



Universitat d'Alacant
Universidad de Alicante

Sistemas de Ayuda a la Decisión Clínica en
Enfermedades de Diagnóstico Complejo

Ana Eddy Monsalve Torra



Tesis **Doctorales**

www.eltallerdigital.com

UNIVERSIDAD de ALICANTE

TESIS DOCTORAL

Ana Eddy Monsalve Torra

Sistemas de Ayuda a la Decisión Clínica en
Enfermedades de Diagnóstico Complejo

Universitat d'Alacant
Universidad de Alicante
Universidad de Alicante

Febrero 2017



Universitat d'Alacant
Universidad de Alicante

Departamento de Tecnología Informática y Computación

Escuela Politécnica Superior

Sistemas de Ayuda a la Decisión Clínica en Enfermedades de
Diagnóstico Complejo

Ana Eddy Monsalve Torra

Doctorado en Tecnologías de la Sociedad de la Información

Tesis presentada para aspirar al grado de
DOCTORA POR LA UNIVERSIDAD DE ALICANTE

Dirigida por:

Dr. Daniel Ruiz Fernández

Dr. Antonio Soriano Payá

Febrero 2017

*Dedicado a:
Mi familia
Mi esposo*



Universitat d'Alacant
Universidad de Alicante

Agradecimientos

Pensando en escribir este apartado llegan a mi mente tantos recuerdos, cada uno de los momentos vividos en estos años de trabajo. A través de ellos, he tenido la oportunidad de conocer y vivir experiencias muy enriquecedoras, no sólo en el aspecto académico, sino también en el personal. Es ahora cuando pienso en los diferentes momentos vividos, los buenos y malos, las dificultades atravesadas, pero también aquellos instantes que sacaron de mi una sonrisa que me invitaba a continuar adelante, las diferentes charlas con un buen café. Por tal motivo quisiera agradecer a las personas que han apoyado este proceso de formación; algunas de carácter institucional, académico, otras con un apoyo incondicional y otras por el simplemente hecho de “estar ahí” en el momento adecuado.

A mi familia, por su comprensión y cariño, a mis padres, por su dedicación y fortaleza al estar lejos de ellos, a mis hermanos por su apoyo incondicional. A la familia Torra Barajas y la familia Jaimes Jeres, gracias por su apoyo a través de la distancia.

A mi amado esposo, no tengo palabras para agradecer todo lo que haces por mí. La paciencia infinita que has tenido durante el proceso de formación, el apoyo incondicional y, sobre todo, por creer en mí; por todo eso y más, solo puedo agradecerte eternamente.

A mis directores de tesis, Dr. Daniel Ruiz Fernández y Dr. Antonio Soriano Payá, su disponibilidad constantes, por sus comentarios y contribuciones metodológicas e intelectuales, por los buenos resultados que hemos construido. Muchas gracias por todo su valioso tiempo para llevar a cabo esta tesis a buen puerto.

A Oscar Marín Alonso, gracias por sus valiosas aportaciones en los artículos y por el apoyo recibido en este proceso.

AGRADECIMIENTOS

A la Dra. Marisol Carreño Jaimes, el Dr. Jaime Camacho Mackenzie del Departamento de Cirugía Cardiovascular de la Fundación Cardioinfantil (Colombia), por facilitar los datos para el desarrollo esta investigación y por sus aportaciones.

A Eddy Triana, Dr. Victor Castillo y el Departamento de Cirugía Cardiovascular Pediátrica de la Fundación Cardiovascular de Colombia, por facilitar los datos para el desarrollo esta investigación y por sus aportaciones.

Al Dr. Vicente Benito Gil por su amistad y por ser quien hizo posible trabajar con mis tutores.

A los amigos y compañeros en el viaje de la vida, Andrés A., Anita Lucia, Encarni, Xiomara, Karla, Maribel (por ayudarme en las gestiones administrativas con Extranjería), Maryam, Vicenta (mi casera), Diego Pini-lla, Deivis, José Fernández (PP), Roberto, Sergio, MariAngeles, Caro y Andrea: simplemente por el hecho de ser, y estar. La familia Uribe Londoño por hacerme sentir siempre parte de la familia. Janeth, Luz Dary, Ariadnedi, César G, Melquis Gutierrez, Elsa María, Hernán Villa-roel, Eddy T, Estelita Martínez Esparza y su familia, gracias por apoyarme desde la distancia.

A Marisol (Mi sol), gracias por tus oraciones, por el apoyo constante y hacerme sentir que siempre estabas ahí.

A Javier Cantoral, gracias por ofrecerme tu amistad incondicional y abrirme las puertas de tu familia.

Y bueno, sé que en el camino se han cruzado muchas más personas, a las cuales hago extensivo mi agradecimiento. A todos y cada uno, de los mencionados y no mencionados me llevo los mejores recuerdos.

Resumen

El trabajo de investigación presentado en esta tesis doctoral aborda el uso de algoritmos de inteligencia artificial y su aplicación en los sistemas de ayuda para la toma de decisiones en enfermedades de diagnóstico complejo. La dificultad en la gestión clínica de dichas enfermedades requiere un trabajo en conjunto (área de salud e ingeniería informática) y un intercambio de conocimientos, experiencias e información que permita obtener mejores resultados en el proceso de atención de los enfermos.

La creación de sistemas que contribuyan en la toma de decisiones tiene como propósito mejorar la calidad de los procesos médico-asistenciales en la predicción y clasificación de enfermedades buscando enriquecer el cuidado de los pacientes. Dichos sistemas de ayuda tienen tres componentes fundamentales: base de datos (información del paciente), algoritmos de inteligencia artificial y usuarios (médicos, personal asistencial, etc.). La integración de los componentes mencionados trae mejoras en la atención de los pacientes, la reducción de errores médicos, diagnósticos y protocolos clínicos más acertados, agilidad en los procesos, prevención de eventos adversos, entre otros.

Dada la gran variedad de datos que se recogen en la historia clínica, extraer conocimiento de ellos es una tarea muy difícil de realizar con los métodos tradicionales. Por este motivo se han incorporado redes neuronales artificiales, redes bayesianas y árboles de decisión en el campo médico, logrando importantes contribuciones. En esta tesis doctoral estos modelos se han utilizado en la clasificación del riesgo en cirugía cardíaca y en la predicción de la mortalidad hospitalaria. Los modelos han sido evaluados teniendo en cuenta su precisión, sensibilidad, especificidad. Tales clasificadores son sensibles cuando los datos presentan un alto contenido de ruido, datos incompletos, duplicidad de las variables, texto, amplio rango de magnitudes, etc. Para solucionar estos problemas, se han utilizado herramientas para el preprocesado de los datos, como la normalización, discretización e imputación de datos. Cuando estas fueron aplicadas se observó una mejora en los resultados. Otra dificultad que

RESUMEN

tienen las bases de datos médicas es la alta dimensionalidad. Por lo tanto, se ha optado por la aplicación de técnicas de selección de características principales y por el análisis de cada una de las variables para evitar la duplicidad de la información. Estas técnicas han mejorado notablemente el desempeño de los clasificadores.

En cuanto a los algoritmos utilizados en este trabajo se ha evidenciado un buen resultado en la clasificación de enfermedades y en la predicción de la mortalidad hospitalaria. Para clasificar el riesgo de cirugía cardíaca pediátrica en tres categorías (alto, medio y bajo) se han implementado el perceptrón multicapa, mapas autoorganizados, redes de función de base radial y árboles de decisión. Los resultados muestran que las redes neuronales presentan un buen desempeño con una precisión entre el 81.7% y 99.8%, mientras que los árboles de decisión produjeron errores más elevados en la clasificación. Una adecuada clasificación del riesgo quirúrgico permite al personal médico y asistencial mejorar el plan de atención y por lo tanto reducir las posibles complicaciones y el riesgo de muerte. Teniendo en cuenta esto, los árboles de decisión no son los más adecuados para la clasificación del riesgo, sin embargo presentan excelentes resultados cuando se trata de un resultado binario.

Los algoritmos usados para la predicción de la mortalidad hospitalaria de los pacientes operados de reparación abierta de aneurisma de aorta abdominal han sido el perceptrón multicapa, las redes de función de base radial y las redes bayesianas. Esta parte de la investigación fue desarrollada en dos fases: en la primera fase, la implementación se llevó a cabo utilizando todas las variables de la base de datos. En la segunda fase, se emplearon técnicas de selección de características principales. Los resultados generales respecto a la precisión de los algoritmos en las dos fases fueron entre 91.2% y 96.4%. La diferencia de las pruebas radica en los valores de sensibilidad y especificidad, pues en la primera fase la sensibilidad y especificidad muestran valores entre 52.1% y 72% y entre 92.6% y 97%, respectivamente. En la segunda fase, en la cual se aplicó un método de selección de características y una combinación de todos los algoritmos, se observó un aumento en la sensibilidad con valores entre el 65.5% y 86.8%, es decir, que los algoritmos mejoraron la capacidad para predecir la mortalidad hospitalaria. El algoritmo que mostró el mejor resultado en todas las pruebas realizadas fue la red bayesiana.

El desarrollo de un sistema de ayuda para la toma de decisiones clínicas basado en los métodos de aprendizaje mencionados en este trabajo podría mejorar la clasificación del riesgo en cirugía y la predicción de mortalidad hospitalaria. Además, el sistema sería útil para que los médicos y el personal asistencial puedan evaluar a tiempo el impacto de las decisiones y dar prioridad a las actividades de prevención de eventos adversos.



Universitat d'Alacant
Universidad de Alicante

Abstract

The research presented in this doctoral thesis deals with the use artificial intelligence algorithms and their application in clinical decision support system for complex disease diagnosis. The difficulty in the medical care of these diseases requires a working together (medical and computer engineering area) and sharing of knowledge, experiences and information that allows to obtain better results in the care process of the patients.

The creation of systems that contribute to decision making is intended to facilitate the prediction and classification of diseases and to improving the quality of medical care. The aforementioned support systems have three basic elements: database (patient information), artificial intelligence algorithms and users (doctors, nursing staff, etc.). Integrating the mentioned elements could bring improvement in patient care, reducing medical errors, more accurate diagnoses and clinical protocols, efficient processes and prevention of unfavourable events, among others. Provided the wide range of data collected in the medical record, it represents a challenge to extract knowledge from by using the traditional methods.

For this reason, we have included the artificial neural networks, bayesian networks and the decision tree are widely used in the medical area and they have achieved significant contributions. Due to this, these models have been used in this doctoral thesis for classifying the risk in cardiac surgery and predicting inhospital mortality. The mentioned models have been evaluated bearing in mind their accuracy, sensitivity and specificity. However, the classifiers present a weakness when they use noisy data, for instance missing data, text, wide range of magnitudes, etc. These classifiers are sensitive when the data have a high noise, for example, missing data, duplication of variables, text, wide range of magnitudes, etc. To solve this drawback, tools for preprocessing data such as normalization, discretization and imputation of data have been used and the results have improved. Another difficulty for medical databases is the high dimensionality. This situation has been solved with the application of feature selection techniques and the analysis of each of the variables to avoid

ABSTRACT

duplication of information. These techniques have notably improved the performance of the classifiers

The artificial neural networks, bayesian networks and the decision tree are widely used in the medical area and they have achieved significant contributions. Due to this, these models have been used in this doctoral thesis for classifying the risk in cardiac surgery and predicting inhospital mortality. The mentioned models have been evaluated bearing in mind their accuracy, sensitivity and specificity. However, the classifiers present a weakness when they use noisy data, for instance missing data, text, wide range of magnitudes, etc. To solve this drawback, tools for preprocessing data such as normalization, discretization and imputation of data have been used and the results have improved. Another difficulty for medical databases is the high dimensionality. This situation has been solved with the application of feature selection techniques and the analysis of each of the variables to avoid duplication of information. These techniques have notably improved the performance of the classifiers.

The algorithms used in this work showed an excellent performance in the classification of diseases and in the prediction of inhospital mortality. There were implemented for classifying the risk in congenital heart surgery (three categories (high, medium and low)), as well as a multi-layer perceptron, self-organizing maps, radial basis function and decision tree. The results showed that artificial neural networks had a good performance with an accuracy of 81.7 % and 99.8 %, whereas the decision trees obtained the highest number of errors in the classification. A suitable classification for the risk in congenital heart surgery can allow medical staff to improve the plan of care and therefore to reduce the risk of complications and death. In addition, although decision trees are not the most suitable for risk classification, they have excellent results when it comes to a binary outcome.

The algorithms used for predicting inhospital mortality of patients undergoing open repair of abdominal aortic aneurysm were the multilayer perceptron, radial basis function and bayesian networks. This part of research was developed in two stages: firstly, we used all the variables and later on, we carried out feature selection techniques. The overall results regarding the accuracy of the algorithms in the two phases were between 91.2% and 96.4%. The difference with the test consists in the sensitivity and specificity values; the sensitivity and specificity in the first stage were 52.1% and 72% and 92.6% and 97% respectively. In the second stage, in which a feature selection process was applied, we could see an increase in the sensitivity with values between 65.5% and 86.8%. This means that the algorithms improve the ability for predicting the mortality. The bayesian

network algorithm showed the best performance in all tests.

The development of a clinical decision support system based on the learning methods mentioned in this work could contribute to improve the classification of risk in cardiac surgery prediction and inhospital mortality. In addition, the system can be useful for doctors and nursing staff to assess the impact of decisions and to prioritize the prevention of adverse events.



Universitat d'Alacant
Universidad de Alicante

Índice

Agradecimientos	III
Resumen	V
Abstract	IX
Índice de Figuras	XV
Índice de Tablas	XVII
1 Introducción	1
1.1 Motivación	1
1.2 Hipótesis	4
1.3 Objetivos	5
1.3.1 Objetivo General	5
1.3.2 Objetivos Específicos	5
1.4 Organización del Documento	6
2 Estado del Arte	7
3 Metodología	15
3.1 Algoritmos de inteligencia artificial	16
3.1.1 Redes Neuronales Artificiales	16
3.1.2 Redes Bayesianas	26
3.1.3 Árboles de Decisión	30
3.2 Evaluación de los clasificadores	33
3.3 Recopilación de datos	38
3.4 Procesado de datos	39
3.4.1 Agrupamiento de datos	39
3.4.2 Transformación de Datos	40
3.4.3 Reducción de la dimensionalidad	43
3.4.4 Desbalance de los datos	45

ÍNDICE

4 Resultados	47
5 Conclusiones	53
Anexo I	57
Anexo II	63
Referencias	67
Acrónimos	81



Universitat d'Alacant
Universidad de Alicante

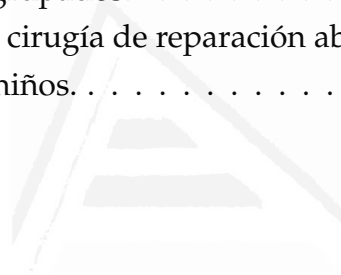
Índice de Figuras

3.1	Ejemplo de un perceptrón multicapa.	19
3.2	Estructura de un mapa autoorganizado.	22
3.3	Estructura de una red de función de base radial.	24
3.4	Grafo de una red bayesiana.	26
3.5	Estructura del clasificador <i>naïve</i> bayes.	27
3.6	Estructura árbol de red aumentada bayes.	29
3.7	Estructura de red bayes aumentada.	30
3.8	Segmento de un árbol de decisión.	31
3.9	Ejemplo de una curva ROC.	37

Universitat d'Alacant
Universidad de Alicante

Índice de Tablas

3.1	Tabla de contingencia	34
3.2	Tabla de contingencia para problemas de dos clases	35
3.3	Ejemplo de datos agrupados.	40
3.4	Tipo de abordaje en cirugía de reparación abierta de AAA.	42
3.5	Rango de pesos en niños.	42



Universitat d'Alacant
Universidad de Alicante

Capítulo 1

Introducción

1.1 Motivación

Todas las enfermedades son importantes, algunas más que otras por su impacto, frecuencia, personas a las que afecta (adultos, niños, bebés, etc.) y la complejidad (enfermedades cardíaca, cáncer, cardiopatías congénitas, etc.). Se entiende por enfermedad compleja aquella patología causada por múltiples factores y su interacción con el ambiente, la alimentación, la actividad física o el consumo de ciertos productos (alcohol, tabaco, etc.), entre otros [1]. Como ejemplo de enfermedad compleja, se pueden mencionar: la hipertensión arterial esencial, enfermedad arteroesclerótica, la diabetes, el cáncer, cardiopatías congénitas o enfermedades del sistema circulatorio y neurológico. El impacto de las enfermedades complejas no sólo afecta la calidad de vida de la persona que la padece, también se ve afectado el entorno familiar y el sistema de sanidad.

En cuanto a las enfermedades que afectan el sistema circulatorio, se pueden destacar las cardiopatías congénitas (CC), la enfermedad coronaria y los aneurismas de aorta abdominal (AAA). Respecto a las CC en Europa recientes trabajos reportados en una base de datos central de 29 registros poblacionales en 16 países mostraron una prevalencia total de 8 por 1000 nacidos vivos [2]. Los cálculos indican que cada año en la Unión Europea nacen 36.000 niños con CC y otros 3.000 son diagnosticados con

CAPÍTULO 1. INTRODUCCIÓN

esta misma dolencia, pero mueren por interrupción del embarazo debido a la anomalía fetal [3]. La incidencia en otros continentes fue reportada como 9.3 en Asia, en África 1.9 y finalmente en América fue del 7.5 por cada 1000 nacidos vivos [4]. En España, se estima que el 5.2 a 12.5 por cada 1.000 nacidos vivos tendrán una CC [5]. En Colombia, se observa que las cardiopatías congénitas tienen una prevalencia entre 7.5 a 9.5 por cada 1000 nacimientos (no discriminan entre nacidos vivos y mortinatos) [6]. Las consecuencias de las enfermedades complejas no sólo se pueden cuantificar en términos de morbilidad y de mortalidad. Su impacto económico directo e indirecto es alto y creciente, provocando importantes consecuencias en el aspecto socioeconómico. En 2006, por ejemplo, las enfermedades cardíacas de los europeos supusieron una pérdida de 41.000 millones de euros en términos de productividad. De este dinero, dos tercios se deben a las muertes prematuras provocadas por estas patologías y otro tercio a los problemas que sufren personas en edad activa [7].

Por otro lado, un AAA se define como el aumento del diámetro de la aorta abdominal en más del 50% de su tamaño original. Afecta de manera significativa a hombres a partir de 65 años. Su prevalencia aumenta paralelamente al incremento de la esperanza de vida y la disminución de la mortalidad cardiovascular. En España, para aneurismas $> 3\text{cm}$ en varones de 65 - 75 años de edad hubo una prevalencia del 4.2% [8]. En estudios poblacionales la prevalencia de aneurismas de 2.9 - 4.9cm varía de un 1.3% en varones de 45 - 54 años a un 12.5% en los de 75 - 84 años y con una relación varón:mujer de 4:1 [9]. Ante el encarecimiento progresivo de la atención médica, el tratamiento quirúrgico con baja morbilidad y mortalidad y de costos controlados, adquiere importancia no sólo para el paciente mismo sino para el sistema de sanidad. En el acceso al tratamiento quirúrgico es importante no sólo para reducir la mortalidad operatoria, sino también la tasa de complicaciones, ya que éstas elevan el costo por la hospitalización prolongada y uso de recursos extraordinarios. De la experiencia del personal médico se deduce que es posible identificar claramente los riesgos pre-operatorios y de esta forma especificar las situaciones reparables. Para ello, realizar un acertado diagnóstico, clasifi-

cación y tratamiento de estas enfermedades es fundamental para mejorar la calidad de vida de los pacientes.

La incorporación de las tecnologías de la información al proceso de diagnóstico y tratamiento esta mejorando la toma de decisiones. En todo caso, las decisiones deben estar garantizadas por criterios de evidencia (medicina basada en la evidencia [10]), guías de práctica clínica y la experiencia del personal médico-asistencial. Para ello, es necesario recopilar y procesar adecuadamente la información referente al estado de salud de los pacientes (historia clínica, pruebas diagnósticas, análisis de laboratorio, etc.). Esta información se concreta en distintos tipos de variables (peso, talla, diagnóstico, edad, género, fecha de nacimiento, presión arterial, etc.). Con estas variables se procede a la aplicación de criterios que permitan llegar a un diagnóstico acertado, la posible evolución de la enfermedad y las posibles complicaciones que se puedan derivar, etc. Considerando la complejidad de la enfermedad, la evaluación de la información debe hacerse desde una perspectiva multivariante, de manera que se consideren al mismo tiempo todas las variables implicadas y se obtenga una generalización adecuada para el tipo de enfermedad [11].

En otro orden, en la práctica médico-asistencial existe un gran interés en conseguir modelos de clasificación y predicción concretos que ayuden a los profesionales de la salud a obtener diagnósticos y tratamientos más acertados. Una de las múltiples alternativas que ofrecen las tecnologías para el análisis de la información son los algoritmos de inteligencia artificial. Las redes neuronales artificiales (RNA), árboles de decisión (AD) y redes bayesianas (RB), son las más usadas en el sector de la salud. En medicina, el uso más común de estos algoritmos es en la clasificación de modelos con el propósito de apoyar al médico en el diagnóstico y tratamiento del paciente y en la predicción de eventos (mortalidad, complicaciones, estancia hospitalaria, etc.). Las RNA se pueden clasificar en base a dos criterios básicos: el modo de aprendizaje y el flujo de información. En una RNA el aprendizaje puede ser supervisado, no supervisado o aprendizaje híbrido. Por otra parte, según el flujo de información que manejan pueden ser: redes alimentadas hacia adelante conocidas como (*redes feedforward*) y

las redes de realimentación total o parcial [12]. Dentro de las redes supervisadas, la estructura más utilizada es el llamado perceptrón multicapa (MLP *multilayer perceptron*). De entre todos los modelos no supervisados los mapas autoorganizados o SOM (*Self-Organizing Maps*) es quizá el más popular. Finalmente, uno de los modelos híbridos más destacados, son las redes de función de base radial o RBF (*Radial Basis Functions*).

Dada la complejidad de las patologías mencionadas, en la presente investigación se propone usar las RNA, los árboles de decisión y las redes bayesianas para analizar las bases de datos (BD) biomédicas de paciente con cardiopatías congénitas (niños) y aneurisma de aorta abdominal (adultos). Estas BD presentan alta dimensionalidad, gran número de variables predictoras y desbalance de datos, haciendo que este problema sea difícil de abordar y solucionar mediante las técnicas estadísticas tradicionales cuando se analizan los datos.

1.2 Hipótesis

La evaluación médica constituye una herramienta primordial para obtener información sobre el estado de salud de los pacientes. De esta forma, se marca el punto de partida de esta investigación que consiste en explorar y conocer si el uso de las máquinas de aprendizaje pueden ser aplicadas en sistemas de ayuda a la toma de decisiones médicas en enfermedades con diagnóstico complejo. Se quiere saber, si el uso de estas metodologías aporta mecanismos que contribuyan a la mejora en el proceso de atención de los pacientes.

Con el avance de las máquinas de aprendizaje ha cambiado la forma de extraer el conocimiento de los datos médicos y por ende, la forma en que el personal médico-asistencial toma las decisiones diagnósticas y terapéuticas que afectan a la salud de los pacientes. Estos métodos también ayudan a que una vez tomada la decisión se pueda conocer el grado influencia alcanzado por dichas decisiones y predecir el comportamiento de las enfermedades.

1.3 Objetivos

Los objetivos de esta investigación se centran en el estudio de algoritmos de inteligencia artificial (máquinas de aprendizaje) y su uso en sistemas de ayuda a la toma de decisiones médicas relacionadas con las enfermedades de diagnóstico complejo.

1.3.1 Objetivo General

El objetivo general de este trabajo de investigación consiste en plantear la fundamentación de un sistema de ayuda a la toma de decisiones médicas utilizando algoritmos de aprendizaje en el contexto de las enfermedades complejas. Para ello, se han tomado datos de la historia clínica de pacientes con cardiopatías congénitas y otra base de datos con los pacientes diagnosticados con aneurisma de aorta abdominal. De esta manera, se estudia la aplicabilidad de éstas metodologías en el apoyo al personal médico-asistencial en la toma de decisiones necesarias en el cuidado de los enfermos.

1.3.2 Objetivos Específicos

Los objetivos específicos de esta investigación son los siguientes:

- Determinar la capacidad que presentan las redes neuronales artificiales y los árboles de decisión para clasificar el riesgo en cirugía cardíaca pediátrica.
- Validar el uso de las redes neuronales artificiales y las redes bayesianas en la predicción de la mortalidad hospitalaria de los pacientes operados de reparación de aneurisma de aorta abdominal.
- Describir la importancia del uso de algoritmos de selección de características principales buscando mejorar el desempeño de los clasificadores.

- Analizar la importancia que tiene el preprocesado de los datos. El tamaño y complejidad de los datos médicos hace necesario la aplicación de estas técnicas para lograr el mejor desempeño de los sistemas.

1.4 Organización del Documento

Este documento de tesis doctoral, está estructurado en cinco capítulos. En el primer capítulo se hace la introducción del tema de investigación, así como la hipótesis, la motivación y los objetivos trazados. En el capítulo dos, se presenta el estado del arte relacionado con el tema a tratar. En el capítulo tres, se detalla la metodología utilizada para el desarrollo de esta investigación. El capítulo cuatro, se centra en los resultados obtenidos con sus respectivos análisis (artículos publicados). Seguidamente, en el capítulo cinco, se comentan las conclusiones y futuras líneas de investigación. Finalmente, se encuentra el apartado de anexos y una lista de los acrónimos utilizados en este trabajo.

Capítulo 2

Estado del Arte

La inteligencia artificial, en particular las redes neuronales artificiales (RNA), los árboles de decisión (AD) y las redes bayesianas (RB) están teniendo un gran desarrollo e impacto en diversas áreas del conocimiento, incluida la medicina. Por esta razón es necesario realizar una revisión del estado actual de los métodos usados en esta investigación y su aplicación en esta área del conocimiento.

Los métodos estudiados se han utilizado en medicina con varios propósitos: clasificación, diagnóstico de enfermedades y predicción de la mortalidad de diversas patologías [13, 14, 15]. En el trabajo de Chimiesk y Ribeiro (2013) [16] se comparan las RNA los AD y las RB en las tareas de clasificación y asociación en pacientes con problemas dermatológicos, columna vertebral y cáncer de mama, analizando cuál es el mejor en cada uno de estos casos. Los resultados mostraron que las RB fueron adecuadas para el diagnóstico de cáncer de mama y en enfermedades de la piel, mientras que los AD mostraron ser mejores para el diagnóstico y clasificación de las patologías de la columna vertebral. Además, en el diagnóstico del melanoma utilizando como entrada el conjunto de imágenes de la lesión cutánea. El sistema analiza la secuencia de imágenes para extraer el área afectada y determina las características que indican el grado de daño. Para ello se proponen varios métodos de clasificación: un perceptrón multicapa, un clasificador bayesiano y el algoritmo K-NN. Estos métodos funcionan inde-

pendientemente y también en combinación haciendo un sistema colaborativo de apoyo a la decisión. Las tasas de clasificación obtenidas son de alrededor del 87% [17].

En el área de hematología se ha usado el perceptrón multicapa (MLP) y las máquinas de vectores (SVM) para la clasificación de la policitemia primaria y secundaria, el desempeño general del sistema de clasificación presentó un 98.9% de acierto [18]. Otro trabajo, presenta la clasificación de las características de la sangre mediante un árbol de decisión C4.5, un clasificador Bayes *naïve* y un perceptrón multicapa para el cribado de la talasemia. El objetivo es clasificar dieciocho clases anómalas de talasemia, que tienen una alta prevalencia en Tailandia. Los resultados indican que el mejor rendimiento de clasificación con una precisión media de 93.23% y 92.60% se logra con el clasificador Bayes *naïve* y el perceptrón multicapa [19]. Un perceptrón multicapa fue utilizado con el objetivo principal de diseñar un método no invasivo para diagnosticar pacientes con cirrosis hepática utilizando únicamente datos de laboratorio. Como resultado del estudio tras comparar las RNA en términos de rendimiento, el modelo fue capaz de reconocer 3 tipos de valores: cirrosis, no cirrosis y los no identificables con una tasa de éxito de casi el 90% [20]. De otra parte, han sido usadas en la clasificación de microcalcificaciones en las mamas y determinar si éstas son de tipo benigno o maligno, los algoritmos utilizados (MLP y SVM) mostraron resultados alentadores, además de mostrar una buena precisión para la clasificación de los diferentes grupos de tumores (82.96% - 90.1%), sin embargo el SVM mostró el mejor rendimiento de la clasificación (90.1%) [21]. Así mismo, han mostrado excelente desempeño en la clasificación del tamaño del feto con el objetivo de mejorar la precisión en la estimación del peso del mismo [22].

En cardiología, se han aplicado RNA para determinar el diagnóstico de la enfermedad coronaria. El modelo que realizó la mejor precisión (93%) fue el MLP que incluyó tanto factores genéticos como no genéticos relacionados con la enfermedad. A pesar de estos buenos resultados, hay que señalar que para algunos modelos la precisión fue inferior al 90% [23]. Las RNA también se han aplicado en otras enfermedades del corazón, tales

como los defectos de las válvulas cardíacas. En este estudio se desarrolló un sistema para la clasificación de las señales cardíacas, obtenidas a partir de 120 pacientes con estenosis pulmonar, mitral y sin estenosis, usando el estetoscopio. La correcta clasificación se logró para el 95% de los diferentes sonidos cardíacos [24]. Otros autores proponen el uso de técnicas de selección de características y posteriormente usan un MLP para identificar la isquemia coronaria. Los resultados mostraron que usando el total de las variables la precisión fue del 70%, mientras que reduciendo a 12 las variables predictoras mejoró el desempeño logrando una precisión del 89.4% [25]. Un MLP es usado con el prototipo de ayudar al dermatólogo en el diagnóstico de seis tipos de enfermedades. Estas enfermedades cutáneas son psoriasis, dermatitis seboreica, liquen plano, pitiriasis rosea, dermatitis crónica y pityriasis rubra pilaris. Realizar un diagnóstico acertado es importante ya que estas enfermedades comparten rasgos clínicos similares de eritema y descamación sin mucha diferencia. Además, una enfermedad puede mostrar la característica de otra enfermedad en las primeras etapas. Los resultados muestran que la red y el uso de selección de características alcanzó el 91.2% de cierto [26].

Diversos autores han hecho estudios que muestran la aplicación de dichos algoritmos en la predicción de enfermedades del corazón, por ejemplo, el MLP ha sido entrenado con el algoritmo de retropropagación para predecir la presencia o ausencia de enfermedad cardíaca en un paciente, el sistema fue probado con diferente número de neuronas en la capa oculta. La precisión más alta fue 98.5% con veinte neuronas en la capa oculta [15]. La investigación de [27] usa la misma RNA y el mismo algoritmo de entrenamiento para predecir la enfermedad coronaria y obtuvo una precisión del 96%, mostrando el buen desempeño de los algoritmos propuestos en comparación con enfoques similares como el propuesto en [28]. Del igual modo han sido aplicadas en el pronóstico de la taquicardia ventricular, fibrilación ventricular y la taquiarritmia ventricular espontánea, los autores presentan tres clasificadores, uno por patología. Los resultados que se obtuvieron son: 82.9% de precisión (71.4% de especificidad) para la taquicardia ventricular, 88.9% de precisión (92.9% de especificidad) en la

fibrilación ventricular, y para la taquiarritmia ventricular fue del 77.3% de precisión (73.8% de especificidad). Para los autores, los resultados observados son adecuados, sin embargo se puede destacar que la predicción de la fibrilación ventricular obtuvo el mejor desempeño [29].

En otras áreas del campo médico se han propuesto las RNA para pronosticar la estancia hospitalaria prolongada de pacientes mayores en el servicio de urgencias, donde el mejor desempeño fue logrado por el MLP con una sensibilidad del 62.5% y especificidad del 96.6% [30]; asimismo, pueden predecir la estancia prolongada en la unidad de cuidados intensivos y la mortalidad con la ayuda de algoritmos como, el vecino más cercano (K-NN), SVM, árboles de clasificación (CART), entre otros. Los datos mostraron que el modelo SVM ofreció los mejores resultados con un valor del área bajo la curva de 0.77 para predecir la estancia y del 0.82 para la predicción de la mortalidad [31]. Por otra parte, en la investigación del cáncer han sido aplicadas por su habilidad para reconocer patrones complejos como, en la predicción de cáncer avanzado de vejiga en pacientes sometidos a cistectomía radical [32], cáncer de mama, el modelo fue evaluado utilizando la matriz de confusión y las curvas ROC. Con los datos de prueba, el modelo RNA implementado dió como resultado una tasa de predicción de la enfermedad del 90.5%. Con estos valores predictivos se concluye que el modelo RNA es rápido, fiable y de gran ayuda en esta especialidad médica [33], además puede predecir la supervivencia después de la recesión hepática por cáncer colorrectal. La investigación comparó el uso de las RNA con el modelo de regresión *C-index* encontrando que a través del análisis de 241 pacientes y de la selección de seis variables de entrada de 28. El modelo RNA predice la supervivencia individual con mayor exactitud que el modelo de regresión de Cox, con de 0.72 frente a 0.66 [34].

Igualmente, han sido utilizadas con gran éxito en predicción de la mortalidad. Por ejemplo, la mortalidad hospitalaria de los pacientes con reparación de aneurisma aórtico abdominal roto, en el cual la RNA mostró mejor desempeño cuando ésta fue comparada con los resultados obtenidos por el método de regresión lineal [35]. En el pronóstico de la mortalidad

de cáncer de vejiga, en este trabajo se utilizaron siete métodos de aprendizaje para predecir la mortalidad a los 5 años después de la cistectomía radical, incluyendo la red neuronal MLP, redes de función de base radial, la máquina de aprendizaje extrema (EML), regularizados ELM (RELM), SVM, las RB *naïve* y el clasificador del vecino más cercano (K-NN). Los resultados experimentales indican que RELM logra la más alta precisión en la predicción con 0.8. Los resultados de la investigación demuestran el potencial de la aplicación de estas técnicas de aprendizaje para apoyar la toma de decisiones clínicas [36]; en cáncer de páncreas el estudio utilizó las RNA para predecir la supervivencia individual a largo plazo de los pacientes sometidos a cirugía radical en el tratamiento de este tipo de cáncer. Los resultados fueron comparados con el modelo de regresión *C-index*, mostrando que las RNA tienen mejor desempeño con 0.79, mientras que el otro modelo mostró un 0.67 de precisión [37]. En pacientes con accidentes cerebrovasculares (isquémico y hemorrágico) para el pronóstico de la mortalidad a 10 días después del evento. La investigación aplicó una red MLP y los métodos estadísticos multivariantes (análisis discriminante multivariado [MDA] y el análisis de regresión logística [LRA]). El MLP tiene una clasificación de 100%, para ambos grupos isquémicos y hemorrágicos en la fase de entrenamiento. Los métodos de LRA y MDA mostraron resultados exitosos, a pesar de que estos resultados son más bajo que el MLP tanto para el grupo hemorrágico y como para el grupo isquémico en la fase de entrenamiento. Sin embargo, cuando se compararon los resultados de los métodos en la fase de prueba, la mejor clasificación, especificidad, sensibilidad y precisión (87.8%) fue alcanzada por el modelo MDA para el grupo hemorrágico. La mejor clasificación, especificidad (75.9%), sensibilidad (85.9%) y precisión (80.9%) para el grupo isquémico fue alcanzada por la técnica de LRA [38]. En [39] proponen el uso de MLP con varios algoritmos de aprendizaje, *quick propagation* (QP), *Levenberg-Marquardt* (LM), *backpropagation* (BP), *quasi-Newton*, *delta bar delta*, y *conjugate gradient descent* (CGD), entre los algoritmos de aprendizaje que fueron entrenados el QP logra los valores más altos con una especificidad (81.3%), sensibilidad (78.4%) y una precisión del 80.7%, y el

área bajo la curva (0.869), mientras que CGD logra los valores más bajos.

Las redes bayesianas se han aplicado en la predicción de la mortalidad a 90 días, 1 y 2 años después del implantar el dispositivo de asistencia ventricular. El modelo inicial cuenta con 298 variables de entrada de las cuales 25 son las que mejor representan la relación causal entre sí y la mortalidad. El modelo mostró una precisión del 91% para la mortalidad a 90 días, 84% de mortalidad a un año y finalmente un 81% de mortalidad a dos años. De lo que se concluye que estas herramientas son adecuadas para predecir la esperanza de vida después del procedimiento [40]. En [41] se construyó un modelo de red bayesiana utilizando una población de 1164 pacientes y posteriormente se validó con 103 pacientes. Los resultados de la validación con un valor predictivo negativo del 91%, demuestran su aplicabilidad para ayudar a los clínicos en la predicción de la angina inestable. Igualmente, las RB se han empleado para predecir las úlceras por presión adquiridas en el hospital y la estancia, el modelo planteado reduce la prevalencia de úlceras diez veces y la estancia en la UCI alrededor de un tercio [42]. Asimismo, diversos modelos usando RB han sido propuestos para la predicción de cáncer [43]. En este trabajo, tres técnicas de clasificación de aprendizaje supervisado se comparan para evaluar las tasas de incidencia y mortalidad por cáncer. A saber, las redes *naïve*, redes bayesianas y el K-NN. Los resultados muestran una precisión para la incidencia del cáncer de 99.90%, 99.21% y 72.45% respectivamente y una precisión de 99.97%, 98.7% y 67.7% para la mortalidad.

Por otro lado, el estudio de Colak y cols (2015) [44] mostró buenos resultados en la predicción de accidentes cerebrovasculares aplicando el MLP, las máquinas de vectores y las RBF, los cuales mostraron una precisión >80%. Las RBF también han mostrado buen desempeño en el diagnóstico de las enfermedades del corazón [45], en el diagnóstico y el análisis de la fibrilación auricular [46]. Debido a su excelente desempeño en el tratamiento de imágenes, las redes SOM han sido aplicadas en la segmentación de imágenes médicas (segmentación de tumores de mamas [47] y cerebrales [48]), además, para la identificación de desordenes como la demencia, la esquizofrenia o el alzheimer [49]. Los AD que tienen entre

otras características su aporte visual para la toma de decisiones, han sido utilizados para identificar los factores de riesgo asociados con la hipertensión arterial. Los modelos fueron evaluados con 9078 sujetos divididos en dos grupos (hipertensos y no hipertensos). Los resultados para el primer modelos fueron 73% (sensibilidad), 63% (especificidad), 77% de precisión. Los valores correspondientes para el segundo modelo fueron 70%, 61%, 74% respectivamente. Según los autores el estudio proporcionó reglas de clasificación fáciles y comprensibles para identificar los factores de riesgo asociados con la hipertensión y esto puede ser útil para desarrollar programas para el manejo de la hipertensión [50]. Como ayuda en la toma de decisiones en paciente con soporte ventricular izquierdo, se estudiaron 183 pacientes con falla ventricular derecha, de los cuales 27 requirieron soporte ventricular izquierdo, los restantes continuaron con manejo médico hasta el transplante o fallecieron. El modelo identificó 8 variables preoperatorias optimas. El árbol de decisión resultante, que constaba de 28 ramas y 15 hojas, identificó a los pacientes que necesitaban el soporte ventricular con una sensibilidad del 85% y un área bajo la curva de 0.87 [51]. En el pronóstico de cáncer pleural [52], además, en la predicción de los síntomas del Parkinson [53].

Como conclusión de este capítulo, indicar que los algoritmos de inteligencia artificial, sus diferentes métodos y herramientas han sido ampliamente usadas en el campo de la medicina con gran éxito. Por otro lado, en esta revisión no se han encontrado trabajos relacionados directamente con el tema de investigación tratado en esta tesis doctoral.

Capítulo 3

Metodología

Este trabajo es fruto de una investigación multidisciplinar que combina diferentes campos de la ciencia como la computación y la salud. Para su desarrollo ha sido necesario el uso de herramientas y algoritmos basados en inteligencia artificial que se muestran a continuación.

La toma de decisiones es un proceso complicado que requiere el manejo de mucha información. Cuando la tasa de errores en este proceso es alto, se convierte en objeto de preocupación por parte de los investigadores [54]. Algunos informes de la Unión Europea y Estados Unidos mencionan que los errores médicos podrían ser en torno a un 10% de los ingresos hospitalarios [55]. Aunque existen numerosas causas de error médico, más allá del error en la toma de decisiones (error de ejecución), se hace necesario la implementación de mecanismos que ayuden a mejorar la actividad médico-asistencial.

Aproximadamente, desde el surgimiento de los primeros ordenadores, se ha intentado que las tecnologías de la información ayuden en la realización del diagnóstico y tratamiento médico. Inicialmente se desarrollaron programas que intentaban emular las habilidades diagnósticas de los clínicos llamados «*sistemas expertos*», con la pretensión de que iban a ser capaces de suplir la actividad deductiva de los médicos. Estos sistemas no tuvieron el éxito esperado porque, aunque demostraban ser capaces de generar hipótesis correctas, han revelado en muchos aspectos, tanto for-

males como prácticos, que condicionan la toma de decisiones en el ejercicio médico [54]. Actualmente, los llamados sistemas de ayuda a la toma de decisiones médicas (en Inglés *Clinical Decision Support Systems* (CDSS)) son menos ambiciosos pero más efectivos. Los CDSS se entienden como un sistema informático diseñado para ayudar a la toma de decisiones clínicas, en los cuales las características individuales de los pacientes se utilizan para generar recomendaciones que son presentadas a los médicos para su consideración [56]. Estos han mostrado mejorar el rendimiento del personal médico-asistencial [57], además de mejorar la calidad de los procesos de atención y el plan del cuidado del paciente.

3.1 Algoritmos de inteligencia artificial

A continuación se describen brevemente los algoritmos empleados en esta investigación.

3.1.1 Redes Neuronales Artificiales

Día a día, el hombre busca nuevas formas para mejorar su calidad de vida. Esto le ha servido para reducir el esfuerzo y el tiempo, sobre todo en aquellas operaciones que requerían un gran sacrificio. Los progresos obtenidos han permitido la construcción de máquinas de todo tipo; desde aquellas dedicadas a procesos industriales, hasta pequeños ordenadores que ayudan a resolver de forma automática y rápida determinadas operaciones que resultan molestas cuando se realizan de manera manual.

Los adelantos actuales se centran en el estudio de las capacidades humanas (inteligencia) y fuentes de nuevas ideas para el desarrollo de nuevos dispositivos (ordenadores, programas, aplicaciones). Así, la inteligencia artificial (IA) es un intento por dar capacidad de razonamiento a máquinas o programas. Este tema se ha venido desarrollando con mayor interés en los últimos años. Teniendo aplicación en campos como el procesamiento de la información, la robótica, la minería de datos, etc. Dentro de la IA, los algoritmos más aplicados en el área de la salud son las redes neuronales

artificiales, tales como: perceptrón multicapa, redes de función de base radial, y mapas autoorganizados, además de las redes bayesianas y los árboles de decisión.

No existe una definición consensuada de redes neuronales artificiales (RNA); existen diferentes conceptos dependiendo del campo de aplicación. El concepto RNA fue asignado por científicos de la computación, matemáticos y expertos en IA. De esta forma se encuentran algunas definiciones:

- Modelo computacional, en paralelo, compuesto de unidades procesadoras adaptativas con altos niveles de conexión entre ellas [58].
- Circuito compuesto de un número elevado de elementos simples de proceso con una base neurológica. Cada elemento funciona sólo con información local. Además, cada elemento opera asíncronamente por lo que no hay un reloj total del sistema [59].
- Una red neuronal es un procesador distribuido y con una estructura paralela que tiene la tendencia natural a almacenar conocimiento experimental, haciéndolo apto para su uso. Se parece al cerebro en dos cosas, el conocimiento es conseguido por la red a través de un proceso de aprendizaje y este conocimiento se almacena en los pesos sinápticos o conexiones entre neuronas [60].
- En general, las RNA están inspiradas en las redes neuronales biológicas. Consisten en un conjunto de elementos simples de procesamiento llamados nodos o neuronas conectadas entre sí. Las conexiones tienen un valor numérico modificable llamado peso. El uso de las RNA ofrece características como: capacidad de adaptación, tolerancia a fallos, robustez y procesado en paralelo, entre otros [61].

En todo modelo de red neuronal artificial interviene los siguientes elementos que conforman su estructura:

- Señales de entradas (x_1, x_2, \dots, x_n) pueden ser datos originales o provenientes de otras neuronas.

- Señales de salida (y_1, y_2, \dots, y_m) que corresponde con los resultados de la red (clasificación o predicción en esta investigación).
- Los pesos sinápticos $(w_{1k}, w_{2k}, \dots, w_{nk})$. A las entradas que provienen de otras neuronas se les asigna un peso, es decir una determinada importancia. El peso es el número que se modifica durante el entrenamiento de la red neuronal, y que permite a la red adaptarse para conseguir la respuesta correcta.
- Regla de propagación. Se realizan diferentes operaciones con las entradas y los pesos. Con esto se busca obtener el valor del potencial que posteriormente se utilizará para realizar el procesamiento. Una de las operaciones más comunes es la suma ponderada. Otra regla puede ser la distancia euclídea, que es utilizada por los mapas autoorganizados y las redes de función de base radial.
- La función de activación. El valor obtenido en la regla de propagación se envía a través de la función de activación y esta define el nuevo estado o la salida de la capa. Según el entrenamiento de la red neuronal se elige la función de activación que pueden ser: lineal, escalon, sigmoidea o Gaussiana.

Las RNA se pueden clasificar en diferentes formas dependiendo de los elementos básicos: según el número de capas pueden ser monocapa y multicapa. Según el tipo de conexión son recurrentes y no recurrentes. Según el mecanismo de aprendizaje supervisadas, no supervisadas e híbridas. Según el tipo de conexión, totalmente conectadas y parcialmente conectadas. Describiremos a continuación con más detalle los tres tipos básicos de aprendizaje de las RNA, esto es, supervisado, no supervisado y el aprendizaje híbrido.

Redes Neuronales Supervisadas - Perceptrón Multicapa

Se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento guiado por un agente externo que determina la respuesta que

debería generar la red a partir de las entradas determinadas. El guía/supervisor controla la salida de la red y en caso de que ésta no corresponda con la deseada, se procederá a modificar los parámetros, con el fin de conseguir que la salida obtenida se aproxime a la deseada. El perceptrón multicapa (MLP) es la arquitectura más usada dentro de este tipo de aprendizaje [39].

El MLP fue descrito por Rosenblatt (1958) [62] y consiste en una RNA formada por múltiples capas. La arquitectura del MLP tiene una capa de entrada, una capa de salida y una o más capas ocultas, (figura 3.1).

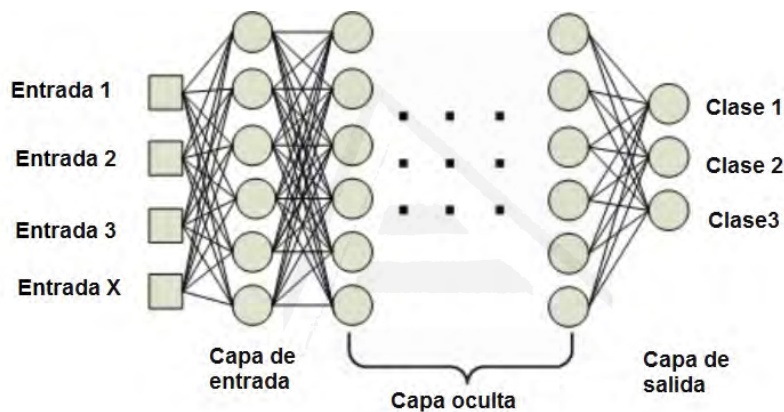


Figura 3.1: Ejemplo de un perceptrón multicapa.

Las neuronas de la capa de entrada se encargan de recibir las señales del exterior y de propagar dichas señales a la capa oculta. Las neuronas de la capa oculta realizan un procesamiento no lineal de las señales recibidas. El número de capas y neuronas en esta capa es la principal preocupación de los investigadores. El estudio de Kavasoğlu (2004) ha demostrado que el número de neuronas en la capa oculta influirá significativamente en la capacidad de la red para generalizar a partir de los datos de entrenamiento [63]. Un número pequeño de neuronas podría dar lugar a un escaso entrenamiento (*under – fitting*) mientras que, demasiadas neuronas pueden dar lugar un sobreentrenamiento (*over – fitting*) [64]. No existe ninguna regla respecto al número óptimo de capas o neuronas ocultas, este debe estar enfocado al tipo de problema que se quiere solucionar. Sin embargo,

Liu y He (2004) han utilizado la fórmula (3.1) para calcular el número de capas ocultas [65]. Para el caso de ésta investigación, el número de capas y neuronas ocultas fueron determinados de forma empírica (ensayo/error) hasta obtener el mejor resultado.

$$N_h = \lfloor J * (K - 1) / (I + J + I) \rfloor \quad (3.1)$$

Donde N_h representa el número de nodos ocultos, J es el número de nodos de salida, I es el número de nodos de entrada y K es el número de ejemplos de entrenamiento.

La última capa actúa como salida de la red, generando la respuesta para los parámetros de entrada. En este trabajo la clasificación del riesgo de cirugía cardíaca pediátrica y la predicción de la mortalidad hospitalaria de los pacientes operados de aneurisma de aorta abdominal.

MLP utiliza el algoritmo de aprendizaje supervisado denominado retropropagación (*Backpropagation*) para entrenarse. Este algoritmo fue introducido por Rumelhart (1986) [66], el cual es un algoritmo de descenso por gradiente que retropropaga las señales desde la capa de salida hasta la capa de oculta optimizando los valores de los pesos sinápticos mediante un proceso que se basa en la minimización de la función de error, y por ello puede dividirse en dos fases:

- Propagación hacia adelante: se propagan las señales desde la capa de entrada hasta la salida, especificándose la salida de la red.
- Propagación hacia atrás: la salida de la red es comparada con la deseada y en función de los errores cometidos en la capa de salida, el algoritmo ajusta los pesos sinápticos que determinan las conexiones entre las neuronas mediante la retropropagación del error desde la capa de salida a la entrada por medio de la capa o capas ocultas. El proceso se repite hasta lograr la salida correcta de la red.

Dado, que el objetivo es que la salida de la red sea lo más próxima a la salida deseada, el aprendizaje de la red se plantea como un problema de minimización de errores de la siguiente forma:

$$\text{Min}(wE) \quad (3.2)$$

siendo w el grupo de parámetros de la red (pesos y umbrales) y E una función error que determina la diferencia entre la salida de la red y la salida deseada. En la mayoría de los casos, la función error se define como:

$$E = \frac{1}{N} \sum_{n=1}^N e(n) \quad (3.3)$$

donde N es el número de patrones y $e(n)$ es el error cometido por la red para el patrón n .

Las redes *Backpropagation* han demostrado su capacidad de trabajar con acierto en una gran cantidad de aplicaciones incluyendo clasificación de imágenes, enfermedades, en el diagnóstico médico y en la predicción de eventos (complicaciones, mortalidad).

Redes Neuronales No Supervisadas - Mapas autoorganizados

En este tipo de aprendizaje, el conjunto de datos de entrenamiento consiste sólo en los patrones de entrada. Por lo tanto, la red es entrenada sin el beneficio de un guía (maestro). La red aprende a adaptarse en función a las experiencias recogidas de los patrones de entrada. Los mapas autoorganizados o (SOM - *SelfOrganizingMap*), son el modelo que mejor representa éste aprendizaje.

Las redes SOM [67] están constituidas por dos capas. La primera capa corresponde a la capa de entrada con n neuronas, y es la encargada de recibir y transmitir los datos a la capa de salida. La segunda capa o capa competitiva, con m neuronas es responsable de llevar a cabo los procesos. A cada neurona de esta capa se le asigna un vector de peso w_j , llamado vector de referencia (o *codebook*) que tiene la misma dimensión que los vectores de entrada. Las dos capas están totalmente interconectadas hacia adelante, como se muestra figura 3.2, además las neuronas se organizan para formar una malla de dos dimensiones.

En general, el entrenamiento de las redes SOM [67] se pueden separar

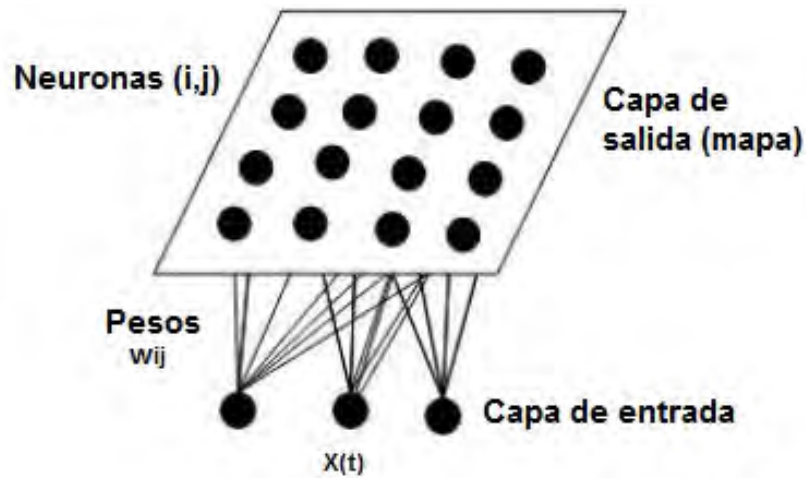


Figura 3.2: Estructura de un mapa autoorganizado.

en diferentes etapas:

1. Inicialización de parámetros y pesos, éste presenta diferentes maneras, asignando como valores iniciales los pesos de unas entradas específicas o determinando éstas de forma aleatoria, normalmente vectores en $[0,1]$.
2. Presentación de un vector de entrada $X_p = x_{p1}, \dots, x_{p2}, \dots, x_{pN}$ y el vector de pesos $W_j : w_{j1}, \dots, w_{ji}, \dots, w_{jN}$
3. Buscar la neurona ganadora: el algoritmo presenta un aprendizaje competitivo, es decir, que las neuronas compiten por ser la más cercana al valor de entrada. La distancia euclidiana (3.4) se utiliza generalmente para medir la similitud entre el valor de entrada y los pesos de las neuronas con el fin de seleccionar la neurona ganadora o (*best matching unit*) [67]. La neurona ganadora será aquella en la que se obtenga la menor medida de comparación.

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (3.4)$$

donde x_{ij} es el valor i -ésimo de la muestra y n es el número de dimensiones el vector.

4. Actualización de pesos: los pesos de la neurona ganadora y sus vecinos son actualizados, según la regla de aprendizaje:

$$\Delta_{ij}(n+1) = \alpha(n)(x_{pi} - w_{j*i}(n)) \quad \text{para } j \in \text{vecin}_j * (n) \quad (3.5)$$

donde n es el número de ciclos, esto es, el número de veces que ha sido presentado y procesado todos los patrones, $\alpha(n)$ es la tasa de aprendizaje, su valor inicial oscila entre 0 y 1. Este valor decrece con el número de iteraciones (n) del proceso de aprendizaje. $\text{vecin}_j * (n)$ es el área de vecindad alrededor de la neurona vencedora j .

5. Volver al paso 2 si no existen más datos o se llegó al número de iteraciones deseadas.

Redes Neuronales Híbridas - Redes de Función de Base Radial

Son redes con conexión hacia adelante (figura 3.3) que tienen la particularidad de usar los dos métodos antes mencionados. Es decir, una de las capas de la red tiene un aprendizaje supervisado y otra tiene un aprendizaje no supervisado. Lo que se busca con esta división es acelerar notablemente el proceso de aprendizaje respecto a otros algoritmos [68]. Estas fueron introducidas en la literatura por Broomhead y Lowe (1998) [69] y presentan una estructura muy similar al MLP. Están compuestas por tres capas: una capa de entrada, una capa oculta y una capa de salida.

- La capa de entrada, es la encargada de transmitir la información a la capa oculta. No lleva pesos asociados y no procesa la información.
- La capa oculta, se compone de un número de unidades de función de base radial. Cada unidad de la capa oculta representa una única función de base radial, la cual determina si se activa para el patrón de entrada o si continua desactivada. La activación esta representada por la siguiente expresión:

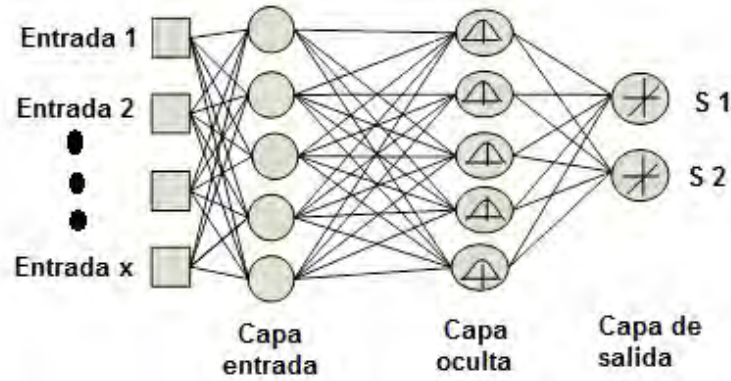


Figura 3.3: Estructura de una red de función de base radial.

$$\phi(n) = \phi(n) \left(\frac{\| X(n) - C_i \|}{d_i} \right) \quad \text{para } k=1.2.3\dots m \quad (3.6)$$

siendo $X_n = x_1(n), x_2(n), \dots, x_p(n)$ el patrón de entrada n , ϕ hace referencia a una función de base radial, C_i equivale a los vectores que forman las coordenadas de los centros i , d_i representa las desviaciones de las funciones y finalmente $\| \cdot \|$ hace referencia a la distancia euclidiana desde el vector de entrada.

Respecto al número de neuronas en la capa oculta, la típica solución consiste en comparar varios modelos con diferentes números de neuronas y escoger aquel que dé los mejores resultados, es decir prueba error. En esta investigación tres neuronas en esta capa fueron utilizadas.

- En la capa de salida, cada elemento de procesamiento calcula su red como una combinación lineal de las salidas de la capa oculta y actúa como salida de red [70]. La función de activación y transferencia es lineal y esta dada por:

$$y_k(n) = \sum_{i=1}^m w_{ij} \phi(n) + u_j \quad \text{para } k=1.2.3\dots r \quad (3.7)$$

siendo w_{ij} el peso de conexión de la neurona oculta i , la neurona de salida j ; $\phi(n)$ es la activación de la neurona oculta i , u_j es el umbral, que está asociado a cada uno de los elementos de procesado de la capa de salida i .

Las RBF cuentan con un proceso de aprendizaje, como todas las redes neuronales. El proceso de aprendizaje se encarga especialmente de determinar los centros, las desviaciones, los pesos de la capa oculta a la capa de salida y los umbrales. Por ello, se pueden determinar dos métodos de aprendizaje: el método híbrido y el totalmente supervisado que se describen a continuación.

1. El método híbrido esta formado por dos fases. La fase no supervisada en la cual los centros y las amplitudes de las neuronas en la capa oculta se determinan, para obtener las salidas más próximas a las salidas deseadas. En esta fase la idea es minimizar la función de error que se mide en la salida de la red y está dada por la función:

$$E = \frac{1}{N} \sum_{i=1}^K e(n) \quad (3.8)$$

Fase supervisada: en la cual se determinan los pesos y los umbrales de la capa de salida, con el propósito de juntar el espacio de entradas en diferentes clases. Para la fijación de los centros de esta fase, el método *K-medias*, es el más utilizado. Este algoritmo fue diseñado por MacQueen (1967) [71] y determina una división del espacio de entrada en k clases. Con éste método se intenta disminuir la distancia euclídea entre los patrones de entrada y su centro más cercano correspondiente, por lo tanto se minimiza la siguiente función:

$$J = \sum_{i=1}^K \sum_{n=1}^N M_{in} \| X(n) - C_i \| \quad (3.9)$$

donde N es el patrón de entrada, $\| \cdot \|$ es la distancia euclídea, $X(n)$ es el patrón de entrada, M_{in} es la función de pertenencia al centro

donde toma valores de 1 si pertenece al centro y 0 en caso contrario. Este tipo de entrenamiento conserva las características locales de la red y es el más utilizado (en esta investigación fue aplicado este método).

2. En el método totalmente supervisado todos los parámetros de la red, centros, amplitudes, pesos y umbrales se realizan completamente guiado por un agente externo. A fin de minimizar el error cuadrático medio en la salida de la red, de esta manera se pierden las características locales de la red.

3.1.2 Redes Bayesianas

Las RB tienen un algoritmo de clasificación descriptivo y predictivo basado en el teorema de Bayes [72]. Son un modelo que se basa en un gráfico acíclico (figura 3.4) en el cual cada uno de los nodos representa una variable y cada arco denota una dependencia y la probabilidad de distribución condicional [73]. Con este modelo se puede hacer la inferencia bayesiana, es decir, calcular la probabilidad posterior de las variables no conocidas, en base a las variables conocidas. Estos modelos tienen aplicaciones en la clasificación [74], predicción [41] y diagnóstico de enfermedades [75].

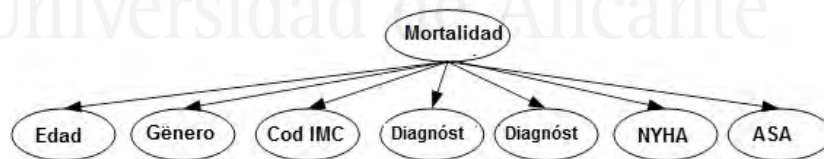


Figura 3.4: Grafo de una red bayesiana.

Una RB provee un sistema de inferencia, donde una vez encontradas nuevas evidencias sobre el estado de ciertos nodos, se modifican las tablas de probabilidades; y a su vez, las nuevas probabilidades son transmitidas al resto de los nodos. La transmisión de probabilidades se reconoce como *inferencia* probabilística, esto significa que la probabilidad de algu-

nas variables puede ser calculada por las evidencias dadas en otras variables. Antes de introducir las evidencias se conocen como probabilidades *a priori*; una vez introducidas las evidencias, se conocen como probabilidades *a posteriori* [76, 77].

Para estimar las probabilidades se han propuesto numerosos clasificadores bayesianos [78] como el árbol de red aumentada bayes (TAN), Red de bayes aumentada (BAN) y el clasificador bayesiano *naïve*. El clasificador TAN está basado en la estructura de la *naïve* Bayes pero el aprendizaje le permite la conexión entre los nodos de un árbol. TAN supone que un conjunto de atributos pueden ser casualmente dependientes [79]. El algoritmo BAN, produce la máxima probabilidad dada la restricción de que cada nodo puede tener como máximo un nodo dependiente, además del nodo raíz [74] y en [80], se plantea el paradigma de que la única limitación entre los arcos de las variables predictoras es que no se formen ciclos.

- Las *Naïve* Bayes [81] son uno de los modelos más simples y más utilizados, en estos modelos la hipótesis que se maneja es que los atributos son condicionalmente independientes de la variable a clasificar.

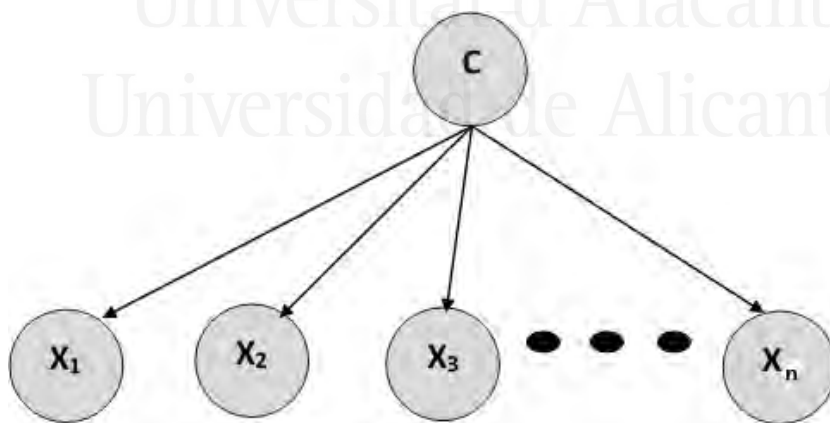


Figura 3.5: Estructura del clasificador *naïve* bayes.

La probabilidad de que j -simo ejemplo pertenezca a la clase i -sima de la variable a clasificar C , puede calcularse aplicando el teorema de

Bayes [72], de la siguiente forma:

$$P(C_i|E) = P(C_i)P(E_1|C_i)P(E_2|C_i)...P(E_n|C_i)/P(E) \quad (3.10)$$

donde n es el número de atributos, C corresponde a la variable de la clase, que esta conectada con los atributos $E_1, E_2...E_n$, $P(C)$ es el vector de probabilidades *a priori* para cada clase, $P(E_i|C)$ es la matriz de probabilidad condicional para cada atributo. Dada la clase los atributos son condicionalmente independientes, de tal manera que no existen arcos entre ellos, como se muestra en la figura 3.5.

Aunque la simplicidad de las *naïve* Bayes y la restricción de que los atributos son independientes entre si, se consideran una debilidad del modelo [82], no obstante esta red consigue un alto grado de aciertos. Muchos investigadores piensan que las *naïve* Bayes están basadas en la misma idea que usan los médicos al realizar un diagnóstico, es decir, independientes dada su clase [83]. En este trabajo de investigación los mejores resultados obtenidos fueron logrados por éste método.

- Árbol de red aumentada Bayes. Para obtener este tipo de configuración se empieza por una estructura de árbol con las variables predictoras, y seguidamente se conecta la variable objetivo con cada una de las variables predictoras.

Este modelo es una extensión de la red *naïve* Bayes en el que los nodos atributo forman un árbol. La propiedad más destacada que debe cumplir una TAN es que cada nodo atributo debe tener a lo sumo dos padres: el nodo clase y algún otro nodo atributo (ver figura 3.6). Uno de los métodos más utilizados para aprender la estructura TAN es el propuesto por Friedman (1997) [74]. Este algoritmo es una variante del método de Chow-Liu (1968) [84], que utiliza la información mutua condicional entre dos atributos según la clase a clasificar. La dependencia se puede medir mediante el cálculo de la información mutua entre pares de atributos $X; Y$ dada la clase C [85]:

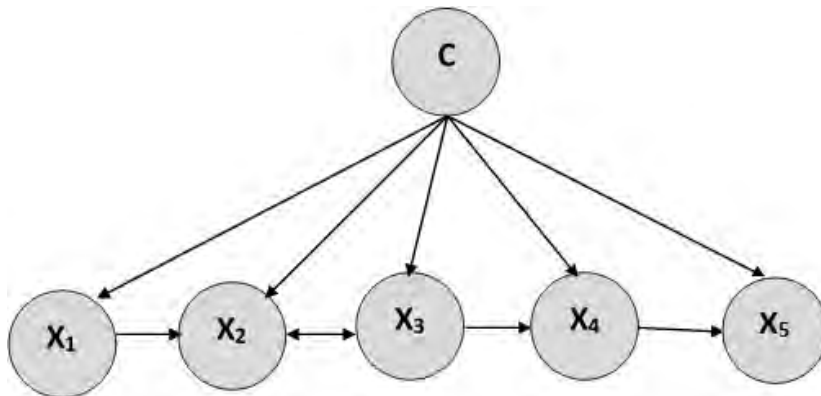


Figura 3.6: Estructura árbol de red aumentada bayes.

$$I_{pr}(X_i, X_j|C) = \sum_{X_i, X_j} Pr(X_i, X_j|C) \text{Log}(Pr(X_i, X_j|C)) / Pr(X_i|C) * (Pr(X_j|C)) \quad (3.11)$$

Se puede decir que la función anterior mide la información que la variable Y asigna sobre la variable X cuando el valor de C es conocido.

- Red de Bayes Aumentada, son una extensión del TAN que permite la creación de un grafo arbitrario [74] en lugar de un árbol como se mencionó anteriormente. A diferencia de la red TAN o *naïve* Bayes, el aprendizaje de éste algoritmo es sin restricciones de padres, un ejemplo de estructura BAN se puede ver en la figura 3.7.

Uno de los problemas a resolver cuando se utiliza la RB es buscar la estructura que mejor se ajusta a los datos. Esta tarea está dada por el aprendizaje automático que presenta dos tipo: el aprendizaje estructural que obtiene la estructura de la red (árboles, poliárboles o redes multiconectadas) a partir de los datos. El aprendizaje paramétrico, el cual obtiene las probabilidades asociadas (tabla de probabilidad de cada nodo raíz y

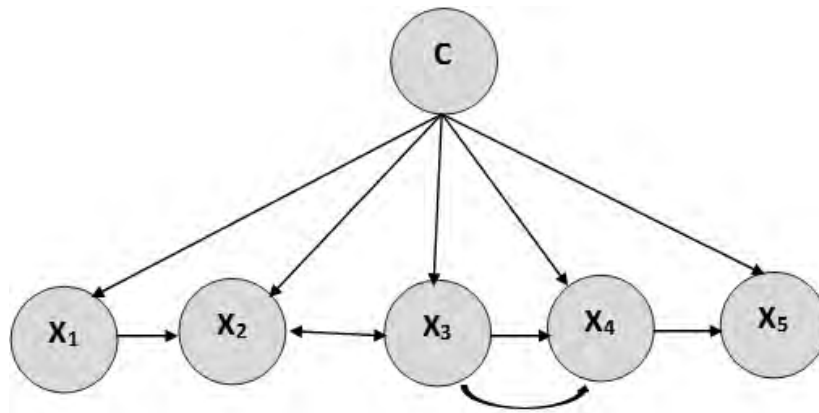


Figura 3.7: Estructura de red bayes aumentada.

las demás variables). Estos dos aprendizajes no se pueden desarrollar de forma completamente independientes, es decir, que para calcular las distribuciones de probabilidad asociadas a cada nodo de la red, es necesario conocer el tipo de red.

3.1.3 Árboles de Decisión

Los árboles de decisión son una de las estructuras más sencillas para la representación del conocimiento adquirido. En los sistemas basados en árboles de decisión, también llamados TDIDT (*Top Down Induction of Decision Trees*), se pueden destacar dos grupos: la familia ID3, los más representativos son el propio algoritmo ID3 propuesto por Quinlan (1986) [86] y el sistema CLS (*Concept Learning Systems*) [87]; y la familia de árboles de regresión, cuyo modelo más importante es *Cart*, desarrollado por Breiman (1984) [88].

Un árbol de decisión (AD) es un gráfico que representa las diferentes alternativas que pueden ocurrir cuando se analiza un problema con el objetivo de determinar la secuencia óptima de las decisiones que se deben realizar. A medida que el algoritmo avanza el árbol va creciendo y los subconjuntos son menos numerosos. El algoritmo ID3 fue mejorado por Quinlan (1993) y pasó a ser el algoritmo C4.5 [89] que resuelve algunas

debilidades del ID3 sobre el uso de atributos continuos. Este modelo se considera un modelo estándar en la clasificación supervisada. En esta investigación el algoritmo C4.5 fue probado.

El AD tiene un conjunto de nodos, hojas y ramas. El nodo principal o raíz es el atributo a partir del cual se inicia el proceso; los nodos internos que corresponden a cada una de las preguntas acerca de un atributo en particular. Cada posible respuesta se representa mediante un nodo hijo. Las ramas que salen de cada uno de estos nodos se encuentran marcados con los posibles valores del atributo [76]. Los nodos finales o nodos hoja corresponden a una decisión (figura 3.8).

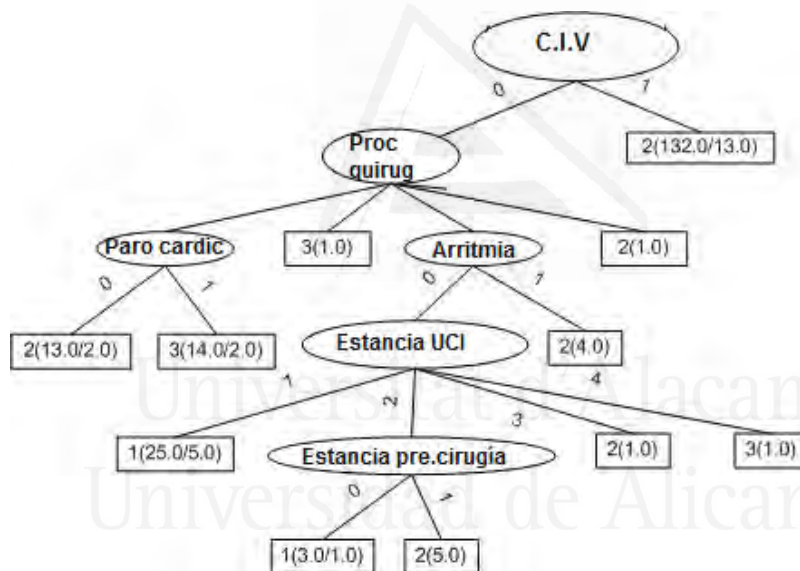


Figura 3.8: Segmento de un árbol de decisión.

C4.5 se basa en el uso de criterios de relación de ganancia (*gain ratio*, que está definido como 3.12), de esta manera se evita que la variable con el mayor número de valores posibles se beneficie en la selección, porque escoge en cada nodo el atributo de mayor ratio de ganancia. Además incorpora un procedimiento de poda para disminuir el tamaño total del árbol y la disminución de la tasa de error estimado [89]. Para la poda utiliza un enfoque heurístico basado en la significación estadística de las divisiones.

$$ratio(X, E) = \frac{ganancia(X, E)}{info_p(X, E)} \quad (3.12)$$

donde

$$info_p(X, E) = - \sum_{i=1}^n \frac{|E_i|}{E} * \log_2 \frac{|E_i|}{E} \quad (3.13)$$

La construcción básica de C4.5 esta dada por:

1. Los nodos de origen, son el nodo superior del árbol. Considera todas las muestras y selecciona los atributos que son más importantes.
2. La información pasa a la sección de nodos siguientes, llamados nodos rama que finalmente terminan en nodos hoja para generar las decisiones.
3. Las reglas surgen para mostrar la trayectoria desde el nodo origen hasta el nodo hoja [90]. En este trabajo, tenemos una alta dimensionalidad y esto hace que sea muy difícil de manejar.

El algoritmo considera todas las posibles pruebas en que se puede dividir el conjunto de datos y selecciona la prueba que genere la mayor ganancia de información. Para cada atributo discreto, se estima una prueba con n resultados, donde n es el número de posibles valores que puede tomar un atributo. Una prueba binaria (1,0) se realiza para los atributos continuos. En cada nodo, el sistema debe decir que prueba se escoge para dividir los datos. Según [91] existen tres tipos de pruebas posibles:

- La prueba *standard* para cada variable discreta, con un resultado y una rama para cada posible valor de la variable.
- Una prueba mas compleja se realiza en una variable discreta, en donde los posibles valores se asignan a un posible grupo, en lugar de a una variable.
- Si una variable X tiene valores numéricos continuos, se realiza la prueba binaria con resultados $X \leq Z$ y $X > Z$, para lo cual debe determinarse el valor del límite Z .

Todas estas pruebas se evalúan mirando el resultado de la proporción de ganancia que resulta de la división de los datos. Es posible adicionar una restricción para evitar las subdivisiones triviales que se tienen en cuenta cuando hay un conjunto muy pequeño.

El algoritmo C4.5, tiene entre otras las siguientes características:

- Permite trabajar con datos de valores continuos para los atributos, separando los probables resultados en dos ramas: $A_i \leq N$ y $A_i > N$.
- Los árboles son menos frondosos, debido a que cada hoja cubre una distribución de clases y no una clase en particular.
- Utiliza el método "divide y vencerás" para generar el árbol inicial, a partir del conjunto de datos de entrenamiento.
- Se basa en la utilización del criterio de proporción de ganancia (*gain ratio*).

3.2 Evaluación de los clasificadores

Un aspecto importante cuando utilizamos clasificadores es cuantificar de alguna forma si su desempeño es bueno o malo de acuerdo al problema estudiado [92]. A la hora de evaluar un clasificador, se deben tener presente diferentes criterios de acuerdo a la interoperatividad con otras herramientas, la interpretación de los resultados, la sencillez y rapidez del modelo; sin embargo, al que más se le presta atención es la precisión (α) o la tasa de error del clasificador. La precisión se refiere al porcentaje de clases (c_i) correctamente clasificadas respecto al conjunto, y esta definida como:

$$\alpha = \frac{TP_i + TN_i}{TP_i + FP_i + TN_i + FN_i} \quad (3.14)$$

Donde FP (falsos positivos), TN (verdaderos negativos), TP (verdaderos positivos), FN (falsos negativos). Los cuatro posibles resultados se

pueden expresar en una tabla de contingencia (Tabla 3.1):

	Resultado Clasificador		
	Si	No	
Estado verdadero	Si	TP_i	FP_i
	No	FN_i	TN_i

Tabla 3.1: Tabla de contingencia

Mediante el uso de los siguientes métodos es posible medir los resultados de los clasificadores:

- Validación cruzada (*Cross validation*): es una técnica utilizada para evaluar los resultados del clasificador y garantizar que son independientes de la partición entre datos de entrenamiento y prueba. La validación cruzada proporciona una estimación casi imparcial, utilizando sólo los datos originales. La validación cruzada resulta estar estrechamente relacionada con la estimación de la tasa de error esta se calcula de la media aritmética de los errores de cada iteración con el cual se alcanza un solo resultado. Si MSE_i indica el error en la iteración i -ésima, entonces el error de la validación cruzada esta dado por [93]:

$$V_{cruzada_n} = \frac{1}{n} \sum_{i=1}^n MSE_i \quad (3.15)$$

En la validación cruzada de N iteraciones o n -fold cross-validation los datos se dividen en N subconjuntos (*folds*). Uno de los subconjuntos se utiliza como datos de prueba y el resto ($n - 1$) como datos de entrenamiento. El proceso es repetido durante n iteraciones con cada uno de los posibles subconjuntos de datos de prueba. Esta técnica ha sido usada en esta investigación con $n = 10$, en todas las pruebas realizadas.

- Dejar-uno-afuera (*Leave-one-out*): Es un tipo de validación cruzada de n folds, donde n es el número de clases del conjunto de datos. En

cada iteración, una de las clases se deja fuera y se entrena el clasificador con el resto de las clases. Los resultados de las n evaluaciones se promedian para determinar la proporción de error. Error que en esta clase de validación es muy bajo, por lo tanto el coste computacional es elevado, lo cual la convierte en una técnica ideal cuando se tiene un conjunto de datos pequeño, porque evalúa todas las posibilidades. Según Witten y Frank, (2011), este procedimiento es interesante por dos razones. Primero, utiliza la mayor cantidad de ejemplos posibles para el entrenamiento y segundo, el procedimiento es determinístico, no se dividen los datos al azar [94].

- Remuestreo (*random subsampling*): En este método se realizan diferentes particiones del conjunto de entrenamiento y prueba, donde el tamaño es definido por el usuario [95]. La precisión del clasificador se realiza a partir de la media obtenida en los distintos conjuntos de prueba.

Dado el problema que se está investigando, algunas veces es necesario conocer otras medidas de evaluación del clasificador. Por ejemplo, se debe predecir mediante un clasificador, si un paciente tiene una enfermedad cardíaca. Si el clasificador comete un error y toma una persona sana como enferma a esto se le llama un *falso positivo*; por el contrario, si el clasificador toma una persona enferma como sana a esto se le conoce como *falso negativo*.

Es importante distinguir entre los diferentes errores, para ello se puede usar la matriz de confusión o tabla de contingencia (tabla 3.1), que muestra los diferentes errores del clasificador. Si se tiene un problema de dos clases (enfermo o sano), se pueden observar los resultados en la tabla 3.2.

	Enfermo	Sano
Enfermo	Verdadero positivo	Falso positivo
Sano	Falso negativo	Verdadero positivo

Tabla 3.2: Tabla de contingencia para problemas de dos clases

A partir de la matriz de confusión (tabla de contingencia), se pueden

construir otras medidas para la evaluación del clasificador. Las medidas más comunes son: sensibilidad, especificidad (introducidos por *Yerushalmy* (1947)) [96].

- Sensibilidad: es la probabilidad de que una prueba sea positiva sí la enfermedad/condición está presente. Se denota entonces la sensibilidad como TPR (*True Positive Rate*) y esta definida como:

$$TPR = \frac{TP}{(TP + FN)} \quad (3.16)$$

- Especificidad: es la probabilidad de que la prueba sea negativa sí la enfermedad/condición no está presente. Entonces la especificidad se denota como TNR (*True Negative Rate*) y esta de la siguiente forma:

$$TNR = \frac{TN}{(TN + FP)} \quad (3.17)$$

Adicionalmente, se puede calcular el valor predictivo positivo y negativo entre los resultados obtenidos de las pruebas.

El valor predictivo del resultado positivo, esta definido como la proporción de los correctos resultados entre los resultados positivos de la prueba:

$$PPV = \frac{TP}{(TP + FP)} \quad (3.18)$$

Mientras que el valor predictivo del resultado negativo será la proporción de los correctos resultados entre los resultados negativos:

$$NPV = \frac{TN}{(TN + FN)} \quad (3.19)$$

En algunas ocasiones (depende del problema) es necesario que las pruebas cumplan las exigencias de tener una sensibilidad y especificidad altas, además de identificar los tipos de errores que comete el clasificador. Para ello, es muy útil tomar las medidas de la curva ROC (*Receiver Operating*

Characteristic), donde se puede ver de manera gráfica la relación entre estos dos conceptos.

La curva ROC es una técnica para la visualización, organización y selección de los clasificadores teniendo en cuenta su rendimiento [97]. Estas han sido ampliamente utilizadas en medicina para el análisis discriminatorio pruebas de diagnóstico [98]. Las curvas ROC son gráficos bidimensionales en donde la tasa *TP* se representa en el eje Y y la tasa de *FP* se representa por el contrario, en el eje X. En la figura 3.9 se muestra un gráfico de la curva ROC.

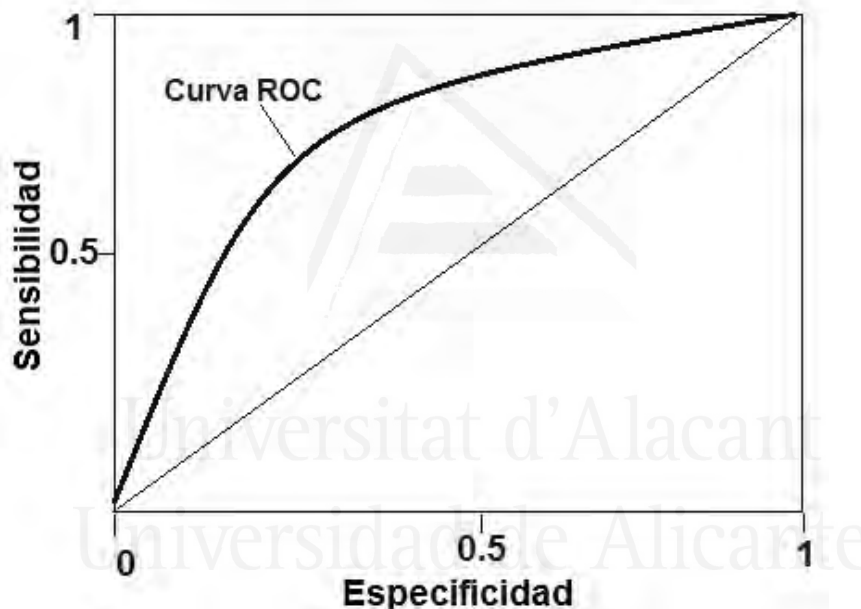


Figura 3.9: Ejemplo de una curva ROC.

La curva ROC permite, evaluar que tan separadas están las distribuciones de la sensibilidad y la especificidad de una prueba. Cuanto mayor sea el área bajo la curva ROC (más alejado de la diagonal principal), mejor será el desempeño del clasificador. La mejor predicción será aquella cuya curva pase por la esquina superior izquierda hasta el punto (0,1), ya que representa un 100% de sensibilidad y un 100% de especificidad (indicando una prueba perfecta).

En síntesis, de acuerdo con Zweig y Campbell (1993) [99], usar las curvas ROC en la evaluación de las pruebas diagnósticas tiene los siguientes beneficios:

- Son curvas fáciles de interpretar.
- No requieren un nivel de decisión particular.
- Son independientes de la prevalencia.

3.3 Recopilación de datos

La recopilación de datos con fines de estudio debe realizarse con suficiente frecuencia que evite la pérdida de los mismos. Por ejemplo, los datos para el seguimiento de la mortalidad de la población deben recopilarse constantemente, mientras que los datos sobre la estancia hospitalaria pueden recopilarse con intervalos de tiempo más largos. La información proveniente de la atención médica de una persona que frecuenta un hospital, centro de salud o una consulta se registra en varios documentos (evolución médica, registro de medicamentos, etc.), la integración de estos documentos se conoce como la historia clínica (HC) [100]. El propósito de la HC es el de garantizar que toda la información de los pacientes este ordenada, accesible, comprensible, disponible y bien conservada. Además, tiene entre otras las siguientes bondades:

- Almacenar los datos de forma segura.
- Evitar errores médicos.
- Disminuir el tiempo en los diferentes procesos (autorizaciones médicas, informe de analíticas, etc.).
- Mejorar la gestión del espacio (archivos físicos).
- Capacitar a los pacientes.

Los datos contenidos en la HC están representados por diferentes variables que formarán parte de las entradas de los algoritmos probados en esta investigación. Dichas variables presentan diferentes valores, magnitudes, rangos y texto. El texto libre ofrece una máxima expresividad pero pocas posibilidades de recuperación de la información. Las variables estructuradas funcionan bien y es preferible que haya cierto grado de estructuración.

Los datos analizados en este proyecto fueron recopilados de la historia clínica de los pacientes llevados a cirugía para la corrección de una cardiopatía congénita (comunicación interauricular, comunicación interventricular, ductus arterioso persistente, estenosis pulmonar, estenosis valvular aórtica, coartación aórtica, tetralogía de Fallot, atresia tricúspide y la transposición de grandes arterias), y los pacientes operados de reparación abierta de aneurisma de aorta abdominal. En los anexos I y II se describen las variables usadas en esta investigación.

3.4 Procesado de datos

El procesado de datos es un paso muy importante para facilitar el análisis de los mismos y su posterior uso en la toma de decisiones. Para ello, es necesario realizar un adecuado procesamiento de datos, en algunos casos este procedimiento no se realiza buscando hacer una comparativa de los resultados antes y después de ser procesados.

Las actuaciones a realizar en el procesado de los datos pueden agruparse en varias técnicas:

3.4.1 Agrupamiento de datos

El agrupamiento de datos permite reunir todos los datos disponibles para la resolución del problema. De otra parte, la distribución en intervalos permite establecer diferentes rangos de valores de una misma variable cuando está presente datos muy dispersos. Un ejemplo de ello es el mostrado en la tabla 3.3 donde se asocia el peso y la altura de un individuo

para obtener el índice de masa corporal.

Intervalo	Descripción	Valor
< de 18	Bajo peso	1
18.1 a 24.9	Normal	2
25 a 26.9	Sobrepeso	3
27 a 29.9	Obesidad I	4
30 a 39.9	Obesidad II	5
< 40	Obesidad mórbida	6

Tabla 3.3: Ejemplo de datos agrupados.

3.4.2 Transformación de Datos

Normalización de datos

Una de las herramientas comúnmente utilizadas en el análisis de datos y con el cual se mejoran los resultados es la normalización. Existen varios tipos de normalización de datos y éste dependerá de los datos y del problema a solucionar, para que el tratamiento de los mismo resulte lo más sencillo posible. En general, los datos no se utilizan en la forma en que se obtienen o los facilita el problema. La mayoría de las veces es necesario hacer un preprocesamiento, para adaptar los datos al modelo usado; por ejemplo, normalizar las salidas [0,1]. También es habitual normalizar las entradas cuando sea oportuno.

A continuación se mostraran brevemente los tipos de normalización más comunes:

- *Normalización estadística (Z-score Normalization):*

$$X_i^{\lambda} = \left[\frac{x_i - \mu}{\sigma} \right] \quad (3.20)$$

Donde: X_i^{λ} es la entrada normalizada de índice i , x_i es la entrada del índice i ; μ es la media de las entradas y finalmente σ es la desviación típica de las entradas.

Este método es más adecuado en los casos donde no se conoce el máximo y mínimo de los datos de entrada o cuando existen valores atípicos que tienen un efecto en el rango de los datos.

- *Normalización Min-Max:*

$$X_i^{\lambda} = (max_{obj} - min_{obj}) * \left[\frac{(x_i - min_{valor})}{(max_{valor} - min_{valor})} \right] + min_{valor} \quad (3.21)$$

Donde: X_i^{λ} es la entrada normalizada del índice i , x_i es la del entrada de índice; max_{obj} es el valor máximo del rango que se aplica a las entradas; min_{obj} es el valor mínimo del rango que se aplica a las entradas; max_{valor} valor máximo de las entradas y min_{valor} valor mínimo de las entradas.

La normalización evita que los atributos con valores muy altos sean beneficiados en el modelo final, que aquellos con valores más bajos [61].

- *Normalización Sigmoidea:*

$$X_i^{\lambda} = \frac{1}{1 + e^{-\frac{(x_i - \mu)}{\sigma}}} \quad (3.22)$$

Donde: X_i^{λ} es la entrada normalizada de índice i , x_i es la entrada de índice i ; μ es la media de las entradas y finalmente σ es la desviación típica de las entradas.

- *Normalización decimal:*

Esta normalización consiste en mover los puntos decimales de los valores del atributo X . El número de puntos decimales movidos depende del máximo valor absoluto de X . Un valor v de X es normalizado a v^{λ} con el cálculo de la siguiente expresión:

$$v^{\lambda} = \frac{v}{10^i} \quad (3.23)$$

Donde i es el centro más pequeño de $Max(|v^{\lambda}|) < 1$

Discretización de los datos

Este proceso transforma datos cuantitativos (categóricos) en datos cualitativos (numéricos) que favorecen el proceso de aprendizaje y mejoran la precisión de la clasificación. Además la discretización suaviza el efecto del ruido y permite modelos más simples; evitando el sobreajuste.

Existen dos tipos de discretización, la dinámica, donde los atributos continuos son discretizados durante el proceso de clasificación. Mientras que en la estática los atributos continuos son previamente discretizados antes de la tarea de clasificación. En las tablas 3.4 y 3.5, se muestran ejemplos de cómo fueron discretizadas las variables abordaje y peso en niños.

Abordaje	Valores
Extraperitoneal	1
Transperitoneal	2
Valores nulos	0

Tabla 3.4: Tipo de abordaje en cirugía de reparación abierta de AAA.

Rango de pesos	Valores
0.6kg - 3kg	1
3.1kg - 6kg	2
6.1kg - 9kg	3
9.1kg - 12kg	4
12.1 kg - 15kg	5
> 15.1 kg	6

Tabla 3.5: Rango de pesos en niños.

Imputación de Datos

La ausencia de datos en una base de datos médica es un problema frecuente y la posible solución radica en utilizar los denominados métodos de imputación. Consisten en reemplazar los valores incompletos por estimaciones basadas en la información disponible en el conjunto de datos. Los métodos de imputación se clasifican en dos tipos:

- **Métodos simples:** estos han sido una de las herramientas más conocidas y aceptadas para el tratamiento de la falta de datos [101]. Consiste en asignar un valor por cada valor faltante basándose en el valor de la propia variable o de otras variables, generando una base de datos completa. Dentro de la imputación simple se encuentran las siguientes:
 - Imputación por media
 - Imputación deductiva
 - Imputación *Cold Deck*
 - Imputación *Hot-Deck*
 - Imputación por regresión
 - Imputación por máxima verosimilitud
- **Métodos múltiples:** esta técnica fue propuesta por Rubin (1978) [102], surge como una alternativa a las técnicas anteriormente mencionadas, con el objetivo de hacer un uso más eficiente de los datos. Este procedimiento consiste en asignar a cada valor faltante varios valores (m) imputados, generando m conjuntos de datos completos. En cada conjunto de datos completo se estiman los parámetros más relevantes y posteriormente se combinan los resultados obtenidos. Éste método tiene tres fases, fase de imputación, fase de análisis y fase de puesta en común. Dentro de la imputación múltiple la más destacada es la siguiente:
 - Imputación Múltiple *Markov Chain Monte Carlo* (MCMC)

3.4.3 Reducción de la dimensionalidad

Otra fase del preprocesado de datos, es descartar los atributos que no son importantes en la toma de decisiones. En la mayoría de los casos de estudio el volumen de datos original suele ser más grande de lo esperado. Además, con mucha frecuencia gran parte de la información es redundante, esto debido a que muchas variables están relacionadas entre si. Por

lo tanto, es importante reducir los atributos sin que la haya una pérdida significativa de la información y que ésta no afecte de forma negativa los resultados y el desempeño de los algoritmos. El objetivo del proceso de selección de características es encontrar el conjunto mínimo de atributos de tal forma que la distribución resultante de las clases sea tan próxima como la distribución original [103]. Esto reduce la dimensionalidad de los datos y permite que los algoritmos de aprendizaje operen de forma más rápida y eficaz.

Algunos estudios han demostrado que los algoritmos de inteligencia artificial más comunes se ven influenciados negativamente por información irrelevante y redundante. Por ejemplo, el *K-nearest neighbors* es sensible a los atributos irrelevantes y su complejidad aumenta significativamente con el número de atributos sin importancia [104, 105]. El clasificador de *naïve Bayes* puede verse afectada negativamente debido a la suposición de que los atributos son independientes dada la clase [106]. Los árboles de decisión como el C4.5 puede algunas veces sobreentrenarse obteniendo como resultado un árbol demasiado grande [107].

Para seleccionar un conjunto de características que dé más información acerca de nuestros datos, existen tres métodos generales que facilitan este proceso: filtros, *wrappers* y *embedded*.

- **Filtros:** este método emplea una función de rango para elegir las mejores características de los atributos. La función de rango da una puntuación de relevancia en base a una secuencia de ejemplos. Una vez que el rango se ha calculado, se crea un conjunto de características con las mejores N características. Como ejemplo de éstos métodos se incluyen la *Chi* cuadrado, la correlación de *Pearson's* y el *mutual information*. Este método es muy útil cuando hay una gran cantidad de datos. El tiempo no es un problema ya que se lleva a cabo un número constante de veces, lo cual disminuye el tiempo de entrenamiento.
- **Wrappers:** este método fue popularizado por Kohavi y Jhon (1997) [107], ofrece una manera simple y poderosa para abordar el pro-

blema de la selección de características, independientemente del método de aprendizaje. Para ello utilizan un algoritmo de inducción que estima el mérito de subconjuntos de características, incorporando un algoritmo eurístico para reducir el espacio de búsqueda.

Wrappers se caracterizan por lograr mejores resultados que los filtros debido a que están sincronizados en una interacción específica entre un algoritmo de inducción y sus datos de entrenamiento. Sin embargo, tienden a ser mucho más lentos que los filtros, ya que deben llamar repetidamente el algoritmo de inducción y se debe repetir cuando se utiliza un algoritmo de inducción diferente.

- *Embedded*: algunos algoritmos de aprendizaje incluyen un método de selección integrado, por lo tanto la selección de características es entonces una parte implícita del proceso de aprendizaje. Por ejemplo, el árbol de decisión *CART* utiliza la función medida de información o pérdida de elegir las mejores características [88].

3.4.4 Desbalance de los datos

El desbalance de los datos es un problema al que se enfrentan los investigadores y que influye de forma adversa en el desempeño de los clasificadores, en particular con los patrones de las clases menos representadas. El desbalance en un conjunto de datos sea grande o pequeño se presenta cuando una clase tiene una gran cantidad de datos (clase mayoritaria), mientras que la otra esta representada por muy pocos datos (clase minoritaria). Con este escenario, es importante que los algoritmos puedan predecir/clasificar correctamente instancias de la clase minoritaria y, al mismo tiempo, no afecte la precisión de los resultados para la clase mayoritaria [108]. El problema de desbalance es complejo, porque no solamente depende de la proporción que existe entre el número de instancias de cada clase; a dicho problema se le conoce como *desbalance entre clases* [109], sino de la complejidad de los datos.

Para mejorar el desempeño de los clasificadores con datos desbalancea-

dos, se han propuesto soluciones que intentan remediar este problema. Las herramientas planteadas son métodos externos o internos que se aplican a los datos en la etapa de preprocesamiento. Otras propuestas, modifican internamente los algoritmos de clasificación con la finalidad de incluir un mecanismo para hacer que las instancias de la clase minoritaria sean consideradas de mayor importancia que el resto [108]. A continuación se describen los métodos más destacados:

- SMOTE (*Syntetic Minority Over-sampling Technique*): es un algoritmo de *oversampling* que produce instancias “sintéticas” para balancear el conjunto de datos basado en la regla del vecino más cercano K - NN , la generación se realiza extrapolando nuevas instancias en lugar de duplicarlas. Para cada una de las instancias minoritarias se buscan las instancias minoritarias más cercanas y se crean N instancias entre la línea que une la instancia original y cada una de las vecinas. El valor de N depende del tamaño del sobreejemplo deseado [110].
- Remuestreo (*Resampling*): es una técnica, que a partir de un modelo genera nuevas muestras simuladas de igual tamaño que la muestra original (remuestras). La técnica más destacada es el *bootstrap*.
 - *Bootstrap* [111]: es una técnica versátil, donde la selección de las muestras se hace por reemplazo. La idea principal, es tratar la muestra como un solo conjunto y a partir de ahí extraer con reposición un gran número de muestras de tamaño n . *Bootstrap* permite estimar estas propiedades aproximando empíricamente la distribución de muestreo estadístico. Además la capacidad general del enfoque ha demostrado su eficiencia en una gran variedad de situaciones, en temas como el análisis de correlación y regresión, el análisis factorial exploratorio y confirmatorio, y el análisis de propiedades de pruebas, entre otros [112]. *Bootstrap* comparte con otras técnicas de remuestreo el procedimiento esencial de usar la simulación por computo intensivo para responder a las preguntas del análisis de datos [113].

Capítulo 4

Resultados

Las máquinas de aprendizaje son útiles para la identificación de conocimiento oculto en bases de datos. Una cuidadosa selección de las variables y la preparación de los datos antes del proceso de análisis son obligatorios para garantizar el éxito de la investigación y su aplicación en la rutina clínica. En este trabajo, se presentan los modelos de ayuda a la toma de decisiones, con el propósito de clasificar el riesgo en cirugía cardíaca pediátrica en tres categorías (bajo, medio y alto) y otro que predice la mortalidad hospitalaria de los pacientes operados de reparación abierta de aneurisma de aorta abdominal.

En primer lugar, se menciona un resumen de los resultados para la clasificación del riesgo en cirugía cardíaca pediátrica. Con el fin de obtener los mejores resultados, se evaluaron cuatro algoritmos diferentes: un perceptrón multicapa, un mapa autoorganizado, una red de función de base radial y finalmente, un árbol de decisión (C4.5). Se aplicó la validación cruzada para evaluar la generalización de la clasificación. La precisión fue similar entre el perceptrón multicapa y las redes de función de base radial con un 99.87% y 95.60%, respectivamente. Por otro lado, la precisión fue menor en el mapa autoorganizado y el árbol de decisión, 87.9% y 80.09%, respectivamente. Estos resultados demuestran que los algoritmos evaluados pueden ser útiles para ayudar a los médicos en las decisiones relacionadas con la estimación del riesgo en las enfermedades cardíacas

congénitas.

Seguidamente, se muestran los resultados de aplicar un perceptrón multicapa, una red de función de base radial y una red bayesiana, para predecir la mortalidad hospitalaria de los pacientes sometidos a reparación abierta de aneurisma de aorta abdominal. En esta prueba se desarrollaron dos tipos de experimentos. En primer lugar, se usaron todas las variables de la base de datos y posteriormente, se llevó a cabo el proceso de selección de características. Es importante mencionar que la precisión de todos los algoritmos probados (en todos los experimentos) fue superior al 91%, lo que significa que a nivel general el rendimiento de los algoritmos fue bueno, sin embargo, de manera individual cada algoritmo mostró notables diferencias entre la sensibilidad y la especificidad. Respecto al primer grupo de experimentos (con todas las variables), se observa, que la sensibilidad (probabilidad de predecir correctamente que un paciente va a morir después de la cirugía) para la red bayesiana fue del 73% y la especificidad (la probabilidad de predecir correctamente que un paciente sobreviva) fue del 92.6%. Para la red de función de base radial la sensibilidad fue del 52.1%, mientras que la especificidad fue del 96.1%. Para el perceptrón multicapa, las mismas medidas son 65.2% y 97%, respectivamente. Los resultados obtenidos para estos algoritmos no son lo suficientemente buenos para ser considerado en un sistema de ayuda a la toma de decisiones. En consecuencia, se continuó con el segundo grupo de experimentos.

En la segunda etapa, se aplicó un proceso de selección de características con el objetivo de reducir la complejidad computacional y el sobreajuste a fin de obtener mejores resultados. A nivel global se observa un mejoramiento de la sensibilidad 65.5%, la especificidad 96.1% y la precisión 95.8%. Adicionalmente, una combinación de algoritmos mejora los resultados para la predicción de la mortalidad en las dos etapas, obteniendo un crecimiento en la sensibilidad (82.6%) cuando se utilizan todas las variables y un ligero aumento con selección de características (87%).

Una amplia explicación de los resultados se muestran en las siguientes publicaciones científicas.

- Ruiz-Fernández, D., Monsalve-Torra, A., Soriano-Payá, A., Marín-Alonso, O., & Triana-Palencia, E. **Aid decision algorithms to estimate the risk in congenital heart surgery**. Volume 126, April 2016, Pages 118-127, ISSN 0169-2607, <http://dx.doi.org/10.1016/j.cmpb.2015.12.021>.

Abstract:

Background and objective: In this paper, we have tested the suitability of using different artificial intelligence-based algorithms for decision support when classifying the risk of congenital heart surgery. In this sense, classification of those surgical risks provides enormous benefits as the a priori estimation of surgical outcomes depending on either the type of disease or the type of repair, and other elements that influence the final result. This preventive estimation may help to avoid future complications, or even death. Methods: We have evaluated four machine learning algorithms to achieve our objective: multilayer perceptron, self-organizing map, radial basis function networks and decision trees. The architectures implemented have the aim of classifying among three types of surgical risk: low complexity, medium complexity and high complexity. Results: Accuracy outcomes achieved range between 80% and 99%, being the multilayer perceptron method the one that offered a higher hit ratio. Conclusions: According to the results, it is feasible to develop a clinical decision support system using the evaluated algorithms. Such system would help cardiology specialists, paediatricians and surgeons to forecast the level of risk related to a congenital heart disease surgery.

- Ana Monsalve-Torra, Daniel Ruiz-Fernandez, Oscar Marin-Alonso, Antonio Soriano-Payá, Jaime Camacho-Mackenzie, Marisol Carreño-Jaimes, **Using Machine Learning Methods for Predicting In-hospital Mortality in Patients Undergoing Open Repair of Abdominal**

Aortic Aneurysm. Volume 62, August 2016, Pages 195-201, ISSN 1532-0464,
<http://dx.doi.org/10.1016/j.jbi.2016.07.007>.

Abstract:

An abdominal aortic aneurysm is an abnormal dilatation of the aortic vessel at abdominal level. This disease presents high rate of mortality and complications causing a decrease in the quality of life and increasing the cost of treatment. To estimate the mortality risk of patients undergoing surgery is complex due to the variables associated. The use of clinical decision support systems based on machine learning could help medical staff to improve the results of surgery and get a better understanding of the disease. In this work, the authors present a predictive system of inhospital mortality in patients who were undergoing to open repair of abdominal aortic aneurysm. Different methods as multilayer perceptron, radial basis function and bayesian networks are used. Results are measured in terms of accuracy, sensitivity and specificity of the classifiers, achieving an accuracy higher than 95%. The developing of a system based on the algorithms tested can be useful for medical staff in order to make a better planning of care and reducing undesirable surgery results and the cost of the post-surgical treatments.

- Zapata-Impata, B. S., Ruiz-Fernandez, D., & Monsalve-Torra, A. **Swarm intelligence applied to the risk evaluation for congenital heart surgery.** 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, 2015, pp. 214-217. doi: 10.1109/EMBC.2015.7318338.
<http://ieeexplore.ieee.org/abstract/document/7318338/>

Abstract:

Particle Swarm Optimization is an optimization technique based on the positions of several particles created to find the best solution to

a problem. In this work we analyze the accuracy of a modification of this algorithm to classify the levels of risk for a surgery, used as a treatment to correct children malformations that imply congenital heart diseases.

- A. Monsalve Torra, D. Ruiz Fernández, A. Soriano Payá, **Clasificación del Riesgo de Cirugía Cardíaca Pediátrica mediante un Perceptrón Multicapa (MLP)**. XXXII Congreso Anual de la Sociedad Española de Ingeniería Biomédica. Noviembre 26-28, 2014. Barcelona. ISBN: 978-84-617-2446-8. www.ibebarcelona.eu/events/Caseib2014.

Abstract:

Las cardiopatías congénitas son una de las causas más comunes de mortalidad en niños especialmente en neonatos. Identificar tempranamente los factores de riesgo en los tratamientos correctivos o paliativos de la cirugía cardíaca es primordial para reducir la mortalidad y las complicaciones derivadas de éstos. El desarrollo de aplicaciones basadas en redes neuronales puede ayudar en la clasificación de dicho riesgo con mayor precisión y prontitud. Las redes neuronales, han mostrado gran capacidad para ser aplicadas como instrumentos de clasificación en el sector médico. En este trabajo de investigación, se ha desarrollado una aplicación, basada en perceptrón multicapa (MLP), que ha dado como resultado tres tipos de ajuste para el riesgo de la cirugía cardíaca congénita (CCC): bajo riesgo, riesgo medio y alto riesgo, con una precisión del 99%.

Capítulo 5

Conclusiones

En este capítulo se describen las conclusiones de este trabajo de investigación junto con las contribuciones y las futuras líneas de investigación.

Las enfermedades complejas están determinadas por la interacción de múltiples factores como los genéticos, los ambientales, nutricionales o el estilo de vida. El impacto de éstas en la calidad de vida de los pacientes que las padecen, sus familias y los costes en los tratamientos, hace que sean objeto de estudio por parte de la comunidad científica.

La atención de los pacientes con enfermedades complejas se ve afectada por variables como el entorno geográfico, la dificultad de acceso a los servicios médicos especializados y las dificultades propias del servicio de salud. Aunque actualmente se han aumentado los recursos destinados a la asistencia e investigación de las mismas, siguen siendo difíciles de tratar. Con la incorporación de los nuevos avances en tecnologías de análisis de los datos, en especial el uso de algoritmos de inteligencia artificial, se ha logrado una mejora en el proceso de diagnóstico y tratamiento de dichas enfermedades.

La investigación desarrollada en este trabajo ha tenido como eje central el uso de algoritmos de inteligencia artificial (redes neuronales artificiales, árboles de decisión y redes bayesianas) para aplicarlos en los sistemas de ayuda a la toma de decisiones en las enfermedades complejas. Estos métodos han hecho importantes aportaciones en el área de la informática para

CAPÍTULO 5. CONCLUSIONES

la salud.

El desarrollo de un sistema de ayuda en la toma de decisiones médicas basado en los métodos que se han probado en esta investigación podría mejorar el proceso de diagnóstico y tratamiento de estas enfermedades. Adicionalmente, el sistema de ayuda también podría ser útil a los médicos y personal sanitarios para evaluar el impacto de las decisiones relacionadas con los procesos antes, durante y después de ser aplicados, con el objetivo de dar prioridad a aquellas actividades que reduzcan la mortalidad, morbilidad y complicaciones de los pacientes.

Las redes neuronales artificiales usadas en esta investigación han demostrado ser una herramienta útil para proporcionar conocimiento en el desarrollo de un clasificador del riesgo en cirugía cardíaca pediátrica. Sin embargo, no todas ofrecen excelentes resultados en las pruebas realizadas. La red que mejor resultado obtuvo fue el perceptrón multicapa con una precisión de 99%. Con esta información se puede concluir que los resultados son favorables para la implementación de sistemas que ayuden a los médicos a lograr un adecuado plan de cuidado para cada paciente según la categoría de riesgo y por lo tanto, hacer una mejor distribución de los recursos.

La alta dimensionalidad es frecuente cuando se trabaja con bases de datos médicas. Para solventar este problema existen diferentes alternativas, entre ellas la selección de características principales. Esta técnica mejora la complejidad computacional, aporta más rapidez en el proceso y aumenta el rendimiento de los clasificadores. En esta investigación se observaron los siguientes resultados en la predicción de la mortalidad hospitalaria antes de aplicar este método; la sensibilidad 52.1%, especificidad 92.6% y la precisión de 94.2%. Posteriormente, tras la aplicación de la selección de características se evidencia una mejora de los valores (sensibilidad 65.3%, especificidad 98.8% y 96.4% en la precisión).

Para predecir la mortalidad hospitalaria se utilizaron las redes bayesianas (utilizadas con frecuencia cuando se espera un resultado binario), un perceptrón multicapa y las redes de función de base radial. Los resultados de las pruebas realizadas mostraron que las redes bayesianas ofrecen

los mejores resultados logrando una precisión del 96.1%, una sensibilidad del 86.8% y una especificidad del 96.8%. Estas redes muestran ventajas sobre las redes neuronales artificiales, debido a que aprenden sobre las relaciones de dependencia y causalidad.

El preprocesado de los datos con técnicas como la discretización, la normalización y los métodos de imputación, ha mostrado ser de gran utilidad cuando se trabaja con datos especialmente sensibles como los datos médicos. Los resultados muestran que aplicar estas técnicas, mejora el desempeño de los clasificadores probados en esta investigación.

Esta tesis ha contribuido al desarrollo de una propuesta basada en el conocimiento de los algoritmos de inteligencia artificial aplicados a las enfermedades complejas integrando las actividades que participan en el proceso de diagnóstico y tratamiento médico y en base a éste desarrollar un plan de cuidado más acertado para cada paciente y de esta manera mejorar el servicio, reducir los costes del tratamiento y mejorar el seguimiento de los pacientes en situación de alto riesgo.

Lineas Futuras de Investigación

Como continuación de este trabajo de investigación se propone profundizar en el estudio de herramientas para el tratamiento del desbalance de los datos. El problema de desbalance en la clasificación o predicción se presenta cuando el conjunto de datos tiene una gran cantidad de datos de un cierto tipo y menos de otro. El problema de desbalance es complejo, y se deben estudiar técnicas como SMOTE, *ramdon forest*, entre otras. Las técnicas mencionadas son unas de muchas que se pueden utilizar, sin embargo, se debe tener en cuenta el tipo de problemas que se aborde para seleccionar la técnica más adecuada.

Con los resultados obtenidos en las diferentes pruebas realizadas, queda abierta la línea para desarrollar un sistema de información que ayude a clasificar el riesgo en cirugía cardiovascular pediátrica. De igual forma se plantearía el desarrollo un sistema que ayude de forma más acertada en

CAPÍTULO 5. CONCLUSIONES

la predicción de la mortalidad hospitalaria de los pacientes operados de reparación abierta y endovascular de aneurisma de aorta abdominal.



Universitat d'Alacant
Universidad de Alicante

Anexo I

Variables base de datos de pediatría

En este anexo se recogen las variables de la base de datos de cardiopatías congénitas.

Variable	Descripción	Valor
Datos básicos		
Edad	Edad de los pacientes, según la clasificación de K, Jenkis [114].	1,2,3
Género	Sexo de los pacientes.	0, 1
Régimen de salud	Empresas prestadoras del régimen de salud.	1 a 7
Fechas de cirugía	Mes en el cual se realizó la cirugía.	1 a 12
Peso	Peso de los pacientes.	1 a 6
Antecedentes médicos		
Asplenia	Ausencia del bazo.	0,1
Poliesplenia	Presencia de varios bazos en el abdomen de un mismo sujeto.	0,1
Síndrome de Down	Trastorno genético causado por la presencia de una copia extra del cromosoma 21.	0,1
Síndrome de Marfan	Enfermedad rara del tejido conectivo que se caracteriza por un aumento inusual de la longitud de los miembros.	0,1

ANEXO I

Síndrome de Di-George	Es una enfermedad causada por la delección de una pequeña parte del cromosoma 22.	0,1
Síndrome dismórfico	Enfermedad congénita particular en la que se encuentran una o más anomalías.	0,1
Soporte vent	Uso de ventilación mecánica preoperatoria.	0,1
Falla renal	Mal funcionamiento los riñones, los cuales no son capaces de filtrar las toxinas y otras sustancias de desecho de la sangre adecuadamente.	0,1
Hipoxia	Deficiencia de oxígeno en la sangre.	0,1
Arritmias	Alteración del ritmo cardíaco.	0,1
Insuficiencia cardíaca congestiva	Desequilibrio de la capacidad del corazón para bombear sangre.	0,1
Dextrocardia	Ubicación del corazón en la mitad derecha del tórax debido a una causa genética.	0,1
Paro cardíaco	Detención del latido cardíaco en un individuo	0,1
Hipertensión pulmonar	Presión arterial anormalmente alta en las arterias de los pulmones.	0,1
Infección preoperatoria	Proceso infeccioso de cualquier tipo que ocurre antes de la cirugía.	0,1
Diagnóstico	Proceso mediante el cual se identifica una enfermedad.	1 a 6
Hospitalizaciones preoperatorias	Si el paciente ha tenido hospitalizaciones previas.	0,1

Intervención Quirúrgica		
Número de diagnósticos	Numero de diagnósticos que tenían los pacientes.	1 a 6
Número de intervenciones	Número de procedimientos quirúrgicos que se le realizaron al paciente.	1 a 6
CIA-Presencia	Presencia de comunicación interauricular.	0,1
CIV-Presencia	Presencia de comunicación interventricular.	0,1
AV-canal	Presencia de canal auriculoventricular.	0,1
Tiempo de cirugía	Tiempo de duración de la intervención quirúrgica (horas).	0,3
Arresto cardíaco	Parada de la circulación extracorpórea en cirugía.	0,1
Circulación extracorpórea	Uso de circulación extracorpórea.	0,1
Otra cirugía	Si se realizó otro procedimiento quirúrgico diferente a la cirugía de corazón.	0,1
Información post-quirúrgica		
T vent meca	Tiempo de uso de ventilación mecánica (horas).	1 a 6
Complicaciones	Presencia de complicaciones en el post-operatorio.	0,1
GRE requerido	Uso de glóbulos rojos requeridos en el post-operatorio.	0,1
Crio-requerido	Uso de criopresipitados requeridos en el postoperatorio.	0,1
Sangrado	Si el paciente presento un sangrado excesivo en el post-operatorio.	0,1
Paro cardíaco	Detención del latido cardíaco en un paciente en el post-operatorio.	0,1
Arritmias	Alteración del ritmo cardíaco en el post-operatorio.	0,1
Bloqueo av completo	Arritmia caracterizada por el bloqueo auriculo-ventricular completo.	0,1

ANEXO I

Marcapasos definitivo	Uso de un dispositivo marcapasos de manera permanente.	0,1
Taponamiento cardíaco	Es la presión sobre el corazón que ocurre cuando se acumula sangre o líquido en el espacio entre el músculo del corazón y el saco exterior.	0,1
Bajo gasto cardíaco	Disminución del volumen de sangre expulsado por un ventrículo en un minuto.	0,1
ICC descompensada	Desequilibrio de la capacidad del corazón para bombear sangre que se presentó en el post-operatorio.	0,1
Crisis HTA post	Aumento incontrolado de presión arterial sistémica en el post-operatorio.	0,1
Trombosis arterial periférica	Es la oclusión de una o varias arterias destinadas a llevar sangre a los miembros inferiores.	0,1
Atelectasia	Es el colapso de una parte o de todo el pulmón.	0,1
Derrame pleural	Es una acumulación de líquido entre las capas de los tejidos que recubren los pulmones.	0,1
Quilotorax	Es la presencia de líquido linfático en el espacio pleural.	0,1
Neumotorax	Es el colapso pulmonar que ocurre cuando el aire escapa del pulmón.	0,1
Edema pulmonar	Acumulación anormal de líquido en los pulmones que lleva a que se presente dificultad respiratoria.	0,1
Enfisema subcutáneo	Complicación que se presenta cuando el aire penetra dentro de los tejidos bajo la piel que cubre la pared torácica.	0,1
Traqueostomía	Procedimiento quirúrgico realizado con objeto de crear una abertura dentro de la tráquea debido a largos periodos de uso de la ventilación mecánica.	0,1
Crisis hipóxicas	Son los cuadros de hipoxemia agudizada en un paciente.	0,1

htp	Presión arterial anormalmente alta en las arterias de los pulmones en el post-operatorio.	0,1
Parálisis diafragmática	Es la pérdida del movimiento del diafragma.	0,1
Infrespiratoria	Infección del sistema respiratorio que se presenta en el post-operatorio.	0,1
Fiebre	Aumento de la temperatura corporal por encima de la normal que manifiesta la reacción del organismo frente a alguna infección (enfermedad).	0,1
Cultivo	Es un método para la multiplicación de microorganismos cuando se presentan procesos agudos de infección en el postoperatorio.	0,1
AB-mas de 5 días	Uso de tratamiento antibióticos por más de 5 días.	0,1
Inf hqca	Proceso infeccioso que se presenta en la herida quirúrgica.	0,1
Mediastinitis	Es la inflamación e irritación de la zona del tórax.	0,1
Endocardi	Es una inflamación de la membrana interna del corazón (endocardio).	0,1
Sepsis	Tiene lugar cuando el cuerpo tiene una respuesta inmunológica a una infección bacteriana muy grave.	0,1
Dialisis	Tratamiento para sustituir el trabajo que realizan los riñones (mal funcionamiento).	0,1
Convulsion	Aparición súbita de una actividad eléctrica anormal en el cerebro.	0,1
Enc hipox	Es el síndrome producido por la disminución del aporte de oxígeno en el flujo sanguíneo cerebral.	0,1
Reinterv sangrado	Reintervención por sangrado post-quirúrgico.	0,1
Plicatura diafragmática	Procedimiento quirúrgico que se realiza cuando el diafragma ha perdido su fuerza y tensión y esto repercute en la salud respiratoria del paciente.	0,1

ANEXO I

Lavado	Lavado mediastinal procedimiento quirúrgico que se realiza para la limpieza del mediastino cuando éste presenta una infección grave.	0,1
Cerclaje	Procedimiento quirúrgico de contención del hueso fracturado del esternón con la ayuda de alambres.	0,1
Reingreso uci	Si el paciente tuvo que reingresar a la unidad de cuidados intensivos, debido a causas derivadas de la cirugía de corazón u otras causas.	0,1
Est preoperatoria	Tiempo que permaneció el paciente hospitalizado antes de ser sometido al procedimiento quirúrgico (agrupados).	1 a 4
Est UCI	Tiempo que permaneció el paciente hospitalizado en la unidad de cuidados intensivos después del procedimiento quirúrgico (agrupados).	1 a 4
Est hosp total pop	Tiempo que permaneció el paciente hospitalizado en planta después del procedimiento quirúrgico (agrupados).	1 a 4
Condición primeros	Condición de sobre el estado del paciente a los 30 días después de ser operado.	0,1
RACHScorregido	Escala de riesgo cardiovascular en cirugía pediátrica según K. Jenkis [114].	1 a 3

Anexo II

Variables base de datos de pacientes con aneurisma de aorta abdominal

En este anexo se recogen las variables de la base de datos de aneurisma de aorta abdominal.

Variable	Descripción	Valor
Información básica del paciente		
Grupo edad	Edad de los pacientes operados clasificados en tres grupos con una diferencia entre ellos de 15 años.	1,2,3
Género	Género del paciente (hombres, mujeres).	0,1
Cod IMC	Índice de masa corporal, medida de asociación entre el peso y la talla de una persona y está clasificada en 6 grupos.	1 a 6
Antecedentes médicos		
Riesgo	Riesgo de muerte según Leiden.	0,1
Electivo	Cirugía que el paciente ha elegido que se le realice.	0,1
NYHA	Es la valoración funcional de insuficiencia cardíaca.	1 a 4

ANEXO II

ASA	Sistema de clasificación que utiliza la <i>American Society of Anesthesiologists</i> (ASA) para estimar el riesgo que plantea la anestesia para los distintos estados del paciente.	1 a 5
Hta	Aumento incontrolado de presión arterial sistémica.	0,1
Diabetes	Presencia de esta patología caracterizada por la presencia de concentraciones elevadas de glucosa en la sangre.	0,1
ICC	Presencia de insuficiencia cardíaca congestiva que se caracteriza por la incapacidad del corazón de bombear sangre en los volúmenes normales.	0,1
ECV	Presencia de enfermedad cerebro vascular. Está patología sucede cuando el flujo de sangre a una parte del cerebro se detiene por diversas causas.	0,1,3
Dislipidemia	La presencia de esta patología se presenta cuando se aumentan las concentraciones de lípidos y lipoproteínas en la sangre.	0,1
IAM	Infarto agudo de miocardio, patología que se caracteriza por la muerte de una porción del músculo cardíaco.	0,1
Angina	Angina inestable es una afección en la cual el corazón no recibe suficiente flujo de sangre y oxígeno, puede ocasionar un ataque cardíaco.	0,1
Tipo angina	Clase de angina que presenta el paciente, puede ser: estable, inestable, sin angina.	0,1,2
Epoc	Enfermedad pulmonar obstructiva crónica, trastorno pulmonar que se caracteriza por la existencia de una obstrucción de las vías aéreas generalmente progresiva.	0,1
Enfermedad renal	Mal funcionamiento los riñones.	0,1

Cx vascular	Antecedente de cirugía vascular previa.	0,1
Cx no cardíaca	El paciente ha tenido otro tipo de cirugía.	0,1
Cx cardíaca	El paciente ha tenido algún tipo de cirugía cardíaca previa.	0,1
RVM previa	Revascularización miocardia previa.	0,1
Angioplastia	Angioplastia previa, consiste en introducir un balón para dilatar una arteria ocluida.	0,1
Tabaquismo	Se presenta cuando un paciente abusa del tabaco.	0,1
Betabloq	Tratamiento del paciente con betabloqueadores.	0,1
Esteroides	Tratamiento del paciente con esteroides.	0,1
Nitratos iv	Tratamiento del paciente con nitratos intravenosos.	0,1
Nitratos oral	Tratamiento del paciente con nitratos por vía oral.	0,1
ASA medica	Tratamiento del paciente con aspirina.	0,1
Anticoagu	Tratamiento del paciente con anticoagulantes.	0,1
IECA	Tratamiento del paciente con inhibidores de la enzima convertidora de angiotensia.	0,1
ARA II	Tratamiento del paciente con antagonistas de sistema angiotensia.	0,1
Diuréticos	Tratamiento del paciente con diuréticos (facilitan la eliminación de orina).	0,1
Clopidogrel	Tratamiento del paciente con un antiplaquetario por vía oral.	0,1
Hipolipemiantes	Tratamiento del paciente con este medicamento para disminuir los niveles de lípidos en sangre.	0,1
Datos quirúrgicos		
Cod FE-preqx	Fracción de eyección previa.	1,2,3,4
Cod vol aneurisma	Volumen de sangre que contenía el aneurisma.	1,2,3,4

ANEXO II

Cod protesis	Tamaño de las prótesis implantadas al paciente	1,2
Cuerpos	Número de prótesis implantada al paciente.	1,2
Cod abordaje	Código del abordaje que puede ser extraperitoneal o transperitoneal.	1,2
TT	Tiempo total de duración de la cirugía (horas).	1,2,3,4
Clampeo	Tiempo total de pinzamiento de la aorta (horas).	1,2,3,4
Sangrado	Cantidad de sangre perdida durante la cirugía.	1,2,3,4
Transfundido	Si el paciente requirió sangre durante la cirugía.	0,1
Comp cx def	Se presentaron complicaciones en cirugía.	0,1
Datos post-quirúrgicos		
Cod transfpop	Uso de transfusión sanguínea en el post-operatorio.	1,2,3,4
Comp neuro def	Presencia de complicación neurológica postoperatorio.	0,1
Comp renal def	Presencia de complicación renal post-operatorio.	0,1
Comp infecc def	Presencia de infecciones en el post-operatorio.	0,1
Comp pulm def	Presencia de complicaciones de tipo respiratorio en el post-operatorio.	0,1
Comp vasc def	Presencia de complicaciones en el sistema vascular periférico.	0,1
Comp cardiac def	Presencia de complicaciones cardíacas.	0,1
Otra comp def	Presencia de otras complicaciones.	0,1
Est UCI	Tiempo de estancia en días en la UCI post-operatoria.	1,2
Totalpop	Tiempo total de estancia en días post-operatorios.	1,2
Est total	Tiempo total de estancia hospitalaria.	1,2
Mort opera	Presencia de mortalidad durante la hospitalización.	0,1

Referencias

- [1] J.M. Soria, J.F. Dilmé, J. Martínez-González, M. Camacho, C. Rodríguez, J.M. Romero, S. Bellmunt, J.R. Escudero, and L. Vila. Métodos de Estudio de las Enfermedades Complejas: Aneurismas de la Aorta Abdominal. *Angiología*, 62(2):58–64, 2010.
- [2] H. Dolk, M. Loane, E. Garne, European Surveillance of Congenital Anomalies (EUROCAT) Working Group, et al. Congenital heart defects in Europe prevalence and perinatal mortality, 2000 to 2005. *Circulation*, 123(8):841–849, 2011.
- [3] M. Burcha and N. Dedieub. Anuario 2012: Cardiopatías congénitas. las revistas de las sociedades nacionales presentan una selección de las investigaciones que han impulsado avances recientes en cardiología clínica. *CorSalud*, 5(1):30–42, 2013.
- [4] D. Van der Linde, E. Konings, M.A. Slager, M. Witsenburg, W.A. Helbing, J.M Takkenberg, and J. Roos-Hesselink. Birth Prevalence of Congenital Heart Disease Worldwide: a Systematic Review and Meta-analysis. *Journal of the American College of Cardiology*, 58(21):2241–2247, 2011.
- [5] D.C. Rodriguez, H.F. Ariza, I. Riano-Galan, C. Moro-Bayon, E. Suarez-Menendez, C. Mosquera-Tenreiro, and E. Garcia-Lopez. The epidemiology of Congenital Heart Disease in Asturias (Spain) During the Period 1990-2004. *Anales de pediatria (Barcelona, Spain : 2003)*, 71(6):502–9, dec 2009.

REFERENCIAS

- [6] E. Baltaxe and I. Zarante. Prevalencia de Malformaciones Cardíacas Congenitas en Colombia. *Archivos de Cardiología de México*, 76(3):263–268, 2006.
- [7] Informe de la Salud Cardiovascular en España en el Contexto Europeo. Technical report, Sociedad Española de Cardiología, 2010.
- [8] J.M. Ortega-Martín, M.C Fernández-Morán, M.I. Alonso-Álvarez, M. García-Gimeno, R. Fernández-Samos, and F. Vaquero-Morillo. Prevalencia de aneurismas de aorta abdominal en una población de riesgo. *Angiología*, 59(4):305–315, 2007.
- [9] A.T Hirsch, Z.J Haskal, N.R Hertzner, C.W Bakal, M.A Creager, J.L Halperin, L.F Hiratzka, W. Murphy, J.W Olin, J.B Puschett, et al. Acc/aha 2005 guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic): A collaborative report from the american association for vascular surgery/society for vascular surgery, society for cardiovascular angiography and interventions, society for vascular medicine and biology, society of interventional radiology, and the acc/aha task force on practice guidelines (writing committee to develop guidelines for the management of patients with peripheral arterial disease). *Journal of the American College of Cardiology*, 47(6):e1–e192, 2006.
- [10] J. González de Dios. De la medicina Basada en la Evidencia a la Evidencia Basada en la Medicina. *Anales de Pediatría*, 55(5):429–439, jan 2001.
- [11] J. Trujillano, J. March, and A. Sorribas. Aproximación metodológica al uso de redes neuronales artificiales para la predicción de resultados en medicina. *Medicina Clinica (Barcelona)*, 122(s1):59–67, 2004.
- [12] N.K. Kasabov. *Evolving Connectionist Systems: Methods and Applications in Bioinformatics, Brain Study and Intelligent Machines*. Springer, London, 1 edition, 2003.

- [13] J. Kuruvilla and K Gunavathi. Lung Cancer Classification Using Neural Networks for CT Images. *Computer Methods and Programs in Biomedicine*, 113(1):202–209, 2014.
- [14] J. Myers, C. R de Souza, A. Borghi-Silva, M. Guazzi, P. Chase, D. Bensimhon, M.A. Peberdy, E. Ashley, E. West, L.P. Cahalin, D. Forman, and R. Arena. A Neural Network Approach to Predicting Outcomes in Heart Failure Using Cardiopulmonary Exercise Testing. *International Journal of Cardiology*, 171(2):265–9, 2014.
- [15] J. S. Sonawane and D. R. Patil. Prediction of Heart Disease Using Multilayer Perceptron Neural Network. In *Information Communication and Embedded Systems (ICICES), 2014 International Conference on*, pages 1–6, Chennai, 2014. IEEE Computer Society.
- [16] F.B. Chimiesk and R.D. Ribeiro F. Association and Classification Data Mining Algorithms Comparison over Medical Datasets. *Journal of Health Informatics*, 5(2):44–51, 2013.
- [17] D. Ruiz, V. Berenguer, A. Soriano, and B. Sánchez. A decision support system for the diagnosis of melanoma: A comparative approach. *Expert Systems with Applications*, 38(12):15217–15223, 2011.
- [18] V.C Korfiatis, P.A. Asvestas, K.K. Delibasis, and G.K. Matsopoulos. A classification System Based on a New Wrapper Feature Selection Algorithm for the Diagnosis of Primary and Secondary Polycythemia. *Computers in Biology and Medicine*, 43(12):2118–26, 2013.
- [19] D. Setsirichok, T. Piroonratana, W. Wongserree, N. Usavanarong, T. and Paulkhaolarn, M. Kanjanakorn, Ch. and Sirikong, Ch. Limwongse, and N. Chaiyaratana. Classification of complete blood count and haemoglobin typing data by a C4.5 decision tree, a naïve Bayes classifier and a multilayer perceptron for thalassaemia screening. In *Biomedical Signal Processing and Control*, volume 7, pages 202–212. 2012.

REFERENCIAS

- [20] V. M. Bostan and B. Pantelimon. Creating a model based on artificial neural network for liver cirrhosis diagnose. In *2015 9th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, pages 295–298, May 2015.
- [21] B.S Khehra and A.P.S Pharwaha. Classification of Clustered Microcalcifications using MLFFBP-ANN and SVM. *Egyptian Informatics Journal*, 17(1):11–20, 2016.
- [22] Y. Cheng, G. Yan, Y. Chiu, F. Chang, C. Chang, and K. Chung. Efficient Fetal Size Classification Combined with Artificial Neural Network for Estimation of Fetal Weight. *Taiwanese journal of Obstetrics & Gynecology*, 51(4):545–53, 2012.
- [23] O. Atkov, S.G. Gorokhova, A.G. Sboev, E.V. Generozov, E.V. Muraseyeva, S.Y. Moroshkina, and N.N. Cherniy. Coronary Heart Disease Diagnosis by Artificial Neural Networks Including Genetic Polymorphisms and clinical Parameters. *Journal of Cardiology*, 59(2):190–4, 2012.
- [24] H. Uğuz. A biomedical system based on artificial neural network and principal component analysis for diagnosis of the heart valve diseases. *Journal of Medical Systems*, 36(1):61–72, 2012.
- [25] K. Rajeswari, V. Vaithyanathan, and T.R. Neelakantan. Feature Selection in Ischemic Heart Disease Identification using Feed Forward Neural Networks. *Procedia Engineering*, 41:1818–1823, 2012.
- [26] S. Abdul-Rahman, Ahmad Khairil Norhan, M. Yusoff, A. Mohamed, and S. Mutalib. Dermatology diagnosis with feature selection methods and artificial neural network. In *2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences*, pages 371–376, Dec 2012.
- [27] M. Durairaj and V. Revathi. Prediction of Heart Disease Using Back-propagation MLP Algorithm. *International Journal of Scientific & Technology Research*, 4(8):235–239, 2015.

- [28] N. Al-Milli. Backpropagation neural network for prediction of heart disease. *Journal of Theoretical and Applied Information Technology*, 56(1):131–135, 2013.
- [29] S. Joo, K. Choi, and S Huh. Prediction of Spontaneous Ventricular Tachyarrhythmia by an Artificial Neural Network Using Parameters Gleaned from Short-term Heart Rate Variability. *Expert Systems with Applications*, 39(3):3862–3866, feb 2012.
- [30] C. P Launay, H. Rivière, A. Kabeshova, and O. Beauchet. Predicting Prolonged Length of Hospital Stay in Older Emergency Department Users: Use of a Novel Analysis Method, the Artificial Neural Network. *European Journal of Internal Medicine*, 2015.
- [31] R. Houthoof, J. Ruysinck, J. Van der Hertten, S. Stijven, I. Couckuyt, B. Gadeyne, F. Ongena, K. Colpaert, J. Decruyenaere, T. Dhaene, and F. De Turck. Predictive Modelling of Survival and Length of Stay in Critically Ill Patients Using Sequential Organ Failure Scores. *Artificial Intelligence in Medicine.*, 63(3):191–207, 2015.
- [32] A.M. Vukicevic, G.R. Jovicic, M.M. Stojadinovic, R.. Prelevic, and N.D. Filipovic. Evolutionary Assembled Neural Networks for Making Medical Decisions with Minimal Regret: Application for Predicting Advanced Bladder Cancer Outcome. *Expert Systems with Applications*, 41(18):8092–8100, 2014.
- [33] I. Saritas. Prediction of Breast Cancer Using Artificial Neural Networks. *Journal of Medical Systems.*, 36(5):2901–2907, 2012.
- [34] L. Spelt, J. Nilsson, R. Andersson, and B. Andersson. Artificial Neural Networks a Method for Prediction of Survival Following Liver Resection for Colorectal Cancer Metastases. *European Journal of Surgical Oncology: the Journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology*, 39(6):648–54, 2013.

REFERENCIAS

- [35] E.S Wise, K.M Hocking, and C.M Brophy. Prediction of In-hospital Mortality After Ruptured Abdominal Aortic Aneurysm Repair Using an Artificial Neural Network. *Journal of Vascular Surgery*, 62(1):8–15, 2015.
- [36] G. Wang, K. Lam, Z. Deng, and K. Choi. Prediction of Mortality After Radical Cystectomy for Bladder Cancer by Machine Learning Techniques. *Computer in Biology and Medicine.*, 63(C):124–132, 2015.
- [37] D. Ansari, J. Nilsson, R. Andersson, S. Regnér, B. Tingstedt, and B. Andersson. Artificial Neural Networks Predict Survival from Pancreatic Cancer After Radical Surgery. *American Journal of Surgery*, 205(1):1–7, 2013.
- [38] G. Çelik, O.K Baykan, Y. Kara, and H. Tireli. Predicting 10-day Mortality in Patients with Strokes Using Neural Networks and Multivariate Statistical Methods. *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association*, 23(6):1506–12, 2014.
- [39] N. Sut and Y. Çelik. Prediction of Mortality in Stroke Patients Using Multilayer Perceptron Neural Networks. *Turkish Journal of Medical Sciences*, 42(5):886–893, 2012.
- [40] N.A. Loghmanpour, M.K. Kanwar, S.H. Bailey, R.L. Benza, J.F. Antaki, and S. Murali. Development of a Novel Predictive Model for Mortality Post Continuous Flow LVAD Implant using Bayesian Networks (BN). *The Journal of Heart and Lung Transplantation*, 33(4):S74, 2014.
- [41] J. Vila-Frances, J. Sanchis, E. Soria-Olivas, A.J Serrano, M. Martinez-Sober, C. Bonanad, and S. Ventura. Expert System for Predicting Unstable Angina Based on Bayesian Networks. *Expert Systems with Applications*, 40(12):5004–5010, 2013.
- [42] I. Cho, I. Park, E. Kim, E. Lee, and D.W. Bates. Using EHR Data to Predict Hospital-Acquired Pressure Ulcers: a Prospective Study of a

- Bayesian Network Model. *International Journal of Medical Informatics*, 82(11):1059–67, 2013.
- [43] S. Kachroo, W.W. Melek, and C.J Kurian. Evaluation of Predictive Learners for Cancer Incidence and Mortality. In *The 4th IEEE International Conference on E-Health and Bioengineering - EHB 2013*, pages 1–6, Iasi, 2013. IEEE.
- [44] C. Colak, E. Karaman, and M.G. Turtay. Application of Knowledge Discovery Process on the Prediction of Stroke. *Computer Methods and Programs in Biomedicine*, 119(3):181–185, 2015.
- [45] M. Alsalamah, S Amin, and J. Halloran. Diagnosis of Heart Disease by Using a Radial Basis Function Network classification technique on patients' medical records. In *RF and Wireless Technologies for Biomedical and Healthcare Applications (IMWS-Bio), 2014 IEEE MTT-S International Microwave Workshop Series on*, pages 1–4, London, 2014. IEEE Computer Society.
- [46] J. Mateo and J.J. Rieta. Radial Basis Function Neural Networks Applied to Efficient QRST Cancellation in Atrial Fibrillation. *Computers in Biology and Medicine*, 43(2):154–63, 2013.
- [47] N. Torbati, A. Ayatollahi, and A. Kermani. An Efficient Neural Network Based Method for Medical Image Segmentation. *Computers in Biology and Medicine*, 44:76–87, 2014.
- [48] P. Mei, C. de Carvalho Carneiro, S.J. Fraser, L. Min, and F. Reis. Analysis of neoplastic lesions in magnetic resonance imaging using self-organizing maps. *Journal of the Neurological Sciences*, 359(1):78–83, 2015.
- [49] A. Ortiz, J.M. Gorriz, J. Ramirez, and D. Salas-Gonzalez. Improving MR Brain Image Segmentation Using Self-organising Maps and Entropy-gradient Clustering. *Information Sciences*, 262:117–136, 2014.

REFERENCIAS

- [50] M. Tayefi, H. Esmaeili, M.S Karimian, A.A. Zadeh, M. Ebrahimi, M. Safarian, M. Nematy, S.M. Parizadeh, G.A. Ferns, and M. Ghayour-Mobarhan. The Application of a Decision Tree to Establish the Parameters Associated with Hypertension. *Computer Methods and Programs in Biomedicine*, 2016.
- [51] Y. Wang, M.A. Simon, P. Bonde, B.U. Harris, J.J. Teuteberg, R.L. Kormos, and J.F. Antaki. Decision Tree for Adjuvant Right Ventricular Support in Patients Receiving a Left Ventricular Assist Device. *The Journal of Heart and Lung Transplantation*, 31(2):140–149, 2012.
- [52] F. Brims, T.M. Meniawy, I. Duffus, D. de Fonseca, A. Segal, J. Creaney, N. Maskell, R.A. Lake, N. de Klerk, and A.K. Nowak. A Novel Clinical Prediction Model for Prognosis in Malignant Pleural Mesothelioma Using Decision Tree Analysis. *Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer*, 11(4):573–82, 2016.
- [53] T.P. Exarchos, A.T. Tzallas, D. Baga, D. Chaloglou, D.I. Fotiadis, S. Tsouli, M. Diakou, and S. Konitsiotis. Using Partial Decision Trees to Predict Parkinson’s Symptoms: A New Approach for Diagnosis and Therapy in Patients Suffering from Parkinson’s Disease. *Computers in Biology and Medicine*, 42(2):195–204, 2012.
- [54] J. Bonis, J.J. Sancho, and F. Sanz. Sistemas Informáticos de Soporte a la Decisión Clínica. *Medicina Clínica*, 122(Supl.1):39–44, 2004.
- [55] A Sanabria. Repercusiones del Error Médico. *Tribuna Médica.*, 103:189–189, 2003.
- [56] T.J. Bright, A. Wong, R. Dhurjati, E. Bristow, L. Bastian, R.R. Coeytaux, G. Samsa, V. Hasselblad, J.W. Williams, M.D. Musty, L. Wing, A.S. Kendrick, G.D. Sanders, and D. Lobach. Effect of clinical decision-support systems a systematic review. *Annals of Internal Medicine*, 157(1):29–43, 2012.

- [57] R.B Haynes, N.L. Wilczynski, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: Methods of a decision-maker-researcher partnership systematic review. *Implementation Science*, 5(1):12, 2010.
- [58] M.H Hassoun. *Fundamentals of Artificial Neural Networks*. MIT Press, 1 edition, 1995.
- [59] A Nigrin. *Neural Networks for Pattern Recognition*. MIT Press, Cambridge, MA, 1 edition, 1993.
- [60] S. Haykin. *Neural Networks: A Comprehensive Foundation*. MacMillan Publishing Company, New York, NY, USA, 1 edition, 1994.
- [61] I.A Basheer and M Hajmeer. Artificial Neural Networks: Fundamentals, Computing, Design and Application. *Journal of Microbiological Methods*, 43(1):3–31, 2000.
- [62] F Rosenblatt. The perceptron: A probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*, 65(6):386–408, 1958.
- [63] T. Kavzoglu. *An investigation of the Design and Use of Feed-forward Artificial Neural Networks in the Classification of Remotely Sensed Images*. PhD thesis, University of Nottingham., 2001.
- [64] G. Panchal, A. Ganatra, Y.P. Kosta, and D. Panchal. Behaviour Analysis of Multilayer Perceptrons with Multiple Hidden Neurons and Hidden Layers. *International Journal of Computer Theory and Engineering*, 3(2):332–337, 2011.
- [65] O.G Liu and S.L He. Determining Automatically Structure Parameters of BP \hat{A} Artificial Neural Network by a Computer. *Computer Engineering and Applications*, 13(146):72–74, 2004.
- [66] D.E Rumelhart and J McClelland. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. MIT Press, Cambridge, MA, 1 edition, 1986.

REFERENCIAS

- [67] T. Kohonen. The Self-organizing Map. *Proceedings of the IEEE*, 78(9):1464–1480, 1990.
- [68] B. Martín del Brío and A. Sanz-Molina. *Redes Neuronales y Sistemas Difusos*. RA-MA, Madrid, 2 edition, 2001.
- [69] D.S Broomhead and D. Lowe. Multivariate Functional Interpolation and Adaptive Networks. *Complex Systems*, 2:321–355, 1988.
- [70] S.A. Hannan, A.V Mane, R.R Manza, and R.J Ramteke. Prediction of Heart Disease Medical Prescription Using Radial Basis Function. In *Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference on*, pages 1–6, Coimbatore, 2010. IEEE Computer Society.
- [71] J. Macqueen. Some methods for classification and analysis of multivariate observations. In *In 5-th Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297, 1967.
- [72] T. Bayes. An Essay Towards Solving a Problem in the Doctrine of Changes. *Philosophical Transactions*, 53:370–418, 1763.
- [73] B. Sierra-Araujo. *Aprendizaje Automático: Conceptos Básicos y Avanzados: Aspectos Prácticos Utilizando el Software Weka*. Pearson Prentice Hall, Madrid, 1 edition, 2006.
- [74] N. Friedman, D. Geiger, and M. Goldszmidt. Bayesian Network Classifiers. *Machine Learning*, 29(2-3):131–163, 1997.
- [75] G. Lindgaard, C. Pyper, M. Frize, and R. Walker. Does Bayes have it? Decision Support Systems in Diagnostic Medicine. *International Journal of Industrial Ergonomics*, 39:524–532, 2009.
- [76] S.J. Russell and P. Norvig. *Artificial Intelligence A Modern Approach*. Prentice Hall, Englewood Cliffs, New Jersey, 1 edition, 2003.

- [77] C. Huang and A. Darwiche. Inference in belief networks: A procedural guide. *International Journal of Approximate Reasoning*, 15(3):225–263, 1996.
- [78] J.A Chaparro, B. Giraldo, and S. Rodon. Evaluacion del Clasificador Naive Bayes como Herramienta de Diagnostico en Unidades de Cuidado Intensivo. *Revista de Tecnologia*, 12(2):87–93, 2013.
- [79] G. Meyfroidt, F. Güiza, J. Ramon, and M. Bruynooghe. Machine Learning Techniques to Examine Large Patient Databases. *Best Practice and Research: Clinical Anaesthesiology*, 23(1):127–143, 2009.
- [80] H. Zhang and C.X. Ling. An improved learning algorithm for augmented naive bayes. In *Bayes, Pacific-Asia Conference on Knowledge Discovery and Data Mining, LNCS*, pages 581–586. Springer Verlag, 2001.
- [81] R.O. Duda and P.E. Hart. *Pattern Classification and Scene Analysis*. Wiley Interscience Publication, 1 edition, 1973.
- [82] E. Frank, M. Hall, and B. Pfahringer. Locally weighted naive bayes. In *Proceedings of the Nineteenth Conference on Uncertainty in Artificial Intelligence, UAI'03*, pages 249–256, San Francisco, CA, USA, 2003. Morgan Kaufmann Publishers Inc.
- [83] F.J. García Castellanos. *Modelos bayesianos para la Clasificación Superpuesta: Aplicaciones al Análisis de Datos de Expresión Genética*. PhD thesis, Universidad de Granada, 2009.
- [84] C. Chow and C. Liu. Approximating discrete probability distributions with dependence trees. *IEEE Transactions Information Theory*, 14(3):462–467, 2006.
- [85] L. Sucar. Redes bayesianas. *BS Araujo, Aprendizaje Automático: conceptos básicos y avanzados*, pages 77–100, 2006.
- [86] J.R. Quinlan. Induction of decision trees. *Machine Learning*, 1(1):81–106, 1986.

REFERENCIAS

- [87] E.B. Hunt, J. Marin, and J.P. Stone. *Experiments in Induction*. Academic Press, 1966.
- [88] L. Breiman, J. Friedman, C.J. Stone, and R.A. Olshen. *Classification and Regression Trees*. The Wadsworth and Brooks-Cole statistics-probability series. Taylor & Francis, 1984.
- [89] J.R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers Inc., San Francisco, 1 edition, 1993.
- [90] A. Shawkat and S.A. Wasimi. *Data mining: Methods and Techniques/A.B.M.* Thomson Victoria, Australia, 1 edition, 2007.
- [91] M. Servente and R. García-Martínez. *Algoritmos TDIDT Aplicados a la Minería de Datos Inteligente*. PhD thesis, Universidad de Buenos Aires, 2002.
- [92] S.M. Weiss and C.A. Kulikowski. How to estimate the True Performance of the Learnign System. In *Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning and Expert Systems (Machine Learning Series)*, chapter 2, pages 17–49. Morgan Kaufmann, 1 edition, 1991.
- [93] B. Efron. Estimating the error rate of a prediction rule: Improvement on cross-validation. *Journal of the American Statistical Association*, 78(382):316–331, 1983.
- [94] I.H. Witten, E. Frank, and M. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, Burlington, USA, 3 edition, 2011.
- [95] M. Baroni, S. Clementi, G. Cruciani, G. Costantino, D. Riganelli, and E. Oberrauch. Predictive ability of regression models. part ii: Selection of the best predictive pls model. *Journal of Chemometrics*, 6(6):347–356, 1992.

- [96] J. Yerushalmy. Statistical Problems in Assessing Methods of Medical Diagnosis, with Special Reference to X-ray Techniques. *Public Health Reports*, 62(40):1432–49, 1947.
- [97] T. Fawcett. An Introduction to ROC Analysis. *Pattern Recognition Letters*, 27(8):861–874, jun 2006.
- [98] D. Faraggi and B. Reiser. Estimation of the Area Under the ROC Curve. *Statistic in Medicine*, 21(20):3093–106, 2002.
- [99] M H Zweig and G Campbell. Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, 39(4):561–577, 1993.
- [100] L.C. Baumann and A. Karel. *Encyclopedia of Behavioral Medicine*, chapter Electronic Health Record, pages 669–670. Springer New York, New York, NY, 2013.
- [101] R.J. A. Little and D.B. Rubin. *Statistical Analysis with Missing Data*. Wiley-Interscience, 2 edition, 2002.
- [102] D.B. Rubin. Multiple imputations in sample surveys a phenomenological bayesian approach to nonresponse. In *Proceedings of the Survey Research Methods Section of the American Statistical Association*, volume 1, pages 20–34. American Statistical Association, 1978.
- [103] J. Han and Jian. P. Kamber, M. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 3 edition, 2012.
- [104] P. Langley and S. Sage. Oblivious decision trees and abstract cases. In *Working Notes of the AAAI-94 Workshop on Case-based Reasoning*, pages 113–117. Seattle, WA, 1994.
- [105] D.W Aha, D. Kibler, and M.K. Albert. Instance-based learning algorithms. *Machine Learning*, 6(1):37–66, 1991.
- [106] P. Langley and S. Sage. Induction of selective bayesian classifiers. In *Proceedings of the Tenth International Conference on Uncertainty in*

REFERENCIAS

- Artificial Intelligence*, pages 399–406. Morgan Kaufmann Publishers Inc., 1994.
- [107] R. Kohavi and G.H. John. Wrappers for feature subset selection. *Artificial Intelligence*, 97(1):273–324, 1997.
- [108] L. Puente-Maury, A. López-Chau, W. Cruz-Santos, and L. López-García. Método rápido de preprocesamiento para clasificación en conjuntos de datos no balanceados. *Research in Computing Science*, 73:129–142, 2014.
- [109] J. Van Hulse and T. Khoshgoftaar. Knowledge discovery from imbalanced and noisy data. *Data & Knowledge Engineering*, 68(12):1513–1542, 2009.
- [110] N.V. Chawla, K.W. Bowyer, L.O. Hall, and W. Kegelmeyer. Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002.
- [111] B. Efron and R. Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, pages 54–75, 1986.
- [112] X. Fan. Using commonly available software for bootstrapping in both substantive and measurement analyses. *Educational and Psychological Measurement*, 63(1):24–50, 2003.
- [113] R. Ledesma. Introducción al Bootstrap. Desarrollo de un ejemplo acompañado de software de aplicación. *Tutorials in Quantitative Methods for Psychology*, 4(2):51 – 60, 2009.
- [114] K.J. Jenkins. Risk Adjustment for Congenital Heart Surgery: the RACHS-1 Method. In *Seminars in Thoracic and Cardiovascular Surgery: Pediatric Cardiac Surgery Annual*, volume 7, pages 180–184. Elsevier, 2004.

Acrónimos

- AAA: Aneurisma de Aorta Abdominal. (*Abdominal Aortic Aneurysm*).
- AD: Árboles de decisión. (*Decision tree*).
- BAN: Clasificador Bayesiano Aumentado con una Red. (*Augmmented Naïve Bayes*).
- BD: Bases de datos. (*Databases*).
- BP: Retropropagación. (*Backpropagation*).
- CC: Cardiopatías Congénitas. (*Congenital Heart Diseases*).
- CDSS: Sistemas de apoyo a la toma de decisiones médicas. (*Clinical decision support systems*).
- HC: Historia Clínica. (*Medical Record*).
- IA: Inteligencia Artificial. (*Artificial Intelligence*).
- K-NN: K-vecinos más cercano. (*K-Nearest Neighbour*).
- MLP: Perceptrón Multicapa (*Multilayer perceptron*).
- RB: Redes Bayesianas. (*Bayesian Networks*).
- RBF: Redes de Función de Base Radial. (*Radial Basis Fucntion*).
- RNA: Redes Neuronales Artificiales. (*Artificial Neural Networks*).
- ROC: Característica Operativa del Receptor. (*Receiver Operating Charac*).
- SOM: Mapas Autoorganizados. (*Self-Organizing Maps*).
- SVMs: Máquinas de Vectores de Soporte. (*Support Vector Machines*).
- TAN: Clasificador Bayesiano Aumentado con un Árbol. (*Tree Augmented Naïve Bayes*).
- UCI: Unidad de Cuidado Intensivo. (*Intensive Care Unit*).