sano (falso negativo), lo que puede retrasar la atención médica del paciente, y en consecuencia, afectar su salud ante una atención tardía, que puede incluso derivar en la muerte.

En resumen, si bien los algoritmos probabilísticos son una gran alternativa para el diagnóstico de la trombosis venosa profunda, también tienen altos costos asociados. Por otra parte, la inteligencia artificial está ganando terreno como una estrategia para estimar diagnósticos, sin embargo, poco se ha abordado en la medicina para el caso del diagnóstico de la trombosis venosa profunda a través de técnicas de bajo costo, siendo el diagnóstico por imagen el que tiene mayor auge. Por ello, se aprecia la necesidad de trabajar en un esquema de diagnóstico inteligente que mejore la exactitud de los algoritmos probabilísticos, tomando en cuenta las características tanto de las redes neuronales como de la naturaleza de los casos médicos que se presentan en la vida real en un hospital.

1.4 Pregunta de investigación

Para definir el abordaje que se le va a dar a este trabajo de investigación, es importante definir la pregunta de investigación, la cual se presenta a continuación:

¿Cuáles son las características de diseño de un modelo de IA que logre una exactitud mayor con al menos la misma sensibilidad que la alcanzada utilizando la Escala de Wells para el diagnóstico de la Trombosis Venosa Profunda?

1.5 Objetivos

1.5.1 Objetivo General

Proponer una nueva técnica de aumentación de datos para generar los datos de entrenamiento con que se alimentará un nuevo modelo de red neuronal artificial basado en el modelo probabilístico de Wells, para la sistematización del diagnóstico de TVP en atención primaria.

1.5.2 Objetivos Específicos

- Generar un conjunto de datos sintéticos utilizando técnicas de aumentación de datos.
- Diseñar la estructura de la red neuronal artificial.
- Entrenar y validar la red neuronal artificial.
- Obtener mejores resultados de Exactitud y Recall que los obtenidos al evaluar el conjunto de datos con la escala de Wells.
- Realizar una prueba con datos reales de un hospital local.

1.6 Delimitación del objeto de estudio

Para tener claridad sobre cómo atender la problemática identificada y así cumplir con los objetivos establecidos en el proyecto de tesis doctoral, es de suma importancia conocer y delimitar los distintos campos del conocimiento involucrados (ver Figura 1.1).

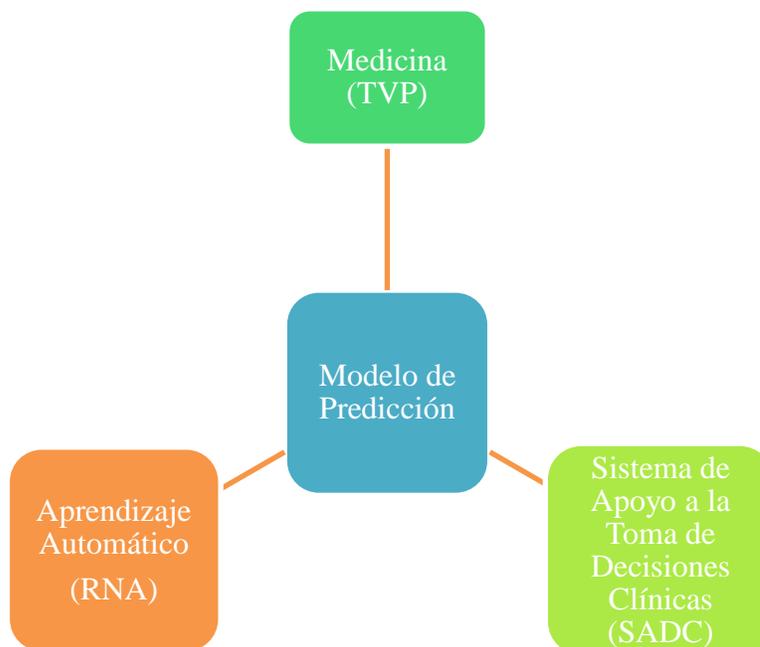


Figura 1.1 Campos del conocimiento.

Por ello, es importante establecer la complejidad del diagnóstico médico, tanto por su naturaleza como por su impacto en la salud de los pacientes. De manera tradicional, es el campo de la medicina quien se encarga de aportar los conocimientos requeridos para conocer el padecimiento que puede tener una persona; para el caso de este proyecto de tesis doctoral, en el que el objetivo esencial es el diagnóstico de la TVP, la medicina aporta algoritmos probabilísticos que pueden por su simplicidad y eficiencia ser estudiados y adaptados en favor de la sistematización del diagnóstico de la TVP. En particular, en este proyecto se analiza el uso de la Escala de Wells (Modi et al., 2016), un algoritmo que permite a los médicos clasificar los pacientes potenciales en virtud del riesgo de padecer una TVP.

Además, como bien se ha mencionado previamente en este capítulo, la Inteligencia Artificial contribuye a la creación de modelos de clasificación con fundamento en el análisis de una base de conocimiento de casos previos. Si bien el aprendizaje automático puede contribuir con distintos algoritmos, se parte de la hipótesis de que las Redes Neuronales Artificiales ofrecen un mejor resultado para el problema a resolver.

Finalmente, es importante mencionar que una vez desarrollado el modelo de clasificación, se requiere de la interpretación del resultado obtenido, por lo que resulta necesario procesar esta información y diseñar interfaces con las que el personal médico se sienta cómodo interactuando. Por lo anterior, se requiere también conocimiento desde el área de los sistemas de apoyo a la toma de decisiones, de manera que se sienten las bases para el desarrollo de aplicaciones de cómputo que puedan ser utilizadas en un hospital, y que faciliten la comprensión del conocimiento generado por el modelo de clasificación. Cabe destacar que por el alcance de la tesis, el cual comprende hasta la generación del modelo de clasificación, el área de Sistemas de Apoyo a la Toma de Decisiones aporta únicamente recomendaciones genéricas que guíen el desarrollo de aplicaciones de cómputo que sean de utilidad para los médicos.

1.7 Conclusiones del Capítulo 1.

En este capítulo se presentó el protocolo de investigación, el cual estableció los lineamientos y alcances de este proyecto de tesis. El objeto de estudio se centra en el diagnóstico automático de la Trombosis Venosa Profunda en el departamento de atención primaria en hospitales públicos de México, derivado de la necesidad que se tiene de contar con personal y equipo especializado para dar un diagnóstico acertado, así como de la carencia de estos recursos en este tipo de atención. Para atender esta problemática, se recurre al uso de algoritmos probabilísticos para el diagnóstico de la Trombosis Venosa Profunda, en particular el de la Escala de Wells (Modi et al., 2016), así como al uso de Redes Neuronales Artificiales, una de las técnicas del aprendizaje automático, que a su vez es un área de conocimiento de la Inteligencia Artificial.

Capítulo 2 Antecedentes.

El Trombo-Embolismo Venoso es una afección mundial, es una condición con altos índices de morbilidad y mortalidad, como se puede apreciar en la Tabla 2.1. Por lo que su detección oportuna permite al especialista determinar el tratamiento adecuado, evitando así que ésta avance y lleve a quien lo padece a un estado de salud que puede ser incluso fatal.

Tabla 2.1 Incidencia del TEV.

Incidencia habitantes	Período	País o Región	TEV		No específica	Ref.
			TVP	EP		
300,000-600,000	anual	Estados Unidos	x	x		(Mozaffarian et al., 2015)
143 y 183 / 100,000	anual	Europa	x	x		(E. Fuentes Camps et al., 2015)
63 / 100,000	anual	España	x	x		(E. Fuentes Camps et al., 2015)
120 / 100,000	No específica	Francia			x	(Moumneh et al., 2018)
12.7 / 100,000	anual	Korea	x			(Hong et al., 2018)
16.6 / 100,000				x		

La TVP se presenta de manera frecuente en el departamento de urgencias; se han estimado cerca de 2 millones de nuevos casos en EE. UU. cada año, con una tasa de incidencia de 1 a 2 casos por cada mil habitantes al año (Benavides Bermúdez et al., 2017). Las estimaciones detalladas del número anual de eventos de TEV son difíciles de obtener puesto que el TEV es difícil de diagnosticar. El TEV a menudo es clínicamente silencioso y, en muchos casos, el primer signo de la enfermedad es una Trombo Embolia Pulmonar repentina y fatal. El TEV a menudo es asintomático, se diagnostica erróneamente y no se reconoce en el momento de la muerte sino a través de rutinas postmortem (P. S. Wells et al., 1997).

2.1 Modelos probabilísticos

Los modelos o algoritmos probabilísticos, clasifican el riesgo de que un paciente presente diagnóstico positivo a una TVP con base en un conjunto de síntomas y factores asociados a la enfermedad. Los principales modelos probabilísticos son la Escala de Wells (P. S. Wells et al., 1997) y la regla de Oudega (Oudega et al., 2005), los cuales se detallan a continuación:

2.1.1 Modelo de Wells

La Escala de Wells (Modi et al., 2016) es el modelo probabilístico más utilizado para guiar el diagnóstico de la TVP como lo mencionan Darwood y Smith en el trabajo “Deep Vein Thrombosis” (Darwood & Smith, 2013), el cual consta de nueve criterios, en donde a cada uno de ellos se les asigna un valor numérico. Wells propone tres diferentes grupos para clasificar la probabilidad de sufrir TVP; dependiendo de la puntuación alcanzada al sumar los valores de los criterios observados en el paciente, la probabilidad de TVP puede clasificarse como: Riesgo Bajo (sumatoria de la puntuación menor a uno), Riesgo Medio (sumatoria de la puntuación entre 1 y 2) y Riesgo Alto (sumatoria mayor o igual a tres). En la Tabla 2.2 se muestran los criterios establecidos, así como el valor asignado.

2.1.2 Regla de Oudega

Un segundo modelo de predicción utilizado para el diagnóstico de la TVP es la escala o regla de Oudega (Oudega et al., 2005), la cual se utiliza en atención primaria para predecir el riesgo de que un paciente sufra TVP. En la Tabla 2.3 se listan los criterios que permiten llevar a cabo la estimación del padecimiento.

Tabla 2.2 Modelo de Wells para la predicción de TVP (Modi et al., 2016).

Criterios considerados	Puntuación
Cáncer activo (en tratamiento o detectado hace menos de 6 meses o en tratamiento paliativo)	1
Parálisis, paresia, inmovilización con yeso del miembro sospechoso	1
Reposo en cama reciente de más de 3 días o cirugía mayor realizada hace menos de 4 semanas	1
Tumefacción de todo un miembro	1
Dolor localizado en el trayecto venoso profundo	1
Tumefacción unilateral de la pantorrilla de más de 3 cm en comparación con la pantorrilla contralateral	1
Edema con fóvea	1
Venas superficiales (no varicosas) colaterales	1
Presencia de un diagnóstico alternativo al menos tan probable como el de trombosis venosa profunda	-2

Tabla 2.3 Escala de Oudega para predicción de TVP (Oudega et al., 2005).

Variables	Puntuación
Hombre	1
Anticonceptivos orales	1
Presencia de malignidad	1
Cirujía reciente	1
Ausencia de trauma en la pierna	1
Distensión de la vena	1
Aumento en la pantorrilla \geq 3cm	2
Dímero-D anormal	6

Al igual que en el modelo de Wells, se suman los criterios seleccionados de la Tabla 2.3 y el resultado obtenido permite conocer la estimación del riesgo para el paciente: riesgo alto (10-13 puntos), riesgo moderado (7-9 puntos), riesgo bajo (5-6 puntos), riesgo muy bajo (0-3 puntos).

Aunque la escala de Wells es la más utilizada en hospitales, en el trabajo “Estudio coste efectividad del proceso diagnóstico de la trombosis venosa profunda desde la atención primaria” (Eva Fuentes Camps et al., 2016) se comenta que ambas escalas presentan “sensibilidad y especificidad similar y no difieren en su validez diagnóstica.”

2.1.3 Protocolos de Atención de la Trombosis Venosa Profunda con diagnóstico por algoritmos probabilísticos.

Una característica notable de los algoritmos probabilísticos, es que deben ser complementadas con algún otro estudio para tener un diagnóstico eficiente. Para ello se han diseñado algunos protocolos de atención al paciente con sospecha de Trombosis Venosa Profunda. En el trabajo de Fox y Bertoglio (Fox & Bertoglio, 2011) se muestra un protocolo (ver Figura 2.1) en el que se establece que los pacientes en los que se presenta una baja probabilidad de sufrir TVP se realiza una prueba sanguínea denominada dímero D (Harder et al., 2018), la cual se presenta con valores elevados producto de la degradación de proteínas en estados TEV. Algunos fabricantes informan sus resultados en FEU (unidades equivalentes de fibrinógeno) y otros en D-DU (unidades de dímero D) con 500 y 230 valores umbral, respectivamente (Zubiate et al., 2019). Cuando el dímero D resulta negativo, se descarta el diagnóstico; por el contrario, ante un dímero D superior al valor del umbral lo siguiente a realizar es una ecografía Doppler. Cuando la probabilidad calculada resulta de moderada a alta el método de diagnóstico inicial recomendado es la ecografía Doppler (Pedraza García et al., 2018).

En la Figura 2.2 se puede observar un segundo protocolo utilizado en pacientes ingresados en el servicio de trauma del centro médico de la institución de los autores (Modi et al., 2016). Al igual que en el protocolo de la Figura 2.1, se evalúa al paciente mediante la escala de Wells. Sin embargo, no se necesita una prueba de Dímero-D; si la escala de Wells muestra una probabilidad baja de TVP, el paciente es reevaluado a las 24 horas, mientras que cuando la probabilidad es moderada o alta, se realiza el diagnóstico ecográfico.

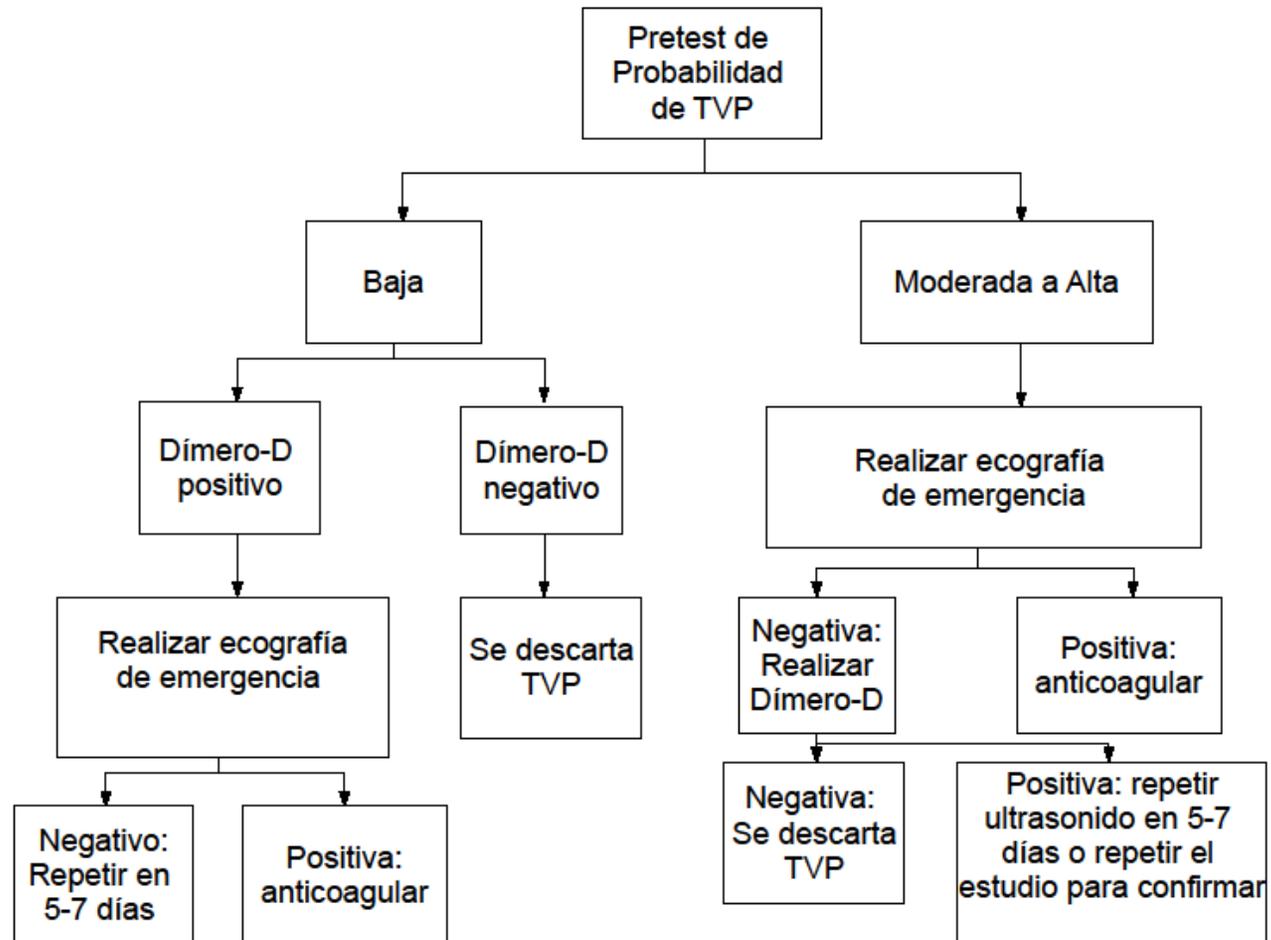


Figura 2.1. Diagrama de flujo del algoritmo clínico para TVP (Fox & Bertoglio, 2011).

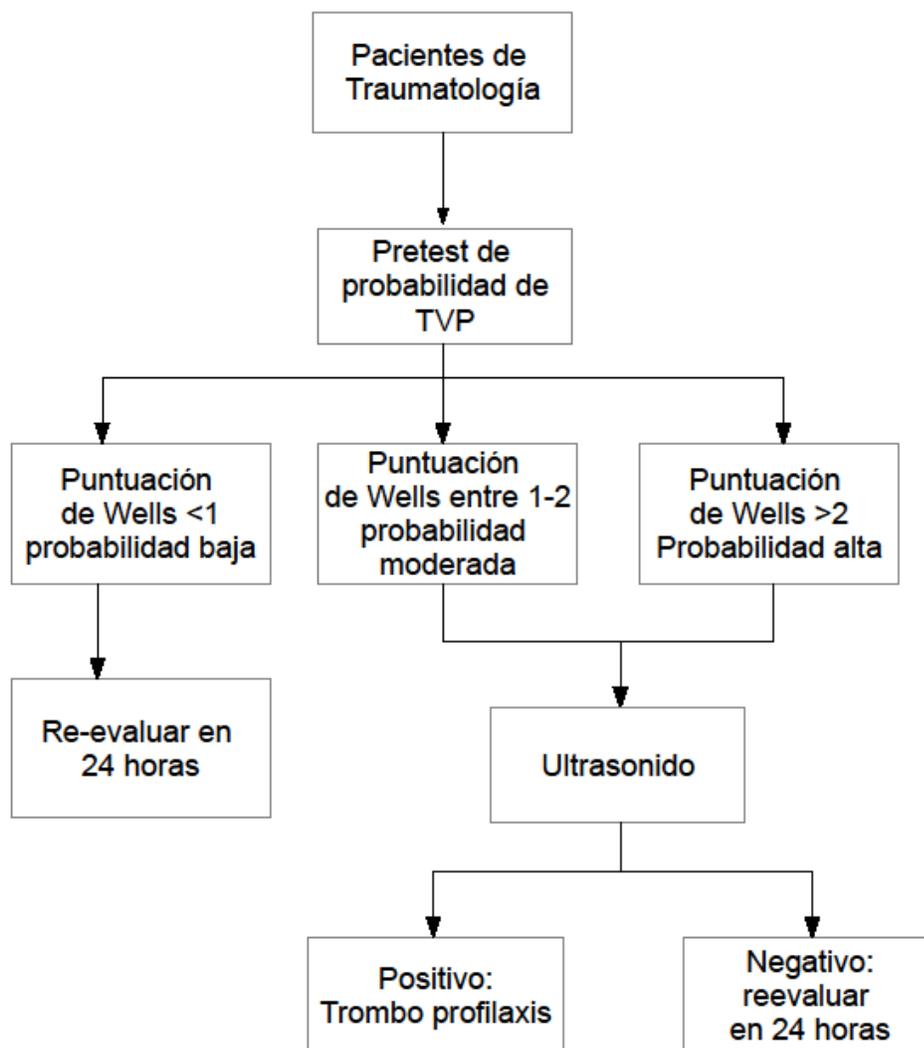


Figura 2.2 Diagrama de flujo del algoritmo clínico para pacientes de traumatología (Modi et al., 2016).

De los protocolos anteriores, debe tenerse en cuenta que existen costos asociados a ellos. Por un lado, el protocolo mostrado en la Figura 2.1 establece la necesidad de realizar el estudio de Dímero-D para tener un diagnóstico con una confianza aceptable. Esto implica un gasto innecesario para los pacientes que no serán diagnosticados con TVP, especialmente los evaluados como de bajo riesgo ya que la tasa de presencia de la enfermedad es realmente baja en este tipo de pacientes.

Además, con respecto al protocolo presentado en la Figura 2.2 se puede observar que los pacientes con riesgo moderado y alto también deben ser evaluados con un estudio ecográfico Doppler, mientras que los pacientes con riesgo bajo permanecen en observación durante 24 horas. Por lo tanto, aunque estos pacientes no tendrán que hacerse cargo de un estudio de laboratorio adicional, es importante tener en cuenta que en condiciones de atención ambulatoria o primaria esto supondría un riesgo para la salud de los pacientes ya que los pacientes tendrían que acudir a su hogar con la posibilidad de tener una TVP, y si esta condición se presenta y no se brinda el tratamiento médico adecuado, puede causar serias complicaciones de salud; sin embargo, al tratarse de un protocolo diseñado para pacientes hospitalizados, la ausencia de un estudio que confirme un diagnóstico que descarte principalmente la TVP no debería generar complicaciones para los pacientes, ya que su salud es monitoreada constantemente.

2.2 Dímero-D

Los Dímeros-D son pequeños fragmentos de proteínas generados por la fibrinólisis de un trombo o coágulo sanguíneo. Los ensayos de Dímero-D se emplean en el diagnóstico de TVP o TEP y se han usado para predecir la probabilidad de Trombo-Embolismo Venoso recurrente (Kirchhof et al., 2014). El Dímero-D opera como un marcador sensible en la detección de TEV y excluye el TEV sin necesidad de pruebas adicionales entre pacientes con baja probabilidad clínica de TEP. Sin embargo, a medida que el Dímero-D aumenta con la edad, los pacientes mayores con mayor frecuencia tienen resultados positivos falsos, lo que disminuye la especificidad de la prueba en estos pacientes. La tasa de falsos positivos puede reducirse utilizando un umbral de Dímero-D ajustado por edad (T et al., 2018).

2.3 Equipo de diagnóstico

Con el paso de los años se han desarrollado diferentes tecnologías capaces de simular y medir la hemodinámica del flujo sanguíneo lo que facilita el diagnóstico de las enfermedades Trombo-Embólicas. A continuación, se describen los principales equipos de diagnóstico de TVP tanto en etapa de investigación, como los que se utilizan actualmente dentro de las unidades de radiología.

2.3.1 Equipo de diagnóstico en etapa de investigación

La termografía (Deng et al., 2015; Francis et al., 1979; Huang et al., 2011; Jin et al., 2012; Kacmaz et al., 2017; Saxena et al., 2018), es una técnica rápida, simple, no invasiva, sin riesgo, con un adecuado costo-beneficio y es altamente sensible en una investigación inicial cuando existe sospecha de TVP. La termografía monitorea la temperatura del cuerpo utilizando cámaras térmicas para obtener correlaciones entre la fisiología térmica y la temperatura de la piel (Lahiri et al., 2012). Una termografía negativa es concluyente en un estudio de TVP. Sin embargo, una termografía positiva tiene un sin número de causas probables, lo que indicaría una evaluación adicional por venografía o Doppler para confirmar o refutar la TVP (Harding, 1995). La termografía también se utiliza en el diagnóstico de enfermedades superficiales (piel, venas, arterias) midiendo la velocidad del flujo, dirección y diámetro de las venas (Saxena et al., 2018).

2.3.2 Equipo de diagnóstico en Unidades de atención médica

2.3.2.1 Venografía

Aunque la venografía es el estándar de oro para el diagnóstico de la TVP (Goldhaber & Bounameaux, 2012) es un método invasivo, costoso y presenta cierto grado de riesgo en la salud del paciente (Cohen et al., 2007), además de ser innecesario en el 50% de los casos; lo que hace limitado su uso en aplicaciones clínicas.

2.3.2.2 Ecografía Doppler

Los equipos basados en el principio del efecto Doppler son una herramienta importante para el diagnóstico vascular no invasivo. A través de las ecografías es posible conocer información como la relación de la permeabilidad arterial y venosa, el sentido del flujo sanguíneo, la presencia de estenosis y fístulas, y el estado de la vasculatura distal y proximal al sitio de examen (Paola Paolinelli Médico Radiólogo, 2004). La ecografía Doppler a color es segura, no invasiva y fácil de usar (Kearon et al., 1998; Landefeld, 2008; Segal et al., 2007). Los equipos ultrasónicos basan su funcionamiento en el principio del efecto Doppler y son una herramienta importante para el diagnóstico vascular. En años

recientes la ecografía Doppler además de ser la técnica más utilizada para analizar el flujo sanguíneo (Muth et al., 2011), es uno de los métodos de diagnóstico fundamental en múltiples disciplinas y especialidades médicas (Fortier O'brien et al., 2018; Stevenson, 2018; Torres Macho et al., 2018), y es el método de diagnóstico inicial recomendado en pacientes con una alta probabilidad de sufrir una TVP en extremidades inferiores.

En la Tabla 2.4 se puede observar una comparación entre las principales herramientas de diagnóstico de TVP utilizadas en hospitales y centros de atención médica.

Tabla 2.4. Comparación de las herramientas de diagnóstico de TVP.

Método de Diagnóstico	Invasivo	Sensibilidad [%]	Especificidad [%]	Ref.
Escala de Wells	No	77-98	37-58	(Darwood & Smith, 2013)
Regla de Oudega	No			(Oudega et al., 2005)
Dímero D	No	94	68	(Darwood & Smith, 2013; Parry et al., 2018; Reardon et al., 2019; Sharif et al., 2019)
Dímero D ajustado a la edad	No	100	66.3	(Resano & Resano, 2018)
Ultrasonido Doppler	No	>95		(Resano & Resano, 2018)
Venografía	Sí	100%		(Resano & Resano, 2018)

2.4 Inteligencia Artificial en el diagnóstico médico

La Inteligencia Artificial se ha introducido en el área médica no para sustituir al personal, sino como un complemento de apoyo en la toma de decisiones. Las Redes Neuronales Artificiales (RNA) y las Máquinas de Soporte Vectorial (MSV) entre otros algoritmos, han aportado a la medicina en conjunto con el personal médico, mayor exactitud en el diagnóstico de enfermedades, mejores sugerencias de tratamientos y dosis de medicamentos.

2.4.1 Máquinas de Soporte Vectorial (MSV)

Las máquinas de soporte vectorial son un grupo de métodos de aprendizaje supervisado que pueden ser aplicados a problemas de clasificación o regresión. Fueron inicialmente introducidos para clasificar clases de objetos linealmente separables. Representan hiperplanos o fronteras de decisión en términos de un pequeño subconjunto de todos los ejemplos de entrenamiento que maximizan la separación entre conjuntos y por lo tanto la capacidad de clasificación entre clases, denominados vectores de soporte (Cano et al., 2017).

2.4.2 K-vecinos más cercanos (KNN)

Los K-vecinos más cercanos conservan su popularidad por ser simples y efectivos a pesar de ser una de las primeras aproximaciones de aprendizaje automático (Gok et al., 2020). KNN se basa en el principio de que las instancias dentro de un conjunto de datos generalmente existirán muy cerca de otras instancias que tienen propiedades similares (Dias Canedo & Cordeiro Mendes, 2020). Las distancias del nuevo objeto con sus vecinos se pueden calcular con funciones como la Euclideana, Manhattan, Minkowski y Chebyshev (Ağbulut et al., 2021). El parámetro k decide cuántos vecinos se elegirán para el algoritmo kNN. La elección adecuada de k tiene un impacto significativo en el rendimiento del diagnóstico del algoritmo kNN (Zhang, 2016).

2.4.3 Árboles de Decisión (AD)

En un árbol de decisión (DT), la estrategia de divide y vencerás se utiliza de tal manera que un proceso complejo de toma de decisiones se divide en una secuencia de decisiones menos difíciles (Subasi et al., 2018). La clasificación AD básicamente divide el espacio en subespacios calculando los límites de decisión para cada nodo, y continúa agregando hijos a los nodos del árbol hasta que no se pueden realizar más mejoras en los resultados de la predicción (Ambagtsheer et al., 2020).

2.4.4 Redes Neuronales Artificiales (RNA)

Las redes neuronales son paradigmas computacionales basados en modelos matemáticos con capacidad de un fuerte patrón de reconocimiento. Una red neuronal es un algoritmo de cálculo que se basa en una analogía del sistema nervioso. Se intenta imitar la capacidad de aprendizaje de este, haciendo que reconozca gradualmente patrones de asociación entre las entradas (variables predictivas) y los estados dependientes de ellas (salidas). La señal progresa desde las entradas, luego por la capa oculta, hasta obtener una respuesta traducida en el nivel de activación de los nodos de salida, los cuales predicen el resultado con base en las variables de entrada (Sprockel et al., 2014).

En la Tabla 2.5 se presentan algunas aplicaciones médicas en las que se ha utilizado la IA para mejorar el diagnóstico médico.

Tabla 2.5. Inteligencia Artificial en aplicaciones médicas.

Algoritmo	Aplicación	Exactitud [%]	Especificidad [%]	Sensibilidad [%]	AUC	Ref.
RNA	Diagnóstico diferencial de enfermedad pulmonar intersticial	85	NA	NA		(Abe et al., 2004)
RNA	Plasma Orina	0.882	76.9	88.2	0.882	(Peña-Bautista et al., 2019)
MSV	Alzheimer Plasma orina		76.9	80.9	0.839	
MSV	Infección del tracto urinario	96.61	92.85	97.77	0.817	(Ozkan et al., 2018)
RNA		98.9	100	97.77	0.664	

Las redes neuronales artificiales son cada vez más importantes en la construcción de herramientas de apoyo al diagnóstico médico. La Figura 2.3 muestra el modelo más simple del funcionamiento de una RNA, que es el perceptrón desarrollado por Frank Rosenblatt (Rosenblatt, 1958).

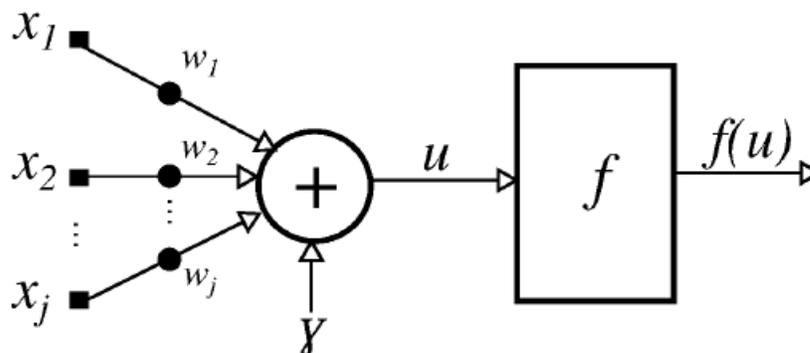


Figura 2.3 Estructura de una neurona artificial (Pano-Azucena et al., 2018).

Las RNA están formadas por neuronas artificiales conectadas entre sí que emulan el funcionamiento de neuronas biológicas en el cerebro de un ser humano. Un conjunto de valores de entrada (x_1, x_2, \dots, x_n) están conectados a una neurona artificial, que se enciende mediante una función de activación. A cada valor de entrada se le asigna un peso (w); los productos de todas las entradas y sus correspondientes pesos se suman antes de pasar a la función de activación. A diferencia del perceptrón simple, el perceptrón multicapa está formado por una o más capas que lo hacen capaz de resolver problemas no lineales (Jain et al., 1996). Aunque existen algunos algoritmos para diseñar la arquitectura de una red, como la regla de la pirámide geométrica (Masters, 1993); la regla basada en el error cuadrático medio (Aylward & Anderson, 1991) y la aplicación de algoritmos evolutivos (Yao, 1999); la configuración de red exitosa depende de la experiencia del diseñador y de la evaluación de diferentes arquitecturas. El número de neuronas en las capas ocultas utilizadas determinará la precisión de la red. Las RNA son algoritmos de aprendizaje automático diseñados para analizar datos sin una hipótesis preexistente sobre las asociaciones que puedan existir (Willan et al., 2019).

Para una clasificación/predicción adecuada, además del diseño de la red, se requiere un algoritmo de entrenamiento. En una RNA, el algoritmo de retropropagación utiliza la diferencia entre el resultado producido y el resultado deseado para cambiar los "pesos" de las conexiones entre las neuronas artificiales (P. Wells, 2020). La importancia de este proceso radica en que, a medida que se entrena la red, las neuronas de las capas intermedias se organizan de tal manera que las distintas neuronas aprenden a reconocer distintas

características del espacio total de entrada. Una RNA de retropropagación funciona bajo el esquema de aprendizaje supervisado, por lo que se requiere un conjunto de datos conocidos para entrenar la red que describe los valores de las variables de entrada (predictores), así como la salida esperada para cada conjunto (Jain et al., 1996).

Una de las principales aplicaciones de las RNA es la clasificación binaria; en medicina, se utiliza principalmente para saber si un paciente debe ser diagnosticado con una enfermedad en particular. Para clasificar, las redes neuronales artificiales requieren un conjunto de datos que se utiliza como base de conocimientos. Una de las características de estos conjuntos de datos es que debe asegurarse que los diferentes tipos estén balanceados; es decir, debe contener un número similar de casos positivos y negativos. Por lo tanto, los conjuntos de datos desbalanceados podrían conducir a resultados de clasificación sesgados y decisiones incorrectas (Hassib et al., 2020) ya que el principal problema con los desbalances de clase es que los clasificadores normales a menudo están sesgados hacia la clase mayoritaria (Bennin et al., 2018). Como resultado, las características de las clases minoritarias no se aprenden adecuadamente (Khan et al., 2018).

2.5 Conclusiones del Capítulo 2

La implementación de la Inteligencia Artificial en sistemas de apoyo a la toma de decisiones clínicas está dando lugar a que se lleven a cabo diagnósticos de enfermedades más acertados y oportunos, lo que permite a los pacientes una mejor calidad de vida o incluso la vida misma ya que en el caso de las enfermedades Trombo-Embólicas, un diagnóstico equivocado puede ser fatal. En el presente trabajo se revisaron las principales herramientas de diagnóstico de TVP, así como algunos algoritmos de Inteligencia Artificial que se han implementado en aplicaciones de diagnóstico médico. Dicho lo anterior, la investigación realizada da lugar a la propuesta de desarrollo de un sistema de toma de decisiones clínico, que opere en conjunto con el médico para la detección de TVP en extremidades inferiores utilizando redes neuronales artificiales; ya que según la literatura mencionada se puede alcanzar mayor exactitud en la clasificación. Es importante enfatizar que no es fácil tener una estadística del número de casos de TVP que suceden en un determinado período de tiempo ya que en la mayoría de las ocasiones los pacientes llegan

al hospital directamente a la Unidad de Urgencias teniendo ya una condición avanzada de TVP, por lo que se sugiere que el modelo de predicción sea implementado en un Sistema de Apoyo en la toma de Decisiones Clínicas (SADC) para ser utilizado como herramienta de apoyo con el primer médico que se entrevista el paciente, y de esta forma detectar de manera oportuna las primeras manifestaciones de este padecimiento. Se sugiere además que el SADC incluya alguno de los modelos probabilísticos utilizados como pre-test puesto que en la mayoría de las ocasiones el primer médico con el cual el paciente tiene contacto no dispone del equipo para realizar estudios de imagen como la venografía o la ecografía Doppler.

Capítulo 3 Metodología

3.1 Análisis del problema

Tradicionalmente, el diagnóstico médico es considerado por los médicos como un arte, y depende en gran medida del conocimiento y experiencia de cada uno de ellos. Se lleva a cabo a través de una combinación de actividades que realizan los médicos, las cuales pueden incluir exploración física, entrevista con el paciente, revisión de historial clínico, interpretación de análisis de laboratorio, entre otros; esta información se utiliza para relacionarla con registros médicos reales y académicos conocidos, antecedentes familiares y, en algunos casos, la opinión de sus colegas.

La ciencia de los datos trae nuevas posibilidades para el diagnóstico médico, ya que permite el análisis de una gran cantidad de información de casos médicos y utilizarlos en el diagnóstico.

En este capítulo se presenta la metodología utilizada en este trabajo de investigación; esta metodología consta de 4 etapas (Figura 3.1), en donde cada una de ellas a su vez se lleva a cabo bajo una metodología específica apropiada al propósito establecido.

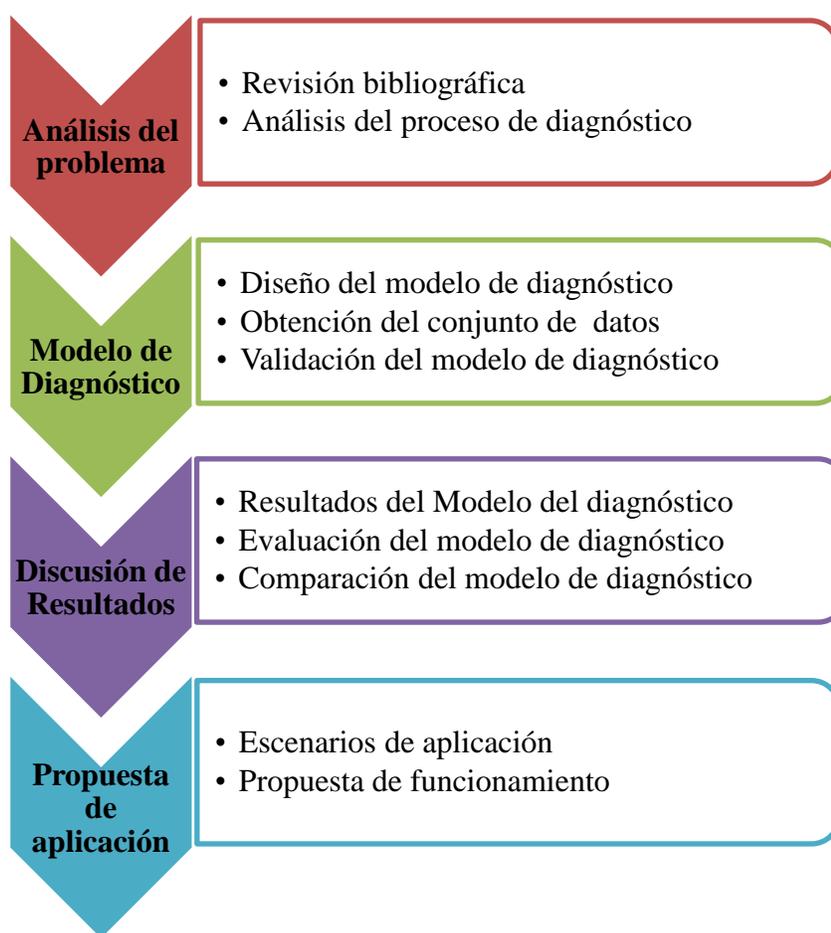


Figura 3.1 Metodología General de Investigación.

3.1.1 Revisión bibliográfica.

En esta etapa se revisan distintas fuentes bibliográficas en editoriales de gran relevancia científica, entre las cuales se destacan Elsevier, IEEE, ACM, MDPI y JAMA. Dado que la mayoría de la literatura relevante está escrita en revistas que únicamente aceptan artículos en idioma inglés, las cadenas de búsqueda se establecen en ese idioma. Las cadenas de búsqueda más relevantes, aunque no las únicas que se utilizan en este trabajo de tesis, son:

- Deep Vein Thrombosis Diagnosis
- Deep Vein Thrombosis Neural Network
- Deep Vein Thrombosis Deep Learning
- Deep Vein Thrombosis Decision Support System
- Deep Vein Thrombosis Primary Care
- Machine Learning Medical Diagnosis
- Artificial Intelligence Medical Diagnosis
- Primary Care Diagnosis

El resultado de estas revisiones sirve para comprender el nicho de oportunidad sobre el cual se debe llevar a cabo la investigación.

3.1.2 Análisis del proceso de diagnóstico de la Trombosis Venosa Profunda.

Otro de los elementos importantes para definir el problema de investigación, es conocer la forma en que ocurre el fenómeno a investigar. En el caso particular de este trabajo de tesis, se requiere conocer a detalle el proceso de diagnóstico de la Trombosis Venosa Profunda en consultorios de atención primaria.

Para ello, se llevan a cabo entrevistas semi-estructuradas breves con 5 médicos de distintos grados de especialidad, las cuales incluyen a dos angiólogos, un médico internista, y dos médicos generales. Estas entrevistas se desarrollan en torno a cuatro preguntas principales:

1. ¿Conoce y utiliza la Escala de Wells?
2. ¿Cuál es su primera opción de diagnóstico para identificar si un paciente está teniendo una trombosis venosa profunda?
3. ¿Puede describir el procesamiento de razonamiento para llevar a cabo un diagnóstico médico sobre una sospecha de trombosis venosa profunda?
4. ¿Cómo maneja la incertidumbre en el diagnóstico de la trombosis venosa profunda?

Derivado del análisis de las respuestas de los médicos, se identifican las siguientes estrategias de diagnóstico:

Estrategia 1. Surge del comentario de un médico, quien utiliza toda la evidencia de casos médicos previos de los que tiene conocimiento, tanto empíricos como académicos, para realizar un diagnóstico de TVP.

Estrategia 2. En este caso, el médico utiliza toda la evidencia de casos médicos previos para realizar un diagnóstico de TVP, pero en caso de incertidumbre el diagnóstico se asume positivo debido a que ante la duda es preferible seguir con el protocolo de TVP.

Estrategia 3. Se establece la estrategia a partir de que un especialista comenta que calcula probabilidades con base en los síntomas y resultados observados en los pacientes a lo largo de su trayectoria médica.

En la sección 3.3.4 se aborda la manera en que estas estrategias son utilizadas como parte del proceso para definir el modelo de diagnóstico.

3.2 Diseño del modelo de predicción de trombosis venosa profunda

Esta etapa es quizás la parte medular de este trabajo de tesis, ya que es aquí donde se gesta el modelo de diagnóstico inteligente de la trombosis venosa profunda. Para lograrlo se requiere pasar por un proceso de obtención de datos y posteriormente diseñar y validar el modelo.

En México es complicado tener acceso a archivos clínicos o datos estadísticos sobre la ocurrencia de ciertos casos médicos, al menos no con el detalle requerido para realizar una investigación como la que se presenta en trabajos previos tales como (P. S. Wells et al., 1997). Para este trabajo de tesis, se recolectó información de 59 pacientes de un hospital público regional en Ensenada, Baja California; dado que estos datos resultan insuficientes para obtener un buen resultado del modelo, se utiliza una estrategia de aumentación de datos basada tanto en los 59 casos recolectados del hospital como en los datos reportados en estudios previos; a partir de la estrategia desarrollada, se genera un

conjunto de datos sintéticos conformado por una matriz con 10,000 casos para utilizarse en el entrenamiento y validación de la red. La validación se lleva a cabo con el algoritmo de la validación cruzada, además de realizar una prueba/validación externa utilizando los 59 registros obtenidos del hospital público. El estudio se realiza de acuerdo con la Declaración de Helsinki, y el protocolo es aprobado el 23 de junio del 2020 por el Comité de Ética de la Universidad Autónoma de Baja California, México con código de identificación POSG/020-1-03.

3.2.1 Configuración de la red neuronal artificial

Una red neuronal está compuesta por varias capas, la capa de entrada, una o más capas ocultas y una capa de salida. Cada neurona o nodo artificial en cada capa está conectado a cada otro nodo en la entrada anterior o la capa siguiente. La capa de entrada puede constar de múltiples nodos, los cuales son utilizados en la predicción, clasificación y/o regresión del modelo. En las capas ocultas se utilizan funciones de activación no lineales que convierten las entradas ponderadas de la capa de entrada (Sohn, 2019).

El objetivo de la RNA es resolver los problemas de la misma manera que el cerebro humano, particularmente aquellos que son difíciles de resolver usando la programación basada en reglas. Para ello, se hace uso de varios algoritmos, entre ellos el de retropropagación. El proceso de retropropagación utiliza la diferencia entre el resultado producido y el resultado deseado para cambiar los "pesos" de las conexiones entre las neuronas artificiales (Werbos, 1974). La importancia de este proceso consiste en que, a medida que se entrena la red, las neuronas de las capas intermedias se organizan a sí mismas de tal modo que las distintas neuronas aprenden a reconocer distintas características del espacio total de entrada.

La red neuronal de retropropagación trabaja bajo el esquema de aprendizaje supervisado, por lo que se requiere de un conjunto de datos conocidos para entrenar a la red, que describan los valores de las variables de entrada, así como la salida esperada para cada conjunto.

La configuración de un esquema de red exitosa depende de la experiencia del diseñador y de la evaluación de distintas estructuras, ya que a la fecha no se encontró una metodología que permita definir el esquema óptimo de una red neuronal para cualquier problema, o al menos, para un modelo de predicción del diagnóstico de la TVP; dado que la cantidad de neuronas utilizadas en las capas ocultas determinan la exactitud de la red, es un proceso que se vuelve muy importante realizarlo adecuadamente.

En el presente trabajo se evalúa el desempeño de más de 100 estructuras de redes neuronales distintas, obteniendo el mejor resultado con la configuración de 11 neuronas en la capa de entrada, tres capas ocultas con 150, 100 y 50 neuronas respectivamente, y una neurona en la capa de salida (Tabla 3.1). Las funciones de activación $\text{tansig}()$ en la capa de entrada, $\text{tansig}()$ en cada una de las capas ocultas, y $\text{purelin}()$ en la capa de salida (Figura 3.2), la cual alimenta a un clasificador binario con un umbral optimizado para el modelo.

Tabla 3.1 Modelo de predicción propuesto para el diagnóstico de TVP.

Capa de entrada 11 predictores	Capas Ocultas	Capa de Salida
Edad, Género, Cáncer, Inmovilización, Cirugía, Sensibilidad, Pierna inflamada, Tobillo inflamado, Edema, Venas superficiales, TVP previa	150-100-50	Diagnóstico TVP

El método de retropropagación de gradiente conjugado (Fletcher–Reeves) se implementa en el modelo de la RNA como algoritmo de aprendizaje. Así mismo, para el entrenamiento de la red se ajustan algunos hiperparámetros de forma manual, entre ellos la tasa de aprendizaje se ajusta para quedar en 1×10^{-2} , se utilizan como máximo 1000 épocas, obteniendo un Mínimo Error Cuadrático de 0.007 (Figura 3.3). La inicialización de los pesos correspondientes a cada neurona se realiza de forma aleatoria y se fija un error mínimo entre 0.001 y 0.09. Se utiliza una computadora con un CPU de 2.9GHz Intel Core i5, memoria RAM de 8GB 1867 MHz DDR3 y una tarjeta de video Intel Iris Graphics 6100 1536 MB. El sistema operativo de la computadora es macOS Sierra y la RNA se desarrolla en Matlab. En la Tabla 3.2 se muestran algunos parámetros del algoritmo de aprendizaje.

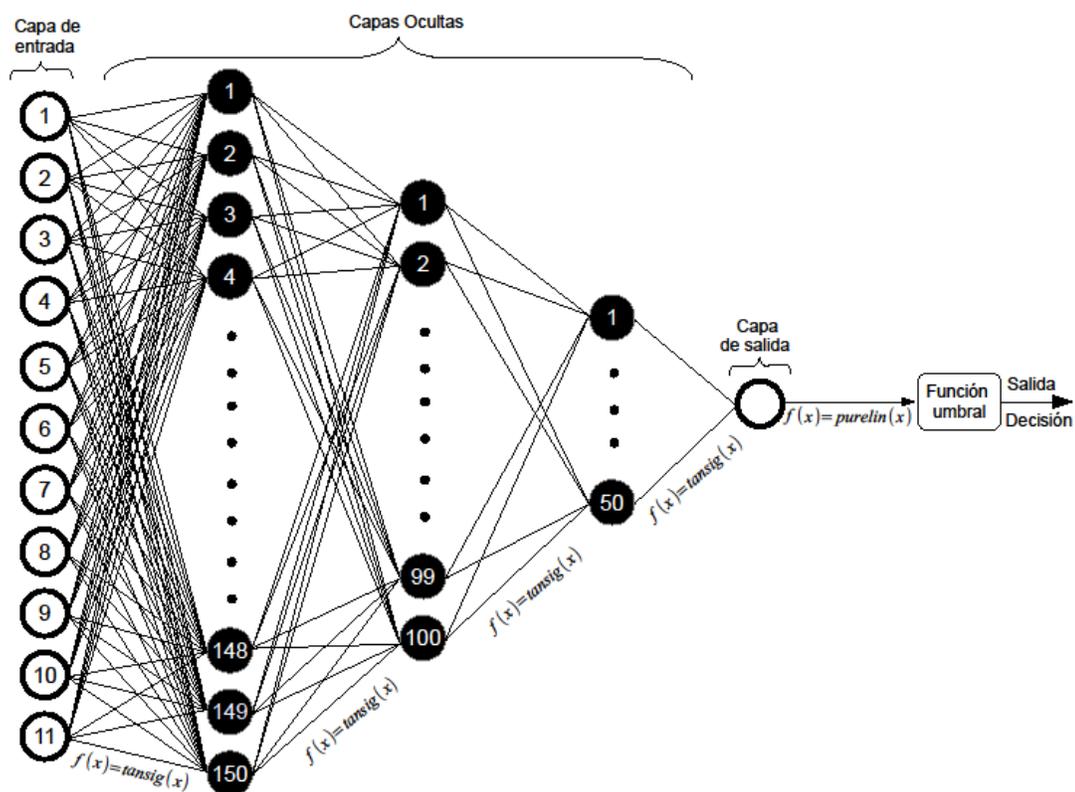


Figura 3.2 Configuración de la red neuronal

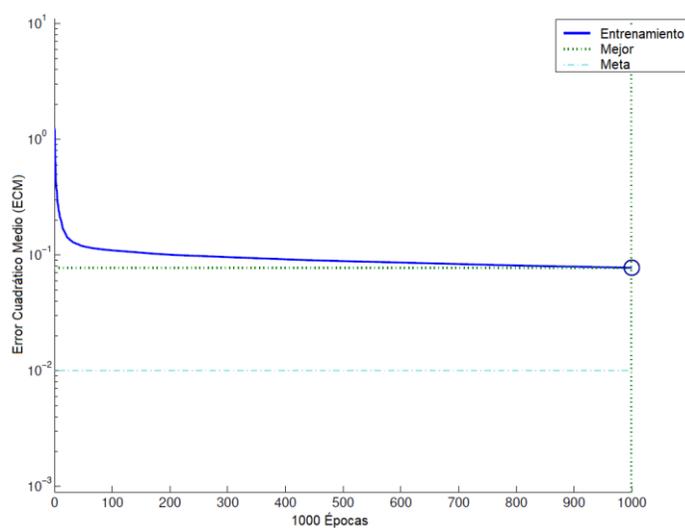


Figura 3.3 Entrenamiento del modelo de RNA propuesto.

Tabla 3.2 Detalles de configuración del algoritmo de retropropagación en Matlab.

Parámetro	Valor	Significado
t.trainParam.epochs	1000	Número máximo de épocas para entrenar
net.trainParam.goal	1e-2	Objetivo de rendimiento
net.trainParam.lr	0.01	Tasa de aprendizaje
net.trainParam.lr_inc	1.05	Proporción para aumentar la tasa de aprendizaje
net.trainParam.lr_dec	0.7	Proporción para disminuir la tasa de aprendizaje
net.trainParam.max_fail	6	Fallos máximos de validación
net.trainParam.max_perf_inc	1.04	Incremento máximo de rendimiento
net.trainParam.mc	0.9	Constante de momento
net.trainParam.min_grad	1e-5	Gradiente de rendimiento mínimo

3.2.2 Algoritmo de aumentación de datos

El principal objetivo del algoritmo de Aumentación de Datos es generar un conjunto de datos sintéticos que se apeguen a los datos observados en la vida real (Gómez-Ríos et al., 2019). Por ello, la primera tarea a realizar es calcular el porcentaje de casos positivos y negativos que se presentan por cada tipo de riesgo de probabilidad de ocurrencia de TVP. Además, se calcula el porcentaje en que se observan cada uno de los factores de la Escala de Wells en los casos reales a los que se tiene acceso. Finalmente, se procede a generar de forma aleatoria cada uno de los casos que comprenderán el conjunto de datos (Figura 3.4).

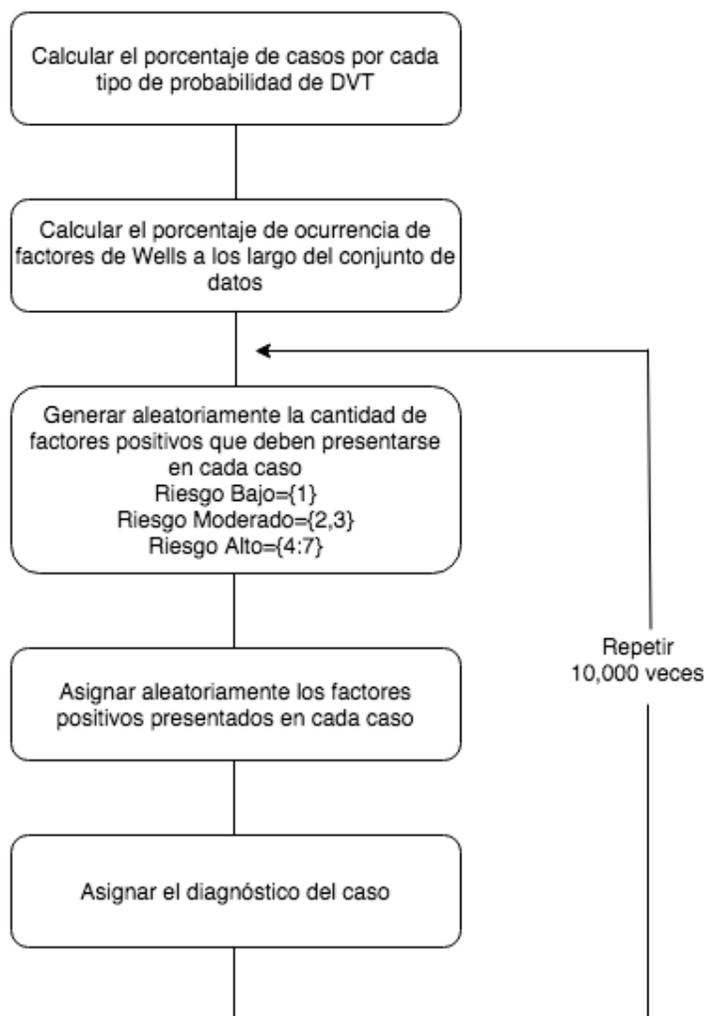


Figura 3.4 Algoritmo para generar el conjunto de datos sintéticos para el entrenamiento y validación de la red neuronal.

La aumentación de datos consiste en aumentar de manera artificial el volumen del conjunto de datos de entrenamiento, aplicando diferentes transformaciones a la información original sin alterar el patrón espacial de las clases. Usualmente, las transformaciones se llevan a cabo durante el entrenamiento, lo cual permite hacerlas sobre la marcha sin grabar la nueva información (Gómez-Ríos et al., 2019). Cuando el conjunto de datos de entrenamiento es insuficiente, la aumentación de datos es una solución elegante a este problema (Cao et al., 2019; Gómez-Ríos et al., 2019; Hu et al., 2020; Yang et al., 2020). El principal objetivo de la aumentación de datos es generar un conjunto de datos sintéticos que

se incorpore a los datos observados en la vida real (Cao et al., 2019; Gómez-Ríos et al., 2019; Hu et al., 2020; Yang et al., 2020).

3.2.2.1 Casos por tipo de probabilidad

Para calcular el porcentaje de casos de sospecha de Trombosis Venosa Profunda en cada uno de los tipos de riesgo que propone Wells, se toman los datos de (P. S. Wells et al., 2006), donde menciona que de todos los casos observados, el 19% fueron diagnosticados como TVP, mientras que el 81% restante concluyeron con un diagnóstico diferente. Así mismo, en la misma fuente se menciona que en los casos detectados como riesgo bajo, únicamente el 5% de los casos fueron diagnosticados como positivos a una TVP, mientras que en riesgo medio se diagnosticaron el 17%, y en riesgo alto se diagnosticó el 53% de los casos. Con esta información, se plantea un modelo de ecuaciones lineales (ver Ecuación 3.1) para determinar el porcentaje de casos que son observados en cada tipo de riesgo. La primera ecuación indica que el 100% de casos de cada uno de los 3 riesgos que establece la Escala de Wells conforman el 100% del conjunto de datos. La segunda ecuación representa el porcentaje de casos positivos que ocurren por cada tipo de riesgo; en donde x representa los casos de riesgo bajo (5%), y los casos de riesgo moderado (17%) y finalmente z , los casos de riesgo alto (53%), los cuales acumulan un 19%. La tercera ecuación representa la distribución de casos negativos del conjunto de datos, indicando que el 95% de casos de riesgo bajo, más el 83% de casos de riesgo medio, adicional al 47% de casos de riesgo alto, conforman el 81% del conjunto de datos que a su vez representa a los casos clínicos positivos.

<i>Riesgo</i> <i>Bajo</i>		<i>Riesgo</i> <i>Moderado</i>		<i>Riesgo</i> <i>Alto</i>		
x	+	y	+	z	=	100
$0.05x$	+	$0.17y$	+	$0.53z$	=	19
$0.95x$	+	$0.83y$	+	$0.47z$	=	81

Ecuación 3.1. Sistema de ecuaciones para determinar el porcentaje de casos que se presentaron de cada tipo de riesgo.

Resolviendo la Ecuación 3.1 se obtiene que $x = 60.9\%$, lo cual representa los casos de riesgo bajo, $y = 13.2\%$ que corresponde al riesgo moderado y, $z = 25.9\%$ identificados como riesgo alto. Posteriormente con esa información, se calcula la distribución de casos de sospecha de TVP diagnosticados como positivos y negativos en cada tipo de riesgo que se utiliza en el conjunto de datos de entrenamiento y validación, los cuales se muestran en la Tabla 3.3. Así, el conjunto de datos se divide en 6,090 casos de riesgo bajo, 1,320 casos de riesgo medio, y 2590 casos de riesgo alto. De esa distribución, se desprende que 305 casos se etiquetan como riesgo bajo con diagnóstico positivo, 5,785 son casos de riesgo bajo con diagnóstico negativo, 224 son de riesgo medio con diagnóstico positivo, 1,096 se clasifican como riesgo medio con diagnóstico negativo, 1,373 con casos de riesgo alto con diagnóstico positivo, y 1,217 son casos de riesgo alto negativo. Como se puede observar, la distribución de datos generada se encuentra altamente desbalanceada, lo que puede ocasionar errores en la clasificación. Sin embargo, según (Brush et al., 2017), la habilidad de los expertos en el diagnóstico no se debe a un proceso de razonamiento distintivo, sino que depende de la capacidad del médico para acceder a su experiencia pasada para poder generar posibles diagnósticos. Por esta razón, es preferible entrenar el modelo de clasificación a partir de un conjunto de datos donde la distribución de sospecha positiva de TVP es consistente con la forma en que ocurre el fenómeno de TVP en la vida real, por consiguiente es necesario tomar en cuenta el impacto que presentan cada uno de los síntomas en el diagnóstico de TVP, justo como lo hacen los médicos.

Tabla 3.3 Distribución de casos en el conjunto de datos para 10,000 instancias.

	Riesgo bajo	Riesgo moderado	Riesgo alto
Diagnóstico de TVP	305	224	1,373
Otro diagnóstico	5,785	1,096	1,217

Para generar un conjunto de datos que tenga una mayor representatividad con respecto a la ocurrencia del fenómeno en la vida real, es de suma importancia tomar en cuenta también en el proceso de aumentación de datos la distribución porcentual de los síntomas de la Trombosis Venosa Profunda que se analizan en la Escala de Wells, y que fueron observados en los casos reales.

3.2.2.2 Distribución de factores de Wells en el Conjunto de Datos

El cálculo de la distribución de factores de Wells en el conjunto de datos está basado en el análisis de los registros médicos obtenidos de un hospital mexicano. Para ello, se contabiliza la observación de todos los factores establecidos en la Escala de Wells que pueden contribuir al diagnóstico de la TVP en cada uno de los registros, y se suma el total de factores, resultando un total de 280 factores observados en los registros médicos (ver Tabla 3.4).

Tabla 3.4 Cantidad de factores establecidos en la Escala de Wells que pueden contribuir al diagnóstico de la TVP (P. S. Wells et al., 2006).

Factores	Ocurrencias
Cáncer	17
Inmovilización MPI	40
Reposo / Cirugía	35
Dolor en Venas	37
Inflamación pierna	32
Inflamación de tobillo	39
Edema	35
Venas superficiales	33
TVP previa	12
Total	280

En los resultados se puede observar que la inmovilización del miembro pélvico inferior es la característica más observada en los pacientes con sospecha de TVP, seguido de inflamación de tobillo, y dolor en las venas. Además, se observa que en la mayoría de los casos no se había presentado un episodio previo de TVP.

Posteriormente se divide el número de observaciones de cada factor sobre el total de factores de riesgo de la escala de Wells que se identifican en el estudio, tal como se muestra en la Ecuación 3.2.

$$\%factor_i = \frac{\text{ocurrencias del factor}_i}{\text{ocurrencias de todos los factores}}$$

Ecuación 3.2 Fórmula para calcular el porcentaje de ocurrencia de cada factor de Wells.

En términos porcentuales, no hubo variación significativa entre la inmovilización del miembro pélvico inferior (14%), tener una cirugía en las semanas previas a la sospecha de TVP (14%), inflamación en el tobillo (14%), y un edema en la pierna (14%). Incluso, el dolor en las venas (13%), venas superficiales (13%) que generan piernas enrojecidas, e inflamación en la pierna (11%) tampoco presentan gran variación porcentual, mientras que en el caso del cáncer (6%) y haber padecido una TVP previamente (4%) si presentan una variación significativa en términos de una presencia menor en los pacientes con sospecha de TVP (Figura 3.5).

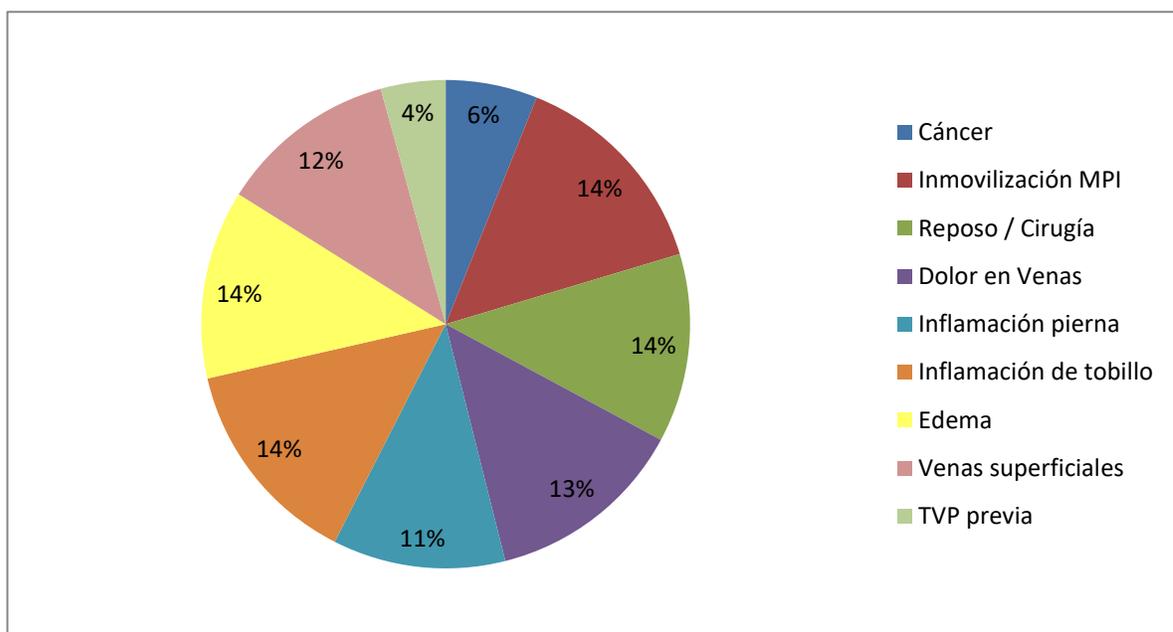


Figura 3.5 Cantidad y porcentaje de ocurrencia de factores de riesgo de Wells observados en los registros médicos del hospital ISSSTE, Ensenada, B.C. (Elaboración propia)

Lo anterior se deberá reflejar en el algoritmo de generación del conjunto de datos, que se presenta en la subsección 3.2.2.3.

3.2.2.3 Algoritmo propuesto para aumentación de datos

La matriz que contiene el conjunto de datos de entrenamiento se genera fila por fila, en donde cada fila representa a un paciente con síntomas de TVP; el pseudocódigo utilizado se puede ver en la Figura 3.6.

```

DataSetGenerator(initial_row, final_row, risk_type)
#El conjunto de datos es almacenado en una matriz de enteros de 10,000 renglones x 12 columnas.
#La función random(base, range) genera números aleatorios del 0 al -1.
#Regresa la suma de la base y el número aleatorio generado.
1: for i=initial_row:final_row
2:   switch(risk_type)
3:     case 1: symptoms_observed=1           # El riesgo bajo tiene un síntoma observable
4:       diagnosis=1                       # risk_type == 1 es un riesgo bajo con diagnóstico positivo
5:     case 2: symptoms_observed=1           # El riesgo bajo tiene un síntoma observable
6:       diagnosis=0                       # risk_type == 2 es un riesgo bajo con diagnóstico negativo
7:     case 3: symptoms_observed=random(2,2) # El riesgo medio tiene 2 o 3 síntomas observables
8:       diagnosis=1                       # risk_type == 3 es un riesgo medio con diagnóstico positivo
9:     case 4: symptoms_observed=random(2,2) # El riesgo medio tiene 2 o 3 síntomas observables
10:      diagnosis=0                       # risk_type == 4 es un riesgo medio con diagnóstico negativo
11:     case 5: symptoms_observed=random(4,3) # El riesgo alto tiene de 4 a 6 síntomas observables
12:       diagnosis=1                       # en este estudio; en la vida real pueden ser hasta 9
13:     case 6: symptoms_observed=random(4,3) # risk_type == 5 es un riesgo alto con diagnóstico positivo
14:       diagnosis=0                       # El riesgo alto tiene de 4 a 6 síntomas observables
15:   end switch                             # risk_type == 6 es un riesgo alto con diagnóstico negativo
16:   for j=1:symptoms_observed
17:     symptom=random(2,9)                 # asignación aleatoria del síntoma observable
18:     dataset[i,0]=random(1,8)           # asignación aleatoria del intervalo de edad del paciente
19:     dataset[i,1]=random(0,2)           # asignación aleatoria del género del paciente
20:     while(dataset[i,symptom]==1)      # si el síntoma ya fue elegido, se asigna el siguiente disponible
21:       symptom++;
22:       if(symptom==11)
23:         symptom=1
24:       end if
25:     end while
26:     dataset[i,symptom]=1               # el síntoma se asigna como positivo (0 es negativo)
27:   end for
28: end for

```

Figura 3.6 Pseudocódigo para generar el conjunto de datos sintéticos.

De manera inicial, se determina el número de “pacientes” (instancias) que serán generados, y se calcula el tipo de riesgo y el diagnóstico correspondiente a cada uno de ellos; para este trabajo se tesis se decide utilizar un conjunto de datos de 10,000 instancias. Así mismo, se calcula la cantidad de ocurrencias de cada uno de los factores de riesgo de la escala de Wells para conocer a cuántos pacientes se asignará ese factor de riesgo.

En cada una de las filas se determina de manera aleatoria la cantidad de factores de riesgo de Wells que serán marcados como “observados”. En el caso de los casos de riesgo bajo, en todas las filas se especifica únicamente un factor, en los casos de riesgo medio las filas pueden contener 2 o 3 factores observados, mientras que en riesgo alto los factores observados varían en un intervalo de 4 a 7.

Una vez determinada la cantidad de factores de riesgo observados, por cada uno de ellos se genera un número aleatorio que se utiliza para indicar el factor de riesgo observado en ese paciente; en caso de que el número aleatorio represente a un factor de riesgo que ya fue marcado como “observado”, se genera un nuevo número aleatorio hasta que se

encuentre alguno que represente a un factor no observado previamente, se marca la columna correspondiente con el valor 1, y se reduce en uno la cantidad de pacientes a los que se asignará dicho factor. Finalmente, se asigna el valor correspondiente al diagnóstico del paciente.

3.2.3 Pre-procesamiento del conjunto de datos

Para entrenar al sistema experto se recolectan datos de registros médicos de hospitales mexicanos, así como de bases de datos obtenidas de artículos médicos (Chen et al., 2019; Hull et al., 2016; Modi et al., 2016; Tang et al., 2018). De cada registro médico se extraen aquellos datos que coinciden con los primeros nueve de los criterios establecidos por Wells para la detección de TVP; además de la edad y el género.

A partir de los datos obtenidos se diseña una matriz, la cual representa los factores de riesgo de un paciente, además del género y un rango de edad. El diagnóstico puede tomar un valor positivo o negativo. Las filas de la matriz representan los casos clínicos o pacientes, mientras que las columnas representan las entradas al sistema experto, es decir, los nueve primeros criterios tomados del modelo de Wells, además del género y la edad del paciente. Cada una de las entradas al sistema se asocia con un valor numérico real, para formar la matriz numérica que sirva de entrenamiento al sistema.

Para identificar el género del paciente se asignan los valores cero y uno como se presenta en la Tabla 3.5.

Tabla 3.5 Asignación por género.

Género	Valor numérico
Femenino	0
Masculino	1

También se agrupan los casos clínicos utilizando intervalos de edades de los pacientes como se muestra en la Tabla 3.6. Cada intervalo es codificado por un valor numérico entre el uno y el nueve.

Tabla 3.6 Asignación por edad.

Edad	Ciclo de vida	Valor numérico
0-5	Infancia	1
6-12	Niñez media	2
13-20	Jóvenes	3
21-39	Adultos jóvenes	4
40-49	Adultos medios	5
50-59	Adultos maduros	6
60-69	Vejez inicial	7
70-84	Vejez intermedia	8
85-120	Vejez avanzada	9

Además, en la Tabla 3.7 se puede observar que a cada criterio de Wells asentido por el paciente o el médico se le asigna el valor de uno, mientras que a la ausencia del mismo se le asigna el valor de cero.

Tabla 3.7 Asignación por factor de riesgo.

Característica Clínica	Presente	Ausente	Forma de recolección de información
-Cáncer activo	1	0	Entrevistar al paciente
-Parálisis, paresia, inmovilización con yeso del miembro sospechoso	1	0	Entrevistar al paciente
-Reposo en cama reciente de más de 3 días o cirugía mayor	1	0	Entrevistar al paciente
-Tumefacción de todo un miembro	1	0	Entrevistar al paciente
-Dolor localizado en el trayecto venoso profundo	1	0	Entrevistar al paciente / Exploración física
-Tumefacción unilateral de la pantorrilla de más de 3 cm	1	0	Exploración física
-Edema con fovea	1	0	Exploración física
-Venas superficiales (no varicosas) colaterales	1	0	Entrevistar al paciente / Exploración física
-Caso previo de TVP	1	0	Entrevistar al paciente

3.2.4 Estrategias de Entrenamiento de la Red Neuronal

El diagnóstico médico no es un proceso universal, ya que depende de la experiencia y conocimientos del médico. Por ello, se considera necesario evaluar el entrenamiento del modelo de predicción desde distintas perspectivas. Así, para entrenar el modelo de clasificación, se utiliza el conjunto de datos sintéticos, representado a través de una matriz de 10,000 x 11 (matriz de entradas), donde cada renglón se considera un caso clínico compuesto por dos variables demográficas y nueve criterios presentes en la escala de Wells (P. S. Wells et al., 2006), así como un vector columna de 10,000 x 1 datos (vector de salidas) que representa el diagnóstico asociado a cada caso clínico.

Es importante destacar que una de las características que presenta el conjunto de datos, similar a como ocurre en la vida real, es que dos o más pacientes con síntomas y datos demográficos similares pueden presentar un diagnóstico diferente. Así, los casos con datos de entrada idénticos y diagnóstico diferenciado los consideramos como casos con incertidumbre. Esto a su vez puede generar confusión en el diagnóstico, tanto a los médicos como a los modelos de clasificación que hayan sido entrenados para esta tarea.

Se encuentra que 2,250 pacientes, distribuidos en 1,248 casos diferentes, tienen 0% de probabilidad de sufrir TVP. Así mismo, 976 de los pacientes distribuidos en 802 casos diferentes presentan una probabilidad de 100% en el diagnóstico positivo de TVP. Esto implica que 6,774 pacientes distribuidos en 550 casos diferentes presentan incertidumbre en el diagnóstico (ver Figura 3.7).

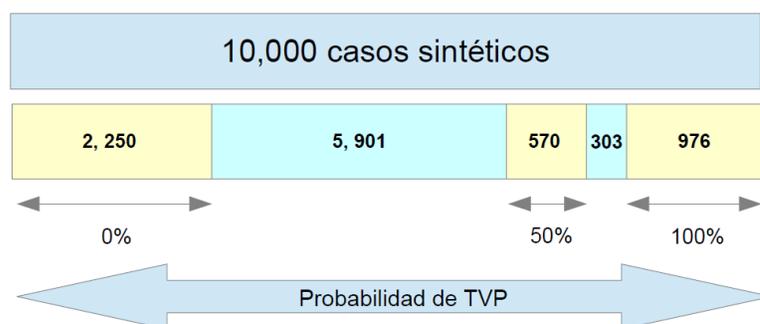


Figura 3.7 Distribución de probabilidad.

Una vez caracterizado el conjunto de datos, se consideran y analizan tres diferentes estrategias de entrenamiento de la red neuronal, los cuales están basados en las estrategias de diagnóstico que se identificaron en entrevistas realizadas a 5 médicos de distintos grados de especialidad.

Por esta razón, en este trabajo se presentan tres estrategias de pre-procesamiento del conjunto de datos para entrenamiento, los cuales son utilizados para buscar reducir el impacto que tiene la incertidumbre en el diagnóstico de los casos en el modelo de clasificación.

Es importante mencionar que independientemente de la estrategia utilizada para el entrenamiento, se prueba la independencia del modelo con respecto a los datos a través de la prueba de la validación cruzada, con un valor $k=10$, y se lleva a cabo una validación externa del modelo con el conjunto de datos reales obtenidos de un hospital público.

Estrategia de Entrenamiento 1. La configuración de este conjunto de datos de entrenamiento, corresponde a la estrategia de diagnóstico identificada en el estudio con los médicos, la cual indica que todos los casos previos son tomados como evidencia para generar un diagnóstico. Entonces, en la primera estrategia, el entrenamiento se realiza con el conjunto de datos original, tal como fueron generados por medio del algoritmo de aumentación. Esto significa que se entrena el modelo con las 10,000 instancias generadas, (ver Tabla 3.8). Al conjunto de datos resultante de aplicar esta estrategia de pre-procesamiento se le llama Dataset1.

Tabla 3.8 Fragmento de datos del Dataset1.

Instancia	Entradas											Salida
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	Y1
1	5	1	0	1	0	0	0	0	0	0	0	1
2	6	0	0	0	1	1	0	1	0	1	0	1
3	4	0	0	1	0	0	1	1	0	1	0	1
4	5	1	1	0	1	1	0	0	0	0	1	0
5	6	0	0	1	0	0	0	0	0	0	0	0
6	8	1	0	0	1	1	0	1	0	1	0	0
7	7	1	0	1	0	0	1	1	0	1	0	0
8	8	0	1	0	1	1	0	0	0	0	1	0
9	6	1	1	1	0	0	0	0	1	0	0	0
10	5	0	0	1	0	0	0	1	0	1	0	1
11	5	0	0	1	0	0	0	1	0	1	0	1
12	6	1	1	0	1	1	0	1	0	0	1	1
13	6	1	1	0	1	1	0	1	0	0	1	1
14	6	1	1	0	1	1	0	1	0	0	1	0
...					
9998	4	1	1	0	1	1	0	0	0	0	1	1
9999	8	1	0	1	0	0	0	0	0	0	0	0
10000	7	0	0	0	1	1	0	1	0	1	0	0

Estrategia de Entrenamiento 2. La configuración de este conjunto de datos de entrenamiento corresponde a la segunda estrategia de diagnóstico identificada, en la que los médicos expresan que ante la duda sobre el diagnóstico, asumen al paciente como positivo en trombosis venosa profunda. En esta estrategia se presentan tres momentos:

En el primer momento, las 10,000 instancias son organizadas por grupos. Cada grupo contiene las instancias que coinciden exactamente en sus factores de riesgo, intervalo de edad y género, es decir, cuyos valores de la matriz de entrada son idénticos, sin importar el valor asociado en el vector de salida. Como resultado, se obtienen 2600 grupos de tamaño no uniforme.

En un segundo momento, se calculan las probabilidades de ocurrencia de diagnóstico positivo en cada uno de los grupos (Ecuación 3.3).

$$Y_i = \frac{\sum_{j=1}^{n_i} \text{Caso}_j \text{Diagnostico}}{n_i}$$

Ecuación 3.3 Fórmula para calcular la probabilidad de TVP positiva en cada grupo.

En donde,

n_i : cantidad de instancias en el grupo i

$i = 1, 2, \dots, 2600$ y,

Caso_jDiagnostico: representa el valor de la variable de salida original ($Y1$) de cada una de las instancias que forman parte de ese grupo, la cual tiene un valor de 0 cuando el diagnóstico de TVP es negativo, y 1 cuando el diagnóstico de TVP es positivo.

A partir del cálculo de Y_i para cada uno de los 2,600 grupos se genera un nuevo vector Y_p de igual dimensión que el vector de la salida original $Y1$ (10,000x1), el cual en lugar de contener valores binarios como $Y1$, contiene la probabilidad de ocurrencia Y_i calculada.

En el tercer momento, se identifican las instancias cuya probabilidad de ocurrencia son mayores a 0.33 y menores a 0.67. Una vez identificadas estas instancias, se genera un nuevo vector de salida Y' , cuyos valores dependen de las instancias que caen dentro del intervalo mencionado, como de las que no lo hacen, es decir, cuando las instancias examinadas quedan fuera del intervalo $0.33 < TVP < 0.67$ el valor que se guarda en $Y1'$ es el mismo que tiene la instancia originalmente en el vector de salida $Y1$. Por el contrario, si las instancias observadas caen dentro del intervalo $0.33 < TVP < 0.67$ se les asigna un 1 en su posición correspondiente del vector $Y1'$ para indicar que el diagnóstico es positivo, sin importar el diagnóstico original $Y1$ que les asigna el algoritmo de aumentación.

En la Tabla 3.9 se puede observar por ejemplo, que en la instancia 14 se modifica el diagnóstico negativo (valor de 0) que tiene en $Y1$ a un diagnóstico positivo (valor de 1) en el vector $Y1'$ debido a que la probabilidad de ocurrencia Y_p es de 0.66, la cual cae dentro del intervalo $0.33 < TVP < 0.67$.

Resumiendo, en el Dataset2 se consideran los 10,000 valores de entradas originales como predictores pero se modifican a 1 (diagnóstico positivo), todos los valores del vector de salida, cuya probabilidad de ocurrencia se encuentre dentro del intervalo $0.33 < TVP < 0.67$. El vector generado $Y1'$ reemplaza vector $Y1$ como vector de salida para el entrenamiento.

Tabla 3.9 Fragmento de datos del Dataset2.

Instancia	Entradas											Salida original	Probabilidad de ocurrencia	Salida modificada
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	Y1	Yp	Y1'
1	5	1	0	1	0	0	0	0	0	0	0	1	1	1
2	6	0	0	0	1	1	0	1	0	1	0	1	1	1
3	4	0	0	1	0	0	1	1	0	1	0	1	1	1
4	5	1	1	0	1	1	0	0	0	0	1	0	0.2	0
5	5	1	1	0	1	1	0	0	0	0	1	0	0.2	0
6	5	1	1	0	1	1	0	0	0	0	1	1	0.2	1
7	5	1	1	0	1	1	0	0	0	0	1	0	0.2	0
8	5	1	1	0	1	1	0	0	0	0	1	0	0.2	0
9	6	1	1	1	0	0	0	0	1	0	0	0	0	0
10	5	0	0	1	0	0	0	1	0	1	0	1	1	1
11	5	0	0	1	0	0	0	1	0	1	0	1	1	1
12	6	1	1	0	1	1	0	1	0	0	1	1	0.66	1
13	6	1	1	0	1	1	0	1	0	0	1	1	0.66	1
14	6	1	1	0	1	1	0	1	0	0	1	0	0.66	1
...					
9998	4	1	1	0	1	1	0	0	0	0	1	1	1	1
9999	8	1	0	1	0	0	0	0	0	0	0	0	0	0
10000	7	0	0	0	1	1	0	1	0	1	0	0	0	0

Estrategia de Entrenamiento 3.

Esta estrategia, corresponde a la tercera estrategia de diagnóstico identificada en las entrevistas a los médicos. En esta estrategia, el entrenamiento se lleva a cabo con el análisis de la probabilidad de que un caso resulte positivo. Para esto, al igual que en el primer momento de la Estrategia 2 se identifican las instancias que coincidan exactamente en sus factores de riesgo, intervalo de edad y género; recordando que se obtienen 2600 grupos diferentes entre sí. En la Tabla 3.10, se observan en color gris las filas que representan diez de los 2600 grupos que se forman. Las filas que no tienen color, pertenecen a alguno de estos grupos debido a la coincidencia en sus entradas.

Por ejemplo, en la Tabla 3.10 las instancias 4-5, 7-8 pertenecen a un primer grupo junto con la instancia 6. Mientras que la instancia 10 forma parte de un segundo grupo con la instancia 11; a su vez las instancias 12 y 14 corresponden en unión con la instancia 13 a un tercer grupo. De ahí que de las 10,000 instancias que se tienen originalmente, se reduzcan a 2,600 instancias para utilizarse en la Estrategia 3.

Por otro lado, el vector de salida se conforma por la probabilidad de que los pacientes que cumplan con las características demográficas y factores de riesgo de cada caso sean diagnosticados con TVP.

Tabla 3.10 División en grupos del conjunto de datos de 10,000 instancias.

Instancia	Entradas											Salida original	Prob. de ocurrencia
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	Yi	Yp
1	5	1	0	1	0	0	0	0	0	0	0	1	1
2	6	0	0	0	1	1	0	1	0	1	0	1	1
3	4	0	0	1	0	0	1	1	0	1	0	1	1
4	5	1	1	0	1	1	0	0	0	0	1	0	0.2
5	5	1	1	0	1	1	0	0	0	0	1	0	0.2
6	5	1	1	0	1	1	0	0	0	0	1	1	0.2
7	5	1	1	0	1	1	0	0	0	0	1	0	0.2
8	5	1	1	0	1	1	0	0	0	0	1	0	0.2
9	6	1	1	1	0	0	0	0	1	0	0	0	0
10	5	0	0	1	0	0	0	1	0	1	0	1	1
11	5	0	0	1	0	0	0	1	0	1	0	1	1
12	6	1	1	0	1	1	0	1	0	0	1	1	0.66
13	6	1	1	0	1	1	0	1	0	0	1	1	0.66
14	6	1	1	0	1	1	0	1	0	0	1	0	0.66
...					
9998	4	1	1	0	1	1	0	0	0	0	1	1	1
9999	8	1	0	1	0	0	0	0	0	0	0	0	0
10000	7	0	0	0	1	1	0	1	0	1	0	0	0

Siendo entonces el conjunto de datos para el entrenamiento la matriz de entradas formada por los 2600 grupos que resultan del primer momento de la estrategia de Entrenamiento 2, y un vector de probabilidades de salida Y_i calculado en el segundo momento de la misma Estrategia de acuerdo a la Ecuación 3.3. Al conjunto de datos resultante de aplicar esta estrategia de pre-procesamiento lo llamamos Dataset3 (Tabla 3.11).

Tabla 3.11 Fragmento de datos del Dataset3.

Instancia	Entradas											Salidas
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	Yi
1	5	1	0	1	0	0	0	0	0	0	0	1
2	6	0	0	0	1	1	0	1	0	1	0	1
3	4	0	0	1	0	0	1	1	0	1	0	1
4	5	1	1	0	1	1	0	0	0	0	1	0.2
5	6	1	1	1	0	0	0	0	1	0	0	0
6	5	0	0	1	0	0	0	1	0	1	0	1
7	6	1	1	0	1	1	0	1	0	0	1	0.66
...					
2598	4	1	1	0	1	1	0	0	0	0	1	1
2599	8	1	0	1	0	0	0	0	0	0	0	0
2600	7	0	0	0	1	1	0	1	0	1	0	0

3.3 Validación del Modelo de predicción de la trombosis profunda

Uno de los aspectos importantes a considerar en la utilización de redes neuronales, es que el modelo sea independiente del conjunto de datos utilizado para su entrenamiento y validación; El método de validación cruzada es quizás el más utilizado para validar el grado de precisión de un modelo de red neuronal con independencia de los datos. Consiste en dividir el conjunto de datos en k segmentos, y durante k veces se elige un segmento distinto para validar el modelo, mientras que los k-1 segmentos restantes se utilizan para entrenar a la red neuronal (ver Figura 3.8). La precisión de los datos se toma del promedio de la exactitud obtenida en cada iteración.

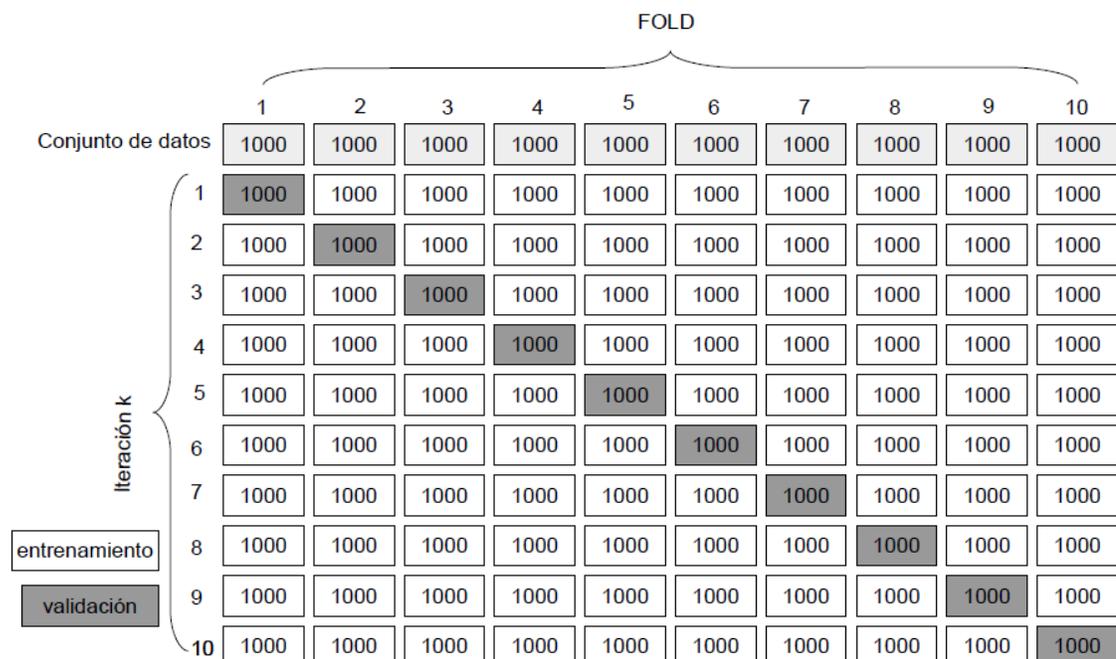


Figura 3.8 Validación Cruzada del conjunto de datos sintético con $k=10$.

De acuerdo a (Rodríguez et al., 2010), el valor de k más recomendado y más comúnmente utilizado es 10. En este trabajo nos apegamos a esa recomendación, por lo que obtuvimos 10 segmentos de 1,000 datos cada uno. Posteriormente, la red neuronal es entrenada y validada, generando resultados de acuerdo a la matriz de confusión que se muestra en la Tabla 3.12, en donde se establecen cuatro categorías de resultados: i) acierto positivo (TP), el cual ocurre cuando tanto la salida del caso a validarse como la salida estimada por la red neuronal coinciden en un diagnóstico positivo de TVP, ii) acierto negativo (TN), el cual ocurre cuando tanto la salida del caso a validarse como la salida estimada por la red neuronal coinciden en un diagnóstico negativo de TVP, iii) falso positivo (FP), que se presenta cuando la salida del caso a validarse es un diagnóstico negativo, mientras que el modelo estima un diagnóstico positivo, y iv) falso negativo (FN), que se presenta cuando la salida del caso a validarse es un diagnóstico positivo, mientras que el modelo estima un diagnóstico negativo.

Tabla 3.12 Matriz de Confusión para Modelo de Redes Neuronales.

Diagnóstico Real	Diagnóstico Estimado	
	TVP Positiva	TVP Negativa
TVP Positiva	TP	FN
TVP Negativa	FP	TN

La importancia de conocer el significado de cada uno de los elementos de la matriz de confusión, así como calcular cuántas ocurrencias hay de cada uno de ellos, está relacionado con el cálculo de las métricas de relevancia para el modelo de predicción.

3.4 Diseño del modelo clasificador

Dentro de las principales aplicaciones de las redes neuronales artificiales, se encuentra la clasificación. En el caso particular del ámbito médico, se utiliza para el diagnóstico de una enfermedad, y en particular en este trabajo se utiliza para definir si un paciente debe ser diagnosticado o no con TVP.

Previamente se estableció la problemática que existe para diseñar un clasificador con un modelo de red neuronal que sea entrenado con datos desbalanceados. Esto se puede apreciar a mayor detalle en la Figura 3.9, en donde se puede ver que en condiciones ideales con datos balanceados (A) es relativamente fácil establecer un umbral para clasificar, ya que debería estar en medio de los valores que arroja el modelo de predicción. Sin embargo, cuando se tienen datos desbalanceados (Figura 3.9 B), tal como en los datos analizados por (P. S. Wells et al., 2006), se debe trabajar en elegir un umbral que sea apropiado para el problema que se está abordando.

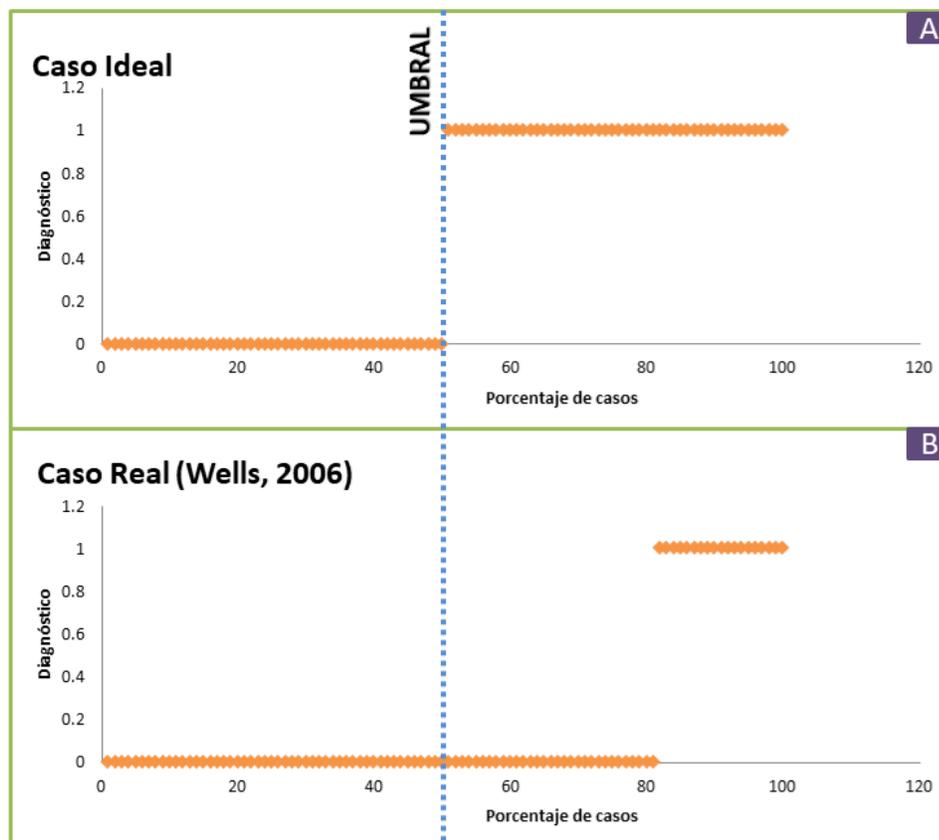


Figura 3.9 Representación de un umbral de clasificación. A. En un caso ideal con datos balanceados. B. En un caso real con datos desbalanceados.

Por ello, se trabaja en diseñar un algoritmo para buscar el umbral que permita obtener el mejor porcentaje con cada modelo. En este punto es importante mencionar que si bien la exactitud es la medida de valor por excelencia en el mundo de las redes neuronales artificiales, no lo es tanto en el trabajo de hospital, ya que hay otros valores que deben ser tomados en cuenta. Por ejemplo, en el conjunto de datos que se utiliza en este trabajo de tesis, se tiene una proporción del 81% de casos negativos con respecto al 19% de casos positivos; si un médico diagnostica que todos los casos son negativos, entonces tiene una exactitud del 81%, ya que la consecuencia de sus diagnósticos es que no falla con ningún paciente negativo, sin embargo, no acierta a ningún caso positivo. Esto a su vez trae la complicación de que entonces, se tiene mayor probabilidad de alcanzar una mejor exactitud cuando se declaran todos los casos como pacientes negativos, aunque las consecuencias en la salud de los pacientes enfermos pueden ser graves e incluso fatales.

Por otro lado, se puede pensar que la solución es asumir que todos los pacientes potenciales están enfermos, de esta forma ningún paciente se queda sin atención médica. Sin embargo, el problema con esta estrategia es que hay costos económicos asociados a la atención médica de los pacientes, además de que también puede haber consecuencias en la salud de los pacientes por someterse a un tratamiento médico no requerido.

Por esa razón, cobra relevancia el uso de más de una métrica para conocer el funcionamiento del modelo de predicción. Por un lado, no se puede dejar de lado la exactitud del modelo, ya que al final de cuentas nos ayuda a conocer la cantidad de aciertos del modelo de predicción. Por otro lado, se vuelve relevante la proporción de casos positivos diagnosticados correctamente.

El algoritmo de la Figura 3.10 describe el enfoque propuesto para mejorar el umbral de clasificación para maximizar ya sea la exactitud o lograr una exactitud superior a la alcanzada por modelos probabilísticos de forma tradicional manteniendo el mismo valor de recall (nombre con el que también se le conoce a la sensibilidad).

```

trainANNModel();
highestAccuracy=-1;
for x=1:1000
    umbral=umbral+0.001;
    tp=0;
    tn=0;
    fp=0;
    fn=0;

    for m=1:10000
        if datasetOutput(m)==1
            if modelOutputs(m)>=umbral
                tp=tp+1;
            else
                fn=fn+1;
            end
        else
            if modelOutputs(m)<umbral
                tn=tn+1;
            else
                fp=fp+1;
            end
        end
    end
    accuracy=((tp+tn)/10000)*100 ;
    if accuracy>highestAccuracy
        optimizedUmbral=umbral;
    end
end
end

```

Figura 3.10 Algoritmo de optimización de umbral de clasificación.

El modelo de clasificación del diagnóstico de la TVP propuesto realiza la clasificación a través de una función umbral, la cual se define mediante un proceso de optimización; el cual consiste en evaluar la exactitud o el recall del modelo de predicción propuesto, con respecto a un conjunto de datos, tomando en cuenta un conjunto de umbrales que tienen una variación entre ellos mismos de 0.001, en un rango de 0 a 0.99. Para cada umbral se calcula la exactitud y el recall. Para alcanzar la máxima Exactitud, el umbral se define en la Ecuación 3.4, la cual indica que cuando la salida de la red neuronal es mayor o igual a 0.484; el diagnóstico de TVP se considera positivo $f(x) = 1$, en caso contrario, es negativo $f(x) = 0$.

$$\text{si } x \geq 0.484 \text{ entonces } f(x) = 1, \quad \text{si no } f(x) = 0$$

Ecuación 3.4 Función umbral para obtener máxima Exactitud.

Al analizar los resultados, se aprecia que al variar el valor del umbral, no se modifica únicamente la exactitud del modelo, sino también el recall. Debido a ello, se establece la necesidad de obtener dos umbrales: uno que priorice la exactitud del modelo, y otro que priorice el recall del modelo, ya que esta última métrica está asociada a clasificar correctamente a un paciente que tiene la enfermedad, evitando en la medida de lo posible generar falsos negativos que puedan afectar la salud del paciente.

Para obtener un recall igual o menor que el presentado por Wells (P. S. Wells et al., 2006), el umbral se define en la Ecuación 3.5, la cual indica que cuando la salida de la red neuronal es mayor o igual a 0.138, el diagnóstico de TVP se considera positivo $f(x) = 1$, en caso contrario, es negativo $f(x) = 0$.

$$\text{si } x \geq 0.138 \text{ entonces } f(x) = 1, \quad \text{si no } f(x) = 0$$

Ecuación 3.5 Función umbral para obtener la máxima exactitud con un Recall igual o menor al presentado por Wells (P. S. Wells et al., 2006).

Capítulo 4 Resultados

En este capítulo se presentan los resultados del entrenamiento del modelo de clasificación desarrollado durante la presente investigación. De inicio, es importante mencionar que la métrica principal en el modelo de RNA es la exactitud, aunque esta se complementa principalmente con el recall. Así, se tiene que la mayor exactitud que alcanza el modelo propuesto es de 90.99%; sin embargo, se debe mencionar que debido al porcentaje de incertidumbre en los datos propia de las enfermedades que requieren de un diagnóstico diferencial, el valor máximo posible a alcanzar con este tipo de trabajos es de 91.85%, lo que significa que el modelo propuesto predice correctamente el 99.06% de los casos posibles.

Por otra parte, para controlar la cantidad de errores negativos (FN), lo que implica que pacientes enfermos sean diagnosticados como pacientes sanos, con los problemas a la salud que eso puede conllevar, se genera un nuevo umbral de clasificación que maximiza la exactitud al mismo tiempo que iguala los errores negativos que presenta Wells, lo que repercute directamente en la métrica recall. La exactitud que se obtiene con ese esquema de clasificación queda de 84.95%.

En ambos casos, es importante mencionar que superan la exactitud calculada al evaluar el conjunto de datos por medio de la Escala de Wells.

Además, en el capítulo se presenta una comparación de resultados con otros algoritmos de aprendizaje, conservándose la ventaja en exactitud para el modelo RNA propuesto.

Finalmente, se presenta una propuesta de escenario de implementación del clasificador de RNA, el cual establece a través de un diagrama de flujo el protocolo a seguir en la unidad médica de atención primaria cuando se utilice el SADC presentado en este trabajo de tesis para generar un diagnóstico de TVP.

4.1 Evaluación de resultados del modelo de predicción de la trombosis venosa profunda

Uno de los elementos más importantes de una investigación científica es medir los resultados, para dimensionar el aporte a realizar. Por ello, es necesario comparar los resultados del modelo propuesto con respecto a los resultados que se obtienen al evaluar el conjunto de datos con la Escala de Wells, así como comparar el desempeño del modelo de clasificación con respecto a otros algoritmos de aprendizaje automático, los cuales se discuten en la sección 4.5.

Además, es necesario utilizar algunos términos que son de utilidad para establecer una mejor ponderación de los resultados. De las métricas elegibles para un modelo de redes neuronales, se toman en cuenta las mostradas en la Ecuaciones 4.1-4.4 (Fu et al., 2017; Sadrawi et al., 2018).

$$\text{Sensibilidad} = \frac{TP}{FN + TP}$$

Ecuación 4.1 Sensibilidad del modelo.

$$\text{Especificidad} = \frac{TN}{FP + TN}$$

Ecuación 4.2 Especificidad del modelo.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Ecuación 4.3 Precisión del modelo.

$$\text{Exactitud} = \frac{TP + TN}{TP + FP + TN + FN}$$

Ecuación 4.4 Exactitud del modelo.

La sensibilidad o recall, es el porcentaje de casos verdaderos que se predicen de forma correcta, con respecto al total de casos positivos existente en el conjunto de datos. En el caso del diagnóstico de la TVP, permite conocer el porcentaje de todos aquellos pacientes que realmente tienen una TVP y que son diagnosticados correctamente. Entonces, mientras mayor sea el recall o la sensibilidad, mayor es la cantidad de pacientes que tendrán un seguimiento médico adecuado.

La especificidad es el porcentaje de casos negativos que se predicen de forma correcta con respecto al total de casos negativos en el conjunto de datos. En el caso del diagnóstico de la TVP, indica el porcentaje de aquellos pacientes sanos que son diagnosticados como tal.

La precisión es el porcentaje de casos que son diagnosticados como positivos, con respecto al total de casos positivos que diagnostica el modelo de precisión. En otras palabras, es el porcentaje de casos positivos diagnosticados correctamente con respecto al total de casos positivos que predice el sistema.

Finalmente se tiene la exactitud, que indica el total de casos que son diagnosticados de forma correcta con respecto al total de casos en el conjunto de datos.

En el Capítulo 3 se menciona la importancia de validar un modelo generado por una red neuronal, en particular en lo referente a la independencia de los datos. Por ello, como parte del proceso metodológico se considera utilizar el algoritmo de la validación cruzada para determinar la independencia del modelo.

En este punto es importante mencionar que al momento de entrenar el modelo, los mejores resultados se obtienen utilizando el conjunto de datos de entrenamiento tres (Dataset3), con una exactitud de 91.69%. Sin embargo, al momento de realizar la validación cruzada se detecta una fuerte dependencia del resultado con respecto a los resultados obtenidos, ya que el resultado promedio arroja un 81.44% (ver Tabla 4.1), lo que genera una variación negativa mayor al 10%.

Tabla 4.1 Resultados de la validación cruzada con el modelo generado con el Dataset3.

Pliegue	Exactitud [%]	Sensibilidad/ Recall [%]	Especificidad [%]	Precisión [%]
1	82.30	18.88	97.76	67.27
2	81.60	16.00	95.52	43.08
3	82.90	51.55	90.45	56.50
4	79.50	20.77	94.83	51.19
5	80.70	17.53	95.91	50.75
6	80.20	19.70	95.61	53.33
7	81.90	22.35	94.10	43.68
8	82.60	23.96	96.53	62.16
9	84.20	26.92	96.94	66.22
10	82.50	28.04	95.19	57.61
Promedio	81.84	24.57	95.28	55.18
Desviación Estándar	1.39	10.24	2.00	8.45

Por lo anterior, se recurre a evaluar el siguiente modelo en exactitud (Dataset1), el cual arroja un resultado de 90.99% y se genera utilizando el conjunto de datos 1. Los resultados de la validación cruzada se muestran en la Tabla 4.2.

Tabla 4.2 Resultados de la validación cruzada con el modelo generado con el Dataset1.

Pliegue	Exactitud [%]	Sensibilidad/ Recall [%]	Especificidad [%]	Precisión [%]
1	91.40	69.39	96.77	83.95
2	91.70	66.29	97.09	82.86
3	91.90	77.32	95.41	80.21
4	91.30	73.43	95.96	82.61
5	91.30	62.89	98.14	89.05
6	90.60	62.03	97.17	83.45
7	91.70	63.64	97.69	85.50
8	91.30	76.92	94.78	78.13
9	90.90	63.74	96.94	82.27
10	83.00	35.20	94.65	61.61
Promedio	90.51	65.08	96.46	80.96
Desviación Estándar	2.67	11.97	1.20	7.40

La validación de este modelo arroja únicamente una diferencia en la exactitud menor al 0.5%, lo que nos permite afirmar que el modelo de diagnóstico de la trombosis venosa profunda obtenido tras entrenar el modelo con el conjunto de datos 1 se considera válido al ser independiente de los datos, y por lo tanto, se utiliza como referencia para la evaluación y comparación con otros modelos de datos.

4.2 Los resultados desde la perspectiva del conjunto de datos

El diagnóstico médico tiene un alto grado de incertidumbre, debido a la complejidad de los sistemas biológicos, que puede llevar a que dos personas con características similares presenten sintomatología diferente ante una misma condición de salud, o por el contrario, manifiesten condiciones de salud opuestas aún con la misma sintomatología. En términos del modelo de clasificación por medio de redes neuronales, podemos decir que lo anterior es similar a que se presenten casos clínicos en los cuales los factores de riesgo que representan las variables de entrada al sistema sean iguales en su totalidad obteniendo como salida del SADC un diagnóstico resultante diferente.

Para contextualizar los resultados del entrenamiento del modelo de predicción, es necesario analizar la naturaleza del conjunto de datos. En este trabajo se utiliza un conjunto de 10,000 datos sintéticos, generados a partir de un proceso de aumentación que combina la proporción de casos reportados en Wells (P. S. Wells et al., 2006) y en 59 casos reales obtenidos de un hospital público. El conjunto de datos (instancias) está formado por dos variables de información personal, como lo son el género y la edad, y los nueve factores propios del modelo de Wells así como el diagnóstico asociado a cada instancia, el cual es positivo en 1902 instancias y negativo en 8,098.

En total se cuenta con 2,600 instancias totalmente diferentes entre sí en sus variables de entrada. De esos casos 1,585 son casos únicos, mientras que 1,015 se presentan en al menos dos pacientes con factores de riesgo (valores de entrada) similares, pudiendo variar el diagnóstico. De lo anterior se aprecia que 8,415 casos se distribuyen en un conjunto de 1,015 combinaciones distintas de variables de entrada que pueden contener diagnósticos diferentes.

Por ello, se considera importante calcular la probabilidad de diagnóstico de TVP positiva, con base en la ocurrencia detectada en el conjunto de datos, a partir de la Ecuación 3.3.

Se encuentra que 2,250 pacientes, distribuidos en 1,248 casos diferentes, tienen 0% de probabilidad de sufrir TVP. Así mismo, 976 de los pacientes distribuidos en 802 casos diferentes presentan una probabilidad de 100% en el diagnóstico positivo de TVP. Esto implica que 6,774 pacientes distribuidos en 550 casos diferentes, tienen riesgo de que se presente un error en su diagnóstico, asumiendo un mismo diagnóstico ante un mismo conjunto de valores de entrada. Más aún, 570 pacientes distribuidos en 261 casos tienen una probabilidad del 50% de TVP positiva, lo que implica a su vez que independientemente del diagnóstico, la mitad de esos pacientes (285) serán evaluados de forma incorrecta.

Lo anterior deja evidencia de la incertidumbre en el diagnóstico de la TVP. Debido a que en este trabajo se presenta un modelo de redes neuronales con una exactitud del 90.99% es importante calcular el valor de exactitud máxima que se puede alcanzar al evaluar estos casos. Para ello, se calcula la exactitud máxima que se alcanzara en el diagnóstico de la TVP con la Ecuación 4.5.

$$Exactitud_{max} = \frac{\sum_{i=1}^{2600} \max(\text{casosPositivos}, \text{casosNegativos})}{TotalDeCasos}$$

Ecuación 4.5 Exactitud máxima que puede alcanzar el modelo.

El análisis arroja que en el mejor de los casos, únicamente en 9,185 casos (91.85%) se puede llevar a cabo un diagnóstico de forma correcta, de lo cual se puede inferir que el modelo de clasificación propuesto en este trabajo es capaz de clasificar el 99.06% de los casos posibles de clasificar sin error, dentro del conjunto de datos.

Sin embargo, es muy importante mencionar que independientemente de la cantidad de datos sintéticos que el modelo puede clasificar satisfactoriamente, cuando el modelo se evalúa con datos reales que no son utilizados de forma directa en el entrenamiento, se obtiene un 98.3% de exactitud.

Otro de los elementos importantes a analizar, es la complicación para determinar el tipo de error que se desea minimizar en el modelo. Como se menciona, el conjunto de datos tiene un poco más del 80% de casos negativos contra el casi 20% de casos positivos, lo que resulta en un conjunto de datos no balanceado, que por sus características tiende a clasificar mejor los casos de diagnóstico negativo.

En la Figura 4.1 se observa la curva PR (así llamada por sus siglas en inglés Precision-Recall) del modelo de clasificación, con un umbral de 0.484, la cual muestra el resultado de variar el pivote que distingue a un caso positivo de un caso negativo. Como se puede apreciar, se debe establecer un compromiso entre reducir los casos de falsos positivos y los casos de falsos negativos. Al momento de minimizar los falsos positivos se incrementan los falsos negativos, y viceversa, por lo que es decisión de diseño determinar la configuración adecuada para el diagnóstico.

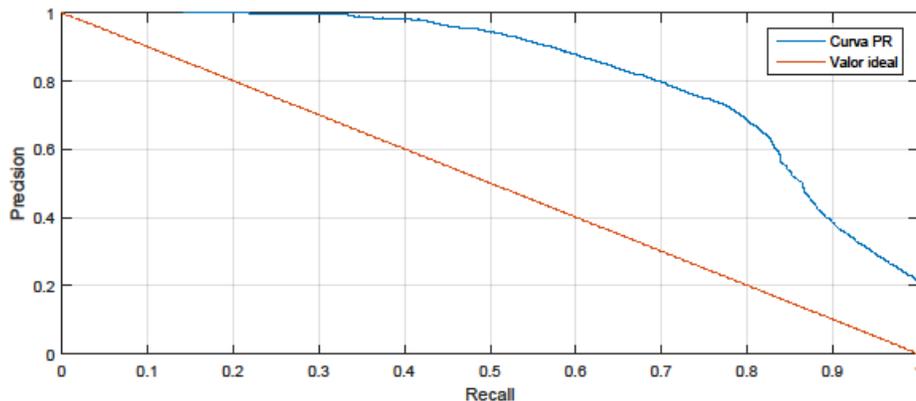


Figura 4.1 Curva PR (Precision-Recall).

En total, el conjunto de datos cuenta con 1902 casos con diagnóstico positivo y 8098 casos con diagnóstico negativo. Analizando los resultados del modelo, en la Figura 4.2 se puede ver que el modelo clasifica correctamente el 68.34% (1,300 de 1,902) de los casos positivos, y el 96.30% (7799 de 8,098) de los casos negativos, con 3.69% de clasificaciones positivas erróneas (299 de 8,098) y 31.65% de clasificaciones negativas erróneas (602 de 1902), mientras que la evaluación con la escala de Wells arroja el 83.96%

(1,597 de 1,902) de los casos positivos, y el 71.43 % (5,785 de 8,098) de los casos negativos, con 28.56 % de predicciones positivas erróneas (2,313 de 8,098) y 16.04% de predicciones negativas erróneas (305 de 1,902).

Los resultados muestran que el modelo propuesto tiene una mejor exactitud que la evaluación cuando se utiliza solamente la escala de Wells (90.99% por 73.82% respectivamente), sin embargo, Wells presenta una mejor predicción de casos positivos (69.93% por 83.96% respectivamente). Por ello, se analiza un nuevo umbral que permita superar la predicción de casos positivos sin perder la ventaja en la exactitud. Adicionalmente, se busca un umbral que supere el recall que presenta el trabajo de Wells (P. S. Wells et al., 2006), al mismo tiempo de maximizar la exactitud del modelo propuesto bajo esas condiciones. Así, con un umbral de 0.138, el modelo clasifica correctamente el 84.17% (1,601 de 1,902) de los casos positivos, y el 86.48% (7,003 de 8,098) de los casos negativos, con 14.83 % de predicciones positivas erróneas (1,201 de 8,098) y 15.98% de predicciones negativas erróneas (304 de 1,902), lo cual representa un mejor recall sin perder la ventaja en la exactitud.

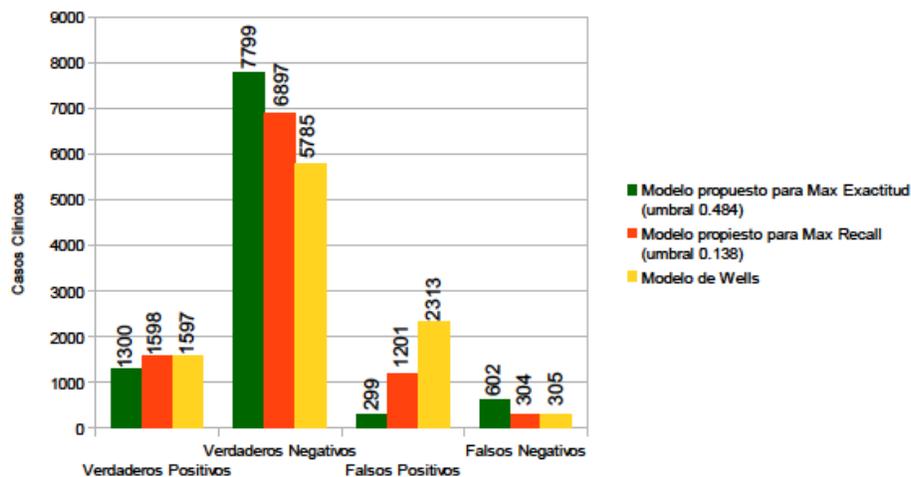


Figura 4.2 Comparación de las matrices de confusión de los modelos propuestos.

En la Figura 4.3 se presenta una comparación entre los porcentajes de la Precisión, la Sensibilidad y la Exactitud obtenidos con el modelo de la RNA propuesto con los umbrales de 0.484 para alcanzar la máxima exactitud del modelo y 0.138 para alcanzar la máxima exactitud que se puede lograr al mejorar la sensibilidad (recall) con respecto al obtenido por Wells (P. S. Wells et al., 2006) en la forma tradicional. Se puede observar que cuando el umbral es de 0.484, la máxima exactitud que alcanza el modelo es 90.99%, la sensibilidad es de 68.35% y la precisión de 81.30%. Sin embargo con un umbral de 0.138, la sensibilidad se incrementa a 84.01%, la exactitud disminuye a 84.95% y la precisión queda en 57.09%; mientras que los porcentajes encontrados por Wells en la forma tradicional son 73.82% de exactitud (menor a los porcentajes alcanzados con el modelo propuesto), sensibilidad de 83.96% y precisión de 40.84%.

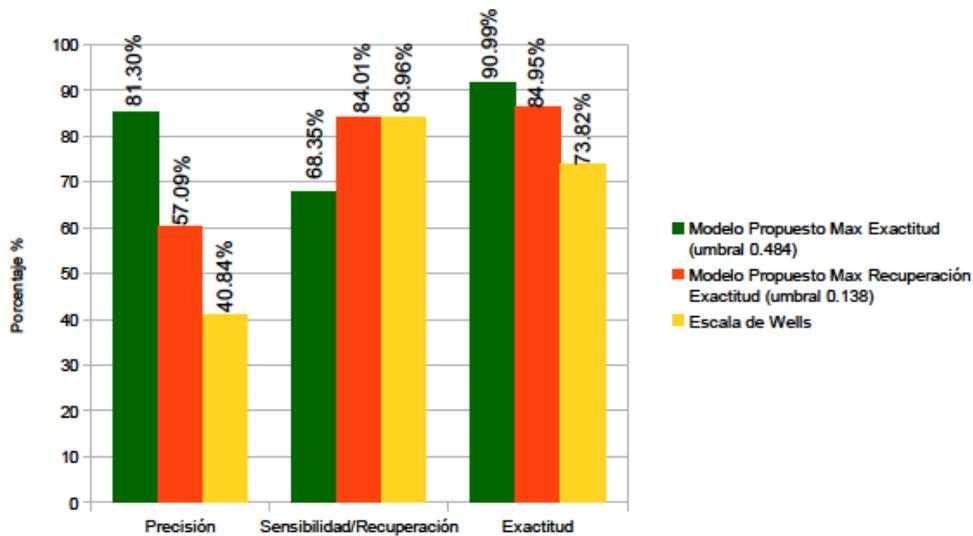


Figura 4.3 Comparación de los porcentajes de precisión, recall y exactitud obtenidos con los modelos propuestos.

En este trabajo de investigación se presenta una alternativa para evaluar el diagnóstico de TVP usando un modelo de RNA basado en el algoritmo de Wells (P. S. Wells et al., 2006) en conjunto con otras dos variables de entrada adicionales (género y edad).

Con el objetivo de determinar cuál o cuáles de las once entradas a la red ejercen una mayor influencia en el desempeño del modelo propuesto, en la Tabla 4.3 se presenta el análisis del impacto que tiene cada una de estas variables de entrada en la salida del modelo. Como puede verse la entrada que más efecto tiene en el desempeño de la red es la edad, dado que eliminando esta variable el porcentaje de la exactitud sufre un decremento hasta llegar a un 84.59%, lo que representa una variación de más del 6% con respecto al valor original obtenido de 90.99% al considerar todas las entradas. Por lo contrario, si alguna de las diez entradas restantes se elimina, el cambio que se produce en el porcentaje de la exactitud es menor ya que el promedio de la exactitud entre las otras entradas es aproximadamente de 88.17% (con una desviación estándar de 0.5%). Lo anterior puede explicarse por la estratificación que tiene cada una de las entradas a la red, a excepción de la edad todas las entradas a la red tienen una segmentación binaria, así que no sorprende que su influencia en el modelo sea similar, mientras que la edad puede asumir cualquiera de nueve diferentes valores, lo que permite que el modelo asigne un mayor peso a esta entrada.

Tabla 4.3 Influencia que ejerce cada una de las variables de entrada en el desempeño de la red.

Factor faltante	TP	TN	FP	FN	Exactitud [%]	Umbral [%]	ECM	Sensibilidad	Especificidad	Precisión
Edad	2338	6121	952	589	84.59	0.49	0.11	79.88	86.54	71.06
Género	1116	7779	319	786	88.95	0.49	0.95	58.68	96.06	77.77
Cáncer	1233	7699	399	669	89.32	0.468	0.08	64.83	95.07	75.55
Inmovilización	1159	7676	422	743	88.35	0.476	0.09	60.94	94.79	73.31
Cirujía	1160	7677	421	742	88.37	0.481	0.09	60.99	94.80	73.37
Dolor	1163	7669	429	739	88.32	0.491	0.09	61.15	94.70	73.05
Inflamación de pierna	641	8230	155	974	88.71	0.452	0.09	39.69	98.15	80.53
Inflamación de tobillo	1053	7793	305	849	88.46	0.501	0.08	55.36	96.23	77.54
Edema	1181	7629	469	721	88.1	0.496	0.08	62.09	94.21	71.58
Venas varicosas	1199	7563	535	703	87.62	0.467	0.09	63.04	93.39	69.15
TVP previa	1253	7652	446	649	89.05	0.489	0.08	65.88	94.49	73.75
Promedio	1226.91	7589.82	441.09	742.18	88.17	0.48	0.17	61.14	94.40	74.24
Desviación Estándar	404.62	518.04	197.13	103.23	1.28	0.01	0.26	9.44	2.90	3.33

4.3 Comparación con otras técnicas de análisis de datos

Para dimensionar el aporte de este trabajo, en esta sección se presenta en la Tabla 4.4 una comparación de los porcentajes de la exactitud, la sensibilidad y la especificidad alcanzados con el modelo de red propuesto; así como los valores obtenidos utilizando únicamente el modelo de Wells de forma tradicional en la detección de la TVP.

Como puede observarse, el porcentaje de la exactitud calculada al implementar el modelo de RNA propuesto es de 90.99% cuando se usa el umbral de 0.484, el cual es superior en un 17.17% con respecto a la exactitud del modelo de Wells al evaluarse de manera tradicional (73.82%). Además, el modelo de RNA también es superior al diagnosticar a los pacientes sanos, dado que la especificidad es de 96.31%, mientras que el modelo tradicional de Wells arroja un 71.43%. Sin embargo, es importante mencionar que este modelo tiene como inconveniente una menor capacidad para diagnosticar correctamente a pacientes enfermos, ya que la sensibilidad, que es la métrica asociada, presenta un valor del 68.35% en el modelo de RNA propuesto, el cual es inferior en un 15.66% al modelo tradicional de Wells.

Por lo anterior, se propone ajustar el umbral para maximizar la exactitud del modelo de predicción con al menos la misma sensibilidad del modelo de Wells cuando se usa de manera tradicional. Así, con un umbral de 0.138 el modelo de RNA arroja una exactitud de 84.95%, que si bien es inferior en un 6.04% al modelo de RNA con el umbral de 0.484, mejora un 11.13% al modelo de Wells tradicional. Además, la sensibilidad es ligeramente mayor (0.05%) al modelo de Wells tradicional, con la ventaja adicional que mejora el diagnóstico de pacientes sanos en un 13.74%.

Tabla 4.4 Desempeño de los modelos de apoyo en la toma de decisiones para el diagnóstico de la TVP.

Trabajo	Exactitud [%]	Sensibilidad/ Recall [%]	Especificidad [%]
Escala de Wells con modelo RNA propuesto para maximizar la exactitud	90.99	68.35	96.31
Escala de Wells con modelo RNA propuesto para maximizar el Recall	84.95	84.01	85.17
Escala de Wells Tradicional (P. S. Wells et al., 2006)	73.82	83.96	71.43

4.4 Validación Externa

La mejor manera de evaluar los resultados del modelo propuesto es ponerlo a prueba en un hospital real. Desafortunadamente, debido a la naturaleza del trabajo médico y las implicaciones del diagnóstico en la salud de las personas, es complicado llevar a cabo la evaluación en un hospital. Por ello, la evaluación del modelo se llevó a cabo utilizando un conjunto de 59 casos reales extraídos de un hospital público, los cuales presentan síntomas atribuibles a una Trombosis Venosa Profunda; 36 casos coinciden con diagnóstico positivo de TVP, mientras que el diagnóstico de los 23 casos restantes corresponde a alguna otra enfermedad.

En la mayoría de los parámetros mostrados en la Tabla 4.5, con el modelo de RNA propuesto se obtienen porcentajes más altos en comparación con los otros modelos de aprendizaje automático. También se puede apreciar que el modelo propuesto en este trabajo de tesis supera en más del 8% la exactitud del modelo generado por el MSV cuadrático, el cual fue el segundo modelo que mejor exactitud logra entre los modelos evaluados.

Tabla 4.5 Comparación con otros algoritmos de aprendizaje automático.

Clasificador	TP	TN	FP	FN	Recall [%]	Especifi- cidad [%]	Preci- sión [%]	Exactitud [%]
RNA propuesta	1300	7799	299	602	96.31	68.35	81.30	90.99
MSV Cuadrático (Doma & Pirouz, 2020; Janghel et al., 2020; M. Liu et al., 2013; Luo et al., 2020)	651	7553	545	1251	93.27	34.23	54.43	82.04
MSV Gaussiano (Janghel et al., 2020; M. Liu et al., 2013)	569	7526	572	1333	92.94	29.92	49.87	80.95
Árboles simples (Doma & Pirouz, 2020; Janghel et al., 2020)	685	7384	724	1217	91.07	31.01	48.62	80.69
Árboles complejos (Doma & Pirouz, 2020; Janghel et al., 2020)	740	7361	737	1162	90.90	38.91	50.10	81.01
KNN (ponderado) (Barbon et al., 2016; Doma & Pirouz, 2020; Janghel et al., 2020)	710	7365	733	1192	90.95	37.33	49.20	80.75
KNN (fino) (Barbon et al., 2016; Doma & Pirouz, 2020; Janghel et al., 2020)	801	6874	1224	1101	84.89	42.11	39.56	76.75
Bosques Aleatorios (BA) (Syed-Abdul et al., 2020)	1265	7658	316	761	96.04	62.44	80.01	89.23
Gradiente de descenso estocástico (SGD) (Syed-Abdul et al., 2020)	1157	7453	295	1095	96.19	51.38	79.68	86.10
Aumento de gradiente extremo (XGBoost) (J. Shen et al., 2020)	1250	7625	312	813	86.07	60.59	80.03	88.75
Árbol de decisión de aumento de gradiente (GBDT) (J. Shen et al., 2020)	1358	7583	294	765	96.27	63.97	82.20	89.41

Así, a partir del modelo de RNA propuesto se evalúan cada uno de los casos reales. Los resultados arrojan que el modelo predice correctamente 35 de los 36 casos positivos, así como los 23 casos con diagnóstico diferente a la TVP (ver Tabla 4.6).

Tabla 4.6 Matriz de confusión de los resultados de evaluar el modelo de RNA con los casos reales.

	Predicción de TVP	Predicción de otra enfermedad
Caso real con TVP	35	1
Caso real sin TVP	0	23

En resumen, la evaluación del modelo propuesto con casos reales presenta una exactitud del 98.30%, con un error del 1.7%, con una especificidad del 100% y una sensibilidad del 97.22%. Así, se aprecia que el modelo funciona con una tasa de error por debajo del máximo tolerado en la venografía, ya mencionada con anterioridad como el estándar de oro en el diagnóstico de la TVP.

Una vez que el modelo de RNA propuesto se entrena y valida con la técnica de validación cruzada, se realiza una validación externa utilizando los 59 casos clínicos registrados durante los años 2017 y 2018 en un hospital regional. La Figura 4.4 muestra los resultados de evaluar los casos reales utilizando el modelo de clasificación que se desarrolla en el presente trabajo. Se puede observar que solo 1 de 36 casos de TVP se diagnostica como negativo cuando en efecto, es positivo; esto representa una exactitud del 98.30% (58 de 59 aciertos) con un error de diagnóstico del 1.70%.

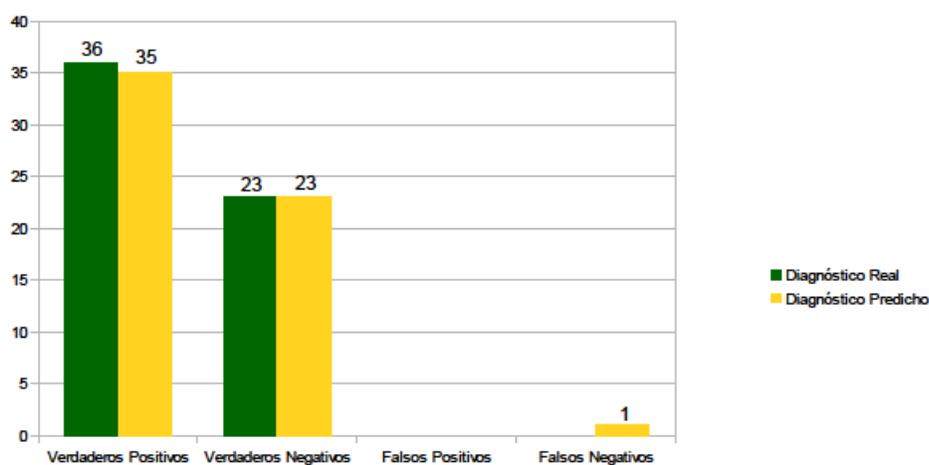


Figura 4.4 Resultados de la evaluación externa.

4.5 Propuesta de aplicación del modelo de predicción de la trombosis venosa profunda

La mejor forma de medir el impacto del resultado de un proyecto de investigación es medir su desempeño en un entorno real. Para el caso de este trabajo de tesis doctoral, el ambiente es un hospital de consulta primaria. Desafortunadamente, la implementación del modelo está fuera del alcance de este proyecto por distintos factores que se discuten en el Capítulo 5 de las Conclusiones.

Sin embargo, es importante dejar asentado la forma en que puede ser implementado en caso de que las condiciones necesarias se cumplan. Para ello, se hace uso del Diseño Basado en Escenarios propuesto por Carroll (Rosson & Carroll, 2002).

Un escenario es la representación del estado del sistema en un momento en particular (Rosson & Carroll, 2002). Así, en este trabajo se presentan dos tipos de escenarios: i) antes de la aplicación del modelo de predicción de la trombosis venosa profunda, y ii) propuesta de aplicación del modelo de predicción de la trombosis venosa profunda.

4.5.1 Escenario de uso

La utilización de un Sistema de Apoyo a las Decisiones Clínicas que implemente el modelo de diagnóstico de TVP basado en la escala de Wells a través de redes neuronales, debe ser complementaria al protocolo que se usa actualmente. Por ello, tal como se ilustra en la Figura 4.5, se propone que el paciente se presente con el médico y le comente los síntomas, el médico realiza la exploración física y alimenta al SADC con los datos clínicos que le proporciona el paciente y los observados durante la exploración, para posteriormente esperar a que el sistema le arroje el resultado de la clasificación. En caso de que tanto el médico como el sistema determinen que el usuario no está sufriendo una TVP, el paciente no recibirá tratamiento. Si el sistema indica que el usuario puede tener una condición de TVP, pero el médico determina que no es así, el paciente debe irse a su casa quedando pendiente de la aparición de nuevos síntomas o de que se agudicen los ya existentes. En caso de que el médico decida que la examinación física sugiere una TVP, pero el sistema

diagnostica lo contrario, el médico toma la decisión, dejando como sugerencia que el paciente debe estar en observación. Finalmente, en caso de que tanto el sistema como el médico establezcan que el paciente está sufriendo una TVP, se debe iniciar el protocolo de la TVP.

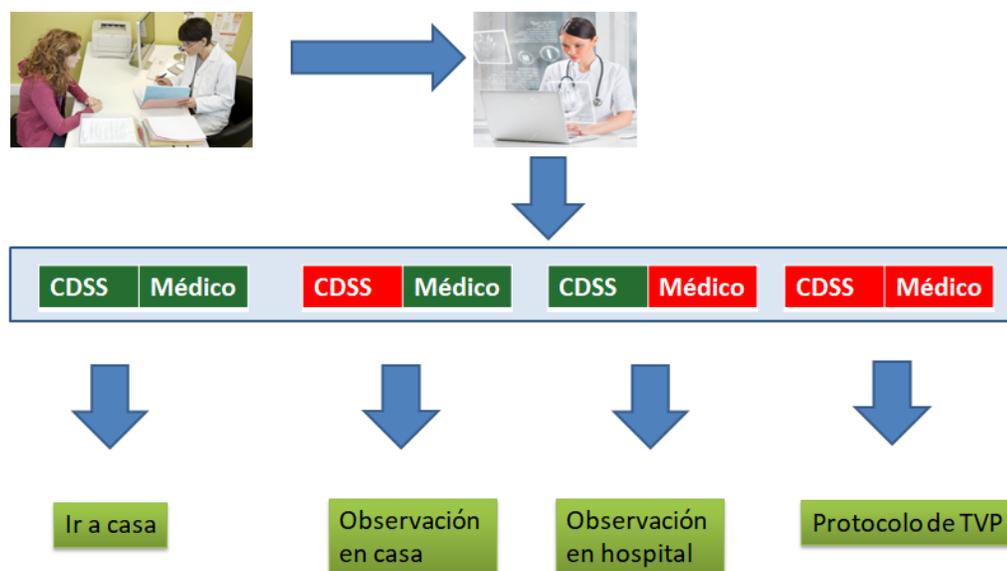


Figura 4.5 Uso sugerido del SADC en módulos de atención primaria.

En la Figura 4.6 se presenta el diagrama a bloques del protocolo sugerido para ser implementado en las unidades de atención primaria, por médicos que utilicen el SADC propuesto en la presente investigación.

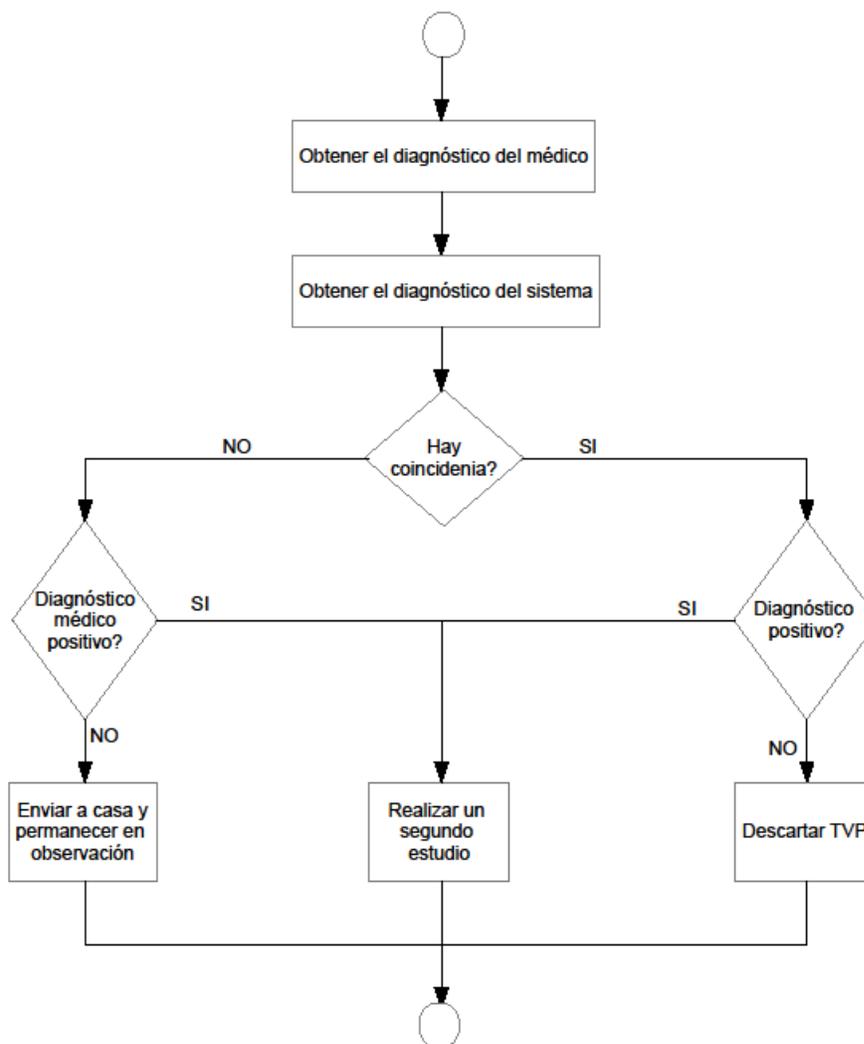


Figura 4.6 Diagrama a bloques del Protocolo sugerido en unidades de atención primaria.

Es importante mencionar que el diagnóstico de la TVP conlleva un alto grado de incertidumbre, lo cual puede generar costos tanto en recursos como en salud de los pacientes. Por ello, es importante la mejora en la exactitud que se logra con el modelo de diagnóstico propuesto, en un 17.58%, particularmente debido a que en países como México, donde en los módulos de atención primaria se trabaja con recursos limitados, es de vital importancia reducir los costos asociados a los errores en el diagnóstico.

4.6 Limitaciones del trabajo de tesis

Los datos utilizados en el entrenamiento y validación del modelo de clasificación son generados de manera aleatoria, y aunque basados en ocurrencias reales, no representan fielmente el comportamiento de la población en riesgo de padecer TVP. Si bien los resultados con datos reales sugieren una alta exactitud, debido al impacto que las decisiones basadas en el modelo de clasificación pueden tener en la salud de los pacientes, se sugiere tomar estos resultados como prueba de la efectividad de las redes neuronales en el diagnóstico de la TVP, y buscar un conjunto de datos de mayor fiabilidad para entrenar el modelo antes de ponerlo en práctica en la vida real.

4.7 Conclusiones del Capítulo 4

En este capítulo se presenta la metodología de trabajo que se sigue en este proyecto de investigación, el cual consta de cuatro etapas principales: i) análisis del problema, ii) definición del modelo de predicción de la trombosis venosa profunda, iii) discusión de los resultados y iv) propuesta de aplicación. Cada una de estas etapas tiene una submetodología en particular, propia de las características de cada etapa.

La etapa que mayor atención requiere en este capítulo es la de la definición del modelo de predicción de la trombosis venosa profunda, ya que el modelo resultante da sustento a la principal aportación de este trabajo.

A destacar entonces queda la necesidad de diseñar y utilizar un algoritmo de aumentación de datos, debido a la dificultad para almacenar y tener acceso al expediente clínico de aquellos pacientes que resultan sospechosos de tener una trombosis venosa profunda. Además, la necesidad de contar con diferentes estrategias para conformar el conjunto de datos de entrenamiento. También se destaca el uso de la técnica de validación cruzada para validar los resultados del modelo, así como la comparación de los resultados obtenidos con respecto a otros modelos de aprendizaje automático para el análisis de datos.

Capítulo 5 Conclusiones

5.1 Conclusiones generales

El diagnóstico de la TVP tiene un alto grado de incertidumbre, así como costos humanos y económicos asociados a las técnicas utilizadas. Los algoritmos probabilísticos surgieron para ayudar en el diagnóstico de enfermedades, pero su nivel de rechazo a pacientes potenciales es mayor al límite aceptado del 2% que ofrece la venografía. En este trabajo se presenta un modelo de diagnóstico de TVP utilizando Redes Neuronales Artificiales con retropropagación. El modelo propuesto está diseñado para mejorar la exactitud aún con conjuntos de datos de entrenamiento altamente desbalanceados. El entrenamiento se lleva a cabo utilizando como entradas 9 factores de riesgo establecidos en el modelo probabilístico de Wells, además del género y la edad. Debido a los pocos registros clínicos que se obtienen de un hospital mexicano de pacientes con TVP se diseña una técnica de aumentación de datos para entrenar el modelo RNA propuesto. Se utilizó la técnica de la validación cruzada con los 10,000 datos sintéticos y una $k=10$, logrando alcanzar una exactitud promedio de 90.51% con respecto a la exactitud del 90.99% alcanzada con el modelo (umbral de 0.484), lo que implica una independencia del modelo con respecto al conjunto de datos de entrenamiento. Además se realiza una validación externa con los 59 registros médicos obtenidos de un hospital regional alcanzando una exactitud del 98.30%. En el diagnóstico médico la clasificación correcta de casos positivos es de vital importancia ya que un diagnóstico de un falso negativo puede afectar la salud de los pacientes. Para ello, se busca un umbral que optimice la clasificación de los casos positivos, pero manteniendo una ventaja en la exactitud del modelo. Así, con un umbral de 0.138, el modelo propuesto mejora la clasificación de casos positivos sin que se vea afectada de manera significativa la clasificación de los casos negativos.

Por otro lado, el modelo propuesto se compara con otros algoritmos de aprendizaje comúnmente utilizados, presentando una exactitud mayor a la alcanzada por dichos algoritmos, se asume que esto se debe a que tales algoritmos no utilizan un umbral optimizado para llevar a cabo la clasificación.

Este modelo presenta un progreso en términos de diagnóstico, así como en términos del ahorro económico que representa con respecto al modelo de Wells, ya que al obtener una exactitud 17.17% mayor que la obtenida por dicho modelo cuando se utiliza de forma tradicional (sin considerar edad ni género) se reducen los costos asociados a la realización innecesaria de la prueba del Dímero-D que se requiere al utilizar el protocolo de la Figura 2.1. Es evidente que el diseño e implementación de un SADC usando RNA en unidades de atención primaria mejora la exactitud en el diagnóstico oportuno de TVP, disminuyendo de esta manera el flujo de pacientes que llegan a las Unidades de Emergencia y teniendo como consecuencia directa la reducción de la tasa de mortalidad asociada a esta condición vascular.

Si bien los resultados no tienen la exactitud de los Sistemas de Apoyo a las Decisiones Clínicas que utilizan imágenes para realizar el diagnóstico, los resultados aquí presentados están acordes al esquema de diagnóstico en el que se basa el modelo, el cual se propone para apoyar al diagnóstico médico en atención primaria, en donde el equipo de imagenología y el personal calificado para manipularlo e interpretar los resultados no está disponible.

Por lo tanto, el principal aporte de esta investigación es complementar el diagnóstico del médico en atención primaria en la detección oportuna de Trombosis Venosa Profunda en extremidades inferiores con base en un sistema de recomendación que utiliza un modelo probabilístico para el diagnóstico de la TVP.

5.2 Aportaciones del trabajo de tesis

La realización de este trabajo de tesis doctoral permite generar un aporte al campo de conocimiento en lo referente al conocimiento y ejemplificación del uso de las redes neuronales para desarrollar modelos de diagnóstico que sean implementados en unidades de atención primaria, aún con las limitantes propias en la recopilación de datos.

- Un nuevo Algoritmo de aumentación de datos que genera un conjunto de datos sintético para entrenar un modelo inteligente de diagnóstico de la trombosis venosa profunda.
- Un nuevo Modelo Inteligente de Diagnóstico de la Trombosis Venosa Profunda, utilizando un clasificador basado en red neuronal artificial.

Además de lo anterior, se tienen como aporte los siguientes productos:

- Artículo en Libro Innovación Científica y Tecnológica para el desarrollo de nuestra comunidad.

Fong-Mata, M. B., Inzunza-González, E. (2016). Dispositivo de ultrasonido Doppler para el diagnóstico computarizado de Trombosis Venosa Profunda en extremidades inferiores. Libro Innovación Científica y Tecnológica para el desarrollo de nuestra comunidad. UABC. 43-47pp. ISBN: 978-0-692-82901-1

- Artículo en Revista de Ciencias Tecnológicas RECIT

Fong-Mata, M. B., Inzunza-González, E., García-Guerrero, E. E., Mejía-Medina, D. A., Morales Contreras, O. A., Gómez-Roa, A. (2020). Trombosis Venosa Profunda en extremidades inferiores: revisión de las técnicas de diagnóstico actuales y su simbiosis con el aprendizaje automático para un diagnóstico oportuno. Revista de Ciencias Tecnológicas, 3(1), 24–35. <https://doi.org/10.37636/recit.v3i12334>

- Artículo Científico en una revista indizada en JCR (Electronics de Editorial MDPI)

Fong-Mata, M. B., García-Guerrero, E. E., Mejía-Medina, D. A., López-Bonilla, O. R., Villarreal-Gómez, L. J., Zamora-Arellano, F., López-Mancilla, D., Inzunza-González, E. (2020) An Artificial Neural Network Approach and a Data Augmentation Algorithm to Systematize the Diagnosis of Deep-Vein Thrombosis by Using Wells' Criteria. *Electronics* 9(11), 1810; <https://doi.org/10.3390/electronics9111810>

5.3 Trabajo Futuro

Una de las características del trabajo científico es que la generación de nuevo conocimiento genera nuevas interrogantes y por ende, la necesidad de conducir trabajo adicional. Por ello, en esta sección se presentan las vertientes que podrían seguirse a partir de los resultados obtenidos en este trabajo de tesis doctoral.

- Obtener una mayor cantidad de casos reales para el modelo de diagnóstico de la trombosis venosa profunda a partir de dispositivos personales y/o tecnología vestible.
- Registro de datos médicos personales en un expediente médico, que permita la generación de conjunto de datos para entrenamiento de modelos de aprendizaje automático aplicados a la salud, particularmente en consultorios de atención primaria.
- Explorar el uso de modelos de inteligencia artificial en el diagnóstico de otras enfermedades, para uso en consultorios de atención primaria.
- Implementar en sistemas embebidos (FPGA, SoC, ASIC, etc.) el modelo de red neuronal obtenido y realizar pruebas de funcionalidad.
- Implementación de dispositivos personales inteligentes de diagnóstico médico.
- Acondicionar sensores al sistema propuesto, para la medición de alguno de los factores de riesgo.
- Desarrollar una interfaz gráfica para el usuario (GUI, Graphic User Interface).
- Desarrollar una App con acceso vía Web para usuarios remotos.
- Identificar los escenarios médicos en donde se requiera el análisis de información con cómputo paralelo para su implementación en modelos de redes neuronales.

Referencias

- Abe, H., Ashizawa, K., Li, F., Matsuyama, N., Fukushima, A., Shiraishi, J., MacMahon, H., & Doi, K. (2004). Artificial neural networks (ANNs) for differential diagnosis of interstitial lung disease: results of a simulation test with actual clinical cases. *Academic Radiology*, *11*(1), 29–37. [https://doi.org/https://doi.org/10.1016/S1076-6332\(03\)00572-5](https://doi.org/10.1016/S1076-6332(03)00572-5)
- Abramowitz, H. B., & Gertz, S. D. (2007). Venous Stasis, Deep Venous Thrombosis and Airline Flight: Can the Seat be Fixed? *Annals of Vascular Surgery*. <https://doi.org/10.1016/j.avsg.2007.03.007>
- Ağbulut, Ü., Gürel, A. E., & Biçen, Y. (2021). Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. *Renewable and Sustainable Energy Reviews*. <https://doi.org/10.1016/j.rser.2020.110114>
- Al-Shayea, Q., & Kadhim Al-Shayea, Q. (2011). Artificial Neural Networks in Medical Diagnosis Efficient Biometric Recognition Approach View project Telemedicine in Jordan View project Artificial Neural Networks in Medical Diagnosis. *IJCSI International Journal of Computer Science Issues*.
- Alfonso Piérola, A., & Páramo Fernández, J. A. (2012). Protocolo de diagnóstico y tratamiento de la enfermedad tromboembólica venosa. *Medicine (Spain)*. [https://doi.org/10.1016/S0304-5412\(12\)70464-X](https://doi.org/10.1016/S0304-5412(12)70464-X)
- Ambagtsheer, R. C., Shafiabady, N., Dent, E., Seiboth, C., & Beilby, J. (2020). The application of artificial intelligence (AI) techniques to identify frailty within a residential aged care administrative data set. *International Journal of Medical Informatics*. <https://doi.org/10.1016/j.ijmedinf.2020.104094>
- AYLWARD, S., & ANDERSON, R. A. Y. (1991). An algorithm for neural network architecture generation. In *8th Computing in Aerospace Conference*. American Institute of Aeronautics and Astronautics. <https://doi.org/doi:10.2514/6.1991-3756>
- Barbon, A. P. A. C., Barbon, S., Mantovani, R. G., Fuzyi, E. M., Peres, L. M., & Bridi, A.

- M. (2016). Storage time prediction of pork by Computational Intelligence. *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2016.06.028>
- Benavides Bermúdez, J. M., Vivas Diaz, J. J., Jaramillo Trujillo, G., & Bernal Torres, W. (2017). Trombosis venosa profunda en un paciente con hipertiroidismo de novo. Presentación de caso. *Repertorio de Medicina y Cirugía*. <https://doi.org/10.1016/j.reper.2017.03.004>
- Bennin, K. E., Keung, J., Phannachitta, P., Monden, A., & Mensah, S. (2018). MAHAKIL: Diversity Based Oversampling Approach to Alleviate the Class Imbalance Issue in Software Defect Prediction. *IEEE Transactions on Software Engineering*. <https://doi.org/10.1109/TSE.2017.2731766>
- Bernardi, E., & Camporese, G. (2018). Diagnosis of deep-vein thrombosis. *Thrombosis Research*, *163*, 201–206. <https://doi.org/https://doi.org/10.1016/j.thromres.2017.10.006>
- Bolado, A. G., Bárcena, M. V., del Cura, J. L., Gorriño, O., & Grande, D. (2011). Indicación de eco-Doppler venosa de extremidades inferiores en el diagnóstico de la enfermedad tromboembólica ante una sospecha de tromboembolismo pulmonar Diagnostic Indication for Venous Echo-Doppler of the Lower Limbs in the Diagnosis of Thromboembolic. *Radiología*, *45*(5), 213–218. [https://doi.org/10.1016/S0033-8338\(03\)77905-3](https://doi.org/10.1016/S0033-8338(03)77905-3)
- Bourouis, A., Feham, M., Hossain, M. A., & Zhang, L. (2014). An intelligent mobile based decision support system for retinal disease diagnosis. *Decision Support Systems*, *59*, 341–350. <https://doi.org/https://doi.org/10.1016/j.dss.2014.01.005>
- Brush, J. E., Sherbino, J., & Norman, G. R. (2017). How Expert Clinicians Intuitively Recognize a Medical Diagnosis. In *American Journal of Medicine*. <https://doi.org/10.1016/j.amjmed.2017.01.045>
- Cano, G., García-Rodríguez, J., Orts, S., García-García, A., Peña-García, J., Pérez-Garrido, A., & Pérez-Sánchez, H. (2017). Predicción de solubilidad de fármacos usando máquinas de soporte vectorial sobre unidades de procesamiento gráfico. *Revista*

Internacional de Metodos Numericos Para Calculo y Diseno En Ingenieria.
<https://doi.org/10.1016/j.rimni.2015.12.001>

Cao, B., Wang, N., Li, J., & Gao, X. (2019). Data Augmentation-Based Joint Learning for Heterogeneous Face Recognition. *IEEE Transactions on Neural Networks and Learning Systems*. <https://doi.org/10.1109/TNNLS.2018.2872675>

Chen, Y., Tsang, Y. S., Chou, X., Hu, J., & Xia, Q. (2019). A lung cancer patient with deep vein thrombosis: a case report and literature review. *BMC CANCER*, 19. <https://doi.org/10.1186/s12885-019-5513-8>

Choquenaira Florez, A. Y., Scabora, L., Amer-Yahia, S., & Rodrigues Junior, J. F. (2020). Augmentation techniques for sequential clinical data to improve deep learning prediction techniques. *Proceedings - IEEE Symposium on Computer-Based Medical Systems*. <https://doi.org/10.1109/CBMS49503.2020.00118>

Christopher, J. J., Nehemiah, H. K., & Kannan, A. (2015). A clinical decision support system for diagnosis of Allergic Rhinitis based on intradermal skin tests. *Computers in Biology and Medicine*, 65, 76–84. <https://doi.org/10.1016/j.combiomed.2015.07.019>

Clarke, M. J., Broderick, C., Hopewell, S., Juszczak, E., & Eisinga, A. (2016). Compression stockings for preventing deep vein thrombosis in airline passengers. In *Cochrane Database of Systematic Reviews*. <https://doi.org/10.1002/14651858.CD004002.pub3>

Cohen, A. T., Agnelli, G., Anderson, F. A., Arcelus, J. I., Bergqvist, D., Brecht, J. G., Greer, I. A., Heit, J. A., Hutchinson, J. L., Kakkar, A. K., Mottier, D., Oger, E., Samama, M.-M., Spannagl, M., & VTE Impact Assessment Group in Europe (VITAE). (2007). Venous thromboembolism (VTE) in Europe. The number of VTE events and associated morbidity and mortality. *Thrombosis and Haemostasis*.

Darwood, R. J., & Smith, F. C. T. (2013). Deep vein thrombosis. In *Surgery (United Kingdom)*. <https://doi.org/10.1016/j.mpsur.2013.02.001>

de Souza, L. A., Passos, L. A., Mendel, R., Ebigbo, A., Probst, A., Messmann, H., Palm,

- C., & Papa, J. P. (2020). Assisting Barrett's esophagus identification using endoscopic data augmentation based on Generative Adversarial Networks. *Computers in Biology and Medicine*. <https://doi.org/10.1016/j.combiomed.2020.104029>
- Deng, F., Tang, Q., Zeng, G., Wu, H., Zhang, N., & Zhong, N. (2015). Effectiveness of digital infrared thermal imaging in detecting lower extremity deep venous thrombosis. *Medical Physics*. <https://doi.org/10.1118/1.4907969>
- Dias Canedo, E., & Cordeiro Mendes, B. (2020). Software Requirements Classification Using Machine Learning Algorithms. *Entropy*, 22(9). <https://doi.org/10.3390/e22091057>
- Doma, V., & Pirouz, M. (2020). A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals. *Journal of Big Data*. <https://doi.org/10.1186/s40537-020-00289-7>
- Drobnic, F., Pineda, A., Escudero, J. R., Soria, J. M., & Souto, J. C. (2015). Guía de práctica clínica para la prevención, diagnóstico y tratamiento de la enfermedad tromboembólica venosa en el deporte. *Apunts Medicina de l'Esport*. <https://doi.org/10.1016/j.apunts.2015.09.001>
- Farhat, F. C. L. G., Gregório, H. C. T., & de Carvalho, R. D. P. (2018). Evaluation of deep vein thrombosis prophylaxis in a general hospital. *Jornal Vascular Brasileiro*. <https://doi.org/10.1590/1677-5449.007017>
- Fortier O'brien, N., Taty, T. M., Moore-Clingenpeel, M., Mabiala, J. B., Pongo, J. M., Ambitapio Musungufu, D., Uchama, M., & Yotebieng, M. (2018). Transcranial Doppler Ultrasonography Provides Insights into Neurovascular Changes in Children with Cerebral Malaria ARTICLE IN PRESS THE JOURNAL OF PEDIATRICS • www.jpeds.com. *The Journal of Pediatrics*. <https://doi.org/10.1016/j.jpeds.2018.07.075>
- Fox, J. C., & Bertoglio, K. C. (2011). Emergency Physician Performed Ultrasound for DVT Evaluation. *Thrombosis*. <https://doi.org/10.1155/2011/938709>

- Francis, J. E., Roggli, R., Love, T. J., & Robinson, C. P. (1979). THERMOGRAPHY AS A MEANS OF BLOOD PERFUSION MEASUREMENT. *Journal of Biomechanical Engineering*. <https://doi.org/10.1115/1.3426253>
- Fu, G. H., Xu, F., Zhang, B. Y., & Yi, L. Z. (2017). Stable variable selection of class-imbalanced data with precision-recall criterion. *Chemometrics and Intelligent Laboratory Systems*. <https://doi.org/10.1016/j.chemolab.2017.10.015>
- Fuentes Camps, E., Del Val García, J. L., Bellmunt Montoya, S., Hmimina Hmimina, S., Gómez Jabalera, E., & Muñoz Pérez, M. (2015). Factores clínicos que influyen en la probabilidad diagnóstica pretest de trombosis venosa profunda en pacientes ambulatorios. *Angiologia*. <https://doi.org/10.1016/j.angio.2015.03.003>
- Fuentes Camps, Eva, Luis Del Val García, J., Bellmunt Montoya, S., Hmimina Hmimina, S., Gómez Jabalera, E., & Muñoz Pérez, M. Á. (2016). Estudio coste efectividad del proceso diagnóstico de la trombosis venosa profunda desde la atención primaria. *Atencion Primaria*. <https://doi.org/10.1016/j.aprim.2015.05.006>
- G., M., R., D., P., K., O., L., R., S., S., R., P., R., S.D., M., K.A., P., M., S., Y., J., E., C., Y., W., Y., B., S., B., C., S., P.L., M., R., C., L.R., P., ... R., A. (2019). Predicting venous thromboembolism risk from exomes in the Critical Assessment of Genome Interpretation (CAGI) challenges. *Human Mutation*.
- Gao, Z., Li, J., Guo, J., Chen, Y., Yi, Z., & Zhong, J. (2019). Diagnosis of Diabetic Retinopathy Using Deep Neural Networks. *IEEE Access*, 7, 3360–3370. <https://doi.org/10.1109/ACCESS.2018.2888639>
- Gok, E. C., Yildirim, M. O., Eren, E., & Oksuz, A. U. (2020). Comparison of Machine Learning Models on Performance of Single- And Dual-Type Electrochromic Devices. *ACS Omega*. <https://doi.org/10.1021/acsomega.0c03048>
- Goldhaber, S. Z., & Bounameaux, H. (2012). Pulmonary embolism and deep vein thrombosis. *The Lancet*. [https://doi.org/10.1016/S0140-6736\(11\)61904-1](https://doi.org/10.1016/S0140-6736(11)61904-1)
- Gómez-Ríos, A., Tabik, S., Luengo, J., Shihavuddin, A. S. M., Krawczyk, B., & Herrera, F.

- (2019). Towards highly accurate coral texture images classification using deep convolutional neural networks and data augmentation. *Expert Systems with Applications*. <https://doi.org/10.1016/j.eswa.2018.10.010>
- Harder, E. M., Desai, O., & Marshall, P. S. (2018). Clinical Probability Tools for Deep Venous Thrombosis, Pulmonary Embolism, and Bleeding. *Clinics in Chest Medicine*, 39(3), 473–482. <https://doi.org/https://doi.org/10.1016/j.ccm.2018.04.001>
- Harding, J. R. (1995). Thermal imaging in the investigation of deep venous thrombosis. *Proceedings of 17th International Conference of the Engineering in Medicine and Biology Society*, 2, 1972–1974. <https://doi.org/10.1109/IEMBS.1995.579899>
- Hardy, T. J., & Bevis, P. M. (2019). Deep vein thrombosis. *Surgery (Oxford)*. <https://doi.org/https://doi.org/10.1016/j.mpsur.2018.12.002>
- Hassib, E. M., El-Desouky, A. I., Labib, L. M., & El-kenawy, E. S. M. (2020). WOA + BRNN: An imbalanced big data classification framework using Whale optimization and deep neural network. *Soft Computing*. <https://doi.org/10.1007/s00500-019-03901-y>
- Hong, J., Lee, J. H., Yhim, H. Y., Choi, W. Il, Bang, S. M., Lee, H., & Oh, D. (2018). Incidence of venous thromboembolism in Korea from 2009 to 2013. In *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0191897>
- Hsu, W.-Y. (2018). A decision-making mechanism for assessing risk factor significance in cardiovascular diseases. *Decision Support Systems*, 115, 64–77. <https://doi.org/https://doi.org/10.1016/j.dss.2018.09.004>
- Hu, T., Tang, T., Lin, R., Chen, M., Han, S., & Wu, J. (2020). A simple data augmentation algorithm and a self-adaptive convolutional architecture for few-shot fault diagnosis under different working conditions. *Measurement: Journal of the International Measurement Confederation*. <https://doi.org/10.1016/j.measurement.2020.107539>
- Huang, C. L., Wu, Y. T. Y. W., Hwang, C. L., Jong, Y. S., Chao, C. L., Chen, W. J., Wu, Y. T. Y. W., & Yang, W. S. (2011). The application of infrared thermography in

- evaluation of patients at high risk for lower extremity peripheral arterial disease. *Journal of Vascular Surgery*. <https://doi.org/10.1016/j.jvs.2011.03.287>
- Hull, C. M., Rajendran, D., & Fernandez Barnes, A. (2016). Deep vein thrombosis and pulmonary embolism in a mountain guide: Awareness, diagnostic challenges, and management considerations at altitude. *Wilderness and Environmental Medicine*. <https://doi.org/10.1016/j.wem.2015.10.010>
- Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. In *Computer*. <https://doi.org/10.1109/2.485891>
- Janghel, R. R., Verma, A., & Rathore, Y. K. (2020). Performance Comparison of Machine Learning Techniques for Epilepsy Classification and Detection in EEG Signal. *Advances in Intelligent Systems and Computing*. https://doi.org/10.1007/978-981-32-9949-8_29
- Jin, C., He, Z., Zhang, S., Qi, M., Sun, Z., Di, D., & Liu, J. (2012). A feasible method for measuring the blood flow velocity in superficial artery based on the laser induced dynamic thermography. *Infrared Physics and Technology*. <https://doi.org/10.1016/j.infrared.2012.07.007>
- Kacmaz, S., Ercelebi, E., Zengin, S., & Cindoruk, S. (2017). The use of infrared thermal imaging in the diagnosis of deep vein thrombosis. *Infrared Physics & Technology*, 86, 120–129. <https://doi.org/https://doi.org/10.1016/j.infrared.2017.09.005>
- Kadhim, M. A. (2018). ScienceDirect FNDSB : A fuzzy-neuro decision support system for back pain diagnosis. *Cognitive Systems Research*, 52, 691–700. <https://doi.org/10.1016/j.cogsys.2018.08.021>
- Kearon, C., Julian, J. A., Newman, T. E., & Ginsberg, J. S. (1998). Noninvasive diagnosis of deep venous thrombosis. McMaster Diagnostic Imaging Practice Guidelines Initiative. *Annals of Internal Medicine*. <https://doi.org/10.7326/0003-4819-128-8-199804150-00011>
- Khan, S. H., Hayat, M., Bennamoun, M., Sohel, F. A., & Togneri, R. (2018). Cost-sensitive

- learning of deep feature representations from imbalanced data. *IEEE Transactions on Neural Networks and Learning Systems*.
<https://doi.org/10.1109/TNNLS.2017.2732482>
- Kirchhof, M. G., Lee, A. Y. Y., & Dutz, J. P. (2014). D-Dimer Levels as a Marker of Cutaneous Disease Activity: Case Reports of Cutaneous Polyarteritis Nodosa and Atypical Recurrent Urticaria. *JAMA Dermatology*, *150*(8), 880–884.
<https://doi.org/10.1001/jamadermatol.2013.9944>
- Kuipers, S., Cannegieter, S. C., Middeldorp, S., Robyn, L., Büller, H. R., & Rosendaal, F. R. (2007). The absolute risk of venous thrombosis after air travel: A cohort study of 8,755 employees of international organisations. *PLoS Medicine*.
<https://doi.org/10.1371/journal.pmed.0040290>
- Lahiri, B. B., Bagavathiappan, S., Jayakumar, T., & Philip, J. (2012). Medical applications of infrared thermography: A review. In *Infrared Physics and Technology*.
<https://doi.org/10.1016/j.infrared.2012.03.007>
- Landefeld, C. (2008). Noninvasive diagnosis of deep vein thrombosis. *JAMA*, *300*(14), 1696–1697. <http://dx.doi.org/10.1001/jama.300.14.1696>
- Lin, L., Hu, P. J. H., & Liu Sheng, O. R. (2006). A decision support system for lower back pain diagnosis: Uncertainty management and clinical evaluations. *Decision Support Systems*. <https://doi.org/10.1016/j.dss.2005.10.007>
- Liu, M., Wang, M., Wang, J., & Li, D. (2013). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators, B: Chemical*. <https://doi.org/10.1016/j.snb.2012.11.071>
- Liu, S., Zhang, F., Xie, L., Wang, Y., Xiang, Q., Yue, Z., Feng, Y., Yang, Y., Li, J., Luo, L., & Yu, C. (2019). Machine learning approaches for risk assessment of peripherally inserted Central catheter-related vein thrombosis in hospitalized patients with cancer. *International Journal of Medical Informatics*.
<https://doi.org/10.1016/j.ijmedinf.2019.06.001>

- Liu, Z., Cao, Y., Li, Y., Xiao, X., Qiu, Q., Yang, M., Zhao, Y., & Cui, L. (2020). Automatic diagnosis of fungal keratitis using data augmentation and image fusion with deep convolutional neural network. *Computer Methods and Programs in Biomedicine*. <https://doi.org/10.1016/j.cmpb.2019.105019>
- Luo, J., Yan, X., & Tian, Y. (2020). Unsupervised quadratic surface support vector machine with application to credit risk assessment. *European Journal of Operational Research*. <https://doi.org/10.1016/j.ejor.2019.08.010>
- Martínez-Murillo, C., Romo-Jiménez, A., Zavala-Hernández, C., Gaminio-Gómez, E., Montaña-Figueroa, E. H., Ramos-Peñafiel, C. O., & Collazo-Jaloma, J. (2010). Trombofilia primaria en México: experiencia de una institución. *Rev Med Hosp Gen Mex*, 73(4), 225–230.
- Martins, T. D., Annichino-Bizzacchi, J. M., Romano, A. V. C., & Maciel Filho, R. (2020). Artificial neural networks for prediction of recurrent venous thromboembolism. *International Journal of Medical Informatics*. <https://doi.org/10.1016/j.ijmedinf.2020.104221>
- Masters, T. (1993). Practical Neural Networks Recipes in C++.pdf. In *Book*.
- Middeldorp, S., & Levi, M. (2007). Thrombophilia: An Update. *Semin Thromb Hemost*, 33(06), 563–572. <https://doi.org/10.1055/s-2007-985752>
- Mishra, A., & Ashraf, M. Z. (2020). Using Artificial Intelligence to Manage Thrombosis Research, Diagnosis, and Clinical Management. *Seminars in Thrombosis and Hemostasis*. <https://doi.org/10.1055/s-0039-1697949>
- Modi, S., Deisler, R., Gozel, K., Reicks, P., Irwin, E., Brunsvold, M., Banton, K., & Beilman, G. J. (2016). Wells criteria for DVT is a reliable clinical tool to assess the risk of deep venous thrombosis in trauma patients. *World Journal of Emergency Surgery*, 11(1). <https://doi.org/10.1186/s13017-016-0078-1>
- Morales, A., Campos, M., Juarez, J. M., Canovas-Segura, B., Palacios, F., & Marin, R. (2018). A decision support system for antibiotic prescription based on local

- cumulative antibiograms. *Journal of Biomedical Informatics*, 84, 114–122.
<https://doi.org/https://doi.org/10.1016/j.jbi.2018.07.003>
- Moumneh, T., Penaloza, A., & Roy, P. M. (2018). Trombosis venosa profunda. *EMC - Tratado de Medicina*, 22(1), 1–6. [https://doi.org/10.1016/S1636-5410\(17\)87867-3](https://doi.org/10.1016/S1636-5410(17)87867-3)
- Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., Das, S. R., de Ferranti, S., Després, J.-P., Fullerton, H. J., Howard, V. J., Huffman, M. D., Isasi, C. R., Jiménez, M. C., Judd, S. E., Kissela, B. M., Lichtman, J. H., Lisabeth, L. D., Liu, S., ... Turner, M. B. (2015). Heart Disease and Stroke Statistics—2016 Update. *Circulation*. <https://doi.org/10.1161/cir.0000000000000350>
- Muth, S., Dort, S., Sebag, I. A., Blais, M.-J., & Garcia, D. (2011). Unsupervised dealiasing and denoising of color-Doppler data. *Medical Image Analysis*, 15(4), 577–588. <https://doi.org/https://doi.org/10.1016/j.media.2011.03.003>
- Myers, D., Lester, P., Adili, R., Hawley, A., Durham, L., Dunivant, V., Reynolds, G., Crego, K., Zimmerman, Z., Sood, S., Sigler, R., Fogler, W., Magnani, J., Holinstat, M., & Wakefield, T. (2020). A new way to treat proximal deep venous thrombosis using E-selectin inhibition. *Journal of Vascular Surgery: Venous and Lymphatic Disorders*. <https://doi.org/10.1016/j.jvsv.2019.08.016>
- Nafee, T., Gibson, C. M., Travis, R., Yee, M. K., Kerneis, M., Chi, G., AlKhalfan, F., Hernandez, A. F., Hull, R. D., Cohen, A. T., Harrington, R. A., & Goldhaber, S. Z. (2020). Machine learning to predict venous thrombosis in acutely ill medical patients. *Research and Practice in Thrombosis and Haemostasis*. <https://doi.org/10.1002/rth2.12292>
- Osman, A., Kara, S., Salli, A., Emin, M., & Günes, S. (2010). *Advances in Engineering Software Medical diagnosis of rheumatoid arthritis disease from right and left hand Ulnar artery Doppler signals using adaptive network based fuzzy inference system (ANFIS) and MUSIC method.* 41, 1295–1301. <https://doi.org/10.1016/j.advengsoft.2010.10.001>
- Oudega, R., Moons, K. G. M., & Hoes, A. W. (2005). Ruling out deep venous thrombosis

in primary care. A simple diagnostic algorithm including D-dimer testing. *Thrombosis and Haemostasis*. <https://doi.org/10.1160/TH04-12-0829>

Ozkan, I. A., Koklu, M., & Sert, I. U. (2018). Diagnosis of urinary tract infection based on artificial intelligence methods. *Computer Methods and Programs in Biomedicine*, *166*, 51–59. <https://doi.org/https://doi.org/10.1016/j.cmpb.2018.10.007>

Pano-Azucena, A., Tlelo-Cuautle, E., Tan, S., Ovilla-Martinez, B., & de la Fraga, L. (2018). FPGA-Based Implementation of a Multilayer Perceptron Suitable for Chaotic Time Series Prediction. *Technologies*. <https://doi.org/10.3390/technologies6040090>

Paola Paolinelli, G. (2013). Principios físicos e indicaciones clínicas del ultrasonido doppler. *Revista Médica Clínica Las Condes*. [https://doi.org/10.1016/S0716-8640\(13\)70139-1](https://doi.org/10.1016/S0716-8640(13)70139-1)

Paola Paolinelli Médico Radiólogo, D. G. (2004). *Ecografía Doppler: Principios y aplicaciones*. 15.

Parry, B. A., Chang, A. M., Schellong, S. M., House, S. L., Fermann, G. J., Deadmon, E. K., Giordano, N. J., Chang, Y., Cohen, J., Robak, N., Singer, A. J., Mulrow, M., Reibling, E. T., Francis, S., Griffin, S. M., Prochaska, J. H., Davis, B., McNelis, P., Delgado, J., ... Kabrhel, C. (2018). International, multicenter evaluation of a new D-dimer assay for the exclusion of venous thromboembolism using standard and age-adjusted cut-offs. *Thrombosis Research*. <https://doi.org/10.1016/j.thromres.2018.04.003>

PC, S., IK, I., SZ, G., G, P., CB, B., & R, K. (2015). Performance of wells score for deep vein thrombosis in the inpatient setting. *JAMA Internal Medicine*, *175*(7), 1112–1117. <https://doi.org/10.1001/jamainternmed.2015.1687>

Pedraza García, J., Valle Alonso, J., Ceballos García, P., Rico Rodríguez, F., Aguayo López, M. Á., & Muñoz-Villanueva, M. del C. (2018). Comparison of the Accuracy of Emergency Department-Performed Point-of-Care-Ultrasound (POCUS) in the Diagnosis of Lower-Extremity Deep Vein Thrombosis. *Journal of Emergency Medicine*. <https://doi.org/10.1016/j.jemermed.2017.12.020>

- Peña-Bautista, C., Durand, T., Oger, C., Baquero, M., Vento, M., & Cháfer-Pericás, C. (2019). Assessment of lipid peroxidation and artificial neural network models in early Alzheimer Disease diagnosis. *Clinical Biochemistry*. <https://doi.org/10.1016/j.clinbiochem.2019.07.008>
- Qatawneh, Z., Alshraideh, M., Almasri, N., Tahat, L., & Awidi, A. (2019). Clinical decision support system for venous thromboembolism risk classification. *Applied Computing and Informatics*, 15(1), 12–18. <https://doi.org/https://doi.org/10.1016/j.aci.2017.09.003>
- Reardon, P. M., Patrick, S., Taljaard, M., Thavorn, K., Nemnom, M. J., Mukarram, M., Kim, S. M., Le Gal, G., Huang, L., & Thiruganasambandamoorthy, V. (2019). Diagnostic Accuracy and Financial Implications of Age-Adjusted D-Dimer Strategies for the Diagnosis of Deep Venous Thrombosis in the Emergency Department. *Journal of Emergency Medicine*. <https://doi.org/10.1016/j.jemermed.2019.01.027>
- Resano, R. B., & Resano, F. J. B. (2018). Estudio básico ante una trombosis venosa profunda. *FMC Formacion Medica Continuada En Atencion Primaria*. <https://doi.org/10.1016/j.fmc.2018.01.006>
- Rodríguez, J. D., Pérez, A., & Lozano, J. A. (2010). Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/TPAMI.2009.187>
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*. <https://doi.org/10.1037/h0042519>
- Rosson, M. B., & Carroll, J. M. (2002). Usability engineering : scenario-based development of human-computer interaction. In *Interface*.
- Sadrawi, M., Sun, W. Z., Ma, M. H. M., Yeh, Y. T., Abbod, M. F., & Shieh, J. S. (2018). Ensemble genetic fuzzy neuro model applied for the emergency medical service via unbalanced data evaluation. *Symmetry*. <https://doi.org/10.3390/SYM10030071>
- Saxena, A., Ng, E. Y. K., & Raman, V. (2018). Thermographic venous blood flow

- characterization with external cooling stimulation. *Infrared Physics and Technology*.
<https://doi.org/10.1016/j.infrared.2018.02.001>
- Segal, J. B., Eng, J., Tamariz, L. J., & Bass, E. B. (2007). Review of the evidence on diagnosis of deep venous thrombosis and pulmonary embolism. In *Annals of Family Medicine*. <https://doi.org/10.1370/afm.648>
- Sharif, S., Eventov, M., Kearon, C., Parpia, S., Li, M., Jiang, R., Sneath, P., Fuentes, C. O., Marriott, C., & de Wit, K. (2019). Comparison of the age-adjusted and clinical probability-adjusted D-dimer to exclude pulmonary embolism in the ED. *American Journal of Emergency Medicine*. <https://doi.org/10.1016/j.ajem.2018.07.053>
- Shen, J., Chen, J., Zheng, Z., Zheng, J., Liu, Z., Song, J., Wong, S. Y., Wang, X., Huang, M., Fang, P. H., Jiang, B., Tsang, W., He, Z., Liu, T., Akinwunmi, B., Wang, C. C., Zhang, C. J. P., Huang, J., & Ming, W. K. (2020). An innovative artificial intelligence-based app for the diagnosis of gestational diabetes mellitus (GDM-AI): Development study. *Journal of Medical Internet Research*. <https://doi.org/10.2196/21573>
- Shen, Y., Yuan, K., Chen, D., Colloc, J., Yang, M., Li, Y., & Lei, K. (2018). An ontology-driven clinical decision support system (IDDAP) for infectious disease diagnosis and antibiotic prescription. *Artificial Intelligence in Medicine*, 86, 20–32. <https://doi.org/https://doi.org/10.1016/j.artmed.2018.01.003>
- Simões, A. S., Maia, M. R., Gregório, J., Couto, I., Asfeldt, A. M., Simonsen, G. S., Póvoa, P., Viveiros, M., & Lapão, L. V. (2018). Participatory implementation of an antibiotic stewardship programme supported by an innovative surveillance and clinical decision-support system. *Journal of Hospital Infection*, 100(3), 257–264. <https://doi.org/https://doi.org/10.1016/j.jhin.2018.07.034>
- Sohn, I. (2019). A robust complex network generation method based on neural networks. *Physica A: Statistical Mechanics and Its Applications*. <https://doi.org/https://doi.org/10.1016/j.physa.2019.02.046>
- Sprockel, J. J., Diaztagle, J. J., Alzate, W., & González, E. (2014). Redes neuronales en el

- diagnóstico del infarto agudo de miocardio. *Revista Colombiana de Cardiología*.
<https://doi.org/10.1016/j.rccar.2013.10.001>
- Srivastava, A. K., Kumar, Y., & Singh, P. K. (2020). Computer aided diagnostic system based on SVM and K harmonic mean based attribute weighting method. *Obesity Medicine*. <https://doi.org/10.1016/j.obmed.2020.100270>
- Stevenson, J. G. (2018). The Development of Color Doppler Echocardiography: Innovation and Collaboration. *Journal of the American Society of Echocardiography*, 31(12), 1344–1352. <https://doi.org/https://doi.org/10.1016/j.echo.2018.08.005>
- Subasi, A., Alzahrani, S., Aljuhani, A., & Aljedani, M. (2018). Comparison of Decision Tree Algorithms for Spam E-mail Filtering. *1st International Conference on Computer Applications and Information Security, ICCAIS 2018*. <https://doi.org/10.1109/CAIS.2018.8442016>
- Sun, H., Zeng, X., Xu, T., Peng, G., & Ma, Y. (2020). Computer-Aided Diagnosis in Histopathological Images of the Endometrium Using a Convolutional Neural Network and Attention Mechanisms. *IEEE Journal of Biomedical and Health Informatics*. <https://doi.org/10.1109/JBHI.2019.2944977>
- Syed-Abdul, S., Firdani, R. P., Chung, H. J., Uddin, M., Hur, M., Park, J. H., Kim, H. W., Gradišek, A., & Dovgan, E. (2020). Artificial Intelligence based Models for Screening of Hematologic Malignancies using Cell Population Data. *Scientific Reports*. <https://doi.org/10.1038/s41598-020-61247-0>
- T, T., N, K., G, L. G., & PS, W. (2018). Venous thromboembolism: Advances in diagnosis and treatment. *JAMA*, 320(15), 1583–1594. <https://doi.org/10.1001/jama.2018.14346>
- Taha, M. A. H., Busuttil, A., Bootun, R., & Davies, A. H. (2019). A systematic review on the use of deep venous stenting for acute venous thrombosis of the lower limb. *Phlebology*. <https://doi.org/10.1177/0268355518772760>
- Takeuchi, H., & Kodama, N. (2017). Time-series data analysis of home blood flow velocity measurements from radial artery in relation to lifestyle factors. *2017 IEEE EMBS*

- International Conference on Biomedical & Health Informatics (BHI)*, 425–428.
<https://doi.org/10.1109/BHI.2017.7897296>
- Tan, J. H., Fujita, H., Sivaprasad, S., Bhandary, S. V., Rao, A. K., Chua, K. C., & Acharya, U. R. (2017). Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. *Information Sciences*.
<https://doi.org/10.1016/j.ins.2017.08.050>
- Tang, H., Zhang, L., Zhao, W., Chen, R., & Xie, M. (2018). A case of IgA nephropathy with deep venous thrombosis in the mesentery and lower extremities. *Quantitative Imaging in Medicine and Surgery*. <https://doi.org/10.21037/qims.2018.11.05>
- Tlelo-Cuautle, E., Rangel-Magdaleno, J. de J., & de la Fraga, L. G. (2016). Engineering applications of FPGAs: Chaotic systems, artificial neural networks, random number generators, and secure communication systems. In *Engineering Applications of FPGAs: Chaotic Systems, Artificial Neural Networks, Random Number Generators, and Secure Communication Systems*. <https://doi.org/10.1007/978-3-319-34115-6>
- Torres Macho, J., García Sánchez, F. J., Garmilla Ezquerro, P., Beltrán Romero, L., Canora Lebrato, J., Casas Rojo, J. M., Arribas Arribas, P., López Palmero, S., Pintos Martínez, S., Cepeda Rodrigo, J. M., Luordo, D., Beltrán López, M., Méndez Bailón, M., Rodilla Sala, E., Manzano Espinosa, L., Zapatero Gaviria, A., & García de Casasola, G. (2018). Positioning document on incorporating point-of-care ultrasound in Internal Medicine departments. *Revista Clinica Espanola*.
<https://doi.org/10.1016/j.rce.2018.02.001>
- Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*.
[https://doi.org/10.1016/S0895-4356\(96\)00002-9](https://doi.org/10.1016/S0895-4356(96)00002-9)
- Uzoka, F. E., Osuji, J., & Obot, O. (2011). *Expert Systems with Applications Clinical decision support system (DSS) in the diagnosis of malaria : A case comparison of two soft computing methodologies*. 38, 1537–1553.
<https://doi.org/10.1016/j.eswa.2010.07.068>

- van der Hulle, T., Dronkers, C. E. A., Huisman, M. V., & Klok, F. A. (2016). Current standings in diagnostic management of acute venous thromboembolism: Still rough around the edges. *Blood Reviews*. <https://doi.org/10.1016/j.blre.2015.07.002>
- Wang, X., Yang, Y. Q., Liu, S. H., Hong, X. Y., Sun, X. F., & Shi, J. hong. (2020). Comparing different venous thromboembolism risk assessment machine learning models in Chinese patients. *Journal of Evaluation in Clinical Practice*. <https://doi.org/10.1111/jep.13324>
- Wells, P. (2020). Predictive analytics by deep machine learning: A call for next-gen tools to improve health care. *Research and Practice in Thrombosis and Haemostasis*. <https://doi.org/10.1002/rth2.12297>
- Wells, P. S., Anderson, D. R., Bormanis, J., Guy, F., Mitchell, M., Gray, L., Clement, C., Robinson, K. S., & Lewandowski, B. (1997). Value of assessment of pretest probability of deep-vein thrombosis in clinical management. *The Lancet*, 350(9094), 1795–1798. [https://doi.org/https://doi.org/10.1016/S0140-6736\(97\)08140-3](https://doi.org/https://doi.org/10.1016/S0140-6736(97)08140-3)
- Wells, P. S., Owen, C., Doucette, S., Fergusson, D., & Tran, H. (2006). Does this patient have deep vein thrombosis? In *Journal of the American Medical Association*. <https://doi.org/10.1001/jama.295.2.199>
- Werbos, P. J. (1974). Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. In *PhD Thesis, Harvard U*. <https://doi.org/10.1.1.41.8085>
- White, R. H. (2003). The epidemiology of venous thromboembolism. In *Circulation*. <https://doi.org/10.1161/01.CIR.0000078468.11849.66>
- Willan, J., Katz, H., & Keeling, D. (2019). The use of artificial neural network analysis can improve the risk-stratification of patients presenting with suspected deep vein thrombosis. *British Journal of Haematology*. <https://doi.org/10.1111/bjh.15780>
- Yang, L., Tao, L., Chen, X., & Gu, X. (2020). Multi-scale semantic feature fusion and data augmentation for acoustic scene classification. *Applied Acoustics*. <https://doi.org/10.1016/j.apacoust.2020.107238>

- Yao, X. (1999). Evolving artificial neural networks. *Proceedings of the IEEE*, 87(9), 1423–1447. <https://doi.org/10.1109/5.784219>
- Zhang, Z. (2016). Introduction to machine learning: K-nearest neighbors. *Annals of Translational Medicine*. <https://doi.org/10.21037/atm.2016.03.37>
- Zubiate, P., Urrutia, A., Zamarreño, C. R., Egea-Urra, J., Fernández-Irigoyen, J., Giannetti, A., Baldini, F., Díaz, S., Matias, I. R., Arregui, F. J., Santamaría, E., Chiavaioli, F., & Villar, I. Del. (2019). Fiber-based early diagnosis of venous thromboembolic disease by label-free D-dimer detection. *Biosensors and Bioelectronics: X*, 100026. <https://doi.org/https://doi.org/10.1016/j.biosx.2019.100026>
- Zuluaga-Gomez, J., Al Masry, Z., Benaggoune, K., Meraghni, S., & Zerhouni, N. (2020). A CNN-based methodology for breast cancer diagnosis using thermal images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*. <https://doi.org/10.1080/21681163.2020.1824685>

APÉNDICE A. CONJUNTO DE REGISTROS CLÍNICOS

edad	genero	cancer	inmovilización	cirugia	dolor	Inf pierna	Inf tobillo	edema	Vena superficial	TVP previa	diagnostico
5	0	0	1	1	1	1	1	1	0	0	1
7	1	0	1	1	1	0	1	1	1	0	1
7	1	0	0	1	1	1	1	1	1	1	1
8	1	0	0	1	1	1	1	1	1	0	1
5	0	1	1	1	1	1	1	0	0	1	1
6	0	0	1	1	1	1	0	1	1	1	1
8	0	0	1	1	1	0	1	1	1	0	1
5	1	1	1	1	1	1	0	0	1	0	1
7	0	0	1	1	1	0	1	0	1	0	1
7	1	1	1	1	0	1	1	1	0	0	1
6	1	0	0	1	1	1	1	1	1	0	1
5	0	0	1	0	1	1	0	1	1	0	1
6	1	0	1	0	0	1	1	1	1	0	1
8	0	0	1	1	1	1	0	0	1	1	1
7	0	1	1	0	1	1	0	0	1	0	1
5	1	0	1	1	1	1	1	1	0	0	1
8	1	0	1	1	1	0	1	0	1	0	1
9	0	0	1	1	1	1	1	1	0	0	1
8	1	0	1	0	1	1	1	1	1	1	1
5	1	1	1	1	1	0	1	1	1	0	1
9	1	0	1	1	0	1	1	0	1	0	1
6	0	0	0	1	1	1	1	1	1	1	1
5	0	0	1	0	1	0	1	1	1	0	1
7	1	1	1	1	0	0	1	0	0	0	1

CONTINUACIÓN DEL CONJUNTO DE REGISTROS CLÍNICOS

edad	genero	cancer	inmovilización	cirugia	dolor	Inf pierna	Inf tobillo	edema	Vena superficial	TVP previa	diagnostico
9	0	0	1	0	1	0	1	0	1	1	1
7	0	1	1	1	0	1	1	1	0	0	1
7	0	0	0	1	0	1	1	1	1	0	1
5	1	0	1	0	1	1	1	1	0	0	1
8	1	1	1	0	1	1	0	1	0	0	1
7	1	0	1	0	1	1	0	1	1	0	1
7	1	0	1	1	0	1	1	1	0	0	1
8	0	0	0	0	1	0	1	1	1	0	1
8	1	1	1	0	1	1	1	1	1	0	1
6	0	0	1	1	0	0	1	1	0	1	1
5	0	1	0	0	1	1	1	1	0	1	1
8	0	0	0	1	1	1	1	1	0	1	1
8	0	0	1	1	0	0	0	0	0	0	0
5	1	1	0	0	0	1	0	1	1	0	0
6	0	0	1	1	1	0	0	1	1	0	0
5	1	1	1	0	0	1	1	0	1	0	0
6	0	1	1	1	1	1	0	1	1	0	0
6	0	1	1	1	1	0	0	0	1	0	0
5	0	0	1	0	0	0	1	0	0	0	0
9	1	0	0	0	0	0	1	1	0	0	0
7	0	0	0	0	0	1	1	0	1	0	0
8	1	1	0	0	0	0	0	0	1	0	0
8	0	0	1	0	1	0	0	0	0	0	0
5	0	0	1	0	0	0	1	0	1	0	0

CONTINUACIÓN DEL CONJUNTO DE REGISTROS CLÍNICOS

edad	genero	cancer	inmovilización	cirugia	dolor	Inf pierna	Inf tobillo	edema	Vena superficial	TVP previa	diagnostico
8	1	0	1	1	0	0	1	0	0	0	0
8	1	0	0	1	0	0	1	0	0	0	0
5	0	1	0	1	0	0	0	1	0	0	0
5	0	1	0	1	0	0	0	0	0	1	0
8	1	0	0	1	1	0	1	0	0	0	0
7	1	0	1	0	1	0	0	0	0	0	0
4	0	0	1	0	1	1	1	1	0	1	0
4	0	0	0	1	1	0	1	0	0	0	0
7	0	0	1	0	0	0	0	1	1	0	0
6	0	0	0	0	1	0	0	1	1	0	0
7	1	0	0	1	0	1	0	0	0	0	0

APÉNDICE B. CÓDIGO DE ENTRENAMIENTO DEL MODELO DE PREDICCIÓN

```

%Importar el Dataset por medio de Matlab. Nombrarlo Dataset
%cargar el modelo de red neuronal final (cambiar ruta a donde se tenga guardado en la computadora donde se
evaluará).
load('C:\Users\bfong\Documents\ECITEC\2019-
2\Investigacion\CDSS\Finales\Matrices\ModeloRedNeuronal.mat')

for i=1:10000
    for j=1:11
        X(i,j)=Dataset(i,j);
    end
    Y(i)=Dataset(i,12);
end

p=X';
y=Y';

% entrenar la red neuronal
net=newff(minmax(p),[11 150 100 50 1],{'tansig', 'tansig', 'tansig', 'tansig', 'purelin'},'traincgf');
net=init(net);
net.trainparam.goal=1e-2;
net=train(net, p, Y');
a=sim(net, X');

% Variables para buscar el AOT
ep=0; %Errores positivos
en=0; %Errores negativos
ap=0; %Aciertos positivos
an=0; %Aciertos Negativos
PAOt=0;
umbral=0;
AOT=0;

%Procedimiento para calcular el AOT
for i=1:1000

    umbral=umbral+0.001;
    ep=0;
    en=0;
    ap=0;
    an=0;
    APf=0;
    ANf=0;
    EPf=0;
    ENf=0;

    for m=1:10000
        if Y(m)==1 %La salida real es positiva
            if a(m)>=umbral %la predicción de la salida es positiva
                ap=ap+1;
            else

```

```
        en=en+1;
    end
else %La salida real es negativa
    if a(m)<umbral %La predicción de la salida es negativa
        an=an+1;
    else
        ep=ep+1;
    end
end
end
end

pac=((ap+an)/10000)*100;
if pac>PAOt
    PAOt=pac;
    AOT=umbral;
    ENf=en;
    APf=ap;
    ANf=an;
    EPf=ep;
end
end

%mostrar exactitud
PAOt
%mostrar umbral óptimo
AOT
%mostrar aciertos positivos
APf
%mostrar aciertos negativos
ANf
%mostrar errores positivos
EPf
%mostrar errores negativos
ENf
```

APÉNDICE C. CÓDIGO DE OPTIMIZACIÓN DEL UMBRAL

```

% Importar el Dataset por medio de Matlab. Nombrarlo Dataset
% cargar el modelo de red neuronal final (cambiar ruta a donde se tenga guardado en la computadora donde se
evaluará).
load('C:\Users\bfong\Documents\ECITEC\2019-
2\Investigacion\CDSS\Finales\Matrices\ModeloRedNeuronal.mat')

for i=1:10000
    for j=1:11
        X(i,j)=Dataset(i,j);
    end
    Y(i)=Dataset(i,12);
end

p=X';
y=Y';

% evaluar los valores del Dataset en el modelo final
a=sim(net, X');

% Variables para buscar el AOT
ep=0; %Errores positivos
en=0; %Errores negativos
ap=0; %Aciertos positivos
an=0; %Aciertos Negativos
PAOt=0;
umbral=0;
AOT=0;

% Procedimiento para calcular el AOT
for i=1:1000

    umbral=umbral+0.001;
    ep=0;
    en=0;
    ap=0;
    an=0;

    for m=1:10000
        if Y(m)==1 %La salida real es positiva
            if a(m)>=umbral %la predicción de la salida es positiva
                ap=ap+1;
            else
                en=en+1;
            end
        else %La salida real es negativa
            if a(m)<umbral %La predicción de la salida es negativa
                an=an+1;
            else
                ep=ep+1;
            end
        end
    end
end
end

```

```
if en<=305
pac=((ap+an)/10000)*100;
if pac>PAOt
PAOt=pac;
AOT=umbral;
ENF=en;
end
end
end
```

```
%mostrar exactitud
PAOt
```

```
%mostrar umbral óptimo
AOT
```

APÉNDICE D. CÓDIGO DEL EVALUADOR CASOS REALES

```

% Importar matriz de casos reales. Nombrar DatasetReal
% cargar el modelo de red neuronal final (cambiar ruta a donde se tenga guardado en la computadora donde se
evaluará).
load('C:\Users\bfong\Documents\ECITEC\2019-
2\Investigacion\CDSS\Finales\Matrices\ModeloRedNeuronal.mat')

for i=1:59
    rY(i)=DatasetReal(i,12);
    for j=1:11
        rX(i,j)= DatasetReal(i,j);
    end
end

% validar la salida de la red con los valores reales
v=sim(net,rX');

% inicializar los valores de las variables resultantes
apv=0;
anv=0;
epv=0;
env=0;

% evaluar las salidas
for m=1:59
    if v(m)>=0.484
        if rY(m)==1
            apv=apv+1;
        else
            epv=epv+1;
        end
    else
        if rY(m)==0
            anv=anv+1;
        else
            env=env+1;
        end
    end
end

% Porcentaje de aciertos PAC
pac=(apv+anv)/59*100
% Porcentaje de error positivo PEP
pep=(epv)/59*100
% Porcentaje de error negativo PEN
pen=(env)/59*100

```

APÉNDICE E. CÓDIGO DE LA VALIDACIÓN C

```
load('C:\Users\bfong\Documents\ECITEC\2019-2\Investigacion\CDSS\Finales\Modelos\ModeloFinal.mat')
load('C:\Users\bfong\Documents\ECITEC\2019-
2\Investigacion\CDSS\Finales\Matrices\matriz10000original.mat')
```

```
xt=X';
yt=Y';
PorcF=0;
% mseF=0;
```

```
p=xt(:,1001);
t=yt(:,1001);
for i=1002:10000
    p=[p xt(:,i)];
    t=[t yt(:,i)];
end
```

```
pp=xt(:,1);
tp=yt(:,1);
for i=2:1000
    pp=[pp xt(:,i)];
    tp=[tp yt(:,i)];
end
```

```
net=newff(minmax(p'),[11 150 100 50 1],{'tansig', 'tansig', 'tansig','tansig','purelin'},'traincgf' );
net=init(net);
net.trainparam.goal=1e-2;
net=train(net, p',t);
```

```
a=sim(net,pp);
```

```
%inicializar los valores de las variables resultantes
```

```
apv=0;
anv=0;
epv=0;
env=0;
%evaluar las salidas
for m=1:1000
    if a(m)>=0.478
        if tp(m)==1
            apv=apv+1;
        else
            epv=epv+1;
        end
    else
        if tp(m)==0
            anv=anv+1;
        else
            env=env+1;
        end
    end
end
end
```

```
apv
anv
epv
env
% Porcentaje de aciertos PAC
pac=(apv+anv)/1000*100
% Porcentaje de error positivo PEP
pep=(epv)/1000*100;
% Porcentaje de error negativo PEN
pen=(env)/1000*100;
```