UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MECÁNICA

CLASIFICACIÓN DEL ESTADO DE RUPTURA DE ANEURISMAS CEREBRALES BASADA EN LA CARACTERIZACIÓN MORFOLÓGICA Y HEMODINÁMICA MEDIANTE MACHINE LEARNING

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MECÁNICO

VALENTÍN ALEJANDRO RODRÍGUEZ BUSTOS

PROFESOR GUÍA:
ÁLVARO VALENCIA MUSALEM

MIEMBROS DE LA COMISIÓN:
ENRIQUE LÓPEZ DROGUETT
FRANCISCO MERY MUÑOZ

SANTIAGO DE CHILE
2020

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERO CIVIL MECÁNICO
POR: VALENTÍN ALEJANDRO RODRÍGUEZ BUSTOS
FECHA: 2020
PROF. GUÍA: ÁLVARO VALENCIA MUSALEM

CLASIFICACIÓN DEL ESTADO DE RUPTURA DE ANEURISMAS CEREBRALES BASADA EN LA CARACTERIZACIÓN MORFOLÓGICA Y HEMODINÁMICA MEDIANTE MACHINE LEARNING

Los aneurismas cerebrales son dilataciones patológicas localizadas de la pared arterial cerebrovascular. Su ruptura puede causar daño permanente e incluso la muerte. La evolución hasta la eventual ruptura es un proceso complejo y poco predecible. A su vez, su intervención conlleva un riesgo para el paciente. Dado esto, es clave reconocer a tiempo aquellos con mayor riesgo de ruptura. Actualmente esto se hace mediante factores de riesgo. Sin embargo, debido a su naturaleza impredecible, no es posible determinar con certeza absoluta su desenlace, ya sea que se intervenga o no. Es por ello que constituyen un desafío importante en medicina.

En este trabajo se generan modelos de machine learning para clasificar el estado de ruptura de aneurismas cerebrales. Para ello se utilizan datos de 9 parámetros morfológicos y 6 parámetros hemodinámicos, junto con el estado de ruptura, como caracterización de 71 aneurismas cerebrales, teniendo 36 no rotos y 35 rotos. Los 15 atributos morfológicos y hemodinámicos se obtienen a partir de geometrías CAD y simulaciones CFD respectivamente.

Para desarrollar los modelos de clasificación binaria se usan 8 algoritmos de machine learning en paralelo. La clasificación se evalúa por validación cruzada con 10 particiones y con la exactitud como métrica. Inicialmente se evalúa el desempeño en 9 conjuntos que incorporan o seleccionan distintos atributos. Luego se prueban 8 transformaciones de datos aplicadas al conjunto de los 15 atributos. A continuación se realiza una búsqueda exhaustiva en las 8 transformaciones y los 15 atributos, para hallar los subconjuntos que maximizan la exactitud. Después se analizan los subconjuntos asociados a los modelos que logran una exactitud mínima de 80 % con reglas de asociación. Luego se hace una optimización de hiperparámetros de los mejores modelos por algoritmo. Finalmente se evalúan estos modelos mediante validación cruzada usual y estratificada con más métricas.

Los resultados validan la selección de atributos según significancia estadística e incorporación de otros atributos como la ubicación y la multiplicidad. El desempeño mejora al usar transformaciones, destacándose la estandarización y exceptuándose la normalización. La búsqueda exhaustiva eleva la exactitud, logrando el máximo con Gradient Boosting y las transformaciones cuantil normal y uniforme. Los subconjuntos hallados sugieren el uso combinado de atributos morfológicos y hemodinámicos. Las reglas de asociación resaltan el valor individual del ángulo de flujo y el uso de ciertos atributos hemodinámicos en conjunto con otros atributos morfológicos y hemodinámicos. Vía optimización de hiperparámetros no se logra una mejora. La mayor exactitud alcanzada es de 88,8 % ± 8,5 % en validación cruzada. Se logra una sensibilidad máxima de 85,0 % ± 15,3 % y una AUC de la curva ROC máxima de 90,2 % ± 16,1 % en la variante estratificada. Esto evidencia el potencial de la caracterización morfológica y hemodinámica para discriminar el estado de ruptura vía machine learning.

A mi familia, Eduardo, Luz, y Emilio.

Agradecimientos

En primer lugar agradezco a Dios, por su amor infinito y por acompañarme y sostenerme en cada momento de mi vida. Agradezco su guía, paciencia, enseñanzas, y misericordia. Le doy gracias por sacrificarse por nosotros y porque pronto volverá para que estemos con Él por siempre. Agradezco por este trabajo, que es fruto de su generosidad. Gracias por amarme a pesar de mí. Eres la causa de que la vida tenga sentido y merezca ser vivida, y el autor de los más hermosos y mejores momentos de mi existencia.

Agradezco a mis papás, Eduardo y Luz, por amarme y entregarme tanto. Son una bendición inmerecida. Agradezco a mi hermano Emilio, por ser alguien invaluable en mi vida. Gracias por tu amor, los buenos ratos, escucharme y ser un refugio. Gracias a los tres por lo que hemos vivido, y permitirme conocer a Dios y vivir a su amparo como familia.

Gracias a mis tatas, tías y tíos, primas y primos, por su apoyo y cariño siempre presente. En especial agradezco a mi tía Marisa por su amor, entrega y apoyo fundamentales en este proceso, y a mis primas Javi, Flo, y Cami.

Doy gracias a mis amigos de la agrupación Adventistas U.Chile y del MUA Santiago, por entregarme tanto y sobre todo, por compartir juntos el amor de Dios. Agradezco a los amigos del Doble Cuarteto Unidos en Cristo, por dejarme ser parte de este ministerio y atesorar hermosas experiencias junto al Señor. Asimismo, agradezco a los amigos del Conjunto Selah.

Agradezco a los hermanos y amigos de las iglesias Las Condes, Príncipe de Gales, Blest Gana, y especialmente, Ñuñoa. Gracias por ser una bendición y una segunda familia en Santiago. En especial, agradezco a la familia Contreras Mayr, por su gran apoyo y cariño. Gracias también a los hermanos y amigos de mi iglesia La Serena Centro.

Gracias a mis compañeros y amigos de la universidad y carrera, por toda la ayuda recibida y lo bueno y lindo del compartir, por lo cual estoy en gran deuda.

Gracias al profesor Valencia, por su gran disposición como profesor guía y por creer en este trabajo. Gracias también al profesor López y al Dr. Mery, por apoyarme siempre. Agradezco también a mis profesores y a los funcionarios que marcaron mi paso por la universidad. Gracias a Nicolás Amigo, cuyo trabajo y apoyo hicieron posible el presente.

Gracias especialmente a Kevin, Polyn, Daniel, Yanara, y Felipe, y tantos otros que llevo en mi corazón. Gracias a todos quienes estuvieron ahí, siendo parte de esta aventura. Finalmente, gracias Antonela por apoyarme, entregarme tanto, y alegrar mi vida.

Tabla de Contenido

Índice de Tablas

Índice de Ilustraciones

Capítulo 1

Introducción

1.1. Antecedentes Básicos Generales

Los aneurismas cerebrales son dilataciones patológicas localizadas de la pared arterial cerebrovascular debilitada [82]. La prevalencia global de esta patología se estima en 3,2 % en virtud de varios estudios [147]. Existen distintos tipos de aneurismas cerebrales de acuerdo a su morfología, siendo los saculares los más comunes, representando entre el 80 % y 90 % del total [82]. Por esta razón, el término aneurisma cerebral suele referirse a esta clase. Los aneurismas saculares tienen la apariencia de un saco que sobresale de la pared de la arteria [42]. Estas dilataciones patológicas se desarrollan en tres etapas consecutivas: formación, crecimiento y ruptura. No obstante, no todos los aneurismas cerebrales que se forman alcanzan las etapas siguientes. Los mecanismos que gatillan estos procesos son complejos y no se encuentran claramente comprendidos [50]. La mayor parte de los aneurismas cerebrales no rotos son pequeños, con un tamaño menor a 10 [mm] y asintomáticos, por lo cual suelen ser detectados de paso mediante neuroimagenología [113, 144, 82]. En caso de crecer, a partir de cierto tamaño estas lesiones comprimen su entorno cerebral, provocando síntomas [82, 2]. La ruptura de un aneurisma cerebral es una complicación grave. Este fenómeno suele causar una hemorragia subaracnoidea, donde la sangre fluye al espacio subaracnoideo [82]. En torno al 85 % de los casos de esta hemorragia se debe a la ruptura de una de estas lesiones, y su incidencia anual se estima en 7-8 por cada 100.000 personas [144, 77]. La hemorragia subaracnoidea posee una carga significativa de mortalidad y morbilidad [44]. Los casos fatales asociados en los primeros 28 días se estiman entre 30 % y 60 % aproximadamente [51, 130, 55]. Como factores de riesgo para la formación de aneurismas cerebrales se encuentran una edad mayor a 30 años, predisposición familiar, ser mujer, fumar y la hipertensión [147, 115, 82]. Los factores de riesgo para la hemorragia subaracnoidea consideran una edad mayor a 60 años, tamaño grande de aneurisma, crecimiento y carácter sintomático, poseer historial previo del evento, asociaciones familiares, e hipertensión [82, 134, 89, 149, 146]. Típicamente la evaluación clínica del riesgo de ruptura de aneurismas cerebrales se hace en base a lo observable, siendo el tamaño el principal indicador [20]. A su vez, se ha identificado la existencia de una relación entre la hemodinámica y los procesos de formación, crecimiento y ruptura [79].

Las simulaciones fluidodinámicas computacionales, abreviadas como CFD (del inglés *Computational Fluid Dynamics*) permiten resolver numéricamente las ecuaciones de movimiento de los fluidos y así estudiar la física de flujos que son inabordables de forma teórica o experimental, principalmente debido a la complejidad de las geometrías [52, 112]. El desarrollo en crecimiento del CFD ha revolucionado los estudios sobre aneurismas cerebrales durante los últimos 20 años, lo cual explica el aumento de la investigación realizada sobre estas lesiones [88]. En este contexto, las simulaciones CFD de aneurismas cerebrales han permitido ahondar en su hemodinámica, habilitando la identificación, caracterización y cuantificación de su relación con la evolución de la patología. A su vez, el avance tecnológico y computacional ha hecho posible obtener reconstrucciones tridimensionales fidedignas de la geometría de aneurismas cerebrales reales de pacientes específicos, a partir de imagenología médica de alta precisión [84, 61]. Mediante la técnica CFD, se intenta comprender la relación entre la hemodinámica y los procesos de formación, crecimiento, y fundamentalmente ruptura de aneurismas cerebrales. Actualmente se reconoce este avance en la materia, pero la opinión médica es que los parámetros hemodinámicos calculados por CFD todavía carecen de la capacidad predictiva requerida para la práctica clínica. A pesar de ello, no se descarta esta posibilidad y se estima que el aumento de la potencia computacional le otorgará mayor cabida [152].

Por otra parte, durante las últimas décadas ha habido una proliferación enorme de datos en prácticamente todas las esferas de actividad humana, creando la necesidad de herramientas que permitan transformar los datos en conocimiento útil y orientado a la resolución de tareas [132]. En este escenario, aparece el KDD (del inglés *Knowledge Discovery in Databases*), como el proceso global de descubrimiento y extracción de conocimiento útil de los datos [37]. El proceso KDD consta de varias etapas, siendo el núcleo la fase de minería de datos. Esta etapa involucra el descubrimiento de patrones no identificados previamente, lo cual se concreta en un modelo utilizado para la comprensión, análisis y predicción, en virtud de los datos [80].

El machine learning o aprendizaje de máquinas es una subárea de la inteligencia artificial usada para explorar los datos y ajustar a estos modelos comprensibles y utilizables por los usuarios. Responde a la pregunta de cómo construir un programa computacional usando datos históricos para resolver un problema dado y mejorar automáticamente su eficiencia con la experiencia. Dentro del machine learning existen varias categorías, siendo el aprendizaje supervisado una de las más importantes [125]. Este se podría considerar el tipo más fundamental de machine learning. Una forma de definir el aprendizaje supervisado, es como un proceso en que una computadora aprende una función que entrega una salida para cierta entrada, utilizando datos que comprenden los valores de entrada y salida [120]. El objetivo del aprendizaje supervisado es adquirir la habilidad de generalizar, correspondiente a la capacidad de conjeturar y entregar una salida apropiada para una entrada no aprendida. Esto se traduce en poder enfrentar situaciones desconocidas aprendiendo únicamente una fracción del conocimiento, de forma automática. El aprendizaje supervisado se ha usado exitosamente en varios problemas reales, como reconocimiento de voz e imágenes y pronóstico del tiempo, entre otros. [131]. Otra forma de machine learning consiste en las reglas de asociación. Estas buscan identificar combinaciones de cosas que ocurren juntas frecuentemente a partir de los datos. Su aplicación más común es en el análisis de transacciones. Las reglas de asociación poseen ciertas medidas de calidad con el fin de distinguir su valor o importancia [102]. El algoritmo Apriori es un método que encuentra todos los conjuntos de ítems frecuentes y las reglas de asociación de datos dados, de manera eficiente [120].

El problema de clasificación es una tarea común en aprendizaje supervisado. Mediante machine learning, consiste en entregar una serie de instancias o ejemplos etiquetados en clases a un algoritmo que aprende a mapear dichas instancias con su clase respectiva [120]. Las instancias en general son descritas por un vector de valores de distintos atributos [120]. El objetivo es generar un programa capaz de asignar o predecir la clase de una nueva instancia cuya clase se desconoce en virtud de sus propiedades [120]. Según la cantidad de clases, existe la clasificación binaria para dos clases, y en caso de tener más categorías, la clasificación multiclase. Matemáticamente este problema trata sobre minimizar una función de pérdida o de costo, que cuantifica el "costo" de una clasificación incorrecta [120]. El conjunto de datos o instancias típicamente se separa en dos partes principales: conjunto de entrenamiento, donde el algoritmo aprende para generar un modelo que clasifique, y conjunto de prueba, donde el modelo asigna las clases a las instancias tenidas y se compara la predicción con la clase verdadera, para evaluar su desempeño [120]. La evaluación se hace mediante métricas que cuantifican el desempeño del modelo. En clasificación binaria existen varias métricas, siendo la exactitud la más intuitiva y clásica, que mide el porcentaje de clasificaciones correctas respecto del total de clasificaciones hechas [70]. Con el fin de elevar el desempeño del modelo aprendido en clasificación, existen diversas técnicas, como seleccionar los atributos más relevantes para discriminar entre clases, transformar o escalar los datos para ingresarlos al algoritmo, de acuerdo a las propiedades de cada uno, y la elección correcta de los algoritmos que aprenden. Si bien se conoce la utilidad de estos métodos, para elevar el desempeño logrado se debe seguir un proceso iterativo en que se evalúan múltiples aproximaciones y técnicas, cuyo resultado no se puede anticipar con certeza.

1.2. Motivación

Los aneurismas cerebrales aparecen como un desafío clínico importante por sus características. Su formación, crecimiento y ruptura conforman un proceso escasamente comprendido y poco predecible [138]. Por otra parte, representan una amenaza potencialmente fatal o de consecuencias permanentes para las personas [116]. Las lesiones de mayor riesgo pueden conjeturarse de buena manera a partir de los factores de riesgo reconocidos actualmente, sin embargo, dado su carácter impredecible, no es posible identificarlas con certeza absoluta. Ejemplo de esto es la ruptura que ocurre en casos inesperados como aneurismas cerebrales pequeños y muy pequeños [73]. Junto con lo anterior, su tratamiento también es riesgoso y en muchos casos no factible. Siendo así, el tratamiento se justifica únicamente cuando su riesgo asociado no supera al de ruptura. En este escenario, lo ideal sería lograr identificar y tratar únicamente aquellos con mayor riesgo de sufrir ruptura [94, 82]. Las reconstrucciones tridimensionales y las simulaciones CFD de aneurismas cerebrales han posibilitado la generación de datos morfológicos y hemodinámicos para su caracterización [61, 33]. Por otro lado, en muchas áreas, entre las cuales se incluye el diagnóstico y tratamiento médico, las máquinas o programas computacionales que aprenden de los datos, ayudan a orientar la toma de decisiones [13]. Por esta razón, explorar el potencial de desarrollo de herramientas de machine learning que permitan asesorar el diagnóstico y tratamiento de esta patología se vislumbra como un camino interesante y prometedor. Esto contribuiría a profundizar la comprensión de la enfermedad, junto a la mejora continua de los servicios de salud y la calidad de vida de las personas.

1.3. Objetivo General

Desarrollar modelos de machine learning que permitan clasificar el estado de ruptura de aneurismas cerebrales basados en la morfología y la hemodinámica.

1.4. Objetivos Específicos

- Utilizar atributos morfológicos y hemodinámicos para la caracterización de aneurismas cerebrales en un problema de clasificación binaria.
- Aplicar y evaluar distintas aproximaciones y técnicas de minería de datos y machine learning con el fin de alcanzar una exactitud cercana o superior a 90 % en la clasificación.
- Reconocer y validar atributos y relaciones entre estos, que sean relevantes para la clasificación del estado de ruptura de aneurismas cerebrales y para la comprensión de la patología.

1.5. Alcances

- Trabajar con atributos de aneurismas cerebrales obtenidos en trabajos previos.
- Agregar atributos adicionales a los atributos originales, obtenidos pero no utilizados en trabajos previos.
- Utilizar herramientas de minería de datos y machine learning existentes y consolidadas.
- Reportar un modelo final que permita clasificar el estado de ruptura de aneurismas cerebrales con la mayor exactitud lograda.

1.6. Estructura del Informe

En lo que sigue, el Capítulo 2 presenta los antecedentes necesarios acerca de aneurismas cerebrales, simulaciones CFD, proceso KDD, machine learning y clasificación con machine learning. El Capítulo 3 describe todos los datos utilizados en este trabajo. El Capítulo 4 constituye una revisión del estado del arte sobre parámetros predictores del estado de ruptura y la aplicación de machine learning para su predicción, a modo de punto de referencia. El Capítulo 5 explica la metodología utilizada en este trabajo, incluyendo los recursos requeridos. El Capítulo 6 exhibe los resultados obtenidos en el análisis exploratorio de datos y la clasificación mediante machine learning abordada en distintos pasos. El Capítulo 7 presenta el análisis de los resultados. Finalmente, el Capítulo 8 plantea las conclusiones obtenidas a partir del trabajo realizado.

Capítulo 2

Antecedentes

2.1. Aneurismas Cerebrales

Los aneurismas cerebrales son dilataciones patológicas localizadas de la pared arterial cerebrovascular, las que ocurren debido a un debilitamiento de dicha estructura [82]. Dado lo anterior, la pared arterial de estas lesiones se halla propensa a sufrir ruptura [44]. La Figura 2.1 muestra una vista esquemática de un aneurisma cerebral. Los aneurismas cerebrales suelen aparecer con mayor frecuencia en la zona del círculo de Willis, ubicada en la base del cerebro y mostrado esquemáticamente en la Figura 2.2 [32, 38]. Su detección puede ocurrir de varias maneras, ya sea por un síntoma o evento asociado a la patología, como también de forma accidental, a partir de la toma de imágenes cerebrales [151]. A medida que las técnicas de imagenología mejoran, los aneurismas cerebrales no rotos se diagnostican con mayor frecuencia [150]. La hemorragia subaracnoidea es causada por la ruptura de estas lesiones. Constituye la complicación más temida asociada a los aneurismas cerebrales no rotos, ya que posee consecuencias potencialmente fatales o de discapacidad permanente [17].

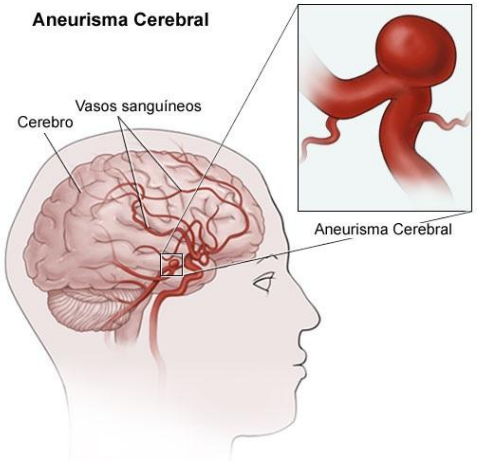


Figura 2.1: Aneurisma cerebral.
Fuente: Adaptada de [97].

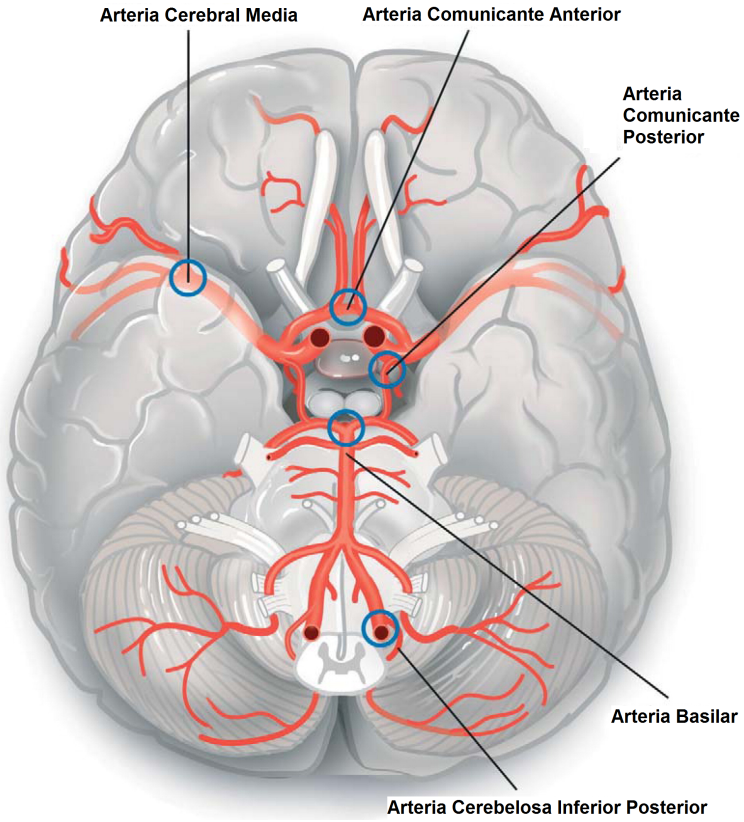


Figura 2.2: Círculo de Willis y arterias.
Fuente: Adaptada de [144].

2.1.1. Epidemiología y Factores de Riesgo

Aneurismas Cerebrales

El porcentaje de individuos que presentan aneurismas cerebrales varía considerablemente según el diseño de los estudios y las poblaciones consideradas, además de las características de cada aneurisma [115]. En virtud de la revisión de varios estudios, la prevalencia de aneurismas cerebrales se estima en 3,2 % de forma global, para una población con edad promedio de 50 años, sin comorbilidad, e igual cantidad de mujeres y hombres [147]. A su vez, dependiendo del estudio, se estiman prevalencias globales desde 0,4 % hasta 6 % [115]. Para determinar la prevalencia de esta patología se requieren estudios de autopsia o investigaciones de poblaciones mediante angiografía [69].

Se estima que la prevalencia de aneurismas cerebrales en mujeres es el doble del valor en hombres [147]. De forma consistente con lo anterior, Imaizumi *et al.* publican un estudio en el año 2018, realizado en un grupo de 4070 adultos japoneses, donde la tasa de detección de aneurismas cerebrales resulta en 6,2 % en mujeres, mientras que en hombres es de 3 %. Por otro lado, la prevalencia es mayor en pacientes con enfermedad renal poliquística autosómica dominante, predisposición familiar, o ateroesclerosis [115]. A su vez, la prevalencia tiende a aumentar con la edad, y es muy baja antes de los 30 años [54, 147]. De acuerdo a un estudio de un grupo de adultos chinos entre 35 y 75 años de edad, la mayor prevalencia de aneurismas cerebrales se da entre las décadas 50 y 60 de una persona [75]. Otros factores de riesgo para el desarrollo de aneurismas cerebrales son la hipertensión y el fumar [82].

Hemorragia Subaracnoidea

Las hemorragias subaracnoideas son causadas por la ruptura de un aneurisma cerebral en el 85 % de los pacientes [144]. La incidencia de la hemorragia subaracnoidea se estima en 7-8 por cada 100.000 personas por año [77]. Según un estudio realizado en Australia y Nueva Zelanda, su tasa de incidencia es mayor en mujeres que en hombres. Asimismo, la incidencia aumenta junto con la edad del paciente, y la edad promedio de casos es de 57 años [6]. Por otra parte, Ingall *et al.* concluyen que la tasa de ataques de la hemorragia subaracnoidea posee alta variación a lo largo de 11 poblaciones de Europa y China, y que no en todas se cumple la noción generalizada de que las mujeres poseen un mayor riesgo de sufrirla [55]. Entre los factores de riesgo de la hemorragia subaracnoidea se encuentran un tamaño grande de aneurisma, historial previo de este tipo de hemorragia, asociaciones familiares, e hipertensión [82, 134, 89]. Wermer *et al.* identificaron una edad mayor a 60 años, el sexo femenino, un tamaño de aneurisma mayor a 5 [mm] y una ubicación en la circulación posterior del cerebro como factores estadísticamente significativos para el riesgo de ruptura. A estos se agrega el carácter sintomático del aneurisma, con el mayor riesgo relativo, y el crecimiento [149, 146].

Los casos fatales por hemorragia subaracnoidea dentro de los primeros 28 días se estiman aproximadamente entre 30 % y 60 %. Felizmente, en las últimas décadas la mortalidad anual por hemorragia subaracnoidea se encuentra en disminución [51, 130, 55]. Un estudio hecho en Dinamarca con 1076 pacientes que sufrieron ruptura de aneurismas cerebrales, luego de 2 años de seguimiento, registra que el 27,5 % de los pacientes está en condiciones normales, el 15,8 % y 9,9 % presenta demencia leve y severa, respectivamente, el 1,3 % resulta vegetativo, y la mortalidad es de 45,5 % [116]. Así, la carga de mortalidad y morbilidad de esta enfermedad es significativa [44]. Por otro lado, el tratamiento de aneurismas cerebrales es riesgoso y en muchos casos no factible, lo cual impele a la investigación de su historia natural elusiva con el fin de identificar y tratar aquellos con mayor riesgo de ruptura [94, 82].

2.1.2. Clasificación y Morfología

Los aneurismas cerebrales se clasifican bajo diversos criterios, como etiología, morfología, tamaño y ubicación. Respecto a la etiología, existen aneurismas cerebrales micóticos, originados por traumatismos, de desarrollo, degenerativos, por flujo sanguíneo, ateroescleróticos, entre otros. Morfológicamente, los aneurismas cerebrales se clasifican en saculares, fusiformes, y disecantes [82]. La Figura 2.3 muestra una representación de los tres tipos de aneurismas cerebrales según morfología. Estos se describen a continuación.

- Aneurismas Cerebrales Saculares

  Los aneurismas cerebrales saculares son el tipo más común, representando entre el 80 % y 90 % de los aneurismas cerebrales [82]. Por esta razón, cuando se habla de aneurismas cerebrales, se suele referir a este tipo en particular[1]. Su apariencia es la de un saco que sobresale de la pared de un vaso sanguíneo. Siendo así, se le llama cuello a la zona en que visualmente el saco aneurismal se une con la pared arterial. Los aneurismas cerebrales saculares se forman en puntos débiles de la pared arterial cerebral, están asociados al crecimiento y ruptura, y son la causa principal de las hemorragias subaracnoideas [42]. La Figura 2.4 muestra la morfología de un aneurisma cerebral sacular.
- Aneurismas Cerebrales Fusiformes

  Los aneurismas cerebrales fusiformes se describen como una dilatación y elongación del lumen o cavidad de las arterias cerebrales [60]. Este tipo de aneurisma representa entre el 3 % y 13 % de los aneurismas cerebrales [3]. Poseen una patología subyacente, hemodinámica, historia natural y tratamiento diferentes a los de tipo sacular [104].
- Aneurismas Cerebrales Disecantes

  Los aneurismas cerebrales disecantes se conocen también como aneurismas falsos, dado que no involucran todas las capas de la pared arterial. Por lo general, abarcan únicamente las capas más externas de esta estructura. Este tipo de aneurisma usualmente es causado por lesiones traumáticas, aunque también puede formarse de manera espontánea [42].

[1]Esto aplica en general para los antecedentes sobre aneurismas cerebrales presentados en este trabajo.

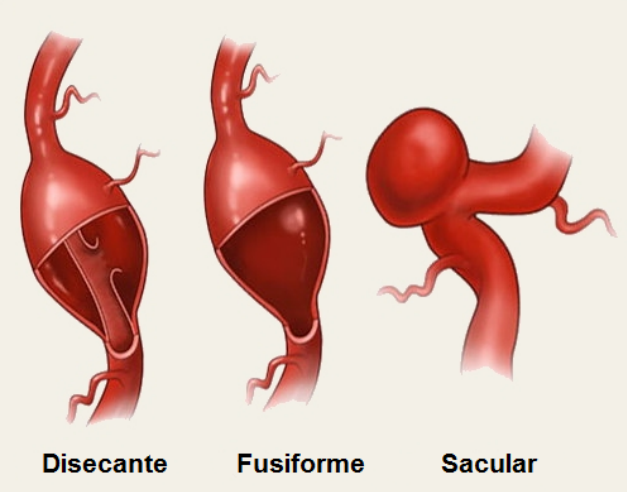


Figura 2.3: Tipos de aneurismas cerebrales según morfología.
Fuente: Adaptada de [8].

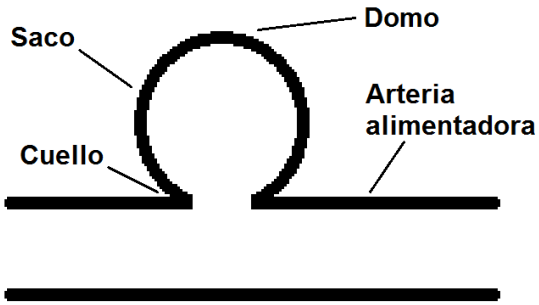


Figura 2.4: Morfología de un aneurisma cerebral sacular.
Fuente: Elaboración propia.

Los aneurismas cerebrales también pueden clasificarse en virtud de su posición respecto de la arteria alimentadora o de origen, como lateral, lateral bifurcación, y terminal [110]. La Figura 2.5 muestra un esquema de los tres tipos de aneurisma. Estas categorías se describen a continuación.

- Lateral

  Este tipo de aneurisma cerebral sobresale por uno de los lados de un segmento recto de la arteria alimentadora.
- Lateral Bifurcación

  Este tipo de aneurisma cerebral se forma en la proximidad de una bifurcación de la arteria alimentadora.
- Terminal

  Este tipo de aneurisma cerebral aparece en un punto de la arteria alimentadora en donde termina otro segmento arterial.

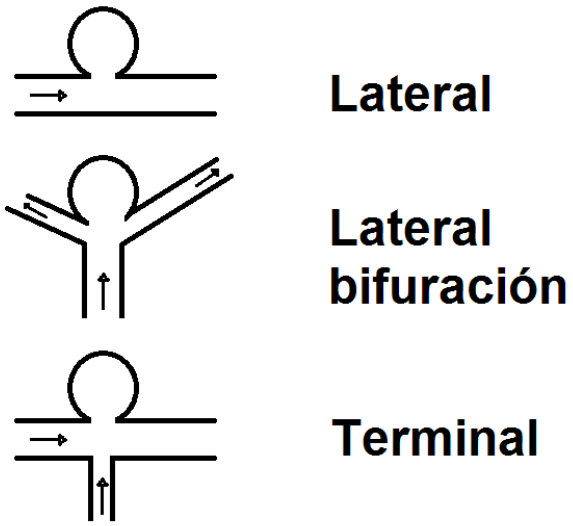


Figura 2.5: Tipos de aneurismas cerebrales según su posición respecto a la arteria alimentadora.
Fuente: Elaboración propia.

El tamaño de los aneurismas cerebrales es otro criterio importante que permite clasificar este tipo de patologías. Las categorías de tamaño para los aneurismas cerebrales no se encuentran del todo unificadas. Una forma conservadora considera las categorías pequeño (<10 [mm]), grande (10-25 [mm]) y gigante (>25 [mm]), en virtud del diámetro máximo del domo del aneurisma [82]. Otras formas de clasificar según esta característica establecen tamaño pequeño (<5 [mm]), mediano (5-10 [mm]), y grande (>10 [mm]) [106]. Junto con esto, el tamaño del cuello del aneurisma cerebral es relevante, el cual se considera grande cuando su diámetro supera los 4 [mm] [82].

Los aneurismas cerebrales aparecen en las ramificaciones arteriales, usualmente en la base del cerebro, ya sea en el círculo de Willis, mayormente en la circulación anterior, o en sitios de ramificación cercanos [144, 2]. Entre el 30 % y 35 % de estas lesiones se desarrolla en la arteria comunicante anterior, un 30 % se forma en la arteria carótida interna y el 20 % en la arteria cerebral media. El 10 % restante se forma en la circulación posterior [122].

Los vasos sanguíneos poseen tres capas en su estructura: íntima, media y externa o adventicia. En los vasos sanguíneos de tamaño medio, la íntima y la media están separadas por la membrana elástica interna, mientras que la media y la externa se hallan separadas por la membrana elástica externa. La capa íntima está compuesta por células endoteliales en el lado del lumen, una capa subendotelial de tejido conectivo, algunas células musculares lisas y por último la membrana elástica interna. La capa media se forma principalmente por células musculares lisas y fibras de elastina. Por su parte, la capa externa se compone mayoritariamente de colágeno [69, 82]. La Figura 2.6 muestra la estructura de la pared arterial de forma esquemática. A diferencia de los vasos sanguíneos de la red vascular sistémica, las arterias cerebrales poseen una capa externa muy delgada, ausencia de la membrana elástica externa, y una densidad reducida de fibras elásticas en la capa media. Estas características únicas de la pared arterial cerebrovascular la hacen más propensa al desarrollo de aneurismas [122].

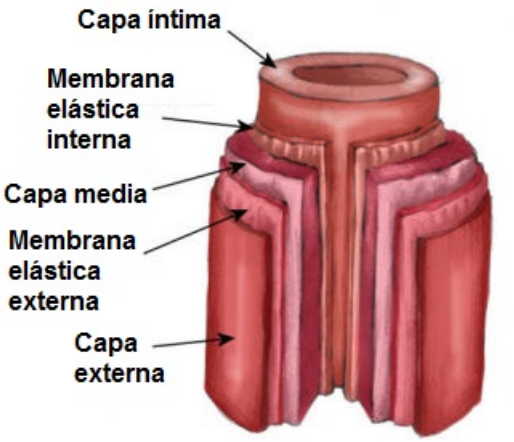


Figura 2.6: Estructura de la pared arterial.
Fuente: Adaptada de [148].

2.1.3. Patogénesis y Hemodinámica

La evolución de los aneurismas cerebrales en el tiempo se lleva a cabo en tres etapas consecutivas, las cuales se señalan a continuación.

1. Formación
2. Crecimiento
3. Ruptura

Luego de su formación, algunos aneurismas cerebrales pueden continuar su evolución hacia las etapas siguientes. Desafortunadamente, a pesar de los avances en el diagnóstico y tratamiento de estas lesiones, no existe una compresión clara sobre los mecanismos de formación, crecimiento y ruptura [50]. La hemodinámica, la biomecánica de la pared arterial, la mecanobiología, y el entorno intracraneal se han identificado como los principales factores involucrados en el proceso de evolución de los aneurismas cerebrales [79]. También hay factores genéticos, hormonales y ambientales que pueden jugar un papel relevante en dicho proceso [29].

Formación

Se cree que la formación de un aneurisma cerebral está relacionada con la interacción entre las fuerzas hemodinámicas de flujos elevados y la pared arterial cerebral [79]. Esto se ilustra de mejor manera al observar que los aneurismas cerebrales aparecen en las uniones arteriales, bifurcaciones, o ángulos abruptos de la red vascular, en donde hay un aumento destacado de los esfuerzos hemodinámicos ejercidos sobre la pared arterial [29]. Este proceso de formación se plantea como un remodelamiento vascular sostenido y localizado. Mientras que el remodelamiento vascular de expansión ocurrido en respuesta a un aumento del flujo sanguíneo representa un proceso adaptativo saludable para regular el esfuerzo de corte en la pared (WSS), la aparición de un aneurisma cerebral es una manifestación patológica de una falla en el intento de mantener la homeostasis bajo una exigencia hemodinámica [50]. El impacto sostenido y localizado de esfuerzos de corte elevados en la pared arterial cerebral ocurrido en segmentos arteriales susceptibles como bifurcaciones o vasos con vasculopatía, genera la activación de las células endoteliales de la capa íntima [82]. Esto se traduce en una disfunción endotelial, la cual se entiende como el cambio de las propiedades del endotelio hacia un fenotipo caracterizado por una vasodilatación alterada y un estado proinflamatorio y protrombótico [129]. Producto de lo anterior, sobreviene una respuesta inflamatoria. De forma simultánea, las células musculares lisas experimentan una modulación fenotípica que las hace evolucionar hacia un estado proinflamatorio [29]. La respuesta inflamatoria total resulta en la degradación patológica de la membrana elástica interna, siendo la estructura responsable en mayor grado de la resistencia mecánica e integridad del vaso sanguíneo. Así, se gatilla una evolución preaneurismal [82]. Junto con esto, se provoca fibrosis, digestión de la matriz extracelular, síntesis anormal de colágeno, y la apoptosis de las células musculares lisas vasculares [29]. Histológicamente, los aneurismas cerebrales no rotos carecen de la capa endotelial, poseen una membrana elástica interna desorganizada, y una capa media disminuida con apoptosis de células musculares lisas. La Figura 2.7 muestra una representación de la pared arterial normal y aneurismal, en donde se observan las diferencias. El incipiente aneurisma cerebral se forma y crece hasta alcanzar un eventual estado de estabilización [82].

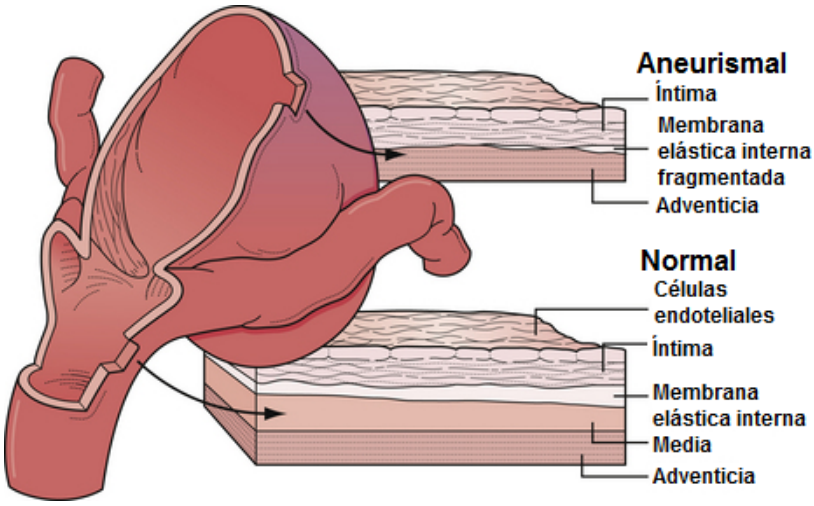


Figura 2.7: Diferencias estructurales de la pared arterial cerebral normal y aneurismal.
Fuente: Adaptada de [35].

Crecimiento y Ruptura

A pesar de que el mecanismo de formación de los aneurismas cerebrales es de común acuerdo en términos generales, no ocurre lo mismo para las etapas siguientes de desarrollo. En el caso de los mecanismos que provocan el crecimiento y la ruptura de estas lesiones en última instancia, existen dos principales corrientes o aproximaciones. Ambas posturas plantean que tanto el crecimiento como la ruptura ocurren debido a un debilitamiento de la pared aneurismal que se genera por la interacción de las células de dicha estructura con el entorno hemodinámico. Por su parte, el crecimiento podría entenderse como un proceso repetitivo en el cual la dilatación cerebrovascular focal cede pasivamente a los efectos de la presión arterial, y luego la pared aneurismal se recupera y engrosa reactivamente, incrementando así el diámetro de la lesión [79]. Cabe decir que la geometría y la hemodinámica del aneurisma cerebral son mutuamente causales. La geometría determina instantáneamente las condiciones de flujo. Por otro lado, el flujo al interior de la dilatación gatilla su remodelamiento y crecimiento, a través de la patobiología, determinando así la geometría futura [83]. El aspecto distintivo entre ambas posturas sobre los procesos de crecimiento y ruptura es el mecanismo responsable del debilitamiento de la pared aneurismal [79]. Luego de 4 décadas, gran parte del interés se centra en el parámetro hemodinámico esfuerzo de corte en la pared, proveniente del análisis realizado mediante simulaciones CFD. Los estudios publicados al respecto reportan hallazgos controversiales y divergentes acerca de la asociación entre el esfuerzo de corte en la pared y la ruptura de un aneurisma cerebral. Así, una postura indica que la ruptura se correlaciona con altos esfuerzos de corte en la pared, mientras que la otra plantea una correlación entre la ruptura y bajos esfuerzos de corte en la pared [158]. Meng *et al.* proponen una teoría unificadora acerca de la controversia entre altos y bajos esfuerzos de corte en la pared. Su artículo plantea que existen dos mecanismos de crecimiento y ruptura en aneurismas cerebrales. En el primer mecanismo, a partir de la expansión de la protuberancia incipiente, el saco aneurismal queda expuesto a menores esfuerzos de corte en la pared. Luego de formarse una zona de recirculación al interior del aneurisma cerebral, este medio queda expuesto a niveles bajos y oscilatorios de esfuerzo de corte en la pared. Este escenario propicia una respuesta inflamatoria del endotelio. Lo anterior resulta en la degradación de la matriz extracelular, lo cual conduce al crecimiento y ruptura. En el segundo mecanismo, luego de formarse la protuberancia incipiente, esta puede permanecer sometida al impacto del flujo sanguíneo. De este modo, el saco aneurismal queda expuesto a esfuerzos de corte en la pared elevados. En este escenario, se genera una respuesta en las células musculares lisas que provoca la degradación de la matriz extracelular, gatillando así el crecimiento y la ruptura [83]. Sobre esta teoría unificadora, Chalouhi *et al.* señalan que los mecanismos propuestos permanecen a nivel especulativo y que se requiere de más investigación con el fin de dar respuesta a la controversia [29]. Finalmente, la ruptura del aneurisma cerebral ocurre cuando la tensión ejercida en la pared aneurismal, producto de la hemodinámica interna, supera su resistencia a la tracción [56].

2.1.4. Presentación Clínica

Aneurismas Cerebrales Rotos

La mayoría de los aneurismas cerebrales permanecen inactivos hasta el momento en que experimentan la ruptura. En la mayoría de los casos, cuando esto ocurre, la sangre fluye hacia el espacio subaracnoideo, aumentando rápidamente la presión intracraneal. Típicamente la ruptura ligada a la hemorragia subaracnoidea se presenta a través de un dolor de cabeza severo y agudo, descrito como "el peor dolor de cabeza de mi vida" por parte de los pacientes [82]. Según J. van Gijn *et al.*, el rasgo característico del dolor de cabeza asociado a la hemorragia subaracnoidea no es la severidad del dolor, sino su comienzo repentino [144]. Este síntoma ha sido denominado "dolor de cabeza centinela" en relación a aneurismas cerebrales no rotos, dado que en ciertos casos precede a la hemorragia subaracnoidea producto de la ruptura de estas lesiones. Sin embargo, también existen casos en que este dolor de cabeza aparece en pacientes que padecen de aneurismas cerebrales no rotos sin evidencia de hemorragia subaracnoidea [31]. Lo anterior se debe a que el "dolor de cabeza centinela" puede ser causado por diversos desórdenes neurológicos [63]. Los resultados de un estudio retrospectivo hecho en 1989 con 109 pacientes que presentaron hemorragia subaracnoidea, indican que las características más comunes son el dolor de cabeza (74 %), las náuseas o el vómito (77 %) y la pérdida de conciencia (53 %). Además, de acuerdo al estudio, el 64 % de los pacientes presenta hallazgos neurológicos y el 35 % sufre rigidez de nuca [41]. Por su parte, los síntomas menos frecuentes asociados a los aneurismas cerebrales rotos son el dolor de cabeza leve, convulsiones, deficiencias neurológicas focales, anormalidades en el electrocardiograma, agitación y confusión [31].

Aneurismas Cerebrales No Rotos

Los aneurismas cerebrales no rotos y asintomáticos son detectados frecuentemente de paso, al realizar procedimientos de neuroimagenología, o screening en individuos en quienes se sospecha su existencia [82]. La mayor parte de los aneurismas cerebrales son asintomáticos y jamás experimentan ruptura. El riesgo de ruptura aumenta con el tamaño, y en torno al 90 % de los aneurismas cerebrales son pequeños, es decir, menores a 10 [mm] [113, 144]. De acuerdo a un estudio, las condiciones más recurrentes que guían hacia el diagnóstico de aneurismas cerebrales no rotos son el dolor de cabeza en el 36 % de los pacientes, enfermedad cerebrovascular isquémica en el 17,6 %, problemas en los nervios craneales en el 15,4 %, y el efecto masa del aneurisma en el 5,7 %, entre otros [101]. El efecto masa ocurre cuando un aneurisma cerebral alcanza su masa crítica, a partir de la cual provoca una compresión de las estructuras nerviosas adyacentes [82]. Sus posibles consecuencias incluyen hemiparesia, defectos del campo visual, convulsiones, y parálisis del tercer par craneal (nervio motor ocular común) [2]. Esto se manifiesta como párpado caído con dilatación de la pupila, intorsión, o depresión leve [82]. Eventualmente, esta compresión podría gatillar un émbolo proveniente del saco del aneurisma, el cual cause un ataque isquémico transitorio o un infarto cerebral por embolización distal [2].

2.1.5. Diagnóstico

Para el diagnóstico de aneurismas cerebrales existen tres técnicas ampliamente utilizadas. La primera consiste en la angiografía convencional o por catéter. En segunda instancia está la angiografía por resonancia magnética, y finalmente la angiografía por tomografía computarizada [123]. Estas técnicas son descritas a continuación.

- Angiografía Convencional

  En la angiografía convencional, se toman imágenes mediante rayos X de los vasos sanguíneos opacados por la inserción de un colorante de contraste. Mediante un catéter insertado en la ingle se inyecta el medio de contraste [100]. Debido a su excelente resolución, permanece como el método a elegir para la detección de aneurismas cerebrales y la determinación de sus características anatómicas. Esta técnica posee un bajo riesgo para el paciente, sin embargo, no es despreciable [123]. La angiografía convencional se establece como el estándar de referencia de los métodos de imagenología para la evaluación de aneurismas cerebrales [82].

- Angiografía por Resonancia Magnética

  La angiografía por resonancia magnética posee la ventaja de basar su funcionamiento en las propiedades magnéticas intrínsecas de los tejidos y la sangre al interior de un campo magnético externo, con el fin de producir una imagen médica [49]. Dado que no requiere de la administración intravascular de material de contraste, esta técnica constituye la forma más conveniente de realizar estudios de diagnóstico y esencialmente no posee ningún riesgo asociado. Si bien ha habido avances, este método no es capaz de detectar ciertos aneurismas cerebrales de tamaño muy pequeño [123].

- Angiografía por Tomografía Computarizada

  La tomografía computarizada es una técnica de imagenología que trabaja mediante rayos X. Durante el proceso, un haz de rayos X se dirige al paciente y rota alrededor del cuerpo, lo cual genera señales que al procesarse en una computadora, generan imágenes de la sección transversal del cuerpo [95]. Recientemente, se ha utilizado la angiografía por tomografía computarizada para la detección de aneurismas cerebrales, con resultados similares a los logrados mediante resonancia magnética [123]. Esta técnica requiere de la inyección de un material de contraste intravenoso para la adquisición de las imágenes. Una de sus ventajas es que guarda la información en un arreglo 3D. Debido a esto, se pueden realizar reconstrucciones tridimensionales que permiten un análisis más versátil [82]. Por su parte, la tomografía computarizada de la cabeza sin contraste es el procedimiento radiológico de elección para establecer el diagnóstico de hemorragia subaracnoidea [82].

2.1.6. Tratamiento

El propósito del tratamiento de aneurismas cerebrales es excluir el saco aneurismal de la circulación cerebrovascular y al mismo tiempo preservar la integridad de la arteria alimentadora [123]. En el caso de aneurismas cerebrales descubiertos de paso e inactivos, se procede a su observación o tratamiento electivo, dependiendo del paciente, el tamaño, y el estado del aneurisma. La observación consiste en el seguimiento mediante la toma de imágenes de rutina de forma periódica [20]. El tratamiento de los aneurismas cerebrales se separa en dos categorías: cirugía y tratamiento endovascular, los cuales se describen a continuación.

Cirugía

La intervención quirúrgica en la red cerebrovascular conlleva riesgos considerables por tratarse de una zona importante y delicada del cuerpo humano. Ejemplo de esto es el provocar daño a otros vasos sanguíneos del cerebro durante la operación [96]. A continuación se describe el principal tipo de cirugía practicada para el tratamiento de aneurismas cerebrales.

- Ligadura Microvascular (Clipping)

  La ligadura microvascular requiere el acceso al aneurisma cerebral por medio de una craneotomía. La intervención consiste en colocar un pequeño clip metálico en el cuello del aneurisma para de esta forma aislar la lesión del flujo sanguíneo de la arteria alimentadora [2]. La Figura 2.8 muestra la representación de un aneurisma cerebral con ligadura microvascular. Aún cuando este procedimiento de neurocirugía se encuentra asociado a mayor mortalidad y morbilidad, posee una baja tasa de reaparición y resangrado [82]. De acuerdo a un estudio de Kotowski *et al.*, a partir de la revisión de los estudios publicados entre los años 1990 y 2011, se encuentra que el clipping de aneurismas cerebrales no rotos exhibe globalmente una mortalidad de 1,7 % y una morbilidad de 6,7 % [67].

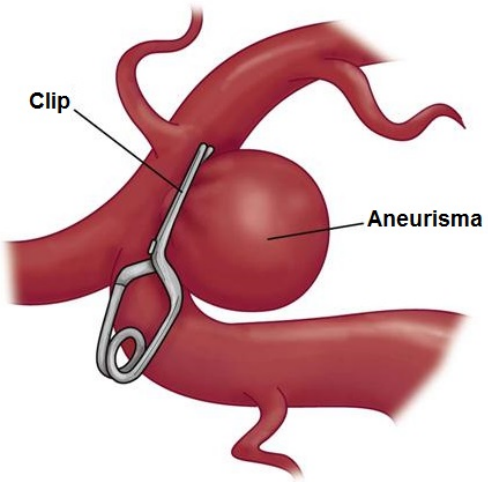


Figura 2.8: Aneurisma cerebral con ligadura endovascular.
Fuente: Adaptada de [140].

Tratamiento Endovascular

Existen dos tipos principales de tratamiento endovascular para los aneurismas cerebrales. A continuación se describen ambos.

- Embolización con Espiral Desmontable (Coiling)

  La forma más común de tratamiento endovascular consiste en la colocación de un espiral desmontable al interior del aneurisma cerebral por medio de un microcatéter. Estos espirales provocan trombosis local y aislamiento del flujo sanguíneo de la arteria alimentadora en la lesión cerebrovascular [2]. La Figura 2.9 muestra un esquema de la embolización con espiral en un aneurisma cerebral. Un estudio en donde se incluyen 188 aneurismas cerebrales no rotos tratados con coiling, con un seguimiento de más de 10 años en el 81 % de los casos, entrega los siguientes resultados: una tasa de ruptura anual de 0,09 %, tratamiento adicional en el 4,8 % de los casos, y 17 pacientes muertos en la cohorte. Por lo anterior, el estudio sugiere que los aneurismas cerebrales con embolización por espiral presentan un bajo riesgo de ruptura para plazos de seguimiento de hasta 20 años [68]. Por otro lado, Fennell *et al.* encuentran en un estudio retrospectivo de una gran base de datos generada entre los años 2009 y 2013, que para 12.400 casos de embolización electiva de aneurismas cerebrales no rotos, la tasa de morbilidad (al menos una complicación) es de 6,4 % y la tasa de mortalidad es de 1,6 % [39].

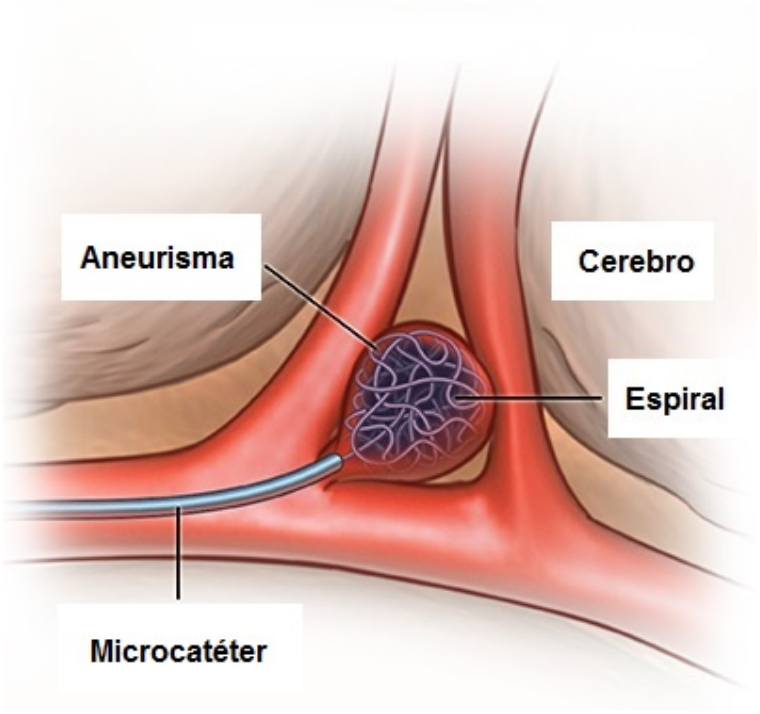


Figura 2.9: Aneurisma cerebral con embolización mediante espiral.
Fuente: Adaptada de [99].

- Dispositivos Desviadores de Flujo

  Los dispositivos desviadores de flujo permiten tratar aneurismas cerebrales de gran tamaño y de cuello extenso, los cuales son candidatos deficientes para otros tipos de tratamiento. Estos dispositivos consisten en stents de malla cilíndricos y metálicos que se posicionan en la arteria alimentadora, extendiéndose a lo largo de la zona del cuello. Así, desvían el flujo de sangre para evitar que se dirija hacia el domo del aneurisma. Yan *et al.* indican en su artículo de revisión y meta-análisis de 26 estudios que el éxito del procedimiento de dispositivos desviadores de flujo es de 96 %. Por su parte, la tasa de oclusión total al final del seguimiento es de 70 %, junto con una tasa de morbilidad global de 20 %. La morbilidad y mortalidad del procedimiento mismo es de 9 % y 4 % respectivamente. Por último, la tasa global de buenos resultados a largo plazo es de 96 % [154].

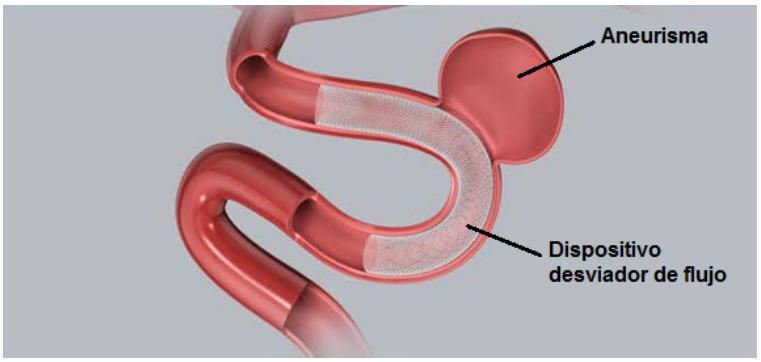


Figura 2.10: Aneurisma cerebral con dispositivo desviador de flujo.
Fuente: Adaptada de [98].

2.2. Simulaciones CFD

La dinámica de fluidos computacional, abreviada como CFD (del inglés *Computational Fluid Dynamics*) es la ciencia de rápida evolución que se encarga de resolver numéricamente las ecuaciones de movimiento de los fluidos para producir predicciones cuantitativas y análisis del fenómeno de flujo. El CFD resulta ideal para llevar a cabo estudios paramétricos o investigaciones acerca de la física del flujo, que serían poco prácticos o imposibles de realizar mediante desarrollos puramente teóricos o experimentales [52]. Los problemas modernos de mecánica de fluidos resultarían inabordables de no ser por el uso de CFD. El espectro de soluciones analíticas para ciertas ecuaciones fundamentales de mecánica de fluidos es muy limitado. Lo usual es que al encontrarse con geometrías de mayor complejidad en esta área de estudio, se debe hacer uso de algún método numérico para la obtención de una solución [112]. Existen dos aproximaciones dominantes en CFD: la formulación de diferencias finitas o FDM (del inglés *Finite Difference Method*) y la de volúmenes o elementos finitos, también conocida como FVM (del inglés *Finite Volume Method*). En ambas formulaciones, el CFD involucra discretizar el dominio espacial en una malla de puntos o elementos y proceder con la resolución numérica en pasos de tiempo discretos [52]. A continuación se entrega una breve descripción de ambos métodos.

- Diferencias Finitas

  En esta formulación, los términos individuales de derivadas en las ecuaciones de movimiento se escriben o aproximan como diferencias de valores del campo determinadas en o entre la malla de puntos, y el sistema de ecuaciones algebraicas resultante se resuelve numéricamente [52].
- Volúmenes Finitos

  En esta formulación, las ecuaciones de movimiento son resueltas dentro de pequeños elementos que en conjunto abarcan el dominio espacial de interés con las condiciones de compatibilidad entre elementos, lo cual genera un sistema de ecuaciones algebraicas que se resuelve numéricamente [52].

2.2.1. Física y Ecuaciones Gobernantes

El CFD encuentra su base en las tres ecuaciones fundamentales que gobiernan la dinámica de fluidos. La primera es la ecuación de continuidad y representa el principio de conservación de masa. La segunda corresponde a la ecuación de momentum, que representa la segunda ley de Newton. En tercer lugar se encuentra la ecuación de energía, que representa el principio de conservación de energía [7]. A continuación se muestran las tres ecuaciones señaladas, obtenidas a partir de la literatura [59].

Ecuación de Continuidad

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) = 0 \tag{2.1}$$

Ecuación de Momentum

$$\frac{\partial \rho \vec{u}}{\partial t} + \nabla \cdot (\rho \vec{u} \vec{u}) = -\nabla p + \nabla \cdot \tau + \vec{f} \tag{2.2}$$

Ecuación de Energía

$$\frac{\partial \rho \mathrm{e}_t}{\partial t} + \nabla \cdot (\rho \vec{u} \mathrm{e}_t) = \nabla \cdot k \nabla T - \nabla \cdot p\vec{u} + \nabla \cdot (\tau \vec{u}) + \vec{f} \cdot \vec{u} \tag{2.3}$$

donde,

ρ: Densidad.

$\vec{u}$: Campo de velocidad.

p: Presión.

τ: Tensor de esfuerzos viscosos.

$\vec{f}$: Fuerzas de cuerpo.

$\mathrm{e}_t = \mathrm{e} + \frac{1}{2}\vec{u} \cdot \vec{u}$

e: Energía interna específica.

k: Conductividad térmica.

T: Temperatura.

2.2.2. Aneurismas Cerebrales

El desarrollo que ha experimentado el CFD, como resultado de las mejoras de software computacional, siendo aplicado a la biomecánica, ha revolucionado los estudios sobre aneurismas cerebrales durante los últimos 20 años. Así, la cantidad creciente de investigación sobre estas lesiones cerebrovasculares se explica por el uso de CFD [88]. Esta técnica ha permitido obtener una descripción tanto cualitativa como cuantitativa acerca de la hemodinámica relativa a los aneurismas cerebrales, lo cual se traduce en resultados como campo de velocidad, presión en el volumen, líneas de corriente, esfuerzo de corte en la pared, entre los principales. Dentro de los diversos factores que determinan la credibilidad de los resultados obtenidos mediante CFD, las condiciones de borde aparecen como las más relevantes [138]. Dependiendo de las fórmulas y supuestos utilizados, las simulaciones se ajustan en mayor o menor grado a la realidad biológica [88]. A pesar de que se reconoce el avance en esta área gracias al uso de CFD, la visión médica plantea que los parámetros hemodinámicos calculados a través de CFD carecen del valor predictivo requerido para la práctica clínica, si bien no se descarta dicha posibilidad. A su vez, señala que el CFD encontrará mayor cabida en la práctica clínica a medida que aumente la potencia computacional [152].

El estudio CFD abarca las ecuaciones de continuidad y momentum. En el caso del flujo sanguíneo, bajo el supuesto de incompresibilidad y despreciando las fuerzas de cuerpo, ambas ecuaciones se simplifican. Las ecuaciones de continuidad y momentum simplificadas, junto con las condiciones de borde, son resueltas numéricamente mediante CFD y se muestran a continuación [4].

$$\nabla \cdot \vec{u} = 0 \tag{2.4}$$

$$\rho \left(\frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} \right) = -\nabla p + \nabla \cdot \tau \tag{2.5}$$

A su vez, para un fluido incompresible, el tensor de esfuerzos viscosos se puede relacionar con la viscosidad del fluido por medio de la siguiente expresión.

$$\tau = \mu \dot{\gamma} \tag{2.6}$$

donde,

μ: Viscosidad dinámica.

$\dot{\gamma}$: Tasa de deformación.

Existe una serie de variables y condiciones de borde a considerar en la aplicación de las simulaciones CFD al estudio de los aneurismas cerebrales. Un punto de partida es la geometría utilizada, la cual posee dos variantes básicas. La primera consiste en el uso de geometrías artificiales o modelos idealizados, los cuales son construidos con software CAD (del inglés *Computer-Aided Design*), que aproximen razonablemente la morfología de un aneurisma cerebral real [124, 53]. De acuerdo a Imai *et al.*, el estudio paramétrico basado en la utilización de modelos idealizados, puede ilustrar de forma clara la relación entre las características geométricas y las propiedades del flujo entrante, mediante CFD [53]. Por otro lado, se encuentra la variante de utilizar modelos específicos de pacientes para la geometría de estas lesiones [25, 90]. Hoy por hoy, debido al desarrollo de técnicas de imagenología médica de alta precisión, la geometría y la estructura de los vasos sanguíneos y eventuales aneurismas formados pueden ser determinadas con un alto grado de exactitud [84]. De esta forma, la información proveniente de angiografías puede ser convertida a una geometría tridimensional para realizar simulaciones computacionales con una representación morfológica fidedigna [61]. Inclusive, se ha llegado a suprimir virtualmente los aneurismas cerebrales en geometrías provenientes de pacientes reales, con el fin de estudiar la hemodinámica de la arteria alimentadora en zonas donde se sabe de la formación de estas lesiones [81].

El flujo sanguíneo de entrada es otro aspecto a considerar en la modelación CFD de aneurismas cerebrales. Análogo al caso de la geometría, esta condición puede ser formulada como un flujo entrante generalizado, o por el contrario, como condiciones específicas de pacientes, las cuales también se obtienen mediante imagenología médica [58]. Específicamente, esta condición de borde de entrada corresponde a la velocidad promedio histórica del ciclo cardíaco. En el caso generalizado, se utilizan aproximaciones artificiales o señales previamente medidas en otros pacientes [138]. Los hallazgos de Jansen *et al.* indican que existen diferencias en los resultados obtenidos sobre las características hemodinámicas y la magnitud del esfuerzo de corte en la pared, al comparar ambas aproximaciones. Dichos autores enfatizan la necesidad de condiciones específicas de pacientes para la condición de entrada de flujo en los aneurismas cerebrales simulados mediante CFD [58]. Si bien esta práctica es la ideal, las señales específicas de pacientes de la velocidad de entrada pueden no estar disponibles en muchos casos [26]. Otro criterio a considerar sobre el modelo del flujo sanguíneo a la entrada, guarda relación con la magnitud de la velocidad promedio en el tiempo. El flujo al interior de las arterias es pulsátil, con una velocidad que varía en función del ciclo cardíaco. Nuevamente, existen dos modelos al respecto. El primero es un flujo estacionario con una velocidad constante a la entrada, que no representa la naturaleza pulsátil del fenómeno. El segundo y más realista pero a la vez de un costo computacional significativamente mayor, considera un flujo transiente, con una velocidad variable periódica representando el flujo pulsátil. Ambas formas se han utilizado para la realización de simulaciones CFD de aneurismas cerebrales [90, 84]. Geers *et al.* concluyen que la condición de flujo estacionario aproxima con alta exactitud el esfuerzo de corte en la pared promediado en el tiempo, luego de comparar los resultados obtenidos con ambas formulaciones. Junto a esto, destacan que las simulaciones CFD estacionarias de aneurismas cerebrales podrían facilitar su introducción en la práctica clínica debido a su bajo tiempo de cómputo [45].

La condición de borde de salida del flujo sanguíneo suele modelarse en términos de la presión [45, 143, 26]. Existen aproximaciones tales como presión cero y constante, cuyo uso ha sido amplio por su simplicidad, si bien no representa correctamente el fenómeno físico, y modelos de resistencia, impedancia y Windkessel, más complejos y que se ajustan mejor a la física involucrada en el flujo sanguíneo [138, 145]. Por su parte, también es posible implementar una señal periódica de presión, de forma análoga al caso del flujo sanguíneo a la entrada. De igual forma, esta señal podría ser específica de cada paciente o generalizada [141]. Los distintos modelos para la condición de borde de salida señalados anteriormente, pueden gatillar errores elevados en la presión calculada en el volumen del aneurisma cerebral, y leves en parámetros como el esfuerzo de corte en la pared, de acuerdo a la tesis de pregrado de Valdivieso [141]. Mediante el uso de la condición de salida de presión, se han llevado a cabo estudios CFD de aneurismas cerebrales en donde se simulan estados de hipertensión e hipotensión arterial, con el fin de analizar cómo dichas condiciones afectan la hemodinámica [143, 121, 72].

Una tercera condición de borde tiene que ver con la naturaleza de la pared cerebrovascular. En la mayoría de las simulaciones CFD esta se trata como una pared rígida, ignorando su distensibilidad. Lo anterior ocurre porque incluir la característica más realista de la capacidad de distensión de la pared arterial es mucho más costoso en términos computacionales [138]. Un avance en las técnicas de simulación computacional de la física de aneurismas cerebrales han sido las simulaciones de interacción fluido-estructura o FSI (del inglés *Fluid-Structure Interaction*) [142]. Este tipo de simulaciones combinan el CFD con el método de elementos finitos o FEM (del inglés *Finite Element Method*), usando este último para la modelación estructural de la pared cerebrovascular [72]. Existen distintos modelos de materiales para la pared cerebrovascular [85]. En la aproximación FSI, el CFD y el FEM trabajan de forma acoplada, lo cual permite obtener un entendimiento acerca del movimiento de la pared aneurismal. Más específicamente, las simulaciones FSI permiten caracterizar los aneurismas cerebrales en términos del esfuerzo de corte en la pared, el esfuerzo de Von Mises en la estructura de la pared y su deformación [72]. De forma lógica, las simulaciones FSI son de un mayor costo computacional que las simulaciones CFD.

La modelación de las propiedades físicas de la sangre también debe considerarse. Típicamente la sangre se modela como un fluido incompresible [84, 152, 108]. Sobre el régimen, mayoritariamente se aplica la formulación laminar, aún cuando existen estudios con modelos de turbulencia, ya que este fenómeno también aparece en los aneurismas cerebrales [36, 135, 47]. La viscosidad es otra variable que posee dos aproximaciones en la modelación CFD de la sangre: fluido newtoniano y no newtoniano. La sangre no puede ser tratada como un fluido newtoniano en general, sin embargo, bajo muchas de las condiciones fisiológicas normales del flujo sanguíneo, puede ser considerada como tal [156]. La reología del flujo de sangre es un fenómeno complejo y no existe un acuerdo generalizado acerca del modelo para simular sus propiedades viscosas. Entre las variantes de modelos para el comportamiento de la sangre bajo el supuesto no newtoniano, se encuentran el de ley de potencia, Casson, y Carreau, que pueden afectar la hemodinámica simulada mediante CFD [117]. Hay estudios que aplican el esquema newtoniano en las simulaciones CFD [90, 84, 25]. Tanaka *et al.* concluyen que la modelación newtoniana de la sangre puede servir para el análisis de aneurismas cerebrales, al encontrar diferencias muy leves entre los resultados hemodinámicos, utilizando ambas formulaciones de la viscosidad [135]. A su vez, hay trabajos que reportan resultados hemodinámicos con errores considerables entre el uso de los modelos newtoniano y no newtoniano [141, 40].

Las simulaciones CFD de aneurismas cerebrales han llegado incluso a aplicarse al estudio de los tratamientos existentes. Byun *et al.* estudiaron las características hemodinámicas de aneurismas cerebrales utilizando modelos geométricos específicos de pacientes antes y después de una cirugía [25]. Esto aplica también para los tratamientos endovasculares. Por su parte, existen estudios hemodinámicos CFD sobre embolización con espiral [87, 74]. A su vez, se ha estudiado el uso de dispositivos desviadores de flujo mediante esta técnica [103].

A modo de resumen, la Figura 2.11 muestra las etapas requeridas del proceso de simulación CFD de aneurismas cerebrales en términos generales.

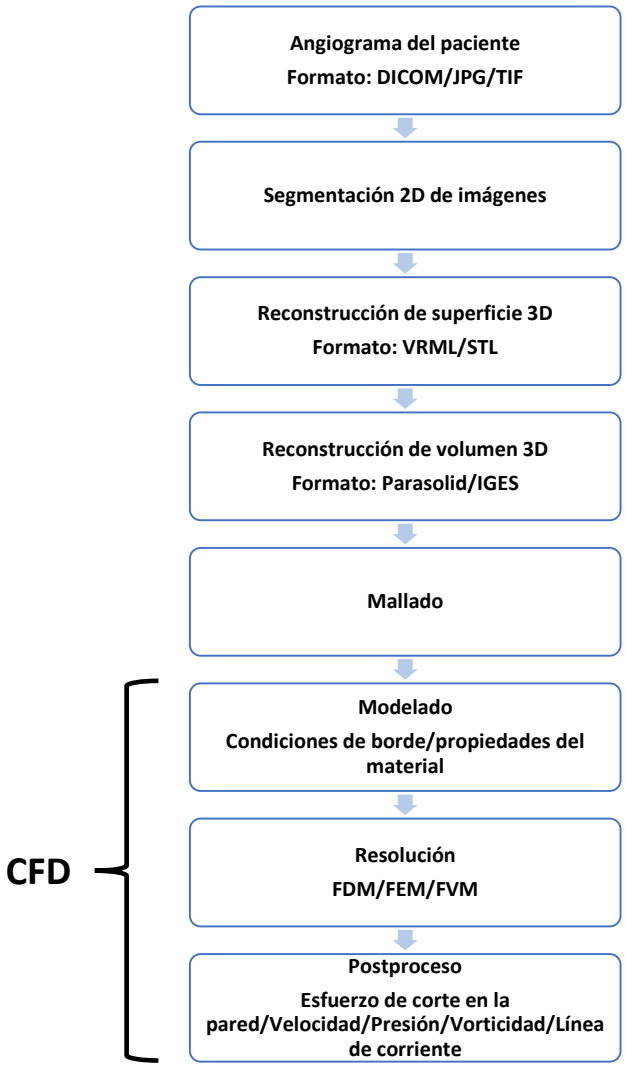


Figura 2.11: Etapas del proceso de simulación CFD de aneurismas cerebrales.
Fuente: Adaptada de [61].

2.3. Proceso KDD

Durante las últimas décadas, ha habido una proliferación enorme de datos y bases de datos en prácticamente todas las esferas de la actividad humana. Esta situación crea la necesidad de contar con herramientas nuevas y potentes que permitan transformar los datos en conocimiento útil y orientado a la resolución de tareas [132]. Dentro del contexto anterior, aparece el concepto de KDD (del inglés *Knowledge Discovery in Databases*), el cual se refiere al proceso global de descubrir y extraer conocimiento útil a partir de los datos. Este representa un proceso interactivo e iterativo, que involucra muchas decisiones por parte del usuario [37]. Otras definiciones de KDD señalan que es el proceso organizado de identificar patrones válidos, novedosos, útiles, y comprensibles en conjuntos de datos extensos y complejos, o que es un análisis exploratorio y modelamiento automático de grandes conjuntos de datos [80]. Así, existe más de una aproximación para la definición de este proceso. A continuación, se detallan las etapas principales del proceso KDD, las cuales se basan en lo planteado por Maimon *et al.* en su manual [80]. La Figura 2.12 muestra un esquema de las etapas básicas del proceso KDD descritas por Fayyad *et al.* en su artículo de 1996 [37].

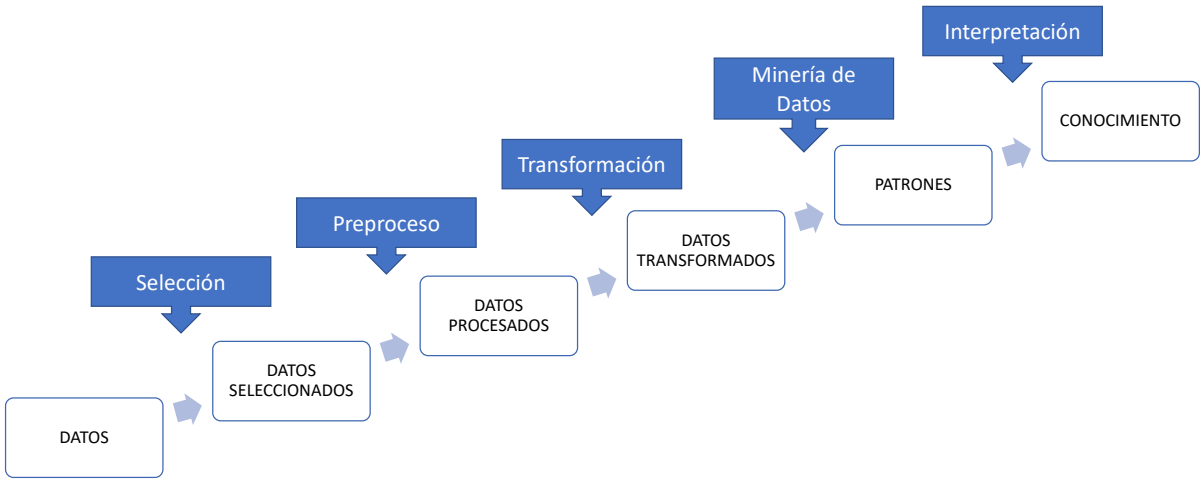


Figura 2.12: Etapas básicas del proceso KDD.
Fuente: Adaptada de [37].

1. Comprensión del Objetivo y Dominio de Aplicación

   Esta fase representa el punto de partida del proceso KDD. Aquí se prepara el escenario para entender cómo abordar las diversas decisiones involucradas en dicho proceso. En esta etapa se requiere definir los objetivos del proceso KDD, así como el entorno en que se lleva a cabo, considerando el conocimiento relevante ya existente. Esta etapa se encuentra sujeta a revisiones y ajustes a medida que el proceso KDD avanza.

2. Selección de Datos

   En esta etapa se busca determinar el conjunto de datos en los cuales se realiza el descubrimiento de conocimiento. Lo anterior involucra el reconocimiento de los datos que se encuentran disponibles y la obtención de datos adicionales. A su vez, la selección de datos abarca definir las instancias o ejemplos y los atributos a considerar. Esta etapa es de gran relevancia, dado que define la materia prima sobre la cual trabajan los algoritmos utilizados en la etapa de minería de datos. Por su parte, la selección de datos es un procedimiento iterativo e interactivo, en virtud de que los datos disponibles pueden variar en función de aspectos como tecnología, investigaciones y costos. Además, los resultados obtenidos en etapas posteriores retroalimentan la selección de datos.

3. Preproceso y Limpieza de Datos

   El objetivo de esta etapa es aumentar la confiabilidad y calidad de los datos seleccionados. Esto implica el manejo de valores faltantes, valores perdidos, y ruido en los datos. Para llevar a cabo estas tareas existen diversos métodos. Eventualmente, esta etapa podría no requerir acción alguna. Por otro lado, en muchos casos representa la etapa de mayor inversión de tiempo.

4. Transformación de Datos

   Esta etapa consiste en la generación de datos mejorados para la fase de minería de datos. Algunas de las operaciones utilizadas contemplan la reducción de dimensionalidad, mediante técnicas como selección y extracción de atributos, y la transformación de atributos, mediante escalado, discretización o funciones. La etapa de transformación de datos en muchos casos es clave y usualmente muy específica en función de cada proyecto. Igualmente, las transformaciones utilizadas se retroalimentan e iteran en función de la minería de datos y las etapas finales del proceso KDD.

5. Minería de Datos

   La minería de datos es el núcleo del proceso KDD, que involucra la inferencia de algoritmos que exploran los datos, desarrollan el modelo, y descubren patrones no identificados previamente. El modelo se utiliza para entender fenómenos a partir de los datos, análisis y predicción.

   - El primer paso consiste en la correcta elección de la tarea específica de minería de datos a realizar. Esto depende en principal medida de los objetivos del proceso KDD y de las etapas anteriores. En términos generales existen dos metas en minería de datos: predicción y descripción. La mayor parte de las técnicas de minería de datos se basan en aprendizaje inductivo, en donde se construye un modelo de forma explícita o implícita generalizando a partir de una cantidad suficiente de instancias o ejemplos de entrenamiento. El supuesto subyacente de la aproximación inductiva es que el modelo entrenado es aplicable a casos futuros.
   - Una vez definida la tarea, se procede con la elección del algoritmo específico de minería de datos a utilizar para la búsqueda de patrones. Esto puede depender de una compensación entre desempeño e interpretabilidad. El meta-aprendizaje se enfoca en explicar qué causa que un algoritmo de minería de datos sea exitoso en un problema particular. Por lo tanto, este acercamiento busca comprender las condiciones bajo las cuales un algoritmo de minería de datos es más apropiado, considerando que cada cual posee distintos parámetros y tácticas de aprendizaje.

- Finalmente se lleva a cabo la aplicación del algoritmo de minería de datos. Esta fase podría requerir la ejecución del algoritmo repetidas veces, hasta lograr un resultado satisfactorio, variando por ejemplo sus hiperparámetros.

6. Evaluación e Interpretación

   En esta etapa se realiza la evaluación e interpretación de los patrones descubiertos en la fase de minería de datos, en relación con los objetivos establecidos en un principio. Aquí se consideran las etapas previas respecto de su efecto en los resultados del algoritmo de minería de datos. El foco de esta etapa se encuentra en la comprensibilidad y utilidad del modelo inducido. Asimismo, considera la documentación del conocimiento descubierto para uso futuro.

7. Aplicación del Conocimiento Descubierto

   En esta etapa final, el conocimiento descubierto se encuentra disponible para incorporarlo en otro sistema, con el fin de ejecutar acciones. Así, este conocimiento se vuelve activo, dado que permite operar cambios en el sistema y medir los efectos. El éxito de esta etapa determina la efectividad de todo el proceso KDD. Además, en esta etapa aparecen muchos desafíos, en particular, la pérdida de las condiciones sobre las cuales se lleva a cabo el trabajo previo. Los datos se vuelven dinámicos, las estructuras pueden cambiar, y el dominio de valores modificarse.

2.4. Machine Learning

El machine learning o aprendizaje de máquinas es una subárea de la inteligencia artificial que ha evolucionado a partir del reconocimiento de patrones, utilizado para explorar la estructura de los datos y ajustar a estos modelos que pueden ser comprendidos y utilizados por los usuarios. Responde a la pregunta de cómo construir un programa computacional usando datos históricos, con el fin de resolver un problema dado, y mejorar de forma automática la eficiencia del programa con la experiencia. Esta disciplina experimenta un crecimiento y desarrollo de forma continua. A grandes rasgos, el machine learning se divide en 4 categorías: aprendizaje supervisado, aprendizaje no supervisado, aprendizaje semi-supervisado, y aprendizaje por refuerzo, entre otras [125]. El tipo de aprendizaje depende de las características de los datos disponibles [131]. En este caso, el interés se concentra mayoritariamente en la primera de estas categorías o tipos de aprendizaje.

2.4.1. Aprendizaje Supervisado

El aprendizaje supervisado sería el tipo más fundamental de machine learning, que considera, a modo de analogía, un estudiante aprendiendo de un supervisor mediante preguntas y respuestas. En el contexto de machine learning, un estudiante corresponde a una computadora y un supervisor corresponde a un usuario de la computadora. Así, la computadora aprende un mapeo de una pregunta a su respuesta a partir de instancias de pares pregunta-respuesta [131]. Otra definición señala que el aprendizaje supervisado se refiere a cualquier proceso de machine learning que aprende una función de una entrada a una salida utilizandos datos, los cuales comprenden ejemplos que contienen tanto los valores de entrada como de salida [120]. El objetivo del aprendizaje es adquirir la habilidad de generalizar, la cual se refiere a la capacidad de conjeturar una respuesta apropiada para una pregunta no aprendida. De este modo, el usuario no necesita enseñar todos los casos o instancias a la computadora, sino que esta puede enfrentar situaciones desconocidas al aprender únicamente una fracción del conocimiento, de forma automatizada. El aprendizaje supervisado ha sido utilizado exitosamente en un amplio espectro de problemas reales, como reconocimiento de escritura manual, voz e imágenes, sistemas de recomendación, pronóstico del tiempo, entre otros [131]. Tradicionalmente, en machine learning el aprendizaje supervisado se divide en dos tipos de tarea dependiendo del carácter continuo o discreto de la respuesta o salida: regresión y clasificación, respectivamente [120].

2.4.2. Reglas de Asociación

Las reglas de asociación buscan identificar combinaciones de cosas que ocurren juntas frecuentemente, conocido como análisis de afinidad. Existen diversos usos de las reglas de asociación. La aplicación más común de esta técnica es el análisis de transacciones, aunque también se ha utilizado las áreas de marketing, negocios, ciencias, medicina, y administración de recursos humanos. En particular, se utiliza en medicina para el diagnóstico de enfermedades. Las reglas de asociación tratan con ítems, que constituyen los objetos de interés. Así, los ítems se agrupan en conjuntos que tienden a ocurrir juntos. Las reglas tienen la forma de un conjunto de ítems a la izquierda, denominado antecedente, con una consecuencia en la derecha. Una de las limitaciones de este análisis es la eventualmente enorme cantidad de combinaciones de ítems posibles [102].

Para poder distinguir entre una regla y otra se requieren algunas medidas de calidad de las reglas. En base a la notación propuesta por Bramer [18], una regla de asociación se escribe como

$$A \Rightarrow C \tag{2.7}$$

donde,

A: Antecedente.

C: Consecuencia.

A continuación se definen 4 variables numéricas que pueden ser determinadas para una regla cualquiera por conteo.

N_A: Cantidad de instancias en que se cumple A.

N_C: Cantidad de instancias en que se cumple C.

$N_{A \cup C}$: Cantidad de instancias en que se cumple tanto A como C.

N_T: Cantidad total de instancias.

Existen 3 medidas para evaluar lo interesante de una regla de asociación, las cuales se definen a continuación.

1. Soporte

   El soporte de una regla es la proporción de instancias en que aparecen tanto el antecedente como la consecuencia.

$$Sop(A \Rightarrow C) = \frac{N_{A \cup C}}{N_T} \tag{2.8}$$

2. Confianza

   La confianza de una regla es la proporción entre las instancias en que aparecen tanto el antecedente como la consecuencia y las instancias en que aparece el antecedente.

$$Conf(A \Rightarrow C) = \frac{Sop(A \Rightarrow C)}{Sop(A)} = \frac{N_{A \cup C}}{N_A} \tag{2.9}$$

3. Interés (Lift)

   El interés de una regla es la tasa entre la proporción de instancias en que aparecen tanto el antecedente como la consecuencia y la proporción en que aparecerían tanto el antecedente como la consecuencia bajo el supuesto de independencia entre ambos [107].

$$Int(A \Rightarrow C) = \frac{Conf(A \Rightarrow C)}{Sop(C)} = \frac{Sop(A \Rightarrow C)}{Sop(A) \cdot Sop(C)} = \frac{\frac{N_{A \cup C}}{N_T}}{\frac{N_A}{N_T} \cdot \frac{N_C}{N_T}} \tag{2.10}$$

   - Si $Int < 1$, A y C aparecen juntos en los datos con menos frecuencia que la esperada bajo el supuesto de independencia, por lo cual se dice que son negativamente interdependientes.
   - Si $Int = 1$, A y C aparecen juntos con la frecuencia esperada bajo el supuesto de independencia, por lo cual se dice que son independientes.
   - Si $Int > 1$, A y C aparecen juntos en los datos con mayor frecuencia que la esperada bajo el supuesto de independencia, por lo cual se dice que son positivamente interdependientes [19].

Algoritmo Apriori

El algoritmo Apriori es un método que entrega todos los conjuntos de ítems frecuentes (según un umbral de frecuencia establecido) y las reglas de asociación de datos dados. El algoritmo encuentra conjuntos de ítems frecuentes mediante una búsqueda en amplitud de lo general a lo específico. Así, genera y prueba conjuntos de ítems en lotes para reducir el costo de acceso a la base de datos. La búsqueda inicia considerando los conjuntos de ítems más generales, es decir, los conjuntos formados por un único ítem (singleton), como candidatos. A continuación, el algoritmo calcula iterativamente las frecuencias de los candidatos y guarda aquellos que son frecuentes. El núcleo del algoritmo se encuentra en la generación de candidatos: en el siguiente nivel (cantidad de ítems por conjunto), los conjuntos de ítems que poseen un subconjunto infrecuente son eliminados, dado que no pueden ser frecuentes. Esto otorga al algoritmo la capacidad de encontrar todos los conjuntos de ítems frecuentes sin la necesidad de invertir mucho tiempo en aquellos que son infrecuentes. Finalmente, el algoritmo Apriori prueba todos las reglas de asociación frecuentes y entrega aquellas que alcanzan el nivel de confianza definido [120].

2.5. Clasificación con Machine Learning

Usualmente se entiende la clasificación como la disposición de cosas en categorías o su agrupación de alguna manera útil. Dentro de la actividad humana, clasificar es algo recurrente, dado que las cosas pertenecientes a un grupo, llamado clase en machine learning, comparten características comunes. Así, se puede saber bastante acerca de un objeto si se conoce su clase. En machine learning, el término clasificación hace referencia a un tipo de aprendizaje supervisado en el cual se le entregan al algoritmo que aprende, una serie de ejemplos o instancias de una o más clases, etiquetadas con su clase respectiva. El algoritmo produce un clasificador que mapea las propiedades de estas instancias a sus etiquetas o clases respectivas [120]. Típicamente las instancias son descritas por los valores de un conjunto finito de atributos, representado por un vector de atributos, en donde cada posición en el vector corresponde a un único atributo [120]. De este modo, el clasificador puede otorgar una clase a un nuevo ejemplo cuya clase se desconoce, en base a sus propiedades [120].

La clasificación puede entenderse como un problema de reconocimiento de patrones de forma supervisada. Las entradas corresponden a vectores $\vec{x}$ de dimensión m (cantidad de atributos) y su clase, representada por un escalar $y \in 1, ..., c$, donde c corresponde a la cantidad de clases. Para entrenar un clasificador, se proveen n instancias de pares entrada-salida $(\vec{x}_i, y_i)_{i=1}^{n}$. Si la regla de clasificación real se denota como la función $y = f(\vec{x})$, la clasificación vía machine learning se puede considerar como un problema de aproximación de una función [131]. El problema de clasificación se denota de distintas formas según la cantidad de clases presentes en el entrenamiento. Las variantes más comunes son la clasificación binaria, donde se tienen 2 clases, y la clasificación multiclase, donde se tienen más de 2 clases. Matemáticamente la clasificación se traduce en un problema de optimización, en donde se busca minimizar una función llamada función de pérdida. Existen diversas funciones de pérdida, teniendo como característica común que su valor de salida disminuye cuando aumentan las clasificaciones correctas. También se le llama función costo, dado que asigna una cantidad numérica que representa el "costo" asociado a una clasificación incorrecta [120].

El conjunto de datos utilizado en un problema de clasificación abordado con machine learning, se separa en dos: conjunto de entrenamiento y conjunto de prueba. El conjunto de entrenamiento corresponde al grupo de instancias que se ingresa a un algoritmo que aprende, el cual lo analiza y se ajusta a este para generar un modelo. Por su parte, el conjunto de prueba está constituido por instancias que se utilizan para evaluar el desempeño del modelo aprendido por el algoritmo con el conjunto de entrenamiento [120]. Más específicamente, luego del ajuste al conjunto de entrenamiento, se genera un modelo que es capaz de entregar una salida o clase cuando se le ingresa una entrada. Seguido a esto, se ingresan al modelo los vectores de atributos del conjunto de prueba, el cual genera una clase para cada uno. Así, el desempeño del modelo aprendido se mide en virtud de la diferencia entre la clase real y generada por el modelo para cada instancia del conjunto de prueba.

Dentro de los problemas en clasificación, se dice que un modelo se sobreajusta en el conjunto de entrenamiento cuando captura atributos que provienen del ruido o la varianza en vez de la distribución subyacente de los datos. El sobreajuste gatilla una disminución de exactitud (ver Subsección 2.5.1) en los datos de prueba, no usados para aprender el modelo [120]. Así, se pierde capacidad de generalizar. La Figura 2.13 muestra una esquema gráfico del sobreajuste.

Figura 2.13: Problemas en el ajuste de modelos de clasificación.
Fuente: Adaptada de [65].

La Figura 2.14 muestra un esquema del conjunto de datos, entrenamiento y prueba, para un problema de clasificación. Típicamente el conjunto de datos se organiza como una tabla, cuyas filas y columnas representan las instancias y atributos, respectivamente, como se observa en la Figura 2.14. Cabe decir que las filas o instancias suelen distribuirse de forma aleatoria entre los conjuntos de entrenamiento y prueba (en la Figura 2.14 aparecen separadas en bloques de filas para facilitar la ilustración). Junto con esto, el conjunto de prueba suele consistir entre el 10 % y 25 % de las instancias totales del conjunto de datos, utilizando el resto para el entrenamiento. Respecto del conjunto de entrenamiento, este suele dividirse a su vez en dos: conjunto de entrenamiento y conjunto de validación. El conjunto de validación actúa como un conjunto de prueba, pero se utiliza para calibrar los hiperparámetros del algoritmo que aprende (ver Subsección 2.5.5). Una vez hecho lo anterior, se evalúa finalmente el desempeño en el conjunto de prueba.

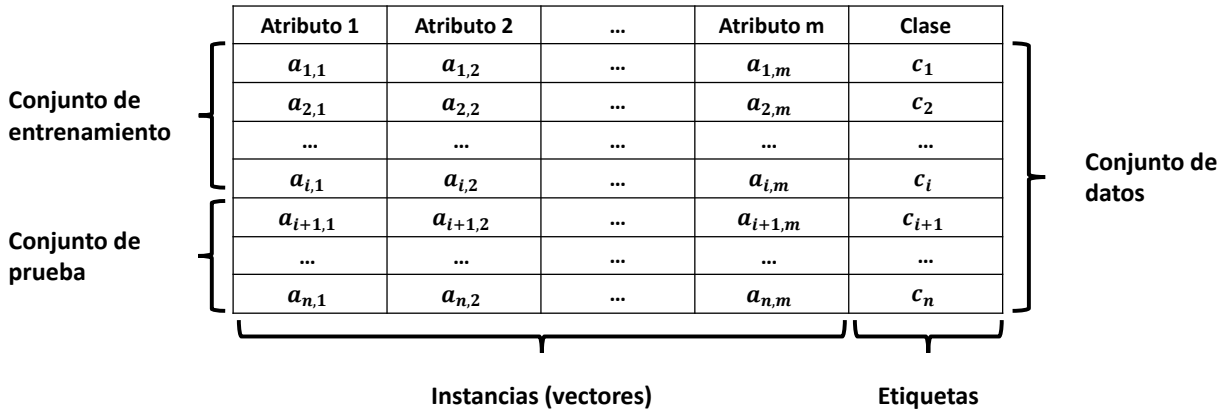


Figura 2.14: Esquema del conjunto de datos, entrenamiento y prueba.
Fuente: Elaboración propia.

2.5.1. Métricas de Evaluación

Las métricas de evaluación utilizadas en un problema de clasificación vía machine learning permiten cuantificar la calidad del modelo logrado. Existen distintas métricas, que pueden representar el desempeño del modelo en menor o mayor grado dependiendo de las condiciones del problema. En clasificación binaria, usualmente se habla de ejemplos positivos y negativos, según las dos clases presentes. A partir de esto, se definen 4 tipos de instancias fundamentales en este tipo de clasificación [70], señaladas a continuación.

- Verdadero Positivo (VP): Instancia etiquetada como positiva y clasificada por el modelo aprendido como positiva.
- Verdadero Negativo (VN): Instancia etiquetada como negativa y clasificada por el modelo aprendido como negativa.
- Falso Positivo (FP): Instancia etiquetada como negativa y clasificada por el modelo aprendido como positiva.
- Falso Negativo (FN): Instancia etiquetada como positiva y clasificada por el modelo aprendido como negativa.

Las instancias anteriores pueden representarse mediante una matriz de confusión, tal como se observa en la Tabla 2.1, donde N_i corresponde a la cantidad de instancias del tipo i.

Tabla 2.1: Matriz de confusión en el caso general.

		Clase Predicha	
		$+$	$-$
Clase Real	$+$	N_{VP}	N_{FN}
	$-$	N_{FP}	N_{VN}

Fuente: Elaboración propia.

A partir de estos tipos de instancias generadas en clasificación binaria, se exponen a continuación las métricas de evaluación del modelo aprendido que aplican en este problema de clasificación específico, en base a lo referido en el libro de Kubat [70].

Exactitud

La exactitud corresponde al porcentaje o fracción de clasificaciones correctas realizadas por el modelo aprendido. En cierto sentido, es la métrica de desempeño más intuitiva y clásica, dado que mide el porcentaje de éxito. Su cálculo viene dado por la siguiente expresión.

$$Ex = \frac{N_{VP} + N_{VN}}{N_{VP} + N_{VN} + N_{FP} + N_{FN}} \tag{2.11}$$

La exactitud es una métrica que conviene utilizar únicamente cuando ambas clases se encuentran balanceadas, es decir, la cantidad de instancias en cada clase es similar. A modo de ejemplo, si se tienen 100 instancias, de las cuales 2 son positivas, un modelo que clasifique todas las instancias como negativas lograría una exactitud de 98 %. A pesar de que el modelo no lograra reconocer ninguna de las instancias positivas, en términos de exactitud, se encontraría cerca del desempeño ideal de 100 %. Lo anterior evidencia que la exactitud es un mal indicador en el caso desbalanceado.

Precisión

La precisión indica el porcentaje de verdaderos positivos dentro de todas las instancias clasificadas como positivas por parte del modelo aprendido. Su cálculo viene dado por la siguiente expresión.

$$Pr = \frac{N_{VP}}{N_{VP} + N_{FP}} \tag{2.12}$$

Esta métrica puede entregar una noción más certera de la calidad del modelo aprendido cuando se tienen clases desbalanceadas.

Sensibilidad

La sensibilidad representa el porcentaje de instancias positivas que el modelo aprendido clasifica como tal. Su cálculo viene dado por la siguiente expresión.

$$Se = \frac{N_{VP}}{N_{VP} + N_{FN}} \tag{2.13}$$

De igual forma, la sensibilidad es un indicador que señala de forma más realista la calidad del modelo aprendido al tener clases desbalanceadas. La sensibilidad es muy importante en el área de diagnóstico médico, dado que el costo de clasificar un paciente enfermo como sano, podría ser mucho mayor que el costo de clasificar un paciente sano como enfermo.

Valor F1

El valor F1 es una métrica que busca combinar la precisión y la sensibilidad en un único valor. De forma general, esta métrica es un caso particular del valor F_β, en donde $\beta \in [0, \infty)$ es un parámetro que permite calibrar la importancia relativa de las dos métricas en cuestión. Cuando $\beta > 1$ se le asigna mayor peso a la sensibilidad, y en caso contrario, el mayor peso se otorga a la precisión. En general, dado un problema de clasificación binaria, no se sabe de antemano si es más relevante la precisión o la sensibilidad, por lo cual se prefiere utilizar el valor neutral $\beta = 1$ con frecuencia. A continuación se muestran las expresiones del caso general y el caso $\beta = 1$.

$$F_\beta = \frac{(\beta^2 + 1) \times Pr \times Se}{\beta^2 \times Pr + Se} \tag{2.14}$$

$$F_1 = \frac{2 \times Pr \times Se}{Pr + Se} \tag{2.15}$$

Matemáticamente, el valor F1 se define como la media armónica de la precisión y la sensibilidad. Siendo así, su valor se ubica entre ambas métricas, acercándose más al valor menor. Para que un modelo posea un valor F1 elevado, requiere tanto de una alta precisión como de una alta sensibilidad [120].

Curva ROC y AUC

Existen muchos clasificadores en donde variar ciertos parámetros puede modificar la cantidad de falsos positivos y falsos negativos, afectando en cierto grado el desempeño del modelo. Un caso podría ser un aumento de sensibilidad teniendo como costo una disminución de la precisión. A modo de ejemplo, se puede considerar un clasificador que trabaja con instancias representadas por un único atributo. El clasificador funciona de la siguiente manera: si el atributo a cumple con $a \geq \theta$, con θ un valor umbral, la instancia se clasifica como un positivo y de no ser así, como un negativo. Si el valor θ es muy pequeño, dentro del rango de valores de a, la mayor parte de las instancias serían clasificadas como positivos. Lo anterior haría que muy probablemente todos los positivos se clasifiquen como tal, pero también muchos negativos se clasificarían como positivos. Esto equivale a tener una alta tasa de verdaderos positivos y de falsos positivos. En caso contrario, si el valor θ es muy grande, dentro del rango de valores de a, la mayoría de las instancias serían clasificadas como negativos. En tal caso, análogamente, se tiene una baja tasa de verdaderos positivos y de falsos positivos. Así, ambas tasas crecen o disminuyen juntas. Un modelo con un buen desempeño se caracteriza por una alta tasa de verdaderos positivos y una baja tasa de falsos positivos. De este análisis, se concluye que existe un valor umbral θ^* que optimiza la compensación entre tasa de verdaderos positivos y tasa de falsos positivos.

La curva ROC (del inglés *Receiver Operating Characteristic*), se construye al variar el umbral θ del ejemplo anterior, calculando para cada valor de dicho umbral la tasa de verdaderos positivos y falsos positivos, y graficando los pares de ambas tasas en el plano. La Figura 2.15 muestra un gráfico de la curva ROC con varios elementos. La curva ROC de referencia (celeste intenso) posee un área bajo la curva (celeste suave), también llamada AUC (del inglés *Area Under Curve*), que constituye una métrica de evaluación para los modelos de clasificación binaria. La recta diagonal (morada) representa una curva ROC generada por un clasificador cuyo desempeño equivale al de una clasificación binaria aleatoria, con un 50 % de probabilidad de predecir cada clase. En este caso, el AUC posee un valor de 0,5 o 50 %. Si la curva ROC se dobla por debajo de la recta diagonal, esto quiere decir que el clasificador en cuestión funciona de peor manera que la clasificación aleatoria. La curva ROC de referencia se dobla por encima de la recta diagonal, en cuyo caso el clasificador funciona mejor que la clasificación aleatoria. La recta de clasificación perfecta (verde) representa un clasificador que alcanza el punto óptimo del plano (0,1), logrando para cierto valor umbral, una tasa de verdaderos positivos de 100 % y una tasa de falsos positivos de 0 %. Para la curva ROC de clasificación perfecta, el AUC es 1 o 100 %. De esta forma, el umbral óptimo de cualquier curva ROC se alcanza en el punto de la curva más cercano al punto (0,1). La AUC es una métrica que se encuentra entre 0 y 1. A mayor AUC, mejor es el desempeño del modelo.

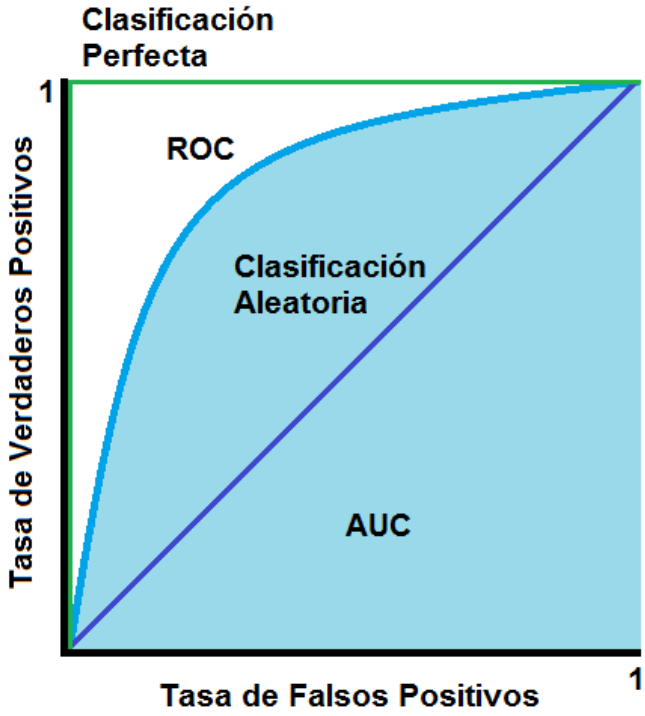


Figura 2.15: Curva ROC y área bajo la curva.
Fuente: Elaboración propia.

2.5.2. Validación Cruzada

La validación cruzada es un proceso que permite generar una distribución de pares de conjuntos de entrenamiento y prueba a partir de un único conjunto de datos. En validación cruzada, las instancias son particionadas en k subconjuntos, $S_1, ..., S_k$. Usualmente las particiones son aproximadamente del mismo tamaño, es decir, poseen una cantidad similar o igual de instancias. El algoritmo que aprende se aplica k veces, de $i = 1$ a $i = k$, utilizando en cada iteración la partición S_i como conjunto de prueba y el resto de las particiones como conjunto de entrenamiento [120]. La Figura 2.16 muestra una representación del proceso de validación cruzada con $k = 5$ particiones como ejemplo. La validación cruzada se emplea generalmente cuando la cantidad de instancias es pequeña [18]. En tales casos, realizar una única separación del conjunto de datos en conjuntos de entrenamiento y prueba podría generar un sesgo considerable dependiendo de las instancias que aleatoriamente queden en ambos conjuntos. Mediante validación cruzada, cada una de las k aplicaciones del algoritmo se ejecuta con conjuntos de entrenamiento y prueba distintos. Debido a lo anterior, como resultado se obtienen métricas de desempeño promedio junto con su desviación estándar, permitiendo así evaluar de forma más robusta el modelo logrado. El conjunto de datos utilizado para realizar la validación cruzada podría ser el conjunto total de datos o el conjunto de entrenamiento completo (entrenamiento y validación).

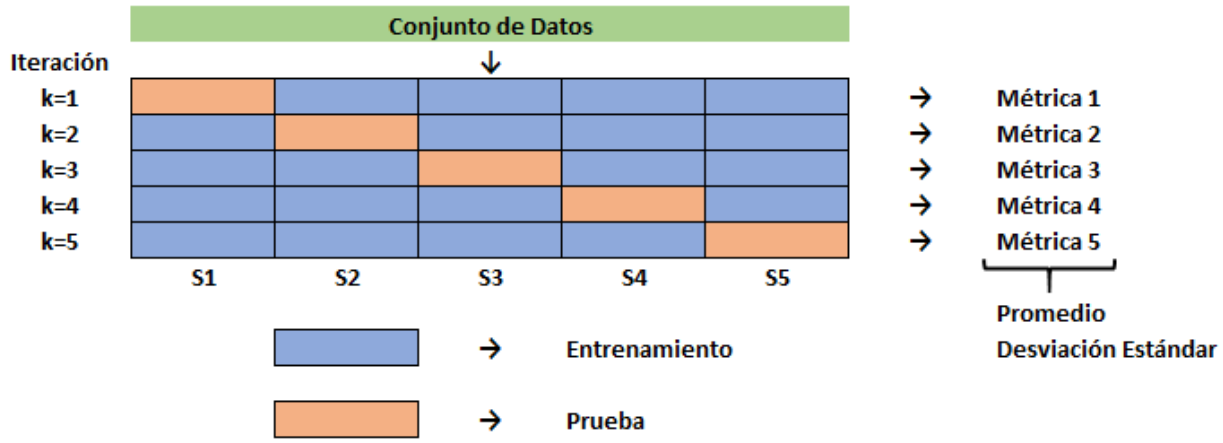


Figura 2.16: Validación cruzada con k=5 particiones.
Fuente: Elaboración propia.

La elección de la cantidad de particiones k constituye un desafío. Esto se debe a que existe una compensación entre sesgo y varianza que depende del número k. Si k es un valor pequeño, los k conjuntos de entrenamiento generados en cada iteración tienden a ser diferentes entre sí en mayor grado, lo cual aumenta el sesgo del modelo. En caso contrario, si k es un valor grande, la cantidad de iteraciones es mayor y los k conjuntos de prueba generados en cada una son de menor tamaño, lo cual genera mayor varianza en los resultados del modelo. James *et al.* señalan en su libro que considerando esta compensación entre sesgo y varianza, típicamente la validación cruzada se ejecuta con $k = 5$ o $k = 10$, dado que se ha mostrado empíricamente que el error del modelo al evaluar en el conjunto de prueba, utilizando dichos valores de k, no sufre de gran sesgo ni varianza [57]. Cabe decir que el método de validación cruzada en un principio ordena las instancias de forma aleatoria para luego realizar la partición. Por lo anterior, posee un estado aleatorio asociado.

Estratificación

Una variante de este procedimiento corresponde a la validación cruzada estratificada. La estratificación corresponde al proceso de reordenar los datos para garantizar que cada partición representa correctamente el total. Como ejemplo, en el caso de clasificación binaria donde cada clase abarca el 50 % de los datos, conviene ordenarlos de tal forma que en cada partición, cada clase conste de mitad de los datos aproximadamente [114].

2.5.3. Selección de Atributos

La selección de atributos es una técnica de reducción de la dimensionalidad. Como tal, apunta a escoger un subconjunto (pequeño) de atributos relevantes a partir del conjunto original de atributos. Esto se lleva a cabo mediante la remoción de atributos irrelevantes, redundantes, o con ruido. La selección de atributos usualmente permite un mejor aprendizaje, mayor exactitud, menor costo computacional, y mejor interpretabilidad del modelo. Los atributos irrelevantes son aquellos que no ayudan a discriminar entre instancias de distintas clases. Por otro lado, los atributos redundantes son los que implican la copresencia de otro atributo. Individualmente, cada atributo redundante puede ser relevante, pero el remover uno de ellos no afecta el aprendizaje. Dicho de otra manera, los atributos redundantes no entregan información adicional entre sí. Por último, los atributos con ruido son un tipo de atributo relevante, que debido a la presencia de ruido ya sea por el proceso de recolección de datos o la naturaleza misma del atributo, no llega a ser relevante para el aprendizaje. Los datos de alta dimensionalidad constituyen un desafío, dado que esta característica tiende a degenerar el desempeño de los algoritmos, y a incrementar significativamente el tiempo y la memoria que estos requieren [120]. En lo que sigue se revisan ciertas técnicas utilizadas para llevar a cabo la selección de atributos.

Correlación

La correlación se refiere a la relación estadística o dependencia existente entre dos variables aleatorias distintas, que en este contexto corresponden a los atributos. Para un problema de clasificación binaria abordado mediante machine learning, la correlación tiene tres utilidades, señaladas a continuación.

- Obtener mayor conocimiento acerca del fenómeno subyacente a los datos mediante las relaciones entre atributos evidenciadas.
- En particular, identificar atributos relevantes para la clasificación, mediante el nivel de correlación entre la etiqueta o clase y el resto de los atributos.
- Reconocer atributos redundantes y mejorar el desempeño de los modelos aprendidos a través de su eliminación.

Existen tres alternativas de correlación entre dos atributos, las cuales se indican a continuación.

- Correlación Positiva: Ambos atributos crecen o decrecen conjuntamente.
- Correlación Neutra: No se observa una relación en el crecimiento o decrecimiento de ambos atributos.
- Correlación Negativa: Ambos atributos crecen o decrecen opuestamente entre sí.

La correlación suele cuantificarse a través de un coeficiente de correlación. Dependiendo de lo sabido acerca de la relación entre atributos y su distribución, se pueden calcular distintos coeficientes. A continuación se describen dos coeficientes de correlación típicamente utilizados.

1. Coeficiente de Correlación de Pearson

   El coeficiente de correlación de Pearson se utiliza para cuantificar el grado de dependencia lineal entre dos atributos. Matemáticamente corresponde a la normalización de la covarianza entre ambos atributos con el fin de otorgar una métrica interpretable. La expresión para su cálculo es la siguiente.

$$r = \frac{cov(X, Y)}{\sigma_X \cdot \sigma_Y} \tag{2.16}$$

   donde,

   r: Coeficiente de correlación de Pearson.

   X, Y: Atributos.

   $cov()$: Covarianza.

   σ_i: Desviación estándar del atributo i.

   Debido a la presencia implícita o explícita de la media y la desviación estándar en la fórmula anterior, los atributos en cuestión deben seguir una distribución normal o aproximadamente normal para el uso de este coeficiente de correlación. El coeficiente que se obtiene cumple con $r \in [-1, 1]$. Cuando $r = 0$ no existe correlación entre los atributos. En el caso que $r = 1$, los atributos siguen una correlación positiva completamente lineal. Contrariamente, cuando $r = -1$, los atributos manifiestan una correlación negativa completamente lineal. En términos generales, se interpreta que si $|r| > 0,5$, entonces la correlación lineal entre ambos atributos es alta o considerable. Si $|r| \leq 0,5$, se concluye que la correlación lineal entre ambos atributos es débil o eventualmente nula.

2. Coeficiente de Correlación de Spearman

   El coeficiente de correlación de Pearson aplica cuando la relación entre los atributos es no lineal y abarca los casos en que las distribuciones de ambos no son normales. En vez de utilizar el supuesto de una dependencia lineal, asume una relación monotónica. Matemáticamente es similar el coeficiente de Pearson, pero utiliza el ranking de valores de cada atributo en vez del valor directo. La expresión para su cálculo es la siguiente.

$$r_s = \frac{cov(rank(X), rank(Y))}{\sigma_{rank(X)} \cdot \sigma_{rank(Y)}} \tag{2.17}$$

   donde,

   r_s: Coeficiente de correlación de Spearman.

   $rank()$: Ranking.

Análogo al caso anterior, el coeficiente cumple con $r_s \in [-1, 1]$. De igual forma, si $r_s = 0$ no existe correlación entre los atributos. Cuando $r_s = 1$, los atributos siguen una correlación positiva monotónica, y en caso que $r_s = -1$, poseen una correlación negativa monotónica. A su vez, nuevamente se utiliza el umbral de $|r_s| = 0,5$ para interpretar la correlación monotónica entre los atributos como fuerte o débil, dependiendo si se encuentra sobre o bajo dicho umbral, respectivamente [23].

Una forma de representación gráfica comúnmente utilizada cuando se realiza el cálculo de coeficientes de correlación es la llamada matriz de correlación. Esta consiste en una matriz cuadrada que posee la lista de atributos o variables tanto en sus filas como en sus columnas. De esta forma, la casilla ubicada en la intersección de una fila y una columna, correspondiente a un par de atributos particular, exhibe el coeficiente de correlación entre dicho par. Dado aquello, la diagonal de la matriz de correlación posee únicamente el valor 1, ya que corresponde a la intersección de fila y columna del mismo atributo. La Figura 2.17 muestra un ejemplo de una matriz de correlación.

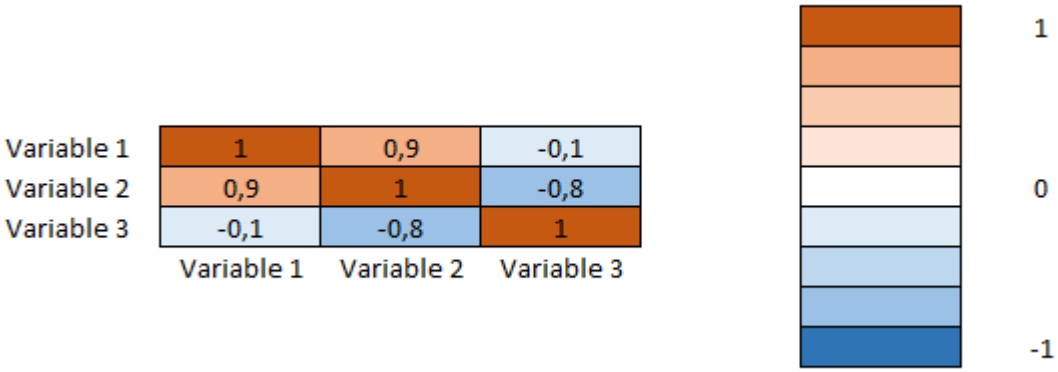


Figura 2.17: Matriz de correlación.
Fuente: Elaboración propia.

Pruebas Estadísticas

Dentro de la estadística existen numerosas pruebas para validar o rechazar ciertas hipótesis relativas a una muestra de datos. Entre ellas, existen pruebas que permiten determinar si la diferencia entre dos grupos, para cierta variable o atributo, es estadísticamente significativa. Lo anterior se utiliza en selección de atributos bajo el contexto de clasificación con machine learning, para identificar atributos que discriminan entre las clases o grupos. Las pruebas estadísticas se componen de dos hipótesis: la hipótesis nula H_0 y la alternativa H_1. El procedimiento consiste en suponer que se cumple la hipótesis nula, lo cual en general conlleva asumir ciertas distribuciones conocidas que permiten calcular probabilidades de forma analítica. De esta forma, en las pruebas estadísticas se calcula la probabilidad de obtener las muestras o datos registrados, bajo la hipótesis nula. Dicha probabilidad recibe el nombre de p-valor o p. Si el p-valor es muy pequeño, es altamente probable que la hipótesis nula sea inválida. Así, se suele establecer un valor umbral α por debajo del cual se rechaza la hipótesis nula. Comúnmente se escoge $\alpha = 0,05$. En consecuencia, si se cumple que $p < \alpha$, se rechaza la hipótesis nula. En virtud de lo anterior, para la selección de atributos, se puede encontrar una diferencia estadísticamente significativa para el mismo atributo, entre dos clases distintas, es decir, la probabilidad del supuesto que propone una distribución idéntica de valores para el atributo particular, en ambas clases, es muy baja. A continuación se describen algunas pruebas estadísticas usadas en selección de atributos.

- Prueba de Shapiro-Wilk

  Se utiliza para probar la normalidad de una muestra de datos. Su hipótesis nula señala que la muestra se obtiene de una distribución normal. Esta prueba no permite evaluar el potencial de un atributo para discriminar entre dos clases, pero confirma un supuesto necesario para otras pruebas o técnicas que se aplican posteriormente para la selección de atributos propiamente tal.

- Prueba de suma de rangos de Wilcoxon

  Se utiliza para probar si dos muestras provienen de la misma distribución, siendo esta su hipótesis nula. La hipótesis alternativa señala que los valores de una muestra tienden a ser mayores que los de la otra. Esta prueba permite identificar atributos relevantes para la clasificación, teniendo como criterio de diferencia entre las clases la distribución de los valores.

- Prueba χ^2

  Se utiliza para probar si dos variables categóricas se encuentran relacionadas o son independientes. Su hipótesis nula plantea la independencia. Esta prueba permite identificar atributos categóricos importantes para la clasificación.

- Prueba exacta de Fisher

  Se utiliza para probar si dos variables categóricas se encuentran relacionadas o son independientes, mediante el análisis de tablas de contingencia[2]. Su hipótesis nula plantea la independencia. La prueba permite identificar atributos categóricos importantes para la clasificación. Se utiliza como alternativa a la prueba χ^2 cuando existen frecuencias esperadas menores a 5 en las tablas de contingencia [14].

Búsqueda Exhaustiva

La búsqueda exhaustiva es un método de optimización que consiste en la prueba de todas las alternativas posibles con el fin de hallar aquella que solucione de mejor forma el problema particular. De esta forma garantiza una solución óptima [136]. Aplicado a la selección de atributos, el problema consiste en encontrar el subconjunto de atributos óptimo para generar el modelo de clasificación. Así, en vez de identificar atributos relevantes para la clasificación mediante distintas técnicas, se genera el espacio de todos los subconjuntos posibles del conjunto de atributos y se evalúa el algoritmo que aprende mediante las métricas de clasificación, para reconocer aquel o aquellos de mejor desempeño. La búsqueda exhaustiva representa el método más sencillo y directo de realizar la selección de atributos, además del más certero por el hecho de trabajar directamente con los resultados del modelo. Por su parte, es el método menos eficiente en cuanto a costo computacional, y para instancias de alta dimensionalidad resulta impracticable [136]. Tomando la teoría de los conjuntos de potencia, el tamaño del espacio de búsqueda para llevar a cabo la búsqueda exhaustiva, al cual se le resta el subconjunto vacío, viene dado por la siguiente expresión.

$$N_{SC} = 2^N - 1 \tag{2.18}$$

[2]Las tablas de contingencia son tablas de frecuencia de 2 variables categóricas, donde las filas y columnas corresponden a las posibles categorías de las variables, respectivamente.

donde,

N_{SC}: Cantidad de subconjuntos en el espacio de búsqueda a probar.

N: Cantidad de atributos o dimensionalidad de los datos.

En general, a partir de cierto valor de N, dependiendo de los recursos computacionales, la búsqueda exhaustiva no es preferible a otro tipo de técnicas de selección de atributos que conllevan tiempos de cómputo mucho menores. En conclusión, es un procedimiento que puede ser realizado solo para dimensionalidades pequeñas.

2.5.4. Transformación de Atributos

La transformación de atributos tiene el potencial de elevar el desempeño de los algoritmos de machine learning utilizados para la tarea de clasificación. A continuación se describen algunas transformaciones usuales, las cuales se extraen de la documentación de la librería Scikit-learn [71, 24, 105].

Estandarización

La estandarización de un conjunto de datos es comúnmente requerida para muchos algoritmos de machine learning, dado que estos pueden exhibir un desempeño deficiente al trabajar con atributos cuya distribución no posea una forma aproximadamente normal estándar, es decir, con media nula y varianza unitaria. Este escalado asume que la distribución de los datos es aproximadamente normal. La estandarización de una variable o atributo se lleva a cabo mediante la ecuación siguiente.

$$\tilde{x}_i = \frac{x_i - \mu}{\sigma} \tag{2.19}$$

donde,

$\tilde{x}_i$: Valor i escalado.

x_i: Valor i.

μ: Media.

σ: Desviación estándar.

Escalado Máx-Abs

El escalado máx-abs escala cada uno de los atributos respecto de su máximo valor absoluto. De esta forma, el máximo valor absoluto de cada atributo cambia a 1. Este escalado se traduce en la siguiente ecuación.

$$\tilde{x}_i = \frac{x_i}{\text{máx}\{|x_i|\}_{i=1}^{n}} \tag{2.20}$$

Escalado Mín-Máx

El escalado mín-máx escala los valores a un rango determinado. En este caso, el rango escogido es [0,1]. La ecuación utilizada se muestra a continuación.

$$\tilde{x}_i = \frac{x_i - \text{mín}\{x_i\}_{i=1}^{n}}{\text{máx}\{x_i\}_{i=1}^{n} - \text{mín}\{x_i\}_{i=1}^{n}} \tag{2.21}$$

Este escalado es uno de los más utilizados, dado que funciona bien cuando los datos no se encuentran distribuidos normalmente. Cuenta con la desventaja de ser sensible a los valores perdidos.

Normalización

Esta transformación normaliza las instancias o vectores de atributos al valor unitario. La normalización se aplica sobre cada instancia con al menos una componente o valor no nulo. Es una transformación comúnmente utilizada en clasificación de texto. A modo de ejemplo, para una instancia de un conjunto de datos con tres atributos x_i, y_i y z_i, la ecuación de normalización para el primer atributo es la siguiente.

$$\tilde{x}_i = \frac{x_{\text{i}}}{\sqrt{x_i^2 + y_i^2 + z_i^2}} \tag{2.22}$$

Transformación Potencia Box-Cox

La transformación Box-Cox pertenece a la familia de las transformaciones potencia. Estas son transformaciones paramétricas y monotónicas utilizadas para acercar los datos a una distribución normal. La transformación Box- Cox trabaja con datos estrictamente positivos. La ecuación para aplicar esta transformación, tomada de [119], se muestra a continuación.

$$y_i^{(\lambda)} = \begin{cases} \frac{y_i^{\lambda}-1}{\lambda}; & \lambda \neq 0 \\ \ln y_i; & \lambda = 0 \end{cases} \tag{2.23}$$

donde,

$y_i^{(\lambda)}$: Valor i transformado.

y_i: Valor i.

λ: Parámetro potencia.

El parámetro λ se optimiza para estabilizar la varianza y minimizar el grado de skewness o asimetría de la distribución, mediante estimación a través de máxima verosimilitud.

Transformación Cuantil Normal y Uniforme

Esta transformación opera utilizando información sobre los cuantiles, y puede llevar los datos a una distribución normal o uniforme. Así, para un atributo dado, la transformación tiende a esparcir los valores más frecuentes. También reduce el impacto de los valores perdidos, lo cual la convierte en un esquema robusto. La transformación se aplica en cada atributo de forma independiente. Primero se hace una estimación de la función de distribución acumulada, para mapear los valores originales a una distribución uniforme. Luego, los valores obtenidos se mapean a la distribución de salida establecida (normal o uniforme) por medio de la función cuantil asociada. La transformación es no lineal, por lo cual puede distorsionar las correlaciones lineales entre variables medidas en la misma escala pero otorga una comparación más directa entre aquellas medidas en escalas distintas. Por su parte, la función cuantil corresponde a la inversa de la función de distribución acumulada, la cual se puede definir de la siguiente forma [126].

$$Q(p) = F^{-1} = \inf\{x : F(x) \geq p\}; \;\; 0 < p < 1 \tag{2.24}$$

donde,

$Q(p)$: Función cuantil.

F: Función de distribución acumulada.

x: Valor.

p: Probabilidad.

Dicho de otra forma, la función cuantil retorna el valor x por debajo del cual la probabilidad entregada por la función de distribución acumulada (normal o uniforme) es p. La cantidad de cuantiles, utilizada para discretizar la función de distribución acumulada estimada, en este caso corresponde a la cantidad de instancias, dado que un número mayor de cuantiles no otorga una aproximación mejor.

Escalado Robusto

El escalado o transformación robusta, realiza el escalado de los atributos mediante el uso de estadísticos que son robustos frente a los valores perdidos. Este escalado anula la mediana y escala los datos respecto a un rango cuantil, siendo por defecto el rango intercuartil, que corresponde a la diferencia entre el tercer y primer cuartil. El escalado robusto es análogo a la estandarización, utilizando la mediana en vez de la media, y el rango intercuartil en vez de la desviación estándar, lo cual permite manejar de mejor forma la influencia negativa de los valores perdidos. La ecuación siguiente permite aplicar el escalado robusto.

$$\tilde{x}_i = \frac{x_i - Q_2(\{x_i\}_{i=1}^n)}{Q_3(\{x_i\}_{i=1}^n) - Q_1(\{x_i\}_{i=1}^n)} \tag{2.25}$$

donde,

Q_1: Primer cuartil.

Q_2: Segundo cuartil o mediana.

Q_3: Tercer cuartil.

2.5.5. Optimización de Hiperparámetros

La optimización de hiperparámetros consiste en la búsqueda de un conjunto de hiperparámetros propios de un algoritmo de aprendizaje que otorgue una buena capacidad de generalización y baja pérdida [120]. La mayor parte de los algoritmos de clasificación poseen hiperparámetros que controlan el proceso de aprendizaje, cuyo desempeño puede variar en función de los tales. A modo de analogía, se puede considerar un polinomio que se ajusta a ciertos datos. El aprendizaje o entrenamiento consistiría en determinar los coeficientes del polinomio, a partir de los datos. Un hiperparámetro correspondería al grado del polinomio, el cual no se determina a partir de los datos, sino que es establecido por el usuario, pudiendo haber grados que sean mejores, peores, o tal vez equivalentes, en términos del desempeño logrado por el modelo.

Para llevar a cabo la optimización de hiperparámetros existen diversos métodos. A continuación se señalan dos de los métodos más simples y de amplia utilización.

- Búsqueda en Grilla

  La búsqueda en grilla consiste en una búsqueda exhaustiva, en donde a cada hiperparámetro se le asigna un conjunto de valores posibles. Así, se prueban todas las combinaciones posibles de hiperparámetros, considerando las alternativas otorgadas a cada uno. Este método permite encontrar una solución óptima, pero, a medida que se amplía el espacio de búsqueda, el tiempo de cómputo puede aumentar considerablemente. La Figura 2.18 muestra un ejemplo esquemático de optimización de hiperparámetros mediante búsqueda en grilla. En este ejemplo, se tienen 3 hiperparámetros para el algoritmo de aprendizaje, a los cuales se les asignan 3, 4 y 6 valores posibles, respectivamente. Así, se tiene un total de 72 combinaciones posibles, las cuales se prueban para encontrar aquella que entrega el mejor desempeño.

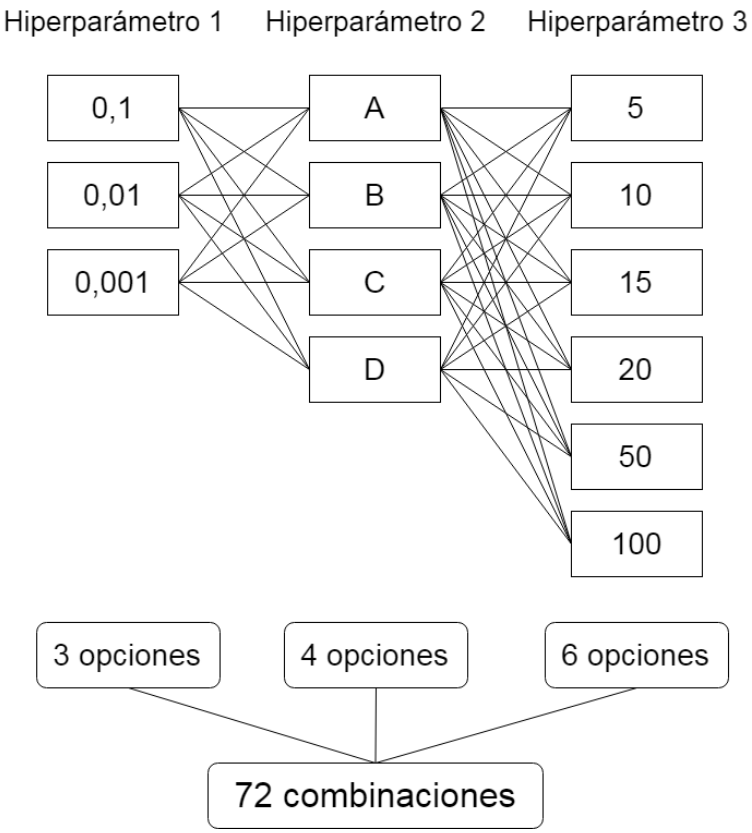


Figura 2.18: Ejemplo esquemático de búsqueda en grilla.
Fuente: Elaboración propia.

- Búsqueda Aleatoria

  La búsqueda aleatoria parte de la misma base que la búsqueda en grilla, en donde se generan todas las combinaciones posibles de hiperparámetros de acuerdo a los posibles valores asignados de cada uno. La diferencia radica en que, en vez de evaluar todas las combinaciones generadas, únicamente se evalúa una subconjunto de estas combinaciones, seleccionado aleatoriamente. Así, este método permite en general disminuir los tiempos de cómputo pero no garantiza el hallazgo de la solución óptima.

2.5.6. Algoritmos

Regresión Logística

La regresión logística es una técnica que permite llevar el modelo de una regresión lineal a problemas de clasificación. Una regresión lineal se expresa mediante la ecuación mostrada a continuación.

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n \tag{2.26}$$

donde,

x_i: Valor del atributo i.

β_i: Peso asignado al atributo i.

Con el fin de mapear la salida al intervalo [0,1] y de esa forma generar un modelo de clasificación binario, se usa la siguiente función, que entrega una curva llamda sigmoide.

$$P(C = 1|x_1...x_n) = \frac{1}{1 + e^{-z}} \tag{2.27}$$

La función definida anteriormente entrega la probabilidad P de que la instancia definida por los valores $\{x_i\}_{i=1}^{n}$ corresponda a la clase $C = 1$. Si se cumple que $P(C = 1|x_1...x_n) < 0,5$, entonces el modelo clasifica la instancia en la clase $C = 0$, y en caso contrario, en la clase $C = 1$. Los coeficientes $\{\beta_i\}_{i=1}^{n}$ se determinan a partir de la estimación de máxima verosimilitud, la cual se lleva a cabo en virtud de los datos de entrenamiento etiquetados. Como resultado, se ajusta una curva sigmoide a los datos de entrenamiento, la cual se utiliza para determinar la probabilidad señalada anteriormente y así llevar a cabo la clasificación [120, 12]. La Figura 2.19 muestra el ajuste de una curva sigmoide a un conjunto de datos etiquetados binariamente, los cuales se representan mediante un único atributo $X1$. Cabe decir que la regresión logística puede extenderse a problemas de clasificación multiclase, sin embargo, por su naturaleza es principalmente aplicada en clasificación binaria. La regresión logística puede volverse inestable cuando las clases se encuentran bien separadas, y cuando se tienen pocas instancias de entrenamiento.

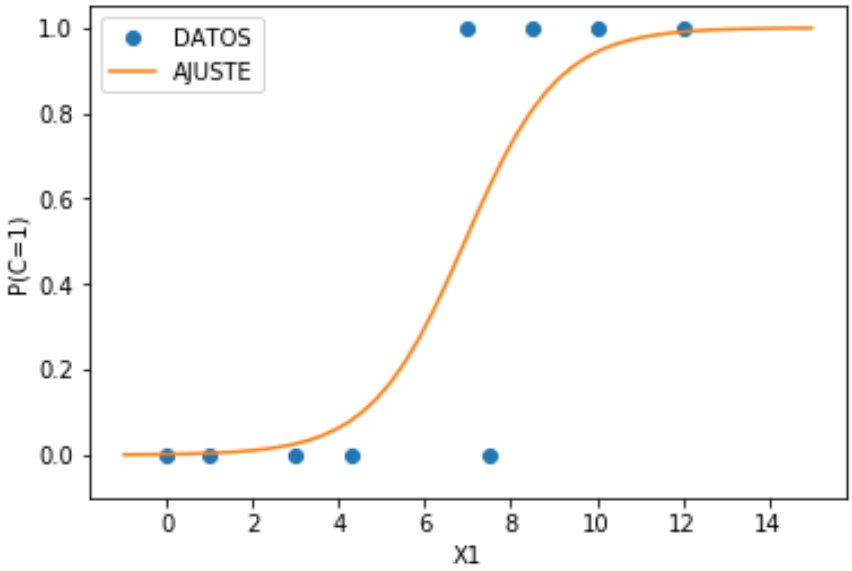


Figura 2.19: Curva sigmoide ajustada a los datos para clasificación binaria.
Fuente: Elaboración propia.

Análisis Discriminante Lineal

Una función discriminante es aquella que toma una variable de entrada $\vec{x}$ y retorna una etiqueta y. Un discriminante lineal utiliza una función lineal de la variable de entrada $f(\vec{x})$. Para el caso de clasificación binaria, lo anterior se traduce en la búsqueda de una proyección a un espacio de menor dimensionalidad, de las entradas $\vec{x}$ en una línea en la dirección de $\vec{w}$, de modo que $f(\vec{x}) = \vec{x} \cdot \vec{w}$. De esta forma, la etiqueta y se puede asignar a partir de un valor umbral de la proyección, por ejemplo, $y = 0$ cuando $f(\vec{x}) \geq C$ y en caso contrario, $y = 1$, para un valor apropiado de C. Mientras la magnitud de $\vec{w}$ es irrelevante, su dirección es crucial. Esta dirección se escoge de tal forma que se maximice la separación entre las clases. Lo anterior se logra en dos frentes. Primeramente, se busca maximizar la diferencia entre las medias proyectadas de ambas clases. En segundo lugar y de forma simultánea, se busca minimizar la varianza de cada clase. Este problema de optimización entrega un vector $\vec{w}$ definido [120]. La Figura 2.20 muestra un ejemplo del funcionamiento del análisis discriminante lineal. En este, se tienen instancias de dos clases (azul y roja), representadas por dos atributos x_1 y x_2. En el gráfico de la izquierda, se aprecia la proyección de los puntos o datos sobre una línea particular. Observando la proyección se ve que no se logra una buena separación entre clases. Por otro lado, al considerar el gráfico de la derecha, se realiza la proyección sobre otra línea, en donde se obtiene una buena separación entre clases.

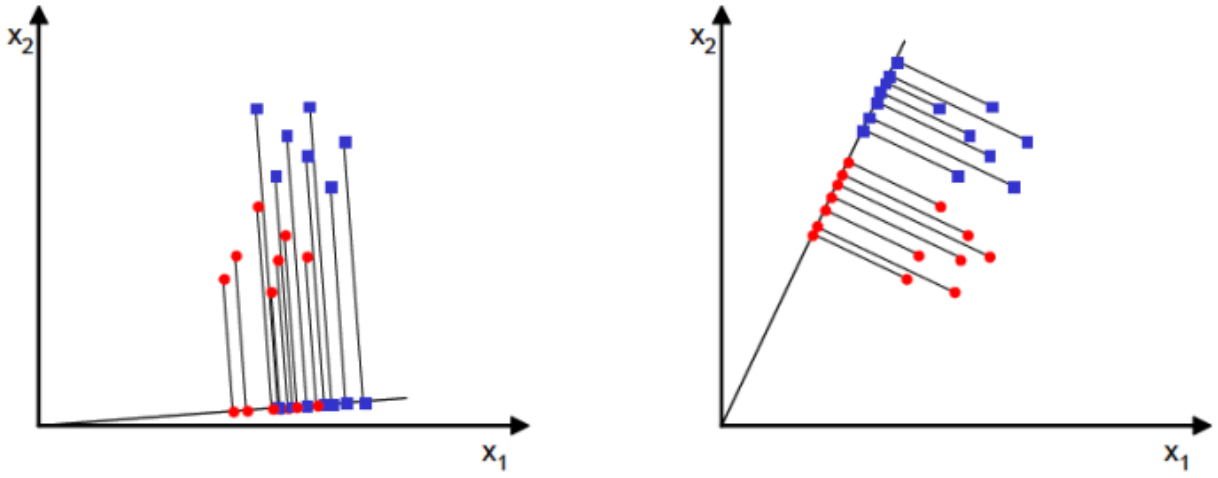


Figura 2.20: Análisis discriminante lineal en un ejemplo bidimensional con dos clases. Fuente: Adaptada de [64].

K Vecinos Más Cercanos

El algoritmo k vecinos más cercanos funciona a partir de un criterio de similitud entre las instancias. Una forma natural de medir la similitud entre dos instancias expresadas como vectores, es la distancia euclidiana, calculada por medio de la siguiente expresión en un espacio de n dimensiones.

$$d(\vec{p},\vec{q}) = d(\vec{q},\vec{p}) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \tag{2.28}$$

El algoritmo k vecinos más cercanos, en su versión más simple, opera en tres pasos, descritos a continuación.

1. Teniendo una nueva instancia a clasificar, identificar las k instancias de entrenamiento más cercanas a esta, de acuerdo a un criterio de similitud.
2. Sea c_i la clase más frecuente entre las k instancias más cercanas.
3. Etiqueta la nueva instancia a clasificar con c_i.

En clasificación binaria, con el fin de evitar un empate, la cantidad k de vecinos cercanos a considerar para decidir la etiqueta debiera ser un número impar. En el caso de un problema de clasificación multiclase, esto no necesariamente garantiza la inexistencia de empates [70]. La Figura 2.21 muestra un ejemplo gráfico de este método. En ella se muestra un gráfico bidimensional de dos atributos $X1$ y $X2$, donde se ubican instancias de entrenamiento de las clases A (roja) y B (verde). En amarillo se indica una nueva instancia a clasificar. En este caso, el algoritmo se implementa con $k = 1$. El vecino más cercano considerado es de la clase A, por lo cual la nueva instancia se clasifica como tal.

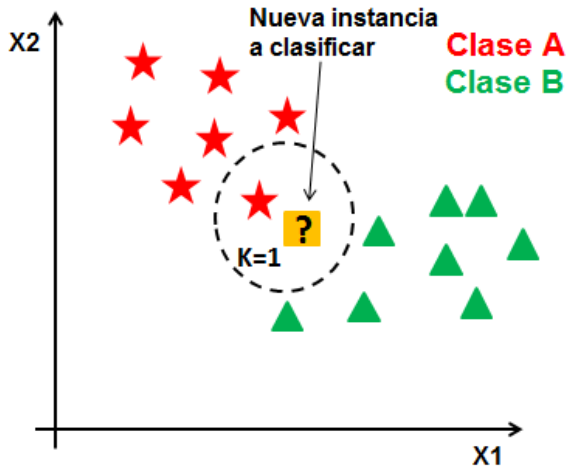


Figura 2.21: K vecinos más cercanos en un ejemplo bidimensional con dos clases.
Fuente: Adaptada de [91].

Árboles de Decisión

Los árboles de decisión se construyen por medio de la división recursiva del conjunto de datos en una secuencia de subconjuntos a través de preguntas del tipo si-entonces. El algoritmo de aprendizaje divide el conjunto de entrenamiento en grupos según los atributos de las instancias, con el objetivo de lograr la mayor uniformidad de clases al interior de cada uno, idealmente una única clase presente en cada grupo. Para esto, el algoritmo elige un atributo y un umbral de dicho atributo a partir del cual se realiza la división de las instancias [139]. Los árboles de decisión se pueden representar gráficamente, por lo cual son muy fáciles de interpretar en cuánto a cómo realizan la clasificación. Los grupos de datos presentes en un árbol de decisión se representan gráficamente como un nodo, de los cuales existen tres tipos, señalados a continuación.

- Raíz: Corresponde al grupo inicial que engloba todas las instancias de entrenamiento, a partir del cual se generan las divisiones.
- Nodo (de decisión): Corresponde a un nodo intermedio, es decir, que recibe una subdivisión de instancias generada previamente y que a su vez divide este grupo, entregando dos subdivisiones.
- Hoja: Corresponde a un nodo terminal, el cual únicamente recibe una subdivisión anterior. Idealmente una hoja debiese agrupar instancias de una sola clase, lo cual no necesariamente ocurre. Por ser un nodo terminal, en clasificación una hoja asigna la clase a la instancia en virtud de la clase más abundante.

El aprendizaje del modelo en un árbol de decisión se lleva a cabo por un algoritmo de inducción de arriba (raíz) hacia abajo (hoja). Para el aprendizaje, se utiliza un criterio de impureza para los nodos. Cuando en un nodo se agrupan instancias de más de una clase, se dice que el nodo es impuro. Existen dos medidas de impureza clásicas, el índice Gini y la entropía de la teoría de la información. Estas vienen dadas por las ecuaciones siguientes [120].

$$Gini(S) = 1 - \sum_{i=1}^{c} \left(\frac{|S_i|}{|S|} \right)^2 \tag{2.29}$$

$$Entropia(S) = - \sum_{i=1}^{c} \frac{|S_i|}{|S|} \cdot log_2 \left(\frac{|S_i|}{|S|} \right) \tag{2.30}$$

donde,

S: Conjunto de instancias de entrenamiento.

S_i: Conjunto de instancias de entrenamiento pertenecientes a la clase c_i.

Un buen atributo divide el conjunto de datos en subconjuntos de la mayor pureza posible. Es por esto que al momento de dividir un nodo se busca seleccionar el atributo que minimice la impureza promedio de la separación. Para un atributo A cualquiera, la impureza promedio se calcula mediante la siguiente expresión, correspondiente a un promedio ponderado [120].

$$Impureza(S, A) = \sum_{t} \frac{|S_t|}{|S|} \cdot Impureza(S_t) \tag{2.31}$$

donde,

S_t: Subconjuntos disjuntos de S inducidos al aplicar el atributo A como criterio de división.

A continuación se explica el funcionamiento del algoritmo de inducción de arriba hacia abajo para el aprendizaje del modelo por parte de los árboles de decisión. El nodo de partida del algoritmo es la raíz.

1. Para un nodo cualquiera, se calcula su impureza.
2. Se prueban todos los atributos como criterio de división del nodo, escogiendo aquel que minimice la impureza promedio de la separación.
3. Si la impureza promedio de la separación es menor que la impureza del nodo, se lleva a cabo la separación. En caso contrario, el nodo considerado se torna una hoja o nodo terminal.
4. Se repiten los pasos 1, 2 y 3 para los nuevos nodos generados.

Los árboles de decisión son proclives al sobreajuste, en donde el modelo se ajusta de tal forma a los datos de entrenamiento que pierde su capacidad de generalizar, exhibiendo un bajo desempeño en el conjunto de prueba. En este caso, el sobreajuste se manifiesta en una estructura altamente compleja del árbol. Con el fin de evitar este problema, se usan técnicas de poda, en donde se restringe la complejidad del árbol de decisión generado [120]. El aprendizaje de un modelo de clasificación a través de un árbol de decisión puede generar más de una solución. Dicho de otra forma, se pueden generar árboles con distintas estructuras a partir del mismo conjunto de entrenamiento. Es por esto que este método posee un estado aleatorio asociado.

La Figura 2.22 muestra un ejemplo de la representación gráfica de un árbol de decisión correspondiente a un modelo de clasificación aprendido. En dicho ejemplo, se tienen 150 instancias en la raíz, pertenecientes a tres clases distintas A, B y C, cada una con 50 instancias. En la raíz se tiene un índice Gini de impureza de 0,667 y se utiliza el atributo $X2$ como criterio de división, de acuerdo a $X2 \leq 2,45$. En base a si se cumple este criterio, se separa la raíz en dos nodos. El nodo de la izquierda queda con 50 instancias de la clase A, por lo cual su impureza es nula y se convierte en una hoja. Así, si una nueva instancia a clasificar cumple con $X2 \leq 2,45$ pasa a esta hoja y por lo tanto se clasifica como clase A. El nodo de la derecha queda con 100 instancias, 50 de la clase B y 50 de la clase C. Con esto, su impureza es de 0,5. Nuevamente, este nodo se divide de acuerdo a $X1 \leq 1,75$. El nodo resultante de la izquierda, donde se cumple el criterio, queda con 54 instancias, 49 de la clase B y 5 de la clase C. Así, su impureza es de 0,168 y se vuelve una hoja que clasifica las instancias como clase B, dado su predominio. El nodo derecho también se convierte en hoja. En este quedan 46 instancias, 1 de la clase B y 45 de la clase C, por lo cual su impureza es de 0,043. Dado el predominio de la clase C, las nuevas instancias que de acuerdo a los criterios de división arriban a esta hoja, se clasifican en dicha clase. Se advierte que la impureza decrece desde la raíz hacia las hojas. A pesar de ello, no todas las hojas generadas son puras.

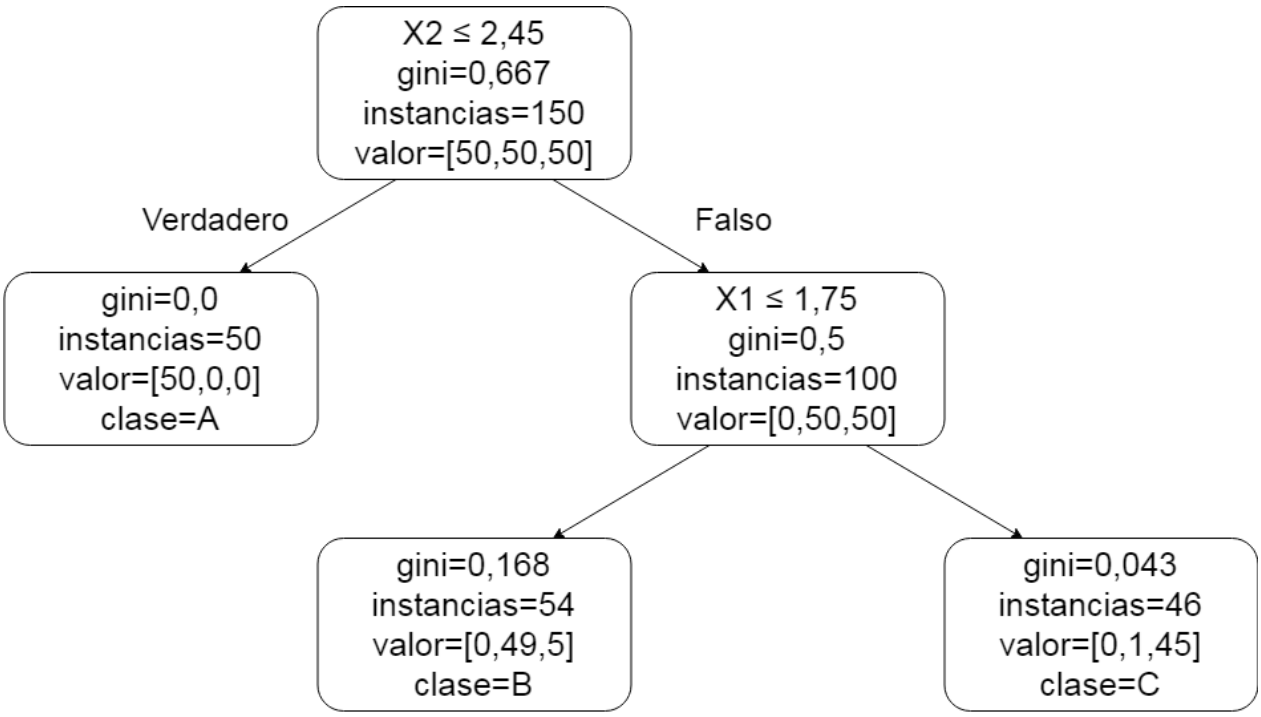


Figura 2.22: Representación gráfica de un modelo de árbol de decisión.
Fuente: Adaptada de [27].

Clasificador Bayesiano Ingenuo

El clasificador bayesiano ingenuo trabaja a través de los principios de la teoría probabilística bayesiana. En términos simples, en esta aproximación se calcula para cada clase la probabilidad de que una instancia específica pertenezca a esta. La instancia se clasifica en la clase que obtiene mayor probabilidad. El teorema de Bayes aplicado a una instancia representada por el vector $\vec{x}$ y una clase particular c_i, se expresa de la siguiente forma.

$$P(c_i|\vec{x}) = \frac{P(\vec{x}|c_i)P(c_i)}{P(\vec{x})} \tag{2.32}$$

Dado que $P(\vec{x})$ es equivalente para cada clase, esta no se considera, y se busca maximizar el numerador $P(\vec{x}|c_i)P(c_i)$ de la fracción. La probabilidad $P(c_i)$ se puede calcular fácilmente a partir de la frecuencia relativa de cada clase en el conjunto de entrenamiento. En cuanto a la probabilidad $P(\vec{x}|c_i)$ de obtener el vector $\vec{x}$ dada la clase c_i, su cálculo no es tan directo. Con el fin de simplificar la determinación de esta probabilidad, se trabaja bajo el supuesto de que todos los atributos son independientes entre sí. Bajo este supuesto, la probabilidad señalada se calcula mediante la expresión siguiente.

$$P(\vec{x}|c_i) = P(x_1|c_i) \times P(x_2|c_i) \times ... \times P(x_n|c_i) = \prod_{j=1}^{n} P(x_j|c_i) \tag{2.33}$$

Debido a que el algoritmo asume el supuesto de independencia entre atributos, recibe el nombre de "ingenuo". En general, este supuesto raramente se justifica, lo cual se puede corroborar a través de una matriz de correlación. A pesar de esto, si bien la probabilidad se estima de forma inexacta cuando el supuesto no se cumple, esto no necesariamente hace que la clasificación sea incorrecta. Se debe tener en cuenta que la etiqueta se asigna para la clase c_i que maximice el producto $P(\vec{x}|c_i)P(c_i)$, por lo cual la inexactitud del primer factor no necesariamente altera la clasificación fruto del producto entero. Es por esto que incluso de no cumplirse la independencia, el asumirlo constituye un supuesto razonable en general [70].

Máquinas de Vectores de Soporte

Las máquinas de vectores de soporte poseen una base matemática más fuerte que otros algoritmos de machine learning. Funcionan generando un hiperplano separador de las clases, el cual a su vez, maximiza su margen entre instancias de las clases a separar. A pesar de que pueden existir múltiples hiperplanos separadores de las instancias en clases, la formulación anterior se traduce matemáticamente en un problema de optimización que entrega como solución un vector $\vec{w}$ que define el hiperplano óptimo. El hecho de maximizar el margen permite una mejor capacidad de generalización del clasificador [120]. Los vectores de soporte corresponden a las instancias (vectores de atributos) que se encuentran más próximas al hiperplano óptimo. Dado lo anterior, definen el hiperplano con su margen [92].

La Figura 2.23 ilustra el funcionamiento de una máquina de vectores de soporte para un caso bidimensional con dos clases. El gráfico de la izquierda muestra tres opciones de hiperplanos separadores, que en este caso corresponden a rectas en el plano. Las rectas azul y anaranjada no logran separar correctamente las instancias de entrenamiento. Por su parte, la recta negra lo logra. Cabe decir que si la recta negra se rotara levemente respecto de su punto medio, resultarían otros planos de separación admisibles. En el gráfico de la derecha se observa la recta negra establecida como la línea de separación óptima, con su margen correspondiente. Los vectores de soporte de ambas clases se indican en el gráfico y consecuentemente se ubican en la frontera del margen. Se advierte que a mayor margen, aumentan las posibilidades de que una nueva instancia a clasificar quede en el lado correcto de la línea de separación y por lo tanto su clase sea bien asignada, lo cual se traduce como la capacidad de generalizar.

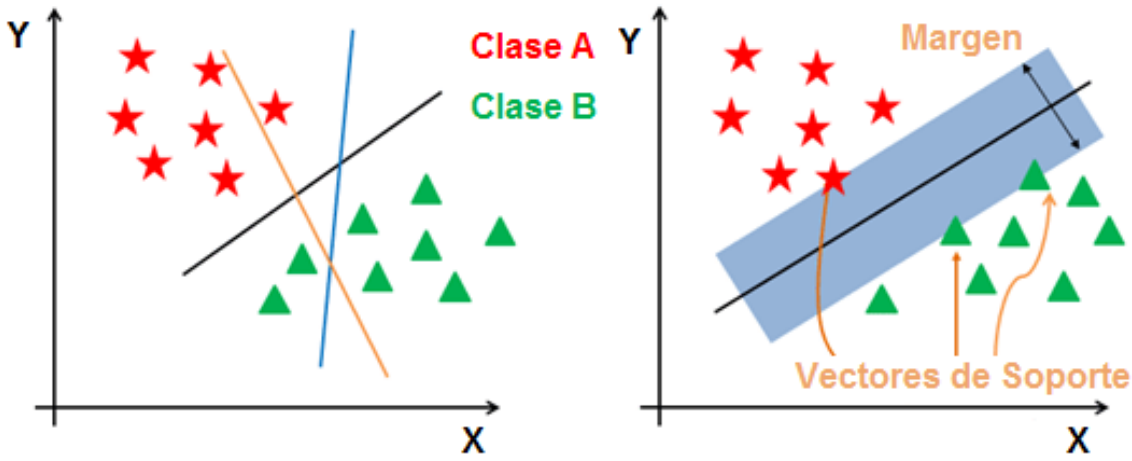


Figura 2.23: Máquina de Vectores de Soporte en un ejemplo bidimensional con dos clases. Fuente: Adaptada de [92].

Existen algunos problemas en donde no se puede utilizar un hiperplano lineal para separar correctamente las clases, tal como se observa en el gráfico de la izquierda de la Figura 2.24. Esta situación se puede resolver llevando las instancias a un espacio de mayor dimensión, donde eventualmente se convertirán en linealmente separables, mediante un hiperplano. En la Figura 2.24, el gráfico de la izquierda muestra las instancias en el plano XY, mientras que el gráfico de la derecha lo hace en el plano XZ, en donde se aprecia que es posible establecer un hiperplano de separación. En primera instancia, para cualquier aumento en la dimensión del espacio, se tendrían que calcular las coordenadas de las instancias en el nuevo espacio dimensional, lo cual es costoso en términos de cálculo. En la práctica, se utiliza el método del kernel para elevar la dimensionalidad. Un kernel es una función que permite operar vectores en una dimensión mayor sin la necesidad de determinar sus coordenadas en el espacio de mayor dimensionalidad, posibilitando el encontrar una dimensión superior donde se pueda llevar a cabo la separación lineal. En clasificación típicamente se utiliza el kernel RBF (del inglés *Radial Basis Function*), que permite mapear un espacio de entrada a dimensiones mayores [92].

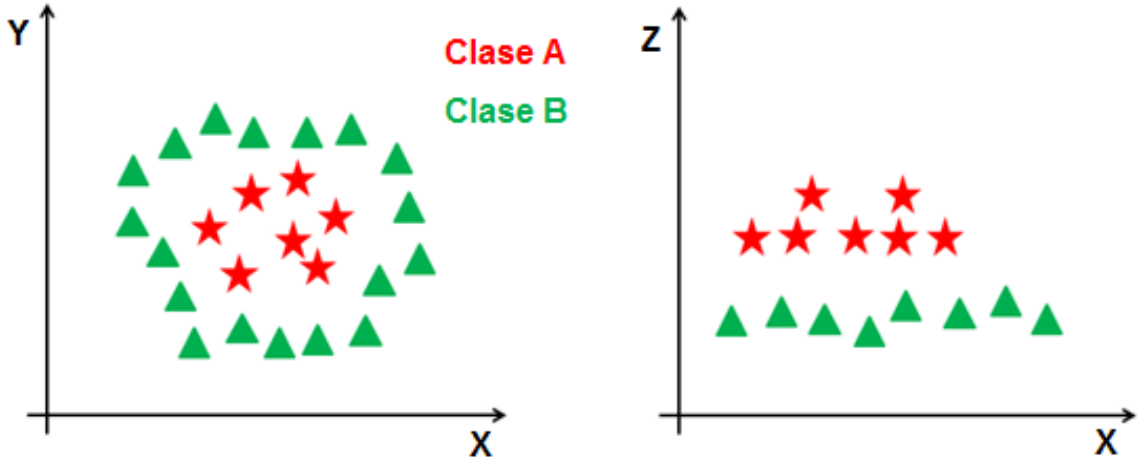


Figura 2.24: Máquina de Vectores de Soporte en un ejemplo no linealmente separable. Fuente: Adaptada de [92].

Bosques Aleatorios

Los bosques aleatorios pertenecen a la categoría de métodos de aprendizaje combinado. En este caso se utilizan árboles de decisión como algoritmo base [120]. El aprendizaje combinado abarca los procedimientos en donde se entrenan múltiples algoritmos de aprendizaje y se combinan los resultados entregados por cada uno, pudiendo considerarse como un comité de toma de decisiones. El principio subyacente es que la decisión del comité, con las predicciones individuales combinadas de forma apropiada, debiera exhibir una mayor exactitud global en promedio, comparado con una predicción individual. Se ha mostrado que los modelos de aprendizaje combinado poseen un mejor desempeño que los modelos tradicionales de un único algoritmo en juego [120]. En el caso de clasificación, la etiqueta se decide por votación al considerar la predicción de los múltiples árboles de decisión del bosque aleatorio. La cantidad de clasificadores individuales corresponde a un hiperparámetro del modelo.

En los bosques aleatorios, cada árbol de decisión individual se construye a partir de muestras bootstrap del conjunto de entrenamiento [120]. El muestreo bootstrap consiste en la generación de una distribución de conjuntos de entrenamiento a partir del conjunto total. Para ello, se define una cantidad de muestras a obtener (conjuntos de entrenamiento), el tamaño de las muestras, y a continuación estas se generan tomando instancias del conjunto total con reposición para construir cada una. Esto permite evaluar el desempeño del modelo aprendido calculando promedios, varianzas e intervalos de confianza [120, 22].

Un aspecto importante a considerar es que los árboles de decisión individuales no se encuentran sujetos a poda luego de construirse, permitiendo así un cierto grado de sobreajuste a sus datos de entrenamiento propios. Con el fin de diversificar los clasificadores individuales, los atributos posibles para generar un criterio de división de un nodo se restringen a un subconjunto aleatorio de los atributos totales. Este subconjunto aleatorio se escoge nuevamente en cada ramificación [120]. La Figura 2.25 muestra una representación del funcionamiento de un bosque aleatorio.

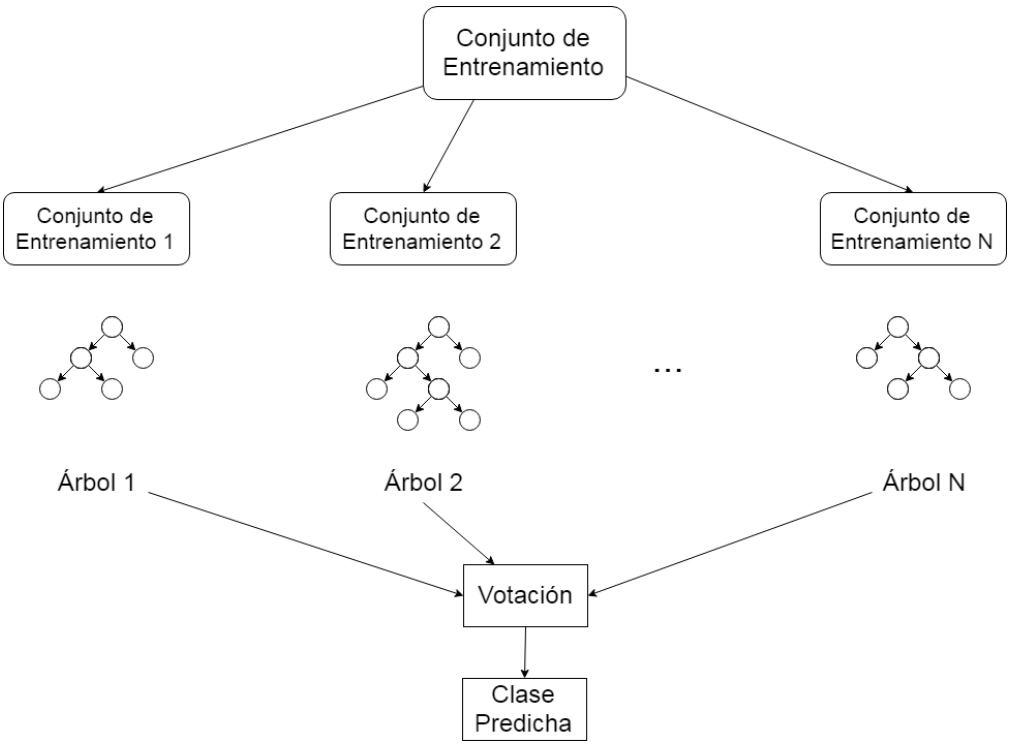


Figura 2.25: Representación de un bosque aleatorio.
Fuente: Elaboración propia.

AdaBoost

AdaBoost es un algoritmo que pertenece a los métodos de aprendizaje combinado [120]. Comúnmente este algoritmo trabaja con árboles de decisión como clasificador base, los cuales poseen un único nivel de profundidad, es decir, la raíz se separa directamente en las hojas, sin nodos intermedios. En cualquier caso, AdaBoost puede utilizar otros algoritmos de machine learning como clasificadores base [71]. La cantidad de clasificadores base corresponde a un hiperparámetro del algoritmo. Típicamente los algoritmos individuales que forman el comité de toma de decisiones, en este caso, árboles de decisión de un nivel de profundidad, no logran aprender un modelo de buen desempeño. AdaBoost entrena los árboles de decisión de forma secuencial, con el fin de lograr mejores modelos iterativamente. Luego de una iteración en donde se entrena un clasificador, se identifican las instancias mal clasificadas y se les asigna una mayor importancia dentro del conjunto de entrenamiento de la iteración siguiente a través de un peso o ponderador. Así, se busca que los clasificadores posteriores compensen los errores cometidos por los anteriores [120]. En cada iteración la ponderación de cada instancia se vuelve a reajustar.

El nombre del algoritmo es una abreviación del concepto de boosting adaptativo. Boosting se refiere al paradigma de crear un clasificador robusto a partir de un conjunto de clasificadores débiles, entendiendo lo último como un clasificador con un mal desempeño [120]. El término adaptativo guarda relación con la asignación iterativa de importancia que el algoritmo asigna a las instancias mal clasificadas. Dicho de otra forma, AdaBoost funciona adaptándose iterativamente para solucionar los errores de clasificación. La etiqueta a asignar a una nueva instancia se define por votación considerando el total de clasificadores que forman el comité de decisión. Sin embargo, en la votación se pondera de forma distinta la predicción hecha por cada clasificador individual, otorgando mayor peso en la decisión a los clasificadores que exhibieron un desempeño mayor [137].

La Figura 2.26 muestra una representación del funcionamiento de un clasificador AdaBoost. En este ejemplo, se tiene un problema de clasificación binario en un espacio de dos dimensiones. La clasificación se lleva a cabo mediante una recta separadora de las clases. El primer clasificador débil se equivoca al clasificar 3 triángulos como círculos. Debido a ello, se le da mayor ponderación a dichos triángulos, y el segundo clasificador débil los etiqueta correctamente, pero ahora clasifica 3 círculos como triángulos. Esto le otorga mayor ponderación a los círculos mal clasificados, y el tercer clasificador débil ahora clasifica bien todas las instancias de mayor ponderación, aunque vuelve a cometer errores. El clasificador robusto se construye teóricamente a partir de los 3 clasificadores débiles.

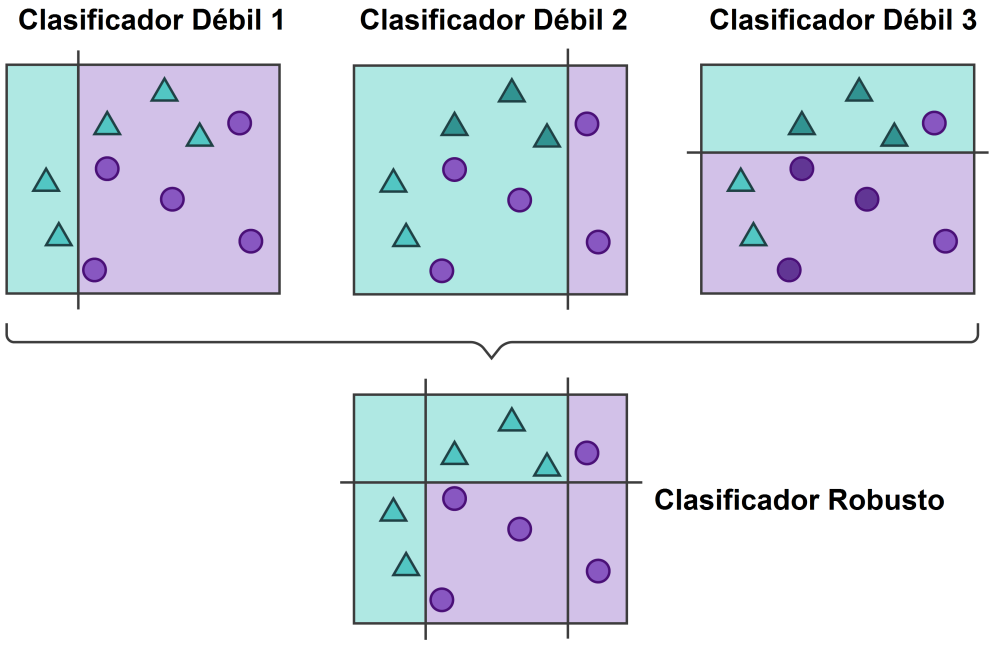


Figura 2.26: AdaBoost en un ejemplo bidimensional con dos clases.
Fuente: Adaptada de [48].

Gradient Boosting

Gradient Boosting pertenece a los algoritmos de aprendizaje combinado, el cual típicamente utiliza árboles de decisión como clasificadores base. La cantidad de clasificadores base o débiles es un hiperparámetro del algoritmo. Como su nombre lo indica, utiliza el boosting para generar un clasificador robusto a partir de una serie de clasificadores débiles, de forma iterativa, de acuerdo a la cantidad de clasificadores débiles.

Gradient Boosting opera minimizando una función de pérdida asociada a la clasificación. Lo anterior se lleva a cabo a través del método del gradiente o de máximo descenso. Las iteraciones se llevan a cabo agregando árboles de decisión que se van acoplando secuencialmente. Cada árbol de decisión se parametriza y luego se modifican sus parámetros en la dirección asignada por el método del gradiente, de modo que cada árbol agregado permite disminuir la función de pérdida. A continuación se describen las etapas de funcionamiento de Gradient Boosting [21, 46, 43].

1. Establecer como modelo una predicción inicial de las clases de las instancias de entrenamiento con una constante.
2. Para cada árbol que se agrega:
   (a) Calcular los pseudo residuos, correspondientes a una forma de cuantificar la diferencia entre la clase real y la predicha de cada instancia.
   (b) Ajustar el árbol agregado para predecir los pseudo residuos generados, con el conjunto de entrenamiento.
   (c) Mediante la técnica de optimización encontrar la modificación de los parámetros del árbol agregado.
   (d) Actualizar el modelo, sumando el árbol optimizado a la predicción inicial.
3. Entregar como modelo final la predicción inicial con la suma de todos los árboles agregados, cuya cantidad se fija como hiperparámetro.

Capítulo 3

Descripción de Datos

3.1. Tabla de Datos

Los datos utilizados en este trabajo provienen de la información recopilada y los resultados generados en la tesis doctoral de Nicolás Amigo [4]. Estos consisten en una serie de atributos o parámetros, que permiten caracterizar aneurismas cerebrales. Dicha información se recibe en forma de tabla de datos, cuya vista esquemática se observa en la Tabla 3.1. Esta tabla posee 71 filas y 16 columnas, donde las filas y columnas representan los aneurismas y los atributos, respectivamente. Los 71 aneurismas cerebrales se dividen en 36 no rotos y 35 rotos. Los atributos se pueden separar en 3 grupos, ordenados de la siguiente manera: morfológicos, hemodinámicos, y estado de ruptura. La descripción de los atributos morfológicos y hemodinámicos[3], según la nomenclatura de la tabla de datos, se entrega a continuación.

Se tienen 9 atributos morfológicos:

- H: Altura del aneurisma.
- AR: Razón de aspecto.
- SR: Razón de tamaño.
- BNF: Factor de cuello de botella.
- NSI: Índice de no esfericidad.
- UI: Índice de ondulación.
- α_A: Ángulo de aneurisma.
- α_F: Ángulo de flujo.
- α_V: Ángulo de arteria.

[3]En este trabajo el conjunto de atributos morfológicos y hemodinámicos se denomina como atributos originales.

Se tienen 6 atributos hemodinámicos:

- $SWSS_n$: Esfuerzo de corte en la pared sistólico normalizado.
- $DWSS_n$: Esfuerzo de corte en la pared diastólico normalizado.
- $TAWSS_n$: Esfuerzo de corte en la pared promediado en el tiempo normalizado.
- OSI: Índice de corte oscilatorio.
- RRT_n: Tiempo de residencia relativo normalizado.
- AFI: Índice de formación de aneurisma.

Tabla 3.1: Esquema de la tabla de datos.

H	AR	SR	BNF	NSI	UI	α_A	α_F	α_V	$SWSS_n$	$DWSS_n$	$TAWSS_n$	OSI	RRT_n	AFI	Ruptura
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Fuente: Elaboración propia.

3.2. Definición de Atributos

En esta sección se presenta la definición de cada uno de los atributos señalados anteriormente, de acuerdo a lo indicado por Amigo en su tesis doctoral [4]. Las definiciones expuestas se extraen a partir de la literatura.

3.2.1. Morfológicos

Altura (H)

El tamaño del aneurisma es un atributo comúnmente utilizado para caracterizar este tipo de lesiones. Se define como la altura del aneurisma, medida desde el cuello hasta el punto más alto de la dilatación.

Razón de Aspecto (AR)

La razón de aspecto da cuenta de la proporción entre el tamaño del aneurisma y su cuello. La expresión para su cálculo viene dada por

$$AR = \frac{H}{N}, \tag{3.1}$$

donde N representa el diámetro del cuello.

Razón de Tamaño (SR)

La razón de tamaño da cuenta de la proporción entre el tamaño del aneurisma y el tamaño de la arteria alimentadora anterior a este. La expresión para su cálculo viene dada por

$$SR = \frac{H}{D}, \tag{3.2}$$

donde D representa el diámetro de la arteria alimentadora en el segmento anterior al aneurisma.

Factor Cuello de Botella (BNF)

El factor cuello de botella da cuenta de la proporción entre el ancho del aneurisma y su cuello. La expresión para su cálculo viene dada por

$$BNF = \frac{W}{N}, \tag{3.3}$$

donde W representa el ancho del aneurisma.

Índice de No Esfericidad (NSI)

El índice de no esfericidad captura la desviación del aneurisma con respecto a un hemisferio perfecto, siendo cero para una semiesfera perfecta. La expresión para su cálculo viene dada por

$$NSI = 1 - (18\pi)^{1/3}\frac{V^{2/3}}{A}, \tag{3.4}$$

donde V representa el volumen del aneurisma y A representa su área.

Índice de Ondulación (UI)

El índice de ondulación captura el grado de concavidad de la superficie del aneurisma. La expresión para su cálculo viene dada por

$$UI = 1 - \frac{V}{V_{ch}}, \tag{3.5}$$

donde V_{ch} representa el volumen convexo.

Ángulo de Aneurisma (α_A)

El ángulo de aneurisma se define como el ángulo entre el cuello del aneurisma y su altura o tamaño.

Ángulo de Flujo (α_F)

El ángulo de flujo se define como el ángulo entre la altura del aneurisma y el eje de la arteria que lo alimenta.

Ángulo de Arteria (α_V)

El ángulo de arteria se define como el ángulo entre el eje de la arteria alimentadora y el plano del cuello del aneurisma.

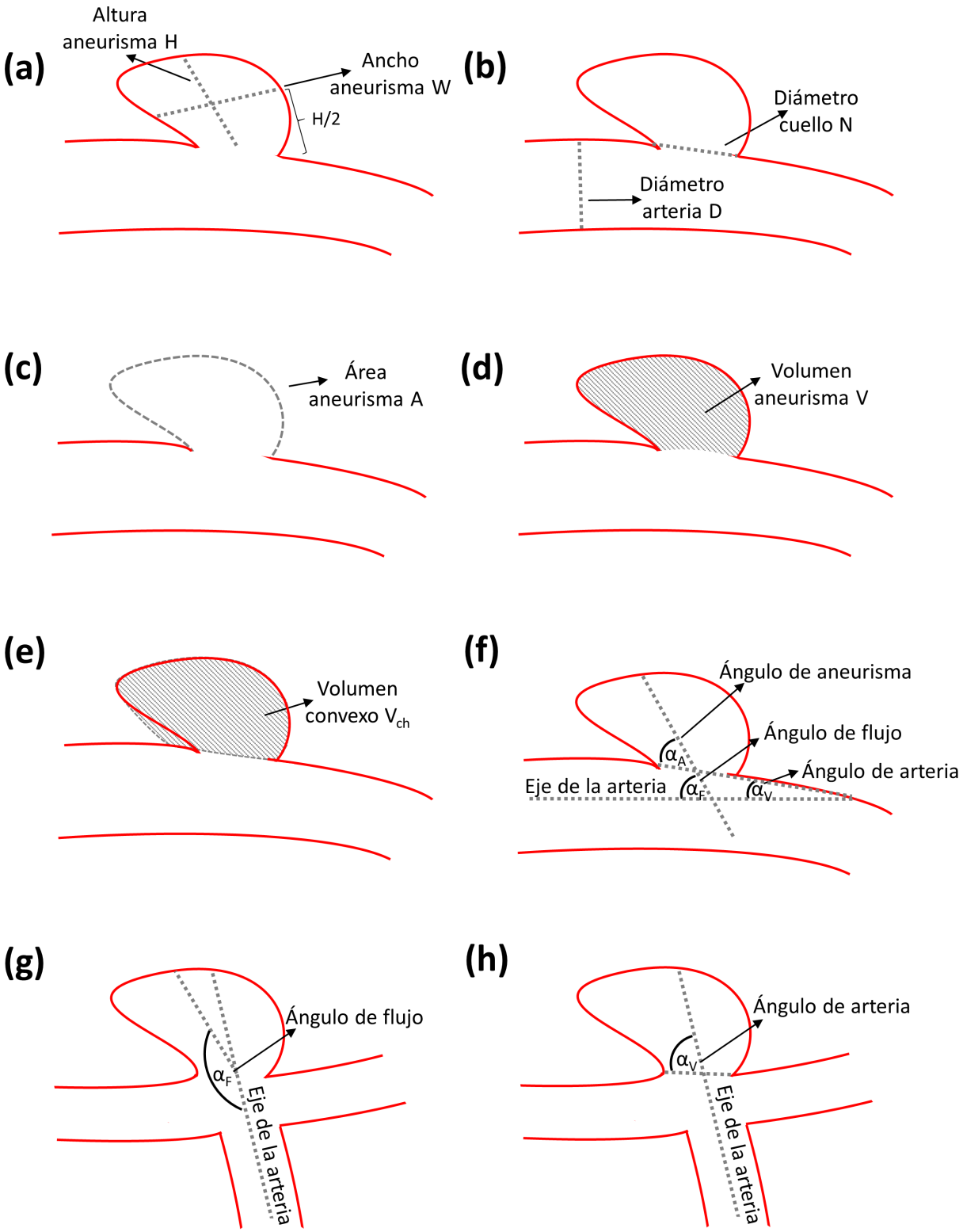


Figura 3.1: Representación esquemática de los atributos morfológicos.
Fuente: Adaptada de [4].

3.2.2. Hemodinámicos

Esfuerzo de Corte en la Pared Diastólico Normalizado ($\mathbf{DWSS_n}$)

Primeramente, se define el esfuerzo de corte en la pared diastólico como el esfuerzo de corte en la pared promediado en la superficie del aneurisma, en el mínimo del ciclo cardíaco, instante en que ocurre la diástole. La expresión para su cálculo viene dada por

$$DWSS = \frac{1}{A_a} \int_{A_a} WSS(\vec{x}, t_{min})\, dA_a, \tag{3.6}$$

donde A_a representa el área de la superficie del aneurisma. A partir de esto, se define el esfuerzo de corte en la pared diastólico normalizado, como la razón entre el esfuerzo de corte en la pared diastólico y el esfuerzo de corte en la pared promediado en el área de la arteria alimentadora anterior al aneurisma, en el instante de la diástole. La expresión para su cálculo viene dada por

$$DWSS_n = \frac{DWSS}{\frac{1}{A_v} \int_{A_v} WSS(\vec{x}, t_{min})\, dA_v}, \tag{3.7}$$

donde A_v representa el área de la arteria alimentadora en el segmento anterior al aneurisma. Esto permite adimensionalizar el esfuerzo de corte en la pared, lo cual lo vuelve un parámetro comparable entre distintos aneurismas.

Esfuerzo de Corte en la Pared Sistólico Normalizado ($\mathbf{SWSS_n}$)

De forma análoga al caso anterior, se define en primer lugar el esfuerzo de corte en la pared sistólico como el esfuerzo de corte en la pared promediado en la superficie del aneurisma, en el máximo del ciclo cardíaco, instante en que ocurre la sístole. La expresión para su cálculo viene dada por

$$SWSS = \frac{1}{A_a} \int_{A_a} WSS(\vec{x}, t_{max})\, dA_a. \tag{3.8}$$

A partir de esto, se define el esfuerzo de corte en la pared sistólico normalizado, como la razón entre el esfuerzo de corte en la pared sistólico y el esfuerzo de corte en la pared promediado en el área de la arteria alimentadora anterior al aneurisma, en el instante de la sístole. La expresión para su cálculo viene dada por

$$SWSS_n = \frac{SWSS}{\frac{1}{A_v} \int_{A_v} WSS(\vec{x}, t_{max})\, dA_v}. \tag{3.9}$$

Esfuerzo de Corte en la Pared Promediado en el Tiempo Normalizado ($\mathbf{TAWSS_n}$)

En primera instancia, se define el esfuerzo de corte en la pared del aneurisma como el esfuerzo de corte en la pared promediado en la superficie del aneurisma en un instante t cualquiera del ciclo cardíaco. La expresión para su cálculo viene dada por

$$WSS_a = \frac{1}{A_a} \int_{A_a} WSS(\vec{x}, t)\, dA_a. \tag{3.10}$$

Seguidamente, se define el esfuerzo de corte en la pared promediado en el tiempo. La expresión para su cálculo viene dada por

$$TAWSS = \frac{1}{T} \int_0^T |WSS_a|\ dt, \tag{3.11}$$

donde T es el tiempo total de un ciclo cardíaco. A partir de lo anterior, se define el esfuerzo de corte en la pared promediado en el tiempo normalizado, como la razón entre el esfuerzo de corte en la pared promediado en el tiempo y el esfuerzo de corte en la pared promediado en el tiempo en la superficie de una sección de la arteria alimentadora anterior al aneurisma. La expresión para su cálculo viene dada por

$$TAWSS_n = \frac{TAWSS}{\frac{1}{T}\int_0^T |WSS_v|\ dt}. \tag{3.12}$$

Índice de Corte Oscilatorio (OSI)

El índice de corte oscilatorio compara el promedio del vector WSS_a con el promedio de la magnitud del mismo vector. La expresión para su cálculo viene dada por

$$OSI = \frac{1}{2}\left(1 - \frac{\left|\int_0^T WSS_a\, dt\right|}{\int_0^T |WSS_a|\ dt}\right). \tag{3.13}$$

Tiempo de Residencia Relativo Normalizado ($\mathbf{RRT_n}$)

El tiempo de residencia relativo es una medida del tiempo que residen las partículas de la sangre en el endotelio. La expresión para su cálculo viene dada por

$$RRT = \frac{1}{(1 - 2 \times OSI) \times TAWSS}. \tag{3.14}$$

A partir de lo anterior, se define el tiempo de residencia relativo normalizado como la razón entre el tiempo de residencia relativo y el tiempo de residencia relativo en la superficie de una sección de la arteria alimentadora anterior al aneurisma. La expresión para su cálculo viene dada por

$$RRT_n = \frac{RRT}{[(1 - 2 \times OSI_v) \times TAWSS_v]^{-1}}, \tag{3.15}$$

donde OSI_v y $TAWSS_v$ representan el índice de corte oscilatorio y el esfuerzo de corte en la pared promediado en el tiempo, respectivamente, en la superficie de una sección de la arteria alimentadora anterior al aneurisma.

Índice de Formación de Aneurisma (AFI)

El índice de formación de aneurisma describe la diferencia entre la orientación local del vector WSS_a instantáneo y la dirección del vector $TAWSS_a$ ignorando su magnitud. La expresión para su cálculo viene dada por

$$AFI = \frac{WSS_a \cdot TAWSS}{|WSS_a|\,|TAWSS|}. \tag{3.16}$$

3.2.3. Estado de Ruptura

El estado de ruptura consiste en la variable objetivo de los aneurismas cerebrales en el contexto del problema de clasificación. Cabe decir que el estado de ruptura no abarca información sobre la evolución en el tiempo, sino que se limita a la caracterización del aneurisma en un instante particular. Para representar el estado de ruptura se utiliza una variable binaria. Los casos no rotos corresponden a los negativos y los casos rotos corresponden a los positivos. A continuación se muestra el significado del valor de esta variable.

- 0: No roto.
- 1: Roto.

3.3. Análisis Estadístico

En esta sección se presentan los resultados del análisis estadístico de los datos referentes a la caracterización morfológica y hemodinámica de los aneurismas cerebrales considerados en la tesis doctoral de Amigo. El análisis realizado consta de tres partes: estadística descriptiva e inferencial univariada, análisis ROC, y estadística multivariada [4].

3.3.1. Estadística Descriptiva e Inferencial Univariada

Los aneurismas cerebrales estudiados se separan en dos grupos de acuerdo al estado de ruptura. En la Tabla 3.2 se presenta el promedio, la desviación estándar, y la significancia estadística según la prueba de la suma de rangos de Wilcoxon de los atributos morfológicos y hemodinámicos.

Tabla 3.2: Media, desviación estándar y significancia estadística de los atributos.

	Media No Rotos	Media Rotos	p-valor
H	4,55±2,48	5,79±1,99	0,026
AR	1,37±0,77	1,66±0,56	0,010
SR	1,53±0,84	2,53±1,05	<0,001
BNF	1,22±0,43	1,63±0,76	0,010
NSI	0,31±0,05	0,29±0,06	0,192
UI	0,14±0,09	0,10±0,06	0,032
α_A	95,56±19,25	94,35±23,17	0,769
α_F	107,76±32,75	128,08±28,51	0,010
α_V	20,40±22,61	34,54±29,09	0,014
$DWSS_n$	0,71±0,44	0,44±0,34	0,004
$SWSS_n$	0,92±0,58	0,59±0,48	0,005
$TAWSS_n$	0,77±0,48	0,51±0,39	0,017
RRT_n	4,92±8,71	6,71±8,00	0,003
OSI	0,01±0,01	0,02±0,01	0,059
AFI	0,98±0,02	0,97±0,02	0,059

Fuente: Adaptada de [4].

Las conclusiones relativas a estos datos se resumen a continuación.

- Los aneurismas rotos poseen una forma más asimétrica que los no rotos, en virtud del promedio de AR, SR, y BNF en ambos grupos.
- El mismo análisis respecto de α_F y α_V indica que los aneurismas rotos se encuentran sometidos a flujos sanguíneos de mayor velocidad.
- $SWSS_n$, $DWSS_n$, y $TAWSS_n$ son menores en promedio en el grupo roto.
- RRT_n es mayor en el grupo roto, de donde se infiere que las partículas de sangre permanecen mayor tiempo en el endotelio.
- De los 15 parámetros en total, 11 son estadísticamente significativos ($p<0,05$). Los parámetros de mayor significancia son SR ($p<0,001$), junto con $SWSS_n$, $DWSS_n$ y RRT_n ($p<0,01$).
- Los atributos de menor significancia estadística son NSI y α_A.

3.3.2. Análisis ROC

Mediante la AUC de las curvas ROC para cada atributo morfológico y hemodinámico, se cuantifica su eficiencia para discriminar el estado de ruptura de los aneurismas cerebrales y se obtiene el valor umbral ideal de cada uno. Estos resultados se exhiben en la Tabla 3.3.

Tabla 3.3: Eficiencia (AUC) y umbral ideal de los atributos.

	AUC	Umbral Ideal
H	65 %	3,70
AR	68 %	1,44
SR	78 %	1,66
BNF	68 %	1,30
NSI	59 %	0,32
UI	65 %	0,14
α_A	52 %	94,64
α_F	68 %	116,28
α_V	67 %	12,55
$\mathrm{DWSS_n}$	70 %	0,75
$\mathrm{SWSS_n}$	69 %	0,99
$\mathrm{TAWSS_n}$	67 %	0,91
$\mathrm{RRT_n}$	71 %	2,68
OSI	63 %	0,01
AFI	63 %	0,99

Fuente: Adaptada de [4].

Las conclusiones relativas a este análisis se resumen a continuación.

- SR alcanza la mayor eficiencia, con una AUC de 78 %.
- En segunda instancia, en cuanto a eficiencia, se encuentran $\mathrm{RRT_n}$, $\mathrm{DWSS_n}$ y $\mathrm{SWSS_n}$, con AUC de 71 %, 70 % y 69 % respectivamente.
- NSI y α_A exhiben la peor eficiencia, con AUC de 59 % y 52 % respectivamente.

3.3.3. Estadística Multivariada

Este análisis se ejecuta mediante la aplicación de una regresión logística multivariada. Para ello, la regresión se realiza en tres casos: atributos morfológicos, atributos hemodinámicos, y el conjunto total de atributos, considerando los atributos significativos en cada caso. Así se generan tres modelos, los cuales se concretan mediante una expresión para el ODD, que representa la fracción entre la probabilidad de ruptura y la probabilidad de no ruptura. Además, se determina un modelo clínico mediante la regresión, utilizando únicamente la altura del aneurisma, dado que el tamaño del aneurisma es la métrica clásicamente utilizada en la práctica clínica. Igualmente, se calcula la AUC de cada modelo. La Tabla 3.4 muestra los resultados de la estadística multivariada.

Tabla 3.4: Modelos para evaluar el riesgo de ruptura y sus eficiencias predictivas.

Nombre	Expresión ODD	AUC
Morfológico puro	$e^{-2,26+1,13SR}$	78 %
Hemodinámico puro	$e^{1,06-1,98DWSS_n}$	70 %
Mixto	$e^{-1,22+1,16SR-1,98DWSS_n}$	82 %
Clínico	$e^{-1,29+0,26H}$	65 %

Fuente: Adaptada de [4].

Las conclusiones relativas al análisis multivariado se señalan a continuación.

- El modelo mixto resulta ser el más eficiente, exhibiendo una AUC de 82 %, reforzando la relación entre la ruptura y una interacción entre la morfología y hemodinámica de los aneurismas cerebrales.
- En el modelo morfológico puro se rescata únicamente el parámetro SR. De la expresión ODD, un aumento unitario en SR implica un aumento del riesgo de ruptura de 3,1 veces.
- Del modelo hemodinámico puro se rescata únicamente DWSS_n. De la expresión ODD, un aumento unitario de este atributo implica una disminución del riesgo de ruptura en 7,24 veces.
- El modelo mixto rescata SR y DWSS_n. El aumento unitario de estos parámetros induce una variación en el riesgo de ruptura prácticamente equivalente al que cada uno induce en los modelos morfológico y hemodinámico puros, respectivamente.
- En el modelo clínico, el aumento unitario de H gatilla un aumento de 1,3 veces en el riesgo de ruptura.

3.4. Atributos Adicionales

Además de los atributos morfológicos y hemodinámicos, los registros del trabajo de Amigo incluyen otros atributos, ya sea de forma directa o indirecta, para cada uno de los aneurismas cerebrales considerados [4]. A continuación se entrega el detalle de estos atributos adicionales.

Edad

Para cada aneurisma cerebral, se tiene la edad del paciente que lo aloja. La edad corresponde a un atributo numérico. Si bien en estricto rigor la edad es un valor continuo, en términos prácticos se utiliza como un valor discreto.

Sexo

Para cada aneurisma cerebral, se tiene el sexo del paciente que lo aloja. El sexo corresponde a una variable categórica binaria. Las posibles categorías para este atributo son: hombre o mujer.

Tipo

El tipo de aneurisma cerebral se refiere a la clasificación de las lesiones según su posición respecto de la arteria alimentadora. Este atributo es de tipo categórico, con 3 posibles clases: latera, lateral bifurcación, o terminal.

Circulación

La circulación corresponde a un atributo categórico, que indica de forma general la ubicación de los aneurismas cerebrales en la red cerebrovascular. Posee 4 posibles categorías: anterior, carótida, media, o posterior. Cada categoría representa la zona específica señalada.

Ubicación (Diagnóstico)

La ubicación de acuerdo al diagnóstico constituye un atributo categórico. Esta variable describe la ubicación de cada aneurisma cerebral dentro de la red vascular de forma detallada. A continuación se indican las 9 posibles categorías.

- AChA: Arteria coroidea anterior.
- ACOM: Arteria comunicante anterior.
- BA: Arteria basilar.
- ICA: Arteria carótida interna.
- MCA: Arteria cerebral media.
- PA: Arteria pericallosa.
- PCA: Arteria cerebral posterior.
- PCOM: Arteria comunicante posterior.
- PICA: Arteria cerebelosa inferior posterior.

Ubicación (Koivisto)

La ubicación según la clasificación de Koivisto indica el sitio de ocurrencia de cada aneurisma cerebral de forma general. Corresponde a un atributo categórico. Koivisto establece una clasificación de la ubicación de las lesiones dentro de la red cerebrovascular, agrupando en 4 categorías posibles los segmentos arteriales usuales [66]. A continuación se detallan las categorías para este atributo, y los segmentos arteriales que cada una agrupa.

- ACA: Arteria cerebral anterior, arteria comunicante anterior, y arteria pericallosa.
- ICA: Arteria carótida interna, arteria oftálmica, arteria comunicante posterior, y arteria coroidea anterior.
- MCA: Arteria cerebral media.
- VBA: Arterias vertebrobasilares.

Multiplicidad

La multiplicidad de cada aneurisma cerebral se determina al considerar la existencia o no de más de un aneurisma en el mismo paciente. Corresponde a un atributo categórico binario. Las posibles categorías son: sí o no.

Capítulo 4

Revisión del Estado del Arte

4.1. Parámetros Predictores de la Ruptura

4.1.1. Morfológicos

Dhar *et al.* realizan un estudio con 25 aneurismas cerebrales no rotos y 20 rotos. En este se encuentra que los parámetros SR, UI, NSI, índice de elipticidad, α_A, y AR poseen una diferencia estadísticamente significativa entre las medias de ambos grupos. Además, se observa que la SR y UI poseen la mayor correlación independiente con los aneurismas rotos. Del análisis mediante curva ROC, se encuentran las mayores AUC para SR y α_A, con 83 % y 85 % respectivamente [33]. Zheng *et al.* publican un artículo en 2016, el cual analiza un grupo de 82 aneurismas no rotos y 68 rotos. Los parámetros que exhiben una correlación independiente de mayor intensidad con el estado roto de los aneurismas son SR, la tasa altura-ancho, α_A, la forma del aneurisma (regular, irregular y con saco hijo) y la ubicación. El análsis ROC indica que SR y α_F entregan las mayores AUC, de 73,5 % y 73 % respectivamente [157]. A su vez, el estudio de Lin *et al.* señala SR, α_F y el ángulo entre las arterias alimentadora y expulsora como parámetros asociados al estado roto de los aneurismas cerebrales [76].

Rahman *et al.* concluyen que SR se correlaciona fuertemente con el estado roto de los aneurismas cerebrales. Además, encuentran que el tamaño promedio máximo y SR son significativamente menores en los aneurismas no rotos [111]. Los resultados obtenidos por Abboud *et al.* en un estudio de 420 aneurismas cerebrales sugieren que la morfología es un predictor de la ruptura independiente. Específicamente, se encuentra que el riesgo de ruptura aumenta progresivamente para los siguientes casos: contorno irregular, existencia de un saco hijo, y saco multilobulado. Por su parte, con la morfología se obtiene una AUC de la curva ROC de 69,3 % con un p-valor menor a 0,001 [1]. Bhogal *et al.* en su estudio de 113 pacientes hallan una diferencia estadísticamente significativa en el tamaño entre aneurismas cerebrales no rotos y rotos, siendo mayor para los últimos. Junto con esto, tanto AR como BNF son significativamente mayores en los aneurismas rotos. A su vez, en los aneurismas no rotos predomina el contorno regular, siendo lo contrario en los rotos [15]. En el año 2018, Duan *et al.* publican un estudio sobre las ubicaciones y los parámetros morfológicos asociados con el estado de ruptura de aneurismas cerebrales pequeños, para un total de 135 aneurismas no rotos y 128 rotos. Los resultados indican que los parámetros de mayor significancia estadística, en orden creciente, son AR, SR, la tasa altura-ancho, y α_F [34]. Igualmente, Skodvin *et al.* encuentran que α_F posee una diferencia estadísticamente significativa en el análisis univariado [128].

4.1.2. Hemodinámicos

De acuerdo al estudio de Xiang *et al.*, el WSS y el OSI aparecen como variables significativas de forma independiente, dentro de otros parámetros hemodinámicos, para discriminar el estado de ruptura de aneurismas cerebrales [153]. Dados los hallazgos controvertidos acerca del WSS y su relación con la ruptura de los aneurismas cerebrales, Zhou *et al.* llevan a cabo una revisión sistemática y meta-análisis de este tópico. Sus hallazgos indican que los aneurismas rotos poseen una tasa significativamente mayor de bajos WSS. A su vez, el WSS promedio es significativamente menor en el grupo roto, considerando los casos de bajo WSS [158]. Según un estudio de Cebral *et al.*, los aneurismas cerebrales rotos son más propensos a presentar patrones de flujo complejos e inestables, flujo entrante concentrado, y regiones de impacto pequeñas, en comparación con los no rotos [28].

Por su parte, Chung *et al.* obtienen como resultados de su estudio que los aneurismas cerebrales rotos poseen significativamente un menor WSS mínimo, un mayor WSS máximo, esfuerzos de corte más oscilantes, mayor velocidad máxima y flujos más complejos, comparados con los no rotos [30]. En el año 2017, Qin *et al.* encuentran que el porcentaje de área de bajo WSS es significativamente mayor cuando existe ruptura, para aneurismas de bifurcación ubicados en la arteria cerebral media [109]. Jiang *et al.* en su análisis de 334 aneurismas cerebrales, con igual cantidad de no rotos y rotos, hallan que tanto OSI como RRT son significativamente mayores en el grupo roto [62]. Asimismo, el artículo de Miura *et al.* señala que mediante análisis univariado, el WSS, su variante normalizada y su gradiente, OSI y AFI son estadísticamente significativos para discriminar el estado de ruptura [86].

Qian *et al.* estudian un paramétro hemodinámico llamado pérdida de energía. Este corresponde a la energía perdida en el flujo producto del aneurisma cerebral, por lo cual se determina al contrastar el escenario sin y con aneurisma cerebral en la arteria considerada. Los resultados de su estudio indican que la pérdida de energía estimada en aneurismas cerebrales que sufrieron ruptura es cerca de 5 veces mayor en promedio que en los aneurismas estables. Es por esto que concluyen que la pérdida de energía podría establecerse como un parámetro hemodinámico útil para la cuantificación del riesgo de ruptura [108].

4.2. Machine Learning y Ruptura

En su estudio reportado el año 2011, Bisbal *et al.* realizan la predicción del estado de ruptura de aneurismas cerebrales utilizando clasificadores de machine learning, considerando atributos hemodinámicos, morfológicos y clínicos. El conjunto de datos empleado consiste en 157 instancias o aneurismas cerebrales y 71 atributos distintos. Se aplican técnicas de extracción de atributos y selección, entre otras. Como algoritmos de clasificación se prueban paralelamente los siguientes: máquina de vectores de soporte, árbol de decisión, clasificador bayesiano ingenuo, reglas difusas, redes neuronales y reglas de asociación predictivas. Para la evaluación se ocupa una validación cruzada con 10 particiones. La exactitud más elevada se obtiene con una máquina de vectores de soporte binaria, logrando 95,5 %. Este modelo alcanza una precisión de 95,2 % y 95,9 % en las clases no roto y roto, respectivamente, junto con una sensibilidad de 96,3 % y 94,7 % en cada una de dichas clases. Además, se reportan las reglas de asociación generadas por medio del algoritmo de reglas de asociación predictivas para un modelo que obtuvo 86 % de exactitud, lo cual constituye una ventaja agregada de este algoritmo, ya que permite explicitar los atributos y valores de atributos que influyen sobre el estado de ruptura. En este caso, se encuentra una relación entre el estado de ruptura y la ubicación, así como con NSI, entre otras [16]. Un estudio elaborado por Niemann *et al.* aborda el problema de clasificación del estado de ruptura de aneurismas cerebrales con machine learning, en base a 22 atributos morfológicos. Teniendo 100 instancias, establecen tres grupos de evaluación: aneurismas laterales (9 rotos de 24), aneurismas de bifurcación (29 rotos de 62) y finalmente el grupo que engloba todos los aneurismas (43 rotos de 100). Se utilizan 10 algoritmos: tres variantes de árboles de decisión, k-vecinos más cercanos, máquina de vectores de soporte con kernel lineal, red neuronal, clasificador bayesiano ingenuo, mínimos cuadrados parciales generalizado, bosque aleatorio, y gradient boosting. A su vez, se aplican 3 transformaciones de datos a modo de preprocesamiento para lidiar con la diferencia de escalas entre atributos. Para la evaluación se lleva a cabo una validación cruzada de 10 particiones estratificada, repetida 5 veces. Además, se realiza una optimización de hiperparámetros mediante búsqueda en grilla, ocupando la exactitud como métrica objetivo. Los mejores resultados de clasificación binaria se obtienen vía máquina de vectores de soporte con kernel lineal en el grupo de aneurismas laterales, con un 80 %±24 % de exactitud y una AUC de 66 % ± 12 %. En el grupo de aneurismas de bifurcación, el algoritmo de mínimos cuadrados parciales generalizado exhibe el mejor desempeño, con 68 %±16 % de exactitud y 68 % ± 2 % de AUC. Finalmente, en el grupo del total de las instancias, el mejor modelo se logra con el algoritmo gradient boosting, obteniendo 69 %±15 % de exactitud, y 70 % ± 2 % de AUC [93].

Liu *et al.* desarrollan una red neuronal artificial para predecir el riesgo de ruptura en base al estado de ruptura de aneurismas cerebrales de la arteria comunicante anterior. Teniendo 594 ejemplos, 54 no rotos y 540 rotos, se diseña una red neuronal de dos capas. Para compensar el desbalance de clases, se utiliza un método para generar más instancias de aneurismas no rotos de forma artificial. Como atributos se incorporaron 13 parámetros morfológicos provenientes de imágenes médicas, dos factores demográficos, y el historial fumador y de hipertensión. Se hace una separación aleatoria de los datos en conjuntos de entrenamiento, validación y prueba, representando el 70 %, 15 % y 15 % respectivamente. La evaluación del desempeño se basa en la AUC de la curva ROC. Así, se obtiene para los conjuntos de entrenamiento, validación, prueba y el total de datos, una AUC de 95,3 %, 93,7 %, 92,8 % y 95 % [78]. Aranda y Valencia publican en el año 2018 un estudio en donde modelan un clasificador mediante una máquina de vectores de soporte radial, valiéndose de un conjunto de 60 aneurismas cerebrales saculares, donde existen 30 no rotos y 30 rotos. Se utilizan 6 atributos o parámetros predictores: género, edad, número de Womersley, TAWSS, AR, y BNF. El conjunto de datos fue separado en 80 % para entrenamiento y 20 % para prueba. Para la validación del modelo se usa validación cruzada de 10 particiones, con 3 repeticiones. La exactitud obtenida en el conjunto de prueba es de 92,86 % [9]. Los mismos autores publican un nuevo estudio el año 2019, en donde generan datos a partir de simulaciones de interacción fluido-estructura, nuevamente para 60 aneurismas cerebrales saculares, siendo la mitad no rotos. A partir del análisis estadístico se identifican AR, BNF, la altura máxima del aneurisma, RRT, el número de Womersley y el esfuerzo de Von Mises como parámetros significativos e importantes para predecir la ruptura. Así, estos atributos se integran en 5 algoritmos de machine learning: modelo bayesiano lineal generalizado, regresión logística, bosque aleatorio, AdaBoost y red neuronal probabilística. El 60 % de los datos se emplean para el entrenamiento, y el resto para prueba. Al ajustar los algoritmos al conjunto de prueba, el mejor modelo se logra con AdaBoost, con una AUC de la curva ROC de 96,7 %. La validación del modelo se realiza con validación cruzada de 10 particiones y 3 repeticiones. Con esto, se logra una AUC de 94,4 % con AdaBoost, siendo el mejor resultado [10].

En un intento por determinar si los modelos de machine learning son capaces de distinguir entre aneurismas cerebrales rotos y no rotos e identificar atributos asociados a la ruptura, Silva *et al.* desarrollan un estudio publicado en 2019. En este trabajo se entrenan tres algoritmos: bosque aleatorio y dos máquinas de vectores de soporte, con kernel lineal y RBF, respectivamente. Por su parte, se tienen 845 aneurismas cerebrales, 309 de los cuales se encuentran en estado roto. Los aneurismas rotos son significativamente de mayor tamaño, con una diferencia promedio en torno a 1 [mm] y con mayor probabilidad de aparecer en la circulación posterior que los no rotos. Como métrica de evaluación se usa la AUC de la curva ROC. Las máquinas de vectores de soporte lineal y RBF obtienen un AUC de 77 % y 78 % respectivamente. En cuanto al bosque aleatorio, este logra un AUC de 81 %. La ubicación y el tamaño del aneurisma cerebral destacan como los dos atributos de mayor contribución al modelo [127]. En Japón, Suzuki *et al.* construyen un modelo de clasificación del estado de ruptura de aneurismas cerebrales, como aparece en su estudio del año 2019. El grupo de investigadores señala que el foco del estudio es un método de predicción de alta exactitud, que sea independiente de la experiencia y expertiz del médico. Para ello, prueban los algoritmos de regresión logística y máquina de vectores de soporte. Para el modelo se cuenta con un total de 338 aneurismas cerebrales, 35 rotos y 303 no rotos. Se aplica un algoritmo para solucionar el desbalance de las clases, junto con un método para seleccionar atributos importantes. Lo anterior permite reducir los 70 atributos originales a 27. Estos últimos se encuentran en una proporción de 40 % pertenenciente a datos médicos y 60 % proveniente de simulaciones del flujo sanguíneo. Con la regresión logística se logra un modelo con una sensibilidad de 63,6 %, una tasa de negativos verdaderos de 84,6 %, y un valor F1 de 43,7 %, mientras que la máquina de vectores de soporte otorga una sensibilidad de 45,5 %, una tasa de negativos verdaderos de 91,2 %, y un valor F1 de 41,7 % [133].

Capítulo 5

Metodología

5.1. Obtención de Datos

El primer paso requerido para el desarrollo del trabajo es la obtención de datos referentes a la caracterización morfológica y hemodinámica de aneurismas cerebrales. Lo anterior se realiza mediante la búsqueda de un conjunto de datos construido en trabajos o estudios previos, que cumpla con la caracterización estipulada. Una vez obtenidos los datos, se procede a realizar algunas tareas iniciales. Estos datos se encuentran en dos archivos que los organizan a manera de tablas, en formatos CSV y XLSX, respectivamente. El archivo CSV incluye, dentro de otros datos, los atributos morfológicos y hemodinámicos, junto con el estado de ruptura de todos los aneurismas cerebrales considerados. Por su parte, el archivo XLSX incluye, dentro de otros datos, atributos adicionales, además de los presentes en el archivo CSV, de todos los aneurismas cerebrales considerados. En primera instancia, se verifica que los tipos de archivo puedan ser leídos mediante las librerías del lenguaje de programación utilizado. Luego de tener los datos cargados y poder manejarlos en el entorno de programación, se quitan ciertas columnas y se eligen otras, dependiendo del archivo, con el fin de estructurar y definir los conjuntos de datos básicos a utilizar en las etapas siguientes. Estos se detallan a continuación, junto con el nombre de la variable utilizada para guardar la información.

- *df*: Tabla de datos que lee y guarda la información del archivo CSV.
- *df_1*: Tabla de datos que lee y guarda la información del archivo XLSX, con el fin de agregar columnas de atributos a X, según corresponda.
- X: Tabla de datos que se extrae de *df*, con 71 filas que representan los aneurismas cerebrales y 15 columnas que abarcan los atributos morfológicos y hemodinámicos. Representa los atributos o variables independientes de cada instancia.
- y: Columna única extraída de *df*, que indica el estado de ruptura, con 71 filas que representan los aneurismas cerebrales. Representa la clase o etiqueta, también llamada variable dependiente de cada instancia.

5.2. Análisis Exploratorio de Datos

Para lograr un primer acercamiento a los atributos y extraer conclusiones preliminares, se realiza un análisis exploratorio de datos. Primeramente, se grafican histogramas normalizados junto con curvas KDE[4] y diagramas de caja y bigote de los atributos morfológicos y hemodinámicos, agrupados según el estado de ruptura, para observar las diferencias en forma de distribución, rangos, valores máximos y mínimos, y la existencia de valores perdidos, por atributo. A esto se suma la prueba de normalidad de Shapiro-Wilk, con el fin de evaluar la normalidad de los datos por atributo, lo cual es relevante para el desempeño de los algoritmos de machine learning a utilizar en etapas siguientes. La prueba de normalidad se ejecuta en tres grupos: aneurismas no rotos y rotos, y el total de aneurismas.

Seguidamente, se grafican matrices de correlación y su p-valor, para identificar atributos correlacionados y su significancia estadística. Esto se hace para la correlación de Spearman (monotónica), dado su carácter no paramétrico, lo cual permite abarcar de forma más robusta el total de atributos considerados. Cabe decir que la correlación de Spearman no es la forma ideal de correlacionar el estado de ruptura, binario, con otros atributos continuos. Sin embargo, por simplicidad en la generación de una única matriz para el total de atributos, donde únicamente el estado de ruptura no es continuo, se procede a su cálculo. El objetivo de esta matriz es identificar correlaciones entre atributos morfológicos y hemodinámicos en primera instancia, y en segundo lugar, entre dichos atributos y el estado de ruptura.

Una vez hecho lo anterior, se analizan los atributos adicionales a incorporar. La edad, correspondiente al único atributo no categórico, se analiza mediante un histograma normalizado junto con una curva KDE y un diagrama de caja y bigote, con el mismo objetivo señalado anteriormente. Sumado a esto, se lleva a cabo la prueba de normalidad de Shapiro-Wilk, de forma análoga a los atributos originales. Para el resto de los atributos adicionales, de carácter categórico, se realizan gráficos de barra para visualizar las diferencias entre categorías. Junto a esto se generan tablas de contigencia para cada uno, con el fin de identificar las frecuencias esperadas y determinar la prueba estadística correcta para evaluar su significancia estadística. En virtud de esto, se lleva a cabo la prueba de la suma de rangos de Wilcoxon para la edad, y en el caso de los atributos adicionales categóricos, la prueba exacta de Fisher, obteniendo así la significancia estadística de los atributos adicionales como discriminantes del estado de ruptura.

[4]La curva KDE (del inglés *Kernel Density Estimation*) es una estimación de la función de densidad de probabilidad de una variable aleatoria.

5.3. Selección de Algoritmos de Clasificación

Para llevar a cabo la clasificación del estado de ruptura de aneurismas cerebrales mediante machine learning, se escoge una serie de algoritmos de clasificación conocidos y ampliamente utilizados. Esto permite una fácil implementación de los algoritmos, lo cual se lleva a cabo mediante la librería Scikit-learn. A su vez, estos algoritmos de clasificación poseen una gran cantidad de literatura y recursos en línea asociados, facilitando así su comprensión y utilización, además de validar su utilidad. El objetivo de probar una serie de algoritmos de clasificación de forma paralela, es aumentar las opciones de alcanzar el nivel de exactitud deseado. Se seleccionan 9 algoritmos de clasificación, señalados a continuación.

- Regresión logística (LR)
- Análisis discriminante lineal (LDA)
- K vecinos más cercanos (KNN)
- Árbol de decisión (DT)
- Clasificador bayesiano ingenuo (NB)
- Máquina de vectores de soporte (SVM)
- Bosque aleatorio (RF)
- AdaBoost (AB)
- Gradient Boosting (GB)

5.4. Clasificación

Esta etapa consiste en la evaluación del desempeño exhibido por los 9 algoritmos de clasificación de forma simultánea, a la hora de predecir el estado de ruptura de aneurismas cerebrales. Para evaluar la capacidad de clasificación se utiliza la técnica de validación cruzada con 10 particiones. Esta técnica se escoge dado que se tienen únicamente 71 instancias o aneurismas cerebrales a clasificar, lo cual representa una baja cantidad. Como métrica de evaluación principal se escoge la exactitud, debido a que los datos se encuentran balanceados por clase. Los estados aleatorios de la validación cruzada y los algoritmos donde aplica, se fijan con el fin de obtener resultados comparables entre todos los casos evaluados.

5.4.1. Conjuntos de Atributos

En primera instancia, se definen distintos conjuntos de atributos a partir de los atributos originales. El objetivo es comparar el desempeño de los algoritmos de clasificación en estos conjuntos de atributos. Para ello se registra la exactitud promedio de la validación cruzada. La descripción de cada uno de estos conjuntos se entrega a continuación.

1. Originales: Contiene todos los atributos morfológicos y hemodinámicos.
2. Morfológicos: Contiene únicamente los atributos morfológicos.

3. Morfológicos-4: Contiene los 4 atributos morfológicos de mayor significancia estadística: SR, BNF, AR y α_F.
4. Hemodinámicos: Contiene únicamente los atributos hemodinámicos.
5. Hemodinámicos-4: Contiene los 4 atributos hemodinámicos de mayor significancia estadística: RRT_n, $DWSS_n$, $SWSS_n$ y $TAWSS_n$.
6. Aumentados: Contiene los atributos originales, a los cuales se agregan los atributos edad y sexo del paciente, tipo de aneurisma, circulación, ubicación según el diagnóstico y según las categorías de Koivisto [66], y multiplicidad de aneurismas. En este caso, los atributos categóricos se transforman a variables tipo dummy[5] dado que originalmente se encuentran como variables categóricas en forma de texto.
7. Aumentados*: Contiene los atributos originales, a los cuales se agregan los atributos adicionales estadísticamente significativos: ubicación según diagnóstico y multiplicidad de aneurismas.
8. Significativos: Contiene los cuatro atributos de mayor significancia estadística: SR, $DWSS_n$, $SWSS_n$ y RRT_n [4, 5].

5.4.2. Transformaciones de Atributos

En esta etapa se aplican 8 transformaciones conocidas al conjunto de atributos originales. Una vez hecho esto, se evalúa el desempeño de los algoritmos de clasificación en los 8 conjuntos de atributos originales transformados. Lo anterior se realiza para modificar las distribuciones de datos de cada atributo, con la expectativa de incrementar el desempeño de los algoritmos de clasificación, ampliando el espacio de búsqueda del nivel de exactitud deseado. Se registra la exactitud promedio de la validación cruzada. Las transformaciones aplicadas se señalan a continuación.

- Estandarización
- Máx-Abs
- Mín-Máx
- Normalización
- Box-Cox
- Cuantil-Normal
- Cuantil-Uniforme
- Robusta

[5]Las variables tipo dummy generan columnas según la cantidad de categorías presentes en un atributo categórico. Usan los valores 0 o 1 para indicar la presencia de la categoría.

5.4.3. Búsqueda Exhaustiva

La búsqueda exhaustiva como forma de selección de atributos se aplica sobre el conjunto de atributos originales y sus 8 transformaciones, para los 9 algoritmos de clasificación empleados. Para ello, se genera el conjunto potencia de los 15 atributos originales, de donde se obtienen 32.767 subconjuntos posibles, identificados mediante un índice. Probando todas los pares algoritmo-transformación sobre los 32.767 subconjuntos, se encuentra la exactitud máxima junto con el o los subconjuntos que la otorgan en cada caso. De esta forma, se maximiza la exactitud vía selección de atributos y se identifican los subconjuntos asociados, de mayor relevancia. Como resultado se registra la exactitud promedio de la validación cruzada. A su vez, para cada uno de los 9 algoritmos se registran las transformaciones y los subconjuntos de atributos que entregan la mayor exactitud promedio en validación cruzada.

5.4.4. Reglas de Asociación

Las reglas de asociación se determinan entre los subconjuntos de atributos hallados mediante búsqueda exhaustiva que generan modelos con una exactitud promedio en validación cruzada mayor o igual a 80 %. Se establece este valor como umbral desde el cual los modelos se consideran con una exactitud suficiente como para considerar el subconjunto de atributos asociado relevante para distinguir el estado de ruptura. Primeramente se calcula el soporte de cada uno de los 15 atributos originales dentro de los subconjuntos que cumplen la condición. Así, se vislumbra su porcentaje de aparición en los mejores subconjuntos, entregando una noción de la importancia individual de cada atributo para diferenciar el estado de ruptura, en base al resultado empírico. Lugo, se identifican las reglas de asociación con un interés mínimo de 3. Esto permite evidenciar qué atributos se asocian dentro de los mejores subconjuntos, en un nivel al menos 3 veces mayor a lo que ocurriría suponiendo independencia entre los atributos. Así, es posible hallar relaciones entre atributos morfológicos y hemodinámicos, o dentro del mismo grupo, en relación al estado de ruptura.

5.4.5. Optimización de Hiperparámetros

La optimización de hiperparámetros se aplica sobre los mejores modelos por algoritmo (con excepción de NB dado que prácticamente no posee hiperparámetros), encontrados mediante búsqueda exhaustiva. Esta se lleva a cabo utilizando una búsqueda en grilla. Los valores incorporados en cada una de las grillas de búsqueda se establecen en función de las especificaciones de los algoritmos de machine learning existentes en Scikit-learn [71]. Para esta fase se registra: la exactitud promedio de validación cruzada de los 9 mejores modelos por algoritmo hallados por búsqueda exhaustiva, su valor entregado directamente en la ejecución de la optimización de hiperparámetros, donde se ejecuta una validación cruzada de características equivalentes, y finalmente dicho valor al volver a aplicar la evaluación previa a la optimización de hiperparámetros, con los 9 modelos ya optimizados. Los dos últimos pasos se realizan de forma separada dado que ambas validaciones cruzadas poseen un componente aleatorio que no se puede compatibilizar.

5.4.6. Evaluación de Modelos Finales

En esta etapa final se realiza la evaluación de los mejores modelos aprendidos, por algoritmo, obtenidos mediante los procesos de búsqueda exhaustiva y optimización de hiperparámetros. En primera instancia, se obtiene la exactitud y la precisión de los 9 modelos finales a través de validación cruzada. Así, se registra el promedio de la exactitud y la precisión junto con su desviación estándar, en los conjuntos de entrenamiento y prueba. En segunda instancia, se realiza la evaluación de los 9 modelos finales utilizando validación cruzada estratificada. Lo anterior permite calcular (correctamente) otras métricas de evaluación del desempeño de los modelos: precisión, sensibilidad, valor F1 y AUC de la curva ROC. Así, se puede interiorizar de forma más completa el desempeño de los modelos finales y se generan más puntos de comparación entre los mismos y con lo reportado en la literatura. Para la validación cruzada estratificada se registra el promedio de la exactitud, precisión, sensibilidad, valor F1 y AUC de la curva ROC, junto con su desviación estándar, en los conjuntos de entrenamiento y prueba.

5.5. Recursos

Para llevar a cabo el trabajo descrito anteriormente, según las etapas de la metodología planteada, se utilizan los siguientes recursos:

- Computadora para llevar a cabo la programación, el manejo de datos, la implementación de los modelos de clasificación, el diseño de gráficos, la realización de cálculos, y la redacción del informe.
- Datos de la caracterización morfológica y hemodinámica de una serie de aneurismas cerebrales, junto con la identificación de su estado de ruptura.
- Entornos de programación Spyder y Google Colaboratory, para la escritura y ejecución de código utilizando el lenguaje Python.
- Librerías de Python utilizadas en machine learning, tales como Pandas, Scikit-learn, Matplotlib, Seaborn, y Scipy, entre otras.
- Software Microsoft Excel, para el manejo y la visualización de datos, la creación de gráficos, y la realización de cálculos.
- Editor de LaTeX en línea Overleaf, para la redacción del informe.

Capítulo 6

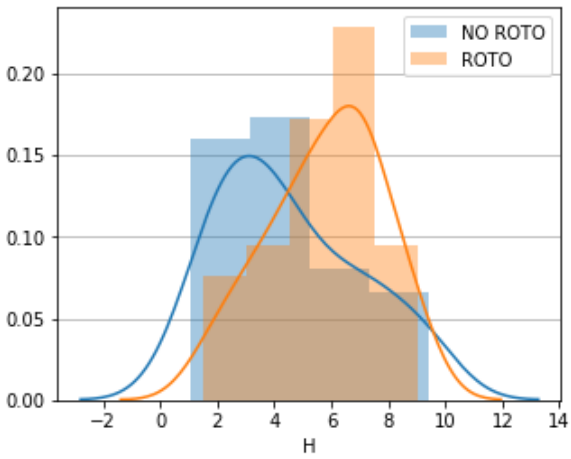
Resultados

6.1. Análisis Exploratorio de Datos

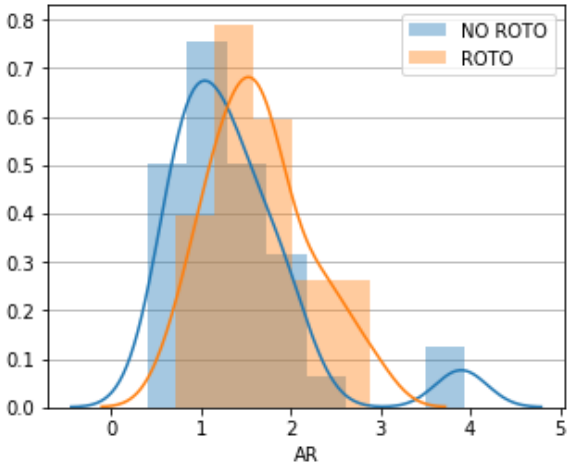
En esta sección se presentan los resultados del análisis exploratorio de datos. Lo anterior se lleva a cabo en dos partes. Primeramente se realiza la exploración de los atributos originales, y en segunda instancia se exploran los atributos adicionales.

6.1.1. Atributos Originales

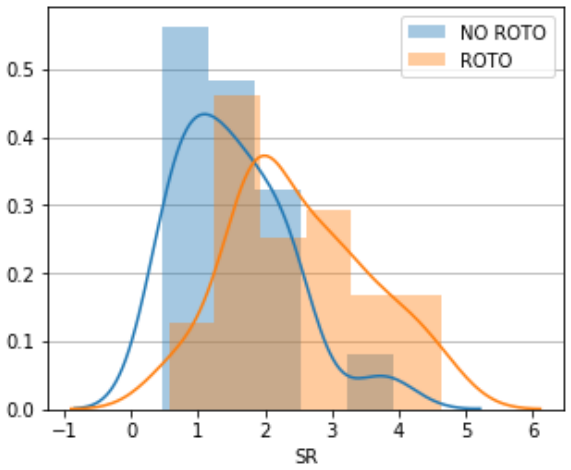
Para el análisis exploratorio de datos de los atributos originales se exhiben histogramas normalizados junto con curvas KDE de los 9 atributos morfológicos en las Figuras 6.1 y 6.2, y de los 6 atributos hemodinámicos en la Figura 6.3. Seguido a esto, se presentan diagramas de caja y bigote de los atributos morfológicos, en las Figuras 6.4 y 6.5, y hemodinámicos en las Figuras 6.6. Posteriormente se muestra en la Tabla 6.1 los p-valores de los atributos originales obtenidos en la prueba de normalidad de Shapiro-Wilk, en los grupos de aneurismas no rotos, rotos y total. Finalmente, se muestra la matriz de correlación de Spearman de los atributos originales y el estado de ruptura en la Figura 6.7, seguida de la matriz de p-valor de dicha correlación en la Figura 6.8.



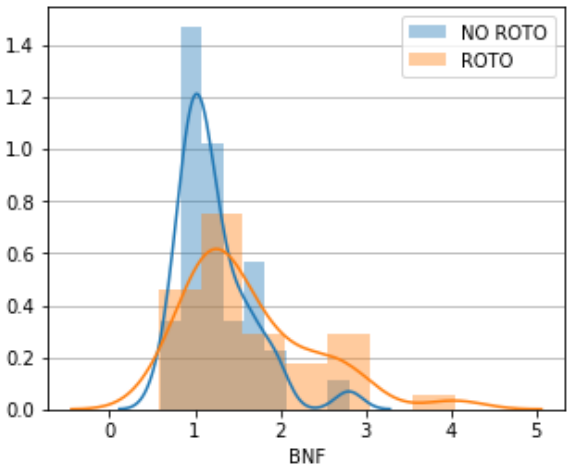
(a) Altura.



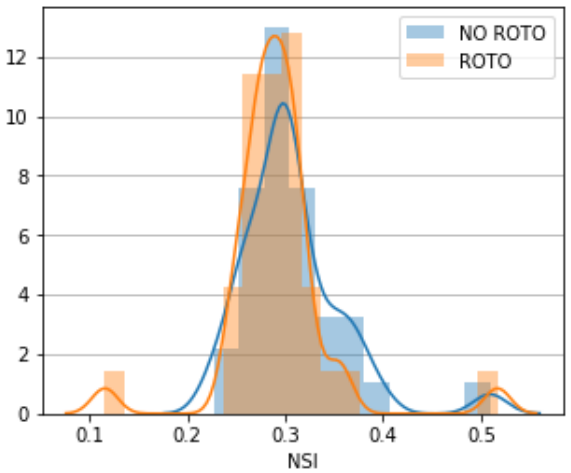
(b) Razón de aspecto.



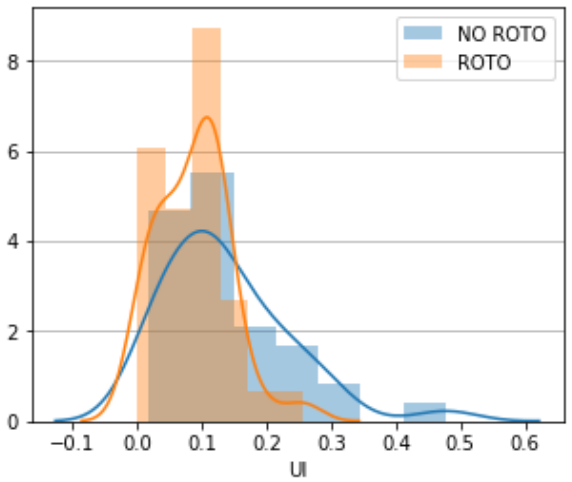
(c) Razón de tamaño.



(d) Factor cuello de botella.

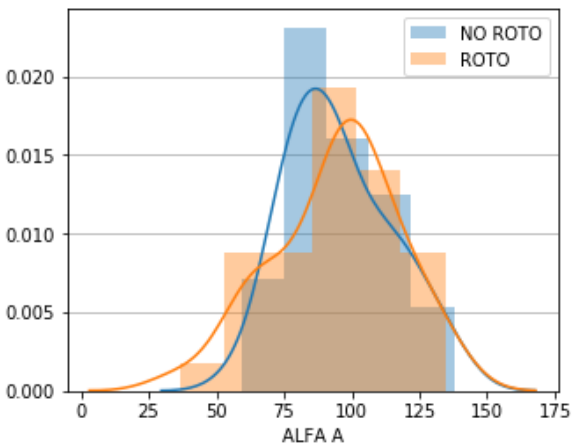


(e) Índice de no esfericidad.

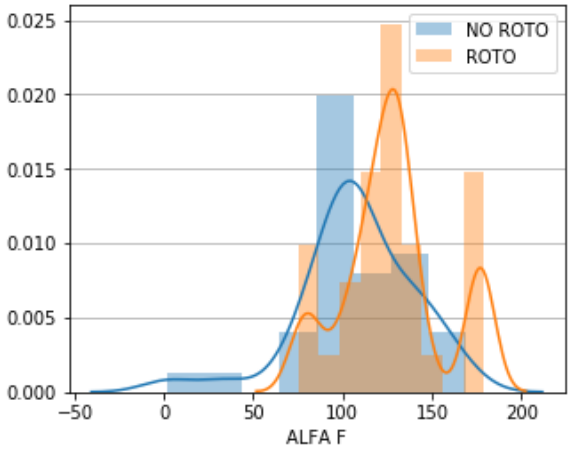


(f) Índice de ondulación.

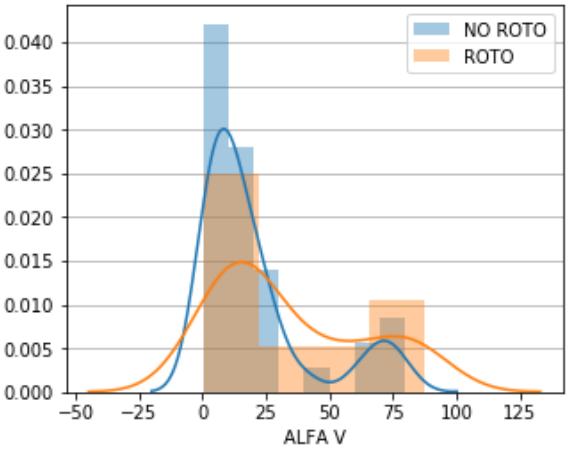
Figura 6.1: Histogramas normalizados y curvas KDE de los atributos morfológicos (1/2). Fuente: Elaboración propia.



(a) Ángulo del aneurisma.

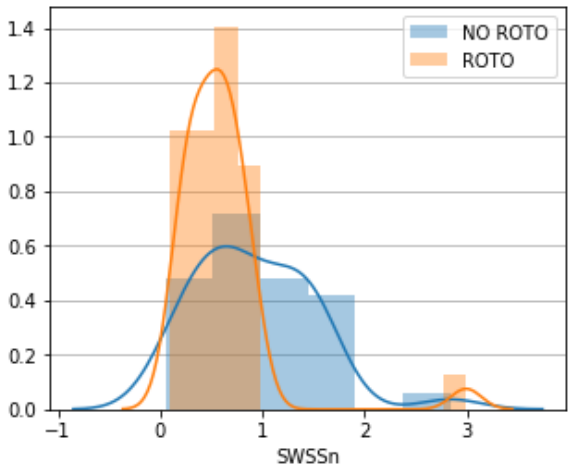


(b) Ángulo de flujo.

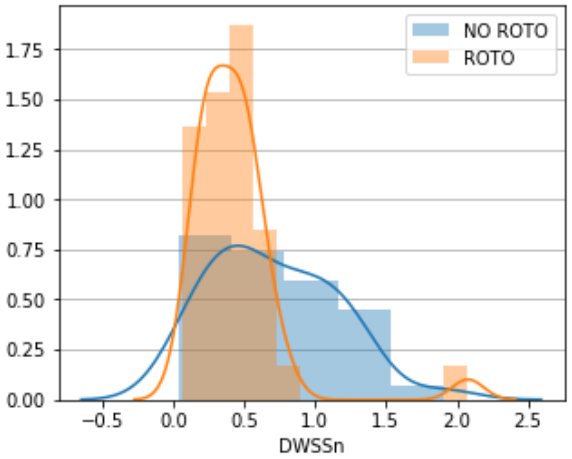


(c) Ángulo de arteria.

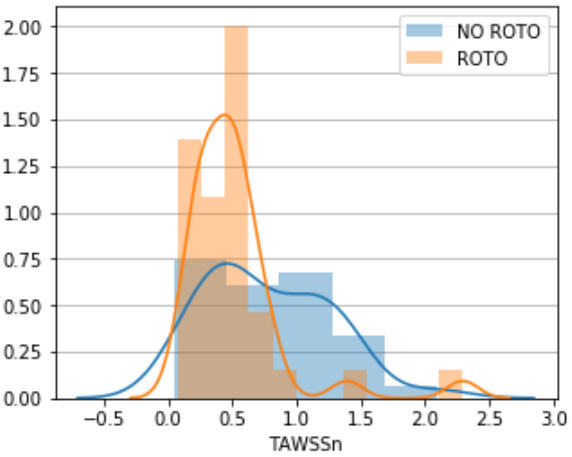
Figura 6.2: Histogramas normalizados y curvas KDE de los atributos morfológicos (2/2). Fuente: Elaboración propia.



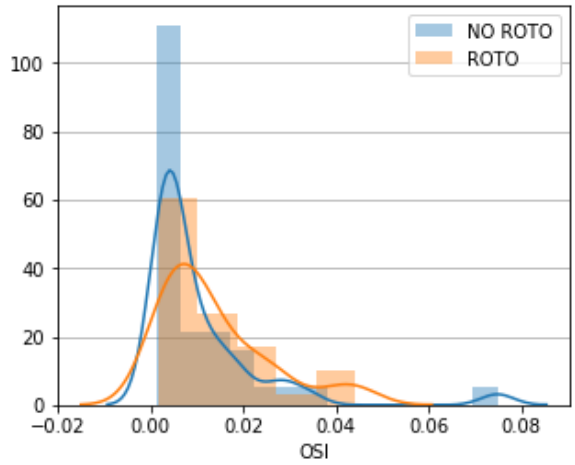
(a) Esfuerzo de corte en la pared sistólico normalizado.



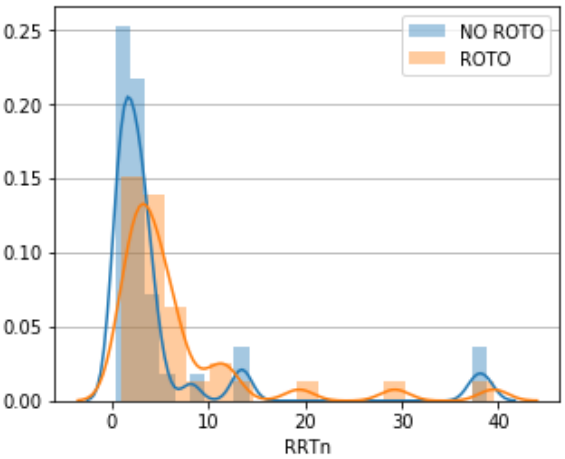
(b) Esfuerzo de corte en la pared diastólico normalizado.



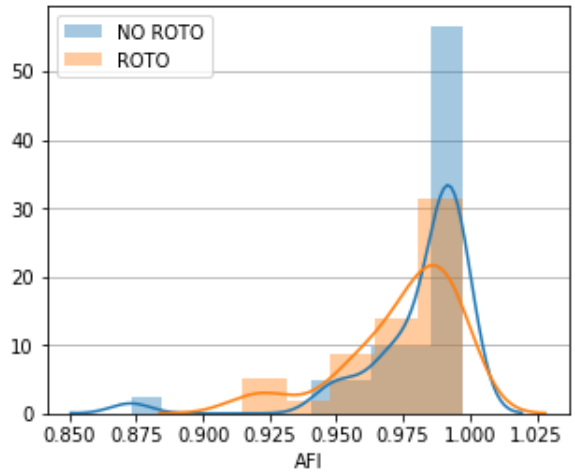
(c) Esfuerzo de corte en la pared promediado en el tiempo normalizado.



(d) Índice de corte oscilatorio.

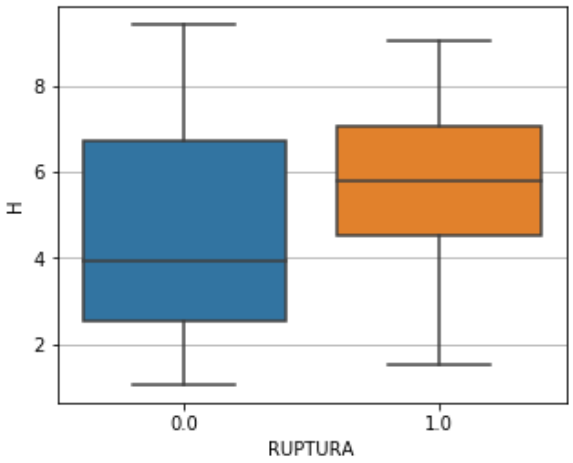


(e) Tiempo de residencia relativo normalizado.

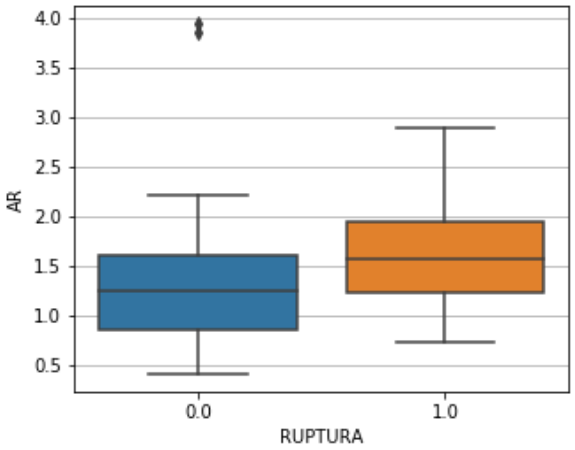


(f) Índice de formación de aneurisma.

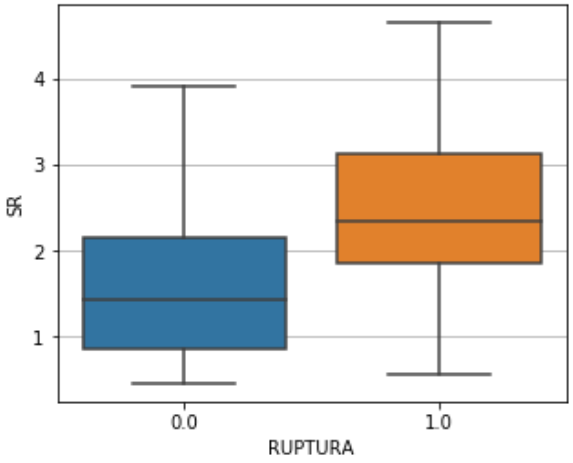
Figura 6.3: Histogramas normalizados y curvas KDE de los atributos hemodinámicos. Fuente: Elaboración propia.



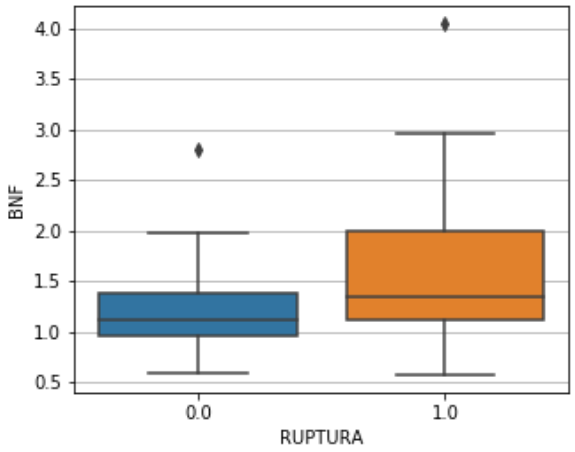
(a) Altura.



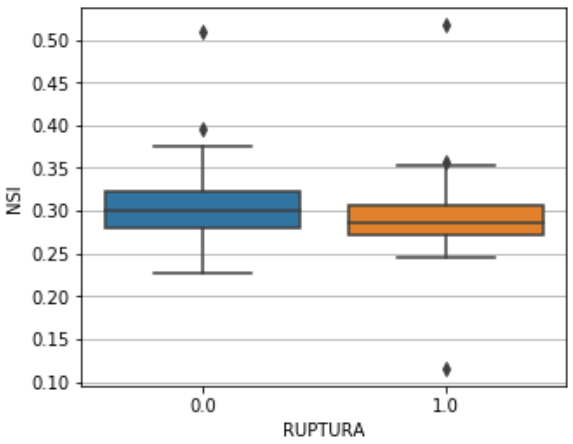
(b) Razón de aspecto.



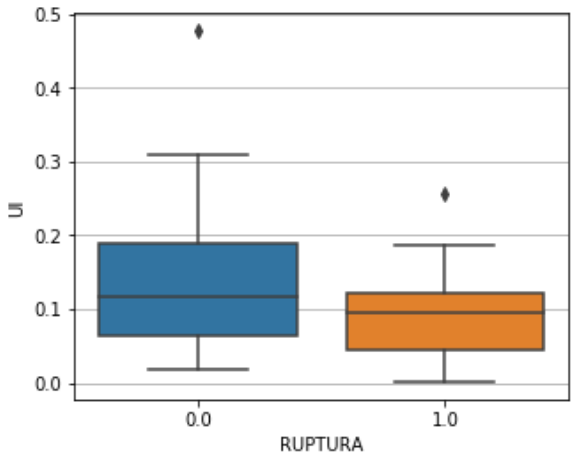
(c) Razón de tamaño.



(d) Factor de cuello de botella.

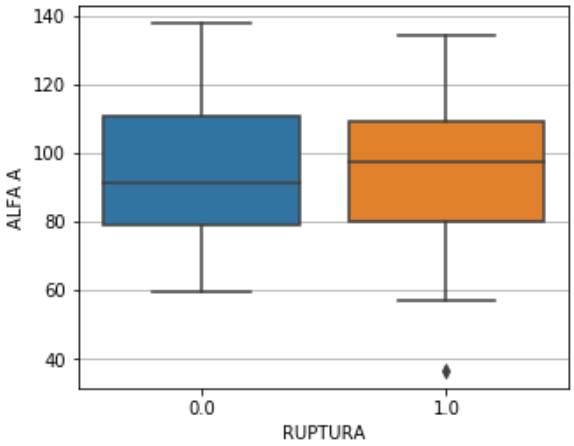


(e) Índice de no esfericidad.

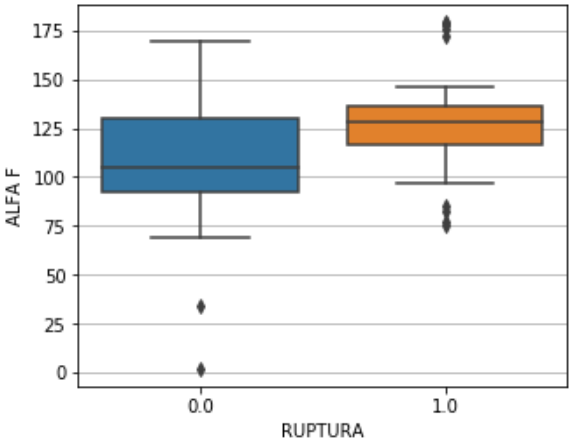


(f) Índice de ondulación.

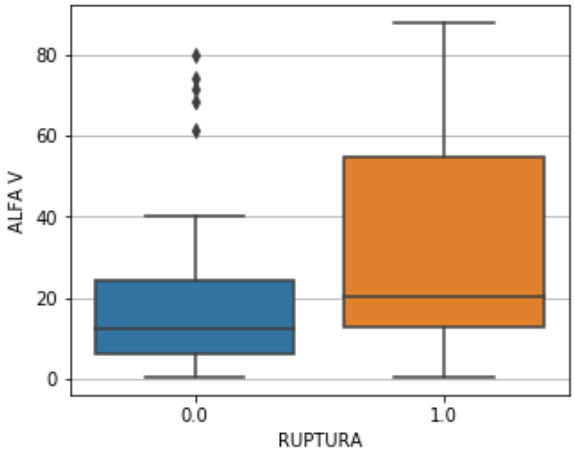
Figura 6.4: Diagramas de caja y bigote de los atributos morfológicos (1/2).
Fuente: Elaboración propia.



(a) Ángulo de aneurisma.

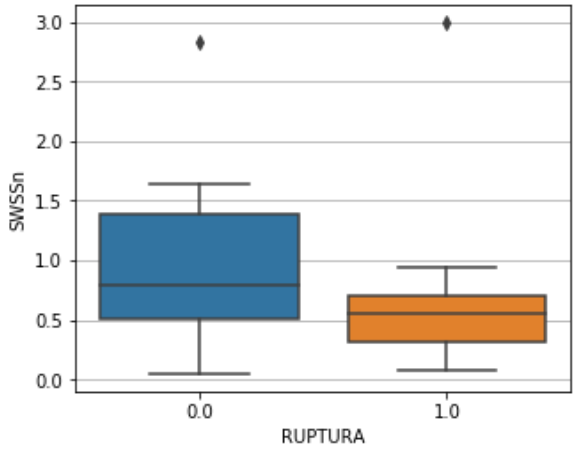


(b) Ángulo de flujo.

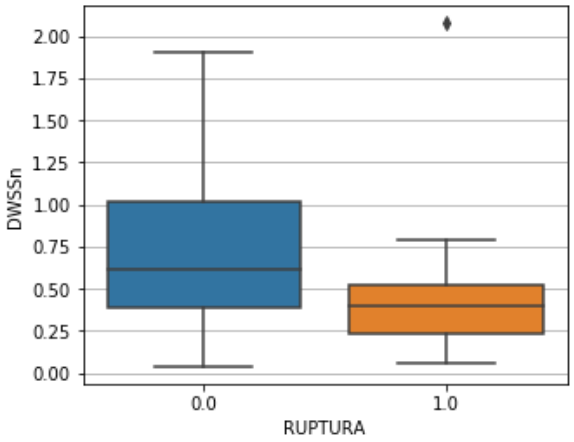


(c) Ángulo de arteria.

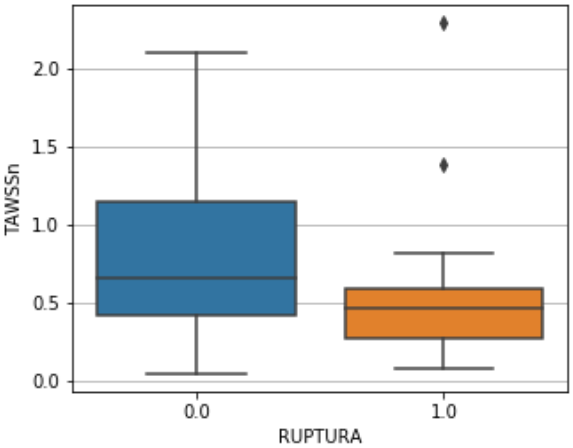
Figura 6.5: Diagramas de caja y bigote de los atributos morfológicos (2/2).
Fuente: Elaboración propia.



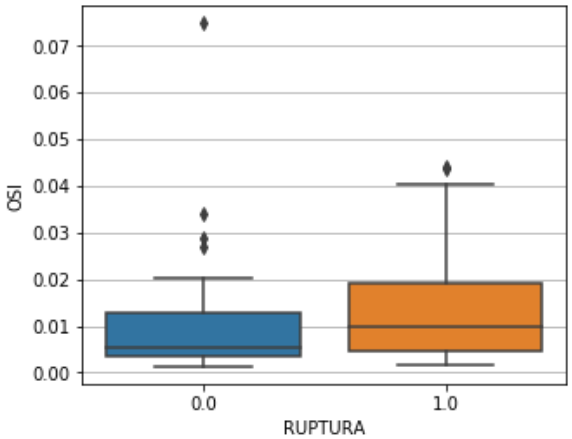
(a) Esfuerzo de corte en la pared sistólico.



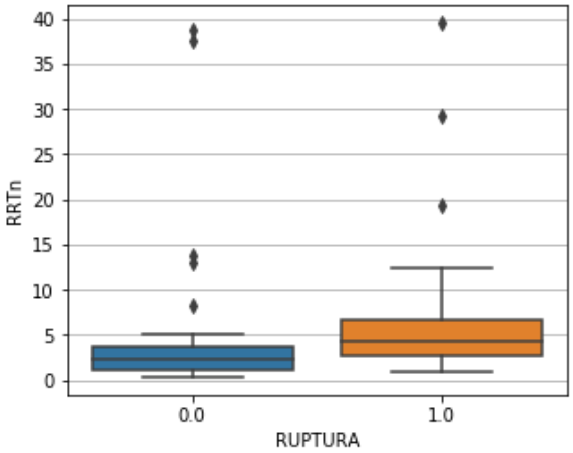
(b) Esfuerzo de corte en la pared diastólico.



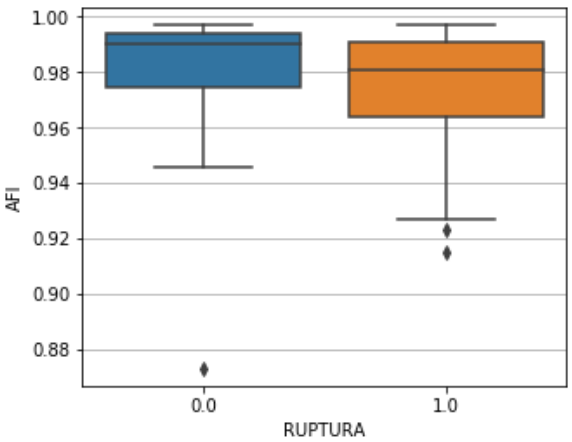
(c) Esfuerzo de corte en la pared promediado en el tiempo.



(d) Índice de corte oscilatorio.



(e) Tiempo de residencia relativo.



(f) Índice de formación de aneurisma.

Figura 6.6: Diagramas de caja y bigote de los atributos hemodinámicos.
Fuente: Elaboración propia.

Tabla 6.1: P-valor de los atributos originales obtenido mediante la prueba de normalidad de Shapiro-Wilk.

Atributo	No Roto	Roto	Total
H	0,02	0,33	0,03
AR	<0,01	0,30	<0,01
SR	0,01	0,20	<0,01
BNF	<0,01	<0,01	<0,01
NSI	<0,01	<0,01	<0,01
UI	<0,01	0,07	<0,01
α_A	0,17	0,45	0,79
α_F	0,03	0,03	0,02
α_V	<0,01	<0,01	<0,01
$SWSS_n$	0,03	<0,01	<0,01
$DWSS_n$	0,16	<0,01	<0,01
$TAWSS_n$	0,09	<0,01	<0,01
OSI	<0,01	<0,01	<0,01
RRT_n	<0,01	<0,01	<0,01
AFI	<0,01	<0,01	<0,01

Fuente: Elaboración propia.

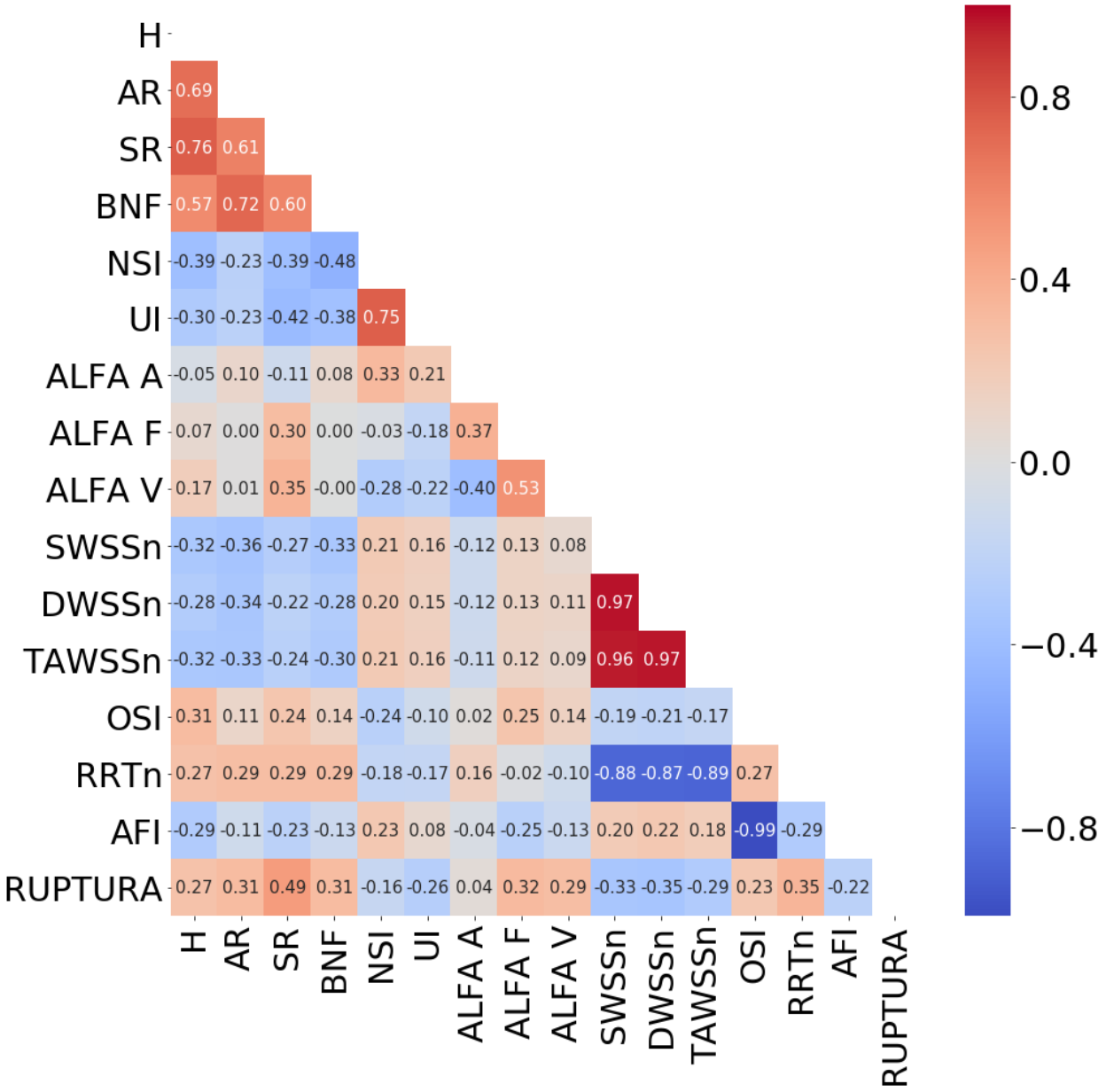


Figura 6.7: Matriz de correlación de Spearman.
Fuente: Elaboración propia.

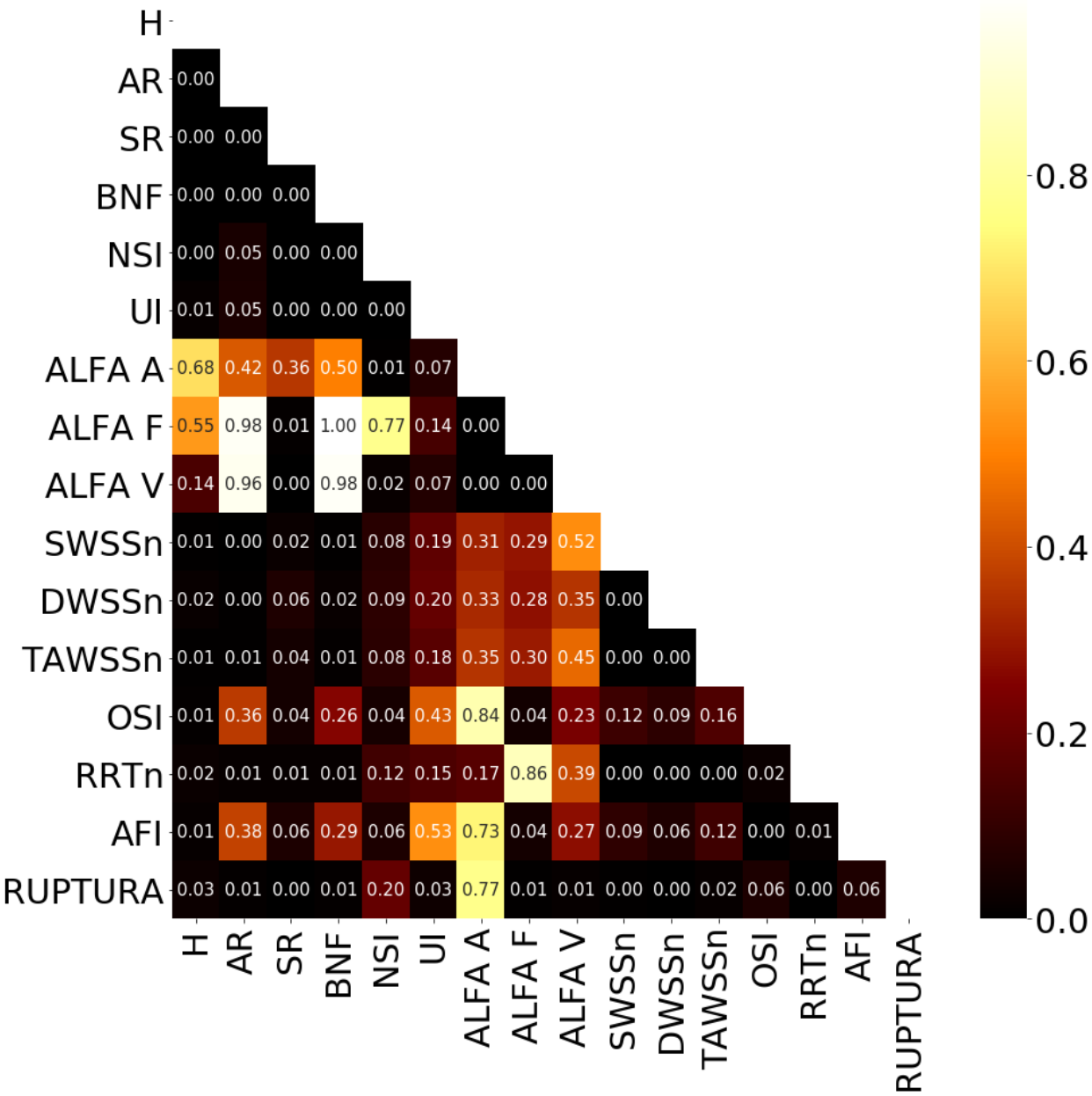
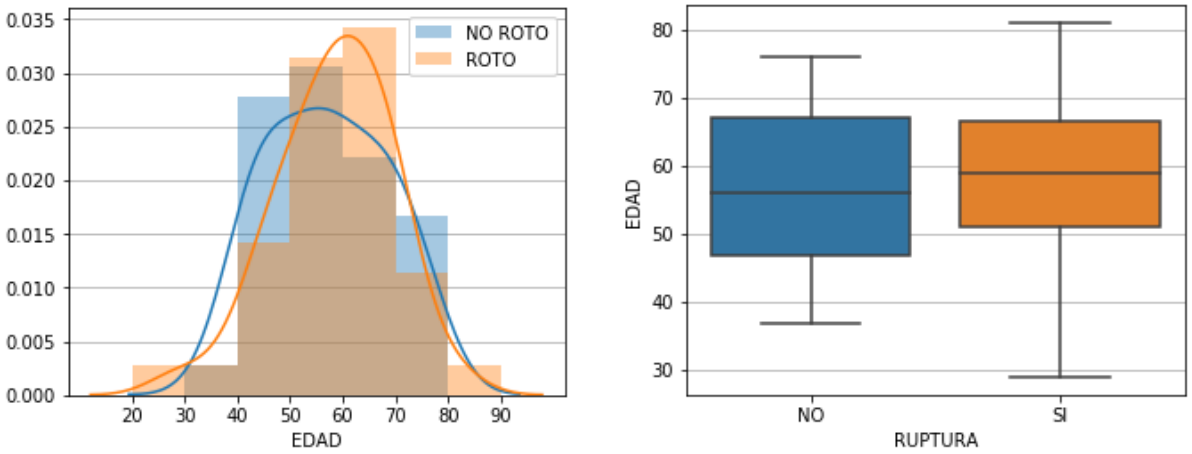


Figura 6.8: Matriz de p-valor de la correlación de Spearman.
Fuente: Elaboración propia.

6.1.2. Atributos Adicionales

Para el análisis exploratorio de datos de los atributos adicionales, en primer lugar se expone un histograma normalizado junto a la curva KDE y un diagrama de caja y bigote de la edad en la Figura 6.9. Luego, se muestra en la Tabla 6.2 los p-valores de la edad obtenidos en la prueba de normalidad de Shapiro-Wilk, en los grupos de aneurismas no rotos, rotos y total. Después, se presentan gráficos de barra por categoría de los 6 atributos adicionales de carácter categórico en la Figura 6.10. Por último, se muestra en la Tabla 6.3 el p-valor de la edad obtenido mediante la prueba de la suma de rangos de Wilcoxon y los p-valores de los atributos categóricos obtenidos por la prueba exacta de Fisher. Las tablas de contingencia con las frecuencias observadas y esperadas se enseñan en las Tablas A.1-A.6 del Apéndice A.



(a) Histograma normalizado y curva KDE.

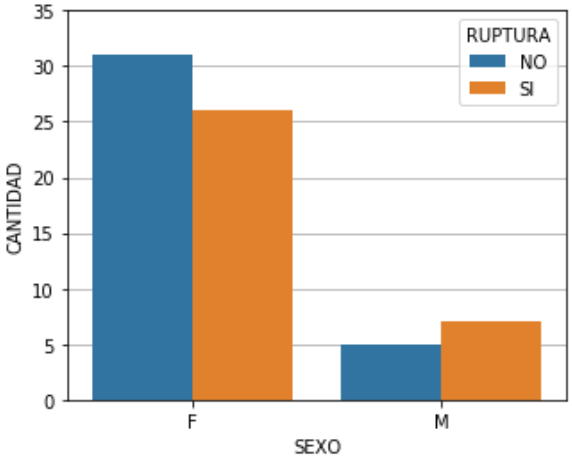
(b) Diagrama de caja y bigote.

Figura 6.9: Análisis gráfico de la edad de los pacientes.
Fuente: Elaboración propia.

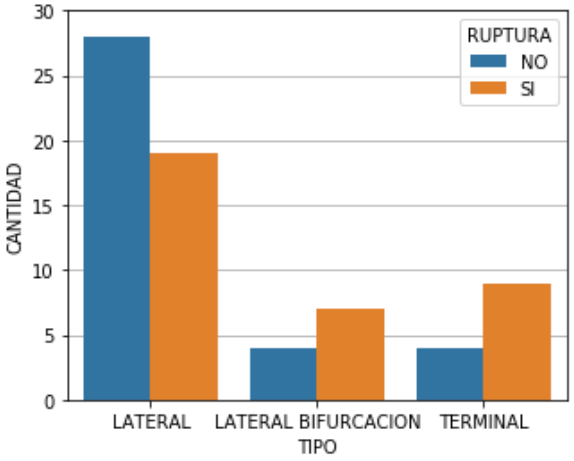
Tabla 6.2: P-valor de la edad del paciente obtenido mediante la prueba de normalidad de Shapiro-Wilk.

Atributo	No Roto	Roto	Total
Edad	0,13	0,79	0,49

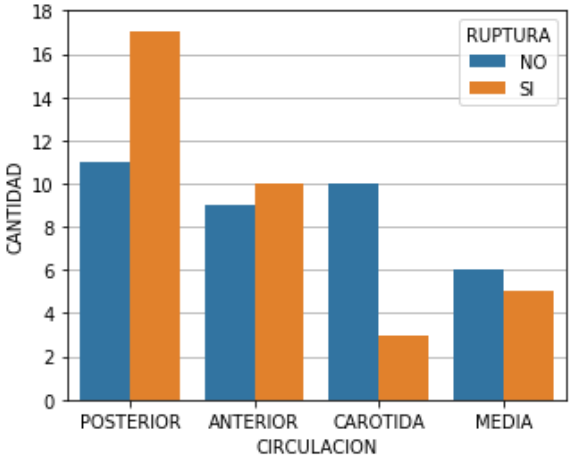
Fuente: Elaboración propia.



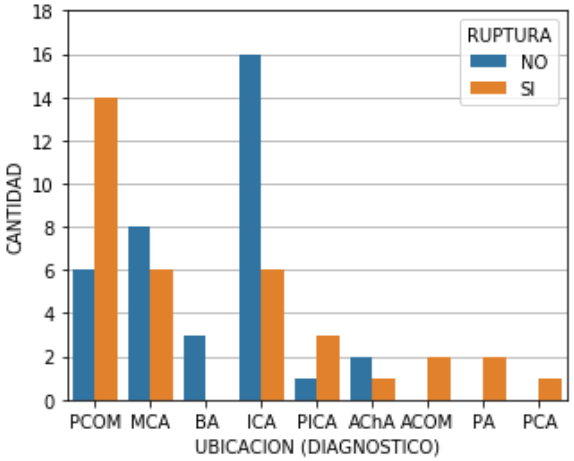
(a) Sexo del paciente.



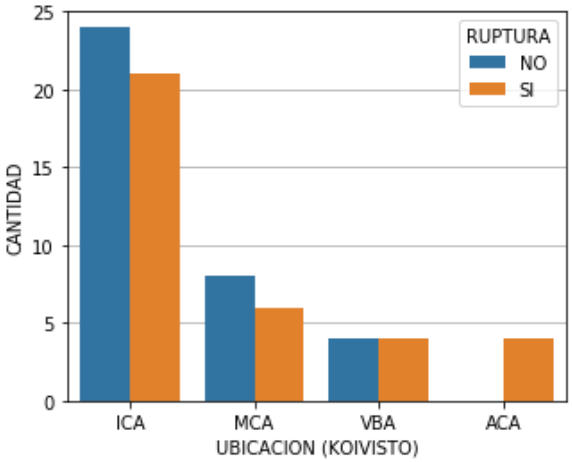
(b) Tipo de aneurisma cerebral.



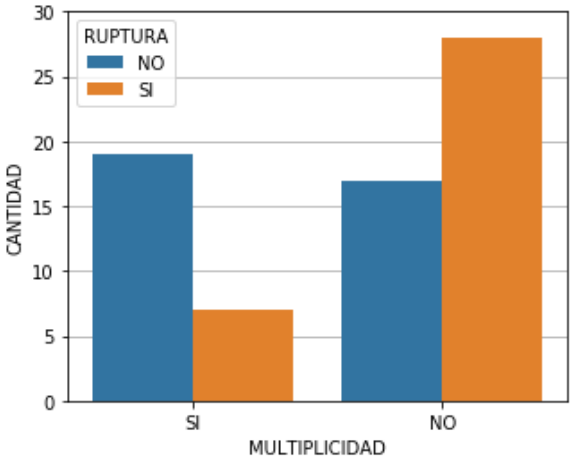
(c) Circulación.



(d) Ubicación según el diagnóstico.



(e) Ubicación según Koivisto.



(f) Multiplicidad de aneurismas.

Figura 6.10: Gráficos de barra de los atributos adicionales categóricos.
Fuente: Elaboración propia.

Tabla 6.3: Significancia estadística de los atributos adicionales.

Atributo	p-valor
Edad	0,53
Tipo	0,18
Circulación	0,16
Ubicación (Koivisto)	0,24
Ubicación (Diagnóstico)	0,01
Sexo	0,53
Multiplicidad	<0,01

Fuente: Elaboración propia.

6.2. Clasificación

En esta sección se presentan los resultados de clasificación del estado de ruptura de aneurismas cerebrales mediante machine learning, a través de la evaluación del desempeño de los modelos generados por medio de distintas métricas.

6.2.1. Conjuntos de Atributos

La Tabla 6.4 muestra la exactitud promedio, obtenida con cada uno de los modelos de clasificación generados, al evaluar en los distintos conjuntos de atributos definidos. Junto con lo anterior, se registra el promedio de exactitud de clasificación por algoritmo de clasificación y por conjunto de atributos.

Tabla 6.4: Exactitud promedio de los modelos de clasificación en los conjuntos de atributos.

Conjunto de atributos	LR	LDA	KNN	DT	NB	SVM	RF	AB	GB	Promedio
Originales	66,1 %	73,6 %	65,0 %	70,7 %	66,3 %	46,6 %	70,5 %	64,8 %	77,5 %	66,8 %
Morfológicos	64,6 %	64,8 %	65,0 %	66,3 %	62,0 %	39,6 %	64,6 %	63,6 %	62,3 %	61,4 %
Morfológicos-4	64,6 %	68,8 %	67,5 %	67,5 %	74,6 %	60,7 %	70,2 %	57,7 %	70,4 %	66,9 %
Hemodinámicos	58,0 %	60,7 %	53,4 %	57,7 %	62,0 %	55,4 %	57,9 %	44,8 %	49,3 %	55,5 %
Hemodinámicos-4	59,3 %	60,9 %	53,4 %	47,9 %	66,1 %	53,9 %	54,8 %	52,0 %	57,9 %	56,2 %
Aumentados	65,7 %	63,4 %	53,6 %	68,9 %	55,4 %	34,1 %	67,7 %	63,4 %	78,9 %	61,2 %
Aumentados*	71,8 %	73,2 %	66,4 %	69,3 %	60,9 %	51,1 %	67,5 %	70,5 %	78,8 %	67,7 %
Significativos	70,4 %	64,6 %	67,3 %	64,6 %	68,9 %	64,8 %	70,4 %	60,2 %	70,4 %	66,8 %
Promedio	65,1 %	66,3 %	61,5 %	64,1 %	64,5 %	50,8 %	65,4 %	59,6 %	68,2 %	62,8 %

Fuente: Elaboración propia.

6.2.2. Transformaciones de Atributos

La Tabla 6.5 muestra la exactitud de clasificación promedio, obtenida con cada uno de los modelos de clasificación generados, al evaluar en el conjunto de atributos originales y sus 8 transformaciones. Junto con lo anterior, se registra el promedio de la exactitud por algoritmo de clasificación y por transformación aplicada.

Tabla 6.5: Exactitud promedio de los modelos de clasificación en los atributos originales con la aplicación de distintas transformaciones de datos.

Transformación	LR	LDA	KNN	DT	NB	SVM	RF	AB	GB	Promedio
Ninguna	66,1 %	73,6 %	65,0 %	70,7 %	66,3 %	46,6 %	70,5 %	64,8 %	77,5 %	66,8 %
Estandarización	67,7 %	73,6 %	67,5 %	70,7 %	66,3 %	70,5 %	70,5 %	64,8 %	77,5 %	69,9 %
Máx-Abs	60,4 %	73,6 %	60,5 %	70,7 %	66,3 %	59,6 %	70,5 %	64,8 %	77,5 %	67,1 %
Mín-Máx	64,6 %	73,6 %	66,1 %	70,7 %	66,3 %	63,4 %	70,5 %	64,8 %	76,1 %	68,5 %
Normalización	49,5 %	65,0 %	65,0 %	63,9 %	69,1 %	38,2 %	62,3 %	65,0 %	62,3 %	60,0 %
Box-Cox	63,4 %	60,7 %	59,1 %	70,7 %	70,5 %	66,3 %	70,5 %	64,8 %	76,1 %	66,9 %
Cuantil-Normal	62,0 %	67,7 %	61,8 %	73,4 %	60,5 %	64,8 %	70,5 %	60,9 %	83,0 %	67,2 %
Cuantil-Uniforme	66,1 %	71,8 %	64,6 %	73,4 %	67,7 %	64,8 %	69,3 %	60,7 %	80,2 %	68,7 %
Robusta	67,7 %	73,6 %	55,0 %	70,7 %	66,3 %	65,0 %	70,5 %	64,8 %	77,5 %	67,9 %
Promedio	63,0 %	70,3 %	62,7 %	70,6 %	66,6 %	59,9 %	69,5 %	63,9 %	76,4 %	67,0 %

Fuente: Elaboración propia.

6.2.3. Búsqueda Exhaustiva

La Tabla 6.6 muestra la exactitud de clasificación promedio, obtenida con cada uno de los modelos de clasificación generados, al evaluar en los subconjuntos de atributos originales y sus 8 transformaciones, hallados mediante búsqueda exhaustiva. Junto con lo anterior, se registra el promedio de la exactitud por algoritmo de clasificación y por transformación aplicada. Seguidamente, las Tablas 6.7 y 6.8 exhiben las transformaciones y subconjuntos de atributos explícitos de los modelos de máxima exactitud promedio por algoritmo de clasificación.

Tabla 6.6: Exactitud promedio de los modelos de clasificación en los subconjuntos de atributos originales hallados por búsqueda exhaustiva y la aplicación de distintas transformaciones de datos.

Transformación	LR	LDA	KNN	DT	NB	SVM	RF	AB	GB	Promedio
Ninguna	81,6 %	80,4 %	78,9 %	83,2 %	83,0 %	76,1 %	83,0 %	83,4 %	87,3 %	81,9 %
Estandarización	80,2 %	80,4 %	81,6 %	83,2 %	83,0 %	81,6 %	83,0 %	83,4 %	87,3 %	82,6 %
Máx-Abs	75,9 %	80,4 %	81,8 %	83,2 %	83,0 %	73,2 %	83,0 %	83,4 %	87,3 %	81,3 %
Mín-Máx	75,9 %	80,4 %	81,8 %	83,2 %	83,0 %	74,6 %	83,0 %	83,4 %	87,1 %	81,4 %
Normalización	52,3 %	78,8 %	76,3 %	81,8 %	81,6 %	45,2 %	80,5 %	78,8 %	80,4 %	72,8 %
Box-Cox	77,3 %	77,5 %	78,9 %	83,2 %	78,8 %	81,6 %	83,0 %	83,4 %	87,1 %	81,2 %
Cuantil-Normal	78,8 %	76,3 %	77,7 %	84,6 %	77,3 %	79,1 %	83,0 %	80,7 %	88,8 %	80,7 %
Cuantil-Uniforme	77,3 %	80,4 %	80,5 %	87,3 %	80,2 %	77,3 %	83,0 %	84,8 %	88,8 %	82,2 %
Robusta	80,2 %	80,4 %	83,0 %	83,2 %	83,0 %	81,4 %	83,0 %	83,4 %	87,3 %	82,8 %
Promedio	75,5 %	79,4 %	80,1 %	83,7 %	81,4 %	74,5 %	82,8 %	82,7 %	86,8 %	80,8 %

Fuente: Elaboración propia.

Tabla 6.7: Transformación, subconjunto de atributos y exactitud de los mejores modelos por algoritmo hallados por búsqueda exhaustiva (1/2).

Algoritmo	Transformación	Subconjunto de atributos	Exactitud
LR	Ninguna	SR NSI UI α_F SWSS_n RRT_n AFI SR NSI UI α_F SWSS_n OSI RRT_n AFI	81,6 %
LDA	Ninguna	AR SR BNF α_F α_V SWSS_n DWSS_n TAWSS_n OSI RRT_n AFI	80,4 %
	Estandarización	AR SR BNF α_F α_V SWSS_n DWSS_n TAWSS_n OSI RRT_n AFI	
	Máx-Abs	AR SR BNF α_F α_V SWSS_n DWSS_n TAWSS_n OSI RRT_n AFI	
	Mín-Máx	AR SR BNF α_F α_V SWSS_n DWSS_n TAWSS_n OSI RRT_n AFI	
	Robusta	AR SR BNF α_F α_V SWSS_n DWSS_n TAWSS_n OSI RRT_n AFI	
KNN	Robusta	AR BNF UI α_F α_V SWSS_n DWSS_n	83,0 %
DT	Cuantil-Uniforme	BNF UI α_F α_V SWSS_n TAWSS_n RRT_n	87,3 %
NB	Ninguna	BNF α_F DWSS_n	83,0 %
	Estandarización	BNF α_F DWSS_n	
	Máx-Abs	BNF α_F DWSS_n	
	Mín-Máx	BNF α_F DWSS_n	
	Robusta	BNF α_F DWSS_n	
SVM	Estandarización	SR BNF UI α_A α_F SWSS_n DWSS_n	81,6 %
	Box-Cox	BNF α_F DWSS_n SR BNF UI α_A α_F SWSS_n	

Fuente: Elaboración propia.

Tabla 6.8: Transformación, subconjunto de atributos y exactitud de los mejores modelos por algoritmo hallados por búsqueda exhaustiva (2/2).

Algoritmo	Transformación	Subconjunto de atributos	Exactitud
RF	Ninguna	BNF α_F SWSS_n	83,0 %
	Estandarización	BNF α_F SWSS_n	
	Máx-Abs	BNF α_F SWSS_n	
	Mín-Máx	BNF α_F SWSS_n	
	Box-Cox	BNF α_F SWSS_n	
	Cuantil-Normal	BNF NSI α_F SWSS_n AR BNF NSI UI α_F SWSS_n AR BNF α_A α_F α_V SWSS_n TAWSS_n	
	Cuantil-Uniforme	BNF NSI α_F SWSS_n SR BNF α_F SWSS_n DWSS_n OSI	
	Robusta	BNF α_F SWSS_n	
AB	Cuantil-Uniforme	H SR α_A α_F α_V SWSS_n OSI H SR α_A α_F α_V SWSS_n OSI AFI	84,8 %
GB	Cuantil-Normal	H BNF UI α_A α_F α_V SWSS_n TAWSS_n AFI	88,8 %
	Cuantil-Uniforme	H BNF UI α_A α_F α_V SWSS_n TAWSS_n AFI	

Fuente: Elaboración propia.

6.2.4. Reglas de Asociación

La Tabla 6.9 muestra el soporte de cada uno de los atributos originales dentro del grupo de subconjuntos de atributos hallados por búsqueda exhaustiva, asociados a modelos que alcanzan una exactitud promedio mínima de 80 %. Por su parte, la Tabla 6.10 presenta las reglas de asociación, junto con su soporte, confianza e interés, que alcanzan un interés mínimo de 3.

Tabla 6.9: Soporte de los atributos en los subconjuntos de atributos hallados por búsqueda exhaustiva de los modelos con exactitud mínima de 80 %.

Atributo	Soporte
H	13,5 %
AR	32,4 %
SR	56,8 %
BNF	81,1 %
NSI	27,0 %
UI	40,5 %
α_A	29,7 %
α_F	91,9 %
α_V	32,4 %
$\mathrm{SWSS_n}$	75,7 %
$\mathrm{DWSS_n}$	43,2 %
$\mathrm{TAWSS_n}$	18,9 %
OSI	18,9 %
$\mathrm{RRT_n}$	18,9 %
AFI	37,8 %

Fuente: Elaboración propia.

Tabla 6.10: Reglas de asociación con interés mínimo de 3 y métricas en los subconjuntos de atributos hallados por búsqueda exhaustiva de los modelos con exactitud mínima de 80 %.

Regla	Soporte	Confianza	Interés
SR, α_V ⇒OSI	10,8 %	80,0 %	4,2
NSI, SR⇒RRT_n	10,8 %	66,7 %	3,5
α_F, SR, α_V ⇒OSI	10,8 %	80,0 %	4,2
α_F, NSI, SR⇒RRT_n	10,8 %	80,0 %	4,2
α_F, SWSS_n, NSI, ⇒RRT_n	10,8 %	57,1 %	3,0
SWSS_n, BNF, α_V ⇒TAWSS_n	10,8 %	57,1 %	3,0
SWSS_n, SR, α_V ⇒OSI	10,8 %	80,0 %	4,2
SWSS_n, NSI, SR⇒RRT_n	10,8 %	66,7 %	3,5
α_F, AFI, SWSS_n, SR⇒OSI	10,8 %	57,1 %	3,0
α_F, AFI, SWSS_n, SR⇒RRT_n	10,8 %	57,1 %	3,0
α_F, SWSS_n, BNF, α_V ⇒TAWSS_n	10,8 %	57,1 %	3,0
α_F, SWSS_n, SR, α_V ⇒OSI	10,8 %	80,0 %	4,2
α_F, SWSS_n, NSI, SR⇒RRT_n	10,8 %	80,0 %	4,2

Fuente: Elaboración propia.

6.2.5. Optimización de Hiperparámetros

Las Tablas 6.11-6.18 muestran los hiperparámetros optimizados mediante búsqueda en grilla de cada uno de los algoritmos de clasificación optimizados, respectivamente. Para cada hiperparámetro se resgistra su valor inicial y su valor optimizado. Las grillas de búsqueda por algoritmo e hiperparámetro se exponen detalladamente en las Tablas B.1 y B.2 del Apéndice B. Luego, la Tabla 6.19 muestra la exactitud promedio de cada algoritmo alcanzada inicialmente, en la optimización de hiperparámetros y evaluando de la misma forma que al inicio luego de la optimización de hiperparámetros.

Tabla 6.11: Hiperparámetros iniciales y optimizados para el modelo LR de mayor exactitud.

Hiperparámetro	Valor inicial	Valor optimizado
C	1	1
Intercepto de ajuste	Verdadero	Verdadero
Método de solución	Liblinear	Liblinear
Iteraciones máximas	100	50
Partida en caliente	Falso	Verdadero
Tolerancia	0,0001	0,001

Fuente: Elaboración propia.

Tabla 6.12: Hiperparámetros iniciales y optimizados para el modelo LDA de mayor exactitud.

Hiperparámetro	Valor inicial	Valor optimizado
Método de solución	SVD	SVD

Fuente: Elaboración propia.

Tabla 6.13: Hiperparámetros iniciales y optimizados para el modelo KNN de mayor exactitud.

Hiperparámetro	Valor inicial	Valor optimizado
Cantidad de vecinos	5	5
Peso	Uniforme	Uniforme

Fuente: Elaboración propia.

Tabla 6.14: Hiperparámetros iniciales y optimizados para el modelo DT de mayor exactitud.

Hiperparámetro	Valor inicial	Valor optimizado
Criterio	Gini	Gini
División	Mejor	Mejor
Profundidad máxima	Ninguna	Ninguna
Instancias mímimas para división	2	2
Instancias mínimas para hoja	1	1
Cantidad máxima de hojas	Ninguna	Ninguna
Disminución mínima de impureza	0	0

Fuente: Elaboración propia.

Tabla 6.15: Hiperparámetros iniciales y optimizados para el modelo SVM de mayor exactitud.

Hiperparámetro	Valor inicial	Valor optimizado
C	1	1
Gamma	Auto	Auto
Tolerancia	0,001	0,001

Fuente: Elaboración propia.

Tabla 6.16: Hiperparámetros iniciales y optimizados para el modelo RF de mayor exactitud.

Hiperparámetro	Valor inicial	Valor optimizado
Cantidad de estimadores	100	200
Profundidad máxima	Ninguna	Ninguna
Instancias mínimas para división	2	2
Instancias mínimas para hoja	1	1
Cantidad máxima de hojas	Ninguna	3
Disminución mínima de impureza	0	0

Fuente: Elaboración propia.

Tabla 6.17: Hiperparámetros iniciales y optimizados para el modelo AB de mayor exactitud.

Hiperparámetro	Valor inicial	Valor optimizado
Cantidad de estimadores	50	50
Tasa de aprendizaje	1	1
Algoritmo	SAMME.R	SAMME.R

Fuente: Elaboración propia.

Tabla 6.18: Hiperparámetros iniciales y optimizados para el modelo GB de mayor exactitud.

Hiperparámetro	Valor inicial	Valor optimizado
Cantidad de estimadores	100	100
Tasa de aprendizaje	0,1	0,1
Submuestra	1	1
Profundidad máxima	3	3
Instancias mínimas para división	2	2
Instancias mínimas para hoja	1	1
Cantidad máxima de hojas	Ninguna	Ninguna
Disminución mínima de impureza	0	0

Fuente: Elaboración propia.

Tabla 6.19: Exactitud obtenida mediante la optimización de hiperparámetros.

Algoritmo	Evaluación Inicial	Optimización de Hiperparámetros	Evaluación Final
LR	81,6 %	81,7 %	81,6 %
LDA	80,4 %	80,3 %	80,4 %
KNN	83,0 %	83,1 %	83,0 %
DT	87,3 %	87,3 %	87,3 %
NB	83,0 %	–	–
SVM	81,6 %	81,7 %	81,6 %
RF	83,0 %	83,1 %	83,0 %
AB	84,8 %	84,5 %	84,8 %
GB	88,8 %	88,7 %	88,8 %

Fuente: Elaboración propia.

6.2.6. Evaluación de Modelos Finales

La Tabla 6.20 muestra la exactitud y la precisión promedio junto con su desviación estándar en los conjuntos de entrenamiento y prueba de los mejores modelos por algoritmo hallados vía búsqueda exhaustiva, evaluados mediante validación cruzada. Finalmente, la Tabla 6.21 exhibe la exactitud, precisión, sensibilidad, valor F1 y AUC de la curva ROC promedio junto con su desviación estándar, en los conjuntos de entrenamiento y prueba de los mejores modelos por algoritmo, hallados por búsqueda exhaustiva, evaluados mediante validación cruzada estratificada.

Tabla 6.20: Evaluación final del desempeño de los mejores modelos por algoritmo hallados por búsqueda exhaustiva mediante validación cruzada.

Modelo	Conjunto	Exactitud	Precisión
LR	Entrenamiento	80,1 % ± 2,9 %	78,9 % ± 2,6 %
	Prueba	81,6 % ± 12,9 %	76,0 % ± 33,6 %
LDA	Entrenamiento	82,2 % ± 2,6 %	82,6 % ± 2,5 %
	Prueba	80,4 % ± 12,9 %	74,7 % ± 38,9 %
KNN	Entrenamiento	85,4 % ± 1,6 %	84,1 % ± 1,8 %
	Prueba	83,0 % ± 15,4 %	73,2 % ± 38,1 %
DT	Entrenamiento	100 % ± 0,0 %	100 % ± 0,0 %
	Prueba	87,3 % ± 10,0 %	78,0 % ± 39,4 %
NB	Entrenamiento	80,7 % ± 2,4 %	82,2 % ± 3,4 %
	Prueba	83,0 % ± 14,0 %	75,0 % ± 38,7 %
SVM	Entrenamiento	85,9 % ± 1,4 %	81,7 % ± 2,5 %
	Prueba	81,6 % ± 15,8 %	75,2 % ± 33,7 %
RF	Entrenamiento	100 % ± 0,0 %	100 % ± 0,0 %
	Prueba	83,0 % ± 10,8 %	72,5 % ± 37,8 %
AB	Entrenamiento	100 % ± 0,0 %	–
	Prueba	84,8 % ± 14,0 %	–
GB	Entrenamiento	100 % ± 0,0 %	100 % ± 0,0 %
	Prueba	88,8 % ± 8,5 %	75,5 % ± 38,8 %

Fuente: Elaboración propia.

Tabla 6.21: Evaluación final del desempeño de los mejores modelos por algoritmo hallados por búsqueda exhaustiva mediante validación cruzada estratificada.

Modelo	Conjunto	Exactitud	Precisión	Sensibilidad	Valor F1	AUC ROC
LR	Entrenamiento	81,2 % ± 2,4 %	79,6 % ± 2,5 %	83,2 % ± 2,9 %	81,3 % ± 2,4 %	85,1 % ± 2,1 %
	Prueba	82,1 % ± 15,7 %	83,2 % ± 18,4 %	85,0 % ± 15,3 %	82,8 % ± 14,1 %	78,5 % ± 16,7 %
LDA	Entrenamiento	81,9 % ± 1,8 %	82,6 % ± 1,8 %	80,0 % ± 3,1 %	81,3 % ± 2,0 %	90,4 % ± 1,4 %
	Prueba	78,2 % ± 19,8 %	78,3 % ± 29,6 %	72,5 % ± 33,3 %	72,8 % ± 29,3 %	83,5 % ± 16,4 %
KNN	Entrenamiento	85,6 % ± 1,6 %	84,3 % ± 3,1 %	87,3 % ± 3,7 %	85,6 % ± 1,6 %	90,1 % ± 2,0 %
	Prueba	81,3 % ± 18,7 %	80,2 % ± 21,0 %	82,5 % ± 28,5 %	79,4 % ± 23,3 %	81,3 % ± 22,6 %
DT	Entrenamiento	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %
	Prueba	79,2 % ± 19,3 %	81,0 % ± 30,4 %	73,3 % ± 33,3 %	73,8 % ± 29,0 %	79,6 % ± 19,2 %
NB	Entrenamiento	82,5 % ± 3,5 %	84,1 % ± 3,3 %	79,7 % ± 7,9 %	81,6 % ± 4,5 %	85,4 % ± 1,5 %
	Prueba	76,7 % ± 15,7 %	78,5 % ± 18,5 %	74,2 % ± 20,9 %	75,2 % ± 17,8 %	80,6 % ± 20,0 %
SVM	Entrenamiento	85,8 % ± 2,1 %	81,0 % ± 2,5 %	93,0 % ± 3,7 %	86,5 % ± 2,0 %	94,4 % ± 1,3 %
	Prueba	75,9 % ± 15,8 %	76,5 % ± 17,1 %	77,5 % ± 19,7 %	75,7 % ± 15,8 %	82,2 % ± 18,0 %
RF	Entrenamiento	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %
	Prueba	82,6 % ± 13,9 %	82,0 % ± 17,1 %	85,0 % ± 21,3 %	81,9 % ± 16,2 %	83,2 % ± 10,9 %
AB	Entrenamiento	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %
	Prueba	71,5 % ± 17,8 %	74,9 % ± 21,8 %	67,5 % ± 27,2 %	68,1 % ± 22,1 %	82,2 % ± 16,4 %
GB	Entrenamiento	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %	100 % ± 0,0 %
	Prueba	82,3 % ± 14,4 %	81,5 % ± 14,7 %	84,2 % ± 26,5 %	80,3 % ± 19,0 %	90,2 % ± 16,1 %

Fuente: Elaboración propia.

Capítulo 7

Análisis de Resultados

7.1. Análisis Exploratorio de Datos

7.1.1. Atributos Originales

Gráficos y Prueba de Normalidad

Los histogramas normalizados y curvas KDE de los atributos morfológicos, mostrados en las Figuras 6.1 y 6.2, exhiben que las distribuciones de valores entre los grupos de aneurismas no rotos y rotos son en general distintas. Visualmente, NSI y α_A poseen las distribuciones más similares entre ambos grupos de aneurismas, concordando con la significancia estadística y el análisis ROC de las Tablas 3.2 y 3.3 provistas por Amigo [4]. La forma de la distribución es similar para los valores de AR entre ambos grupos, sin embargo sus modas y rangos son claramente diferentes. En cuanto a los atributos hemodinámicos, los histogramas normalizados y curvas KDE expuestos en la Figura 6.3, hacen ver que la distribución de valores entre ambos grupos de aneurismas es muy distinta en los atributos $\mathrm{SWSS_n}$, $\mathrm{DWSS_n}$ y $\mathrm{TAWSS_n}$. Visualmente, esta diferencia se atenúa para los atributos OSI, $\mathrm{RRT_n}$, y AFI. En el caso de $\mathrm{RRT_n}$, la diferencia es mayor considerando las modas y el rango de valores, además de la mayor presencia de posibles valores perdidos, en comparación con OSI y AFI. Esto a su vez es consistente con los resultados estadísticos señalados previamente.

Los diagramas de caja y bigote de los atributos morfológicos, como aparecen en las Figuras 6.4 y 6.5, muestran que la mayoría posee valores perdidos, siendo α_F y α_V donde más abundan. Por su parte, únicamente H y SR se encuentran libres de valores perdidos. Junto con esto, se advierte que en el caso de SR se da la intersección más pequeña entre los rangos intercuartiles de los grupos no roto y roto, comparando en proporción con el resto de atributos, lo cual es consistente con los resultados estadísticos de las Tablas 3.2 y 3.3 hallados por Amigo [4]. Para el caso de los atributos hemodinámicos, cuyos diagramas de caja y bigote se muestran en la Figura 6.6, se ve que todos presentan valores perdidos, siendo $\mathrm{RRT_n}$ y OSI aquellos que exhiben más. Visualmente, OSI y AFI poseen una mayor intersección entre los rangos intercuartiles de los grupos no roto y roto, comparando en proporción, siendo acorde a los resultados previamente mencionados. La cantidad significativa de valores perdidos plantea como alternativa su eliminación. Esto podría ayudar al desarrollo de mejores modelos de clasificación, sin embargo, tener una mayor cantidad de instancias posibilita la generación de mejores modelos. Considerando la baja cantidad de instancias a disposición, se opta por no eliminar los valores perdidos para conservarla. Otra opción sería reemplazar los valores perdidos de forma apropiada.

A partir de los p-valores mostrados en la Tabla 6.1, se tiene que en el grupo de aneurismas no rotos, los valores de α_A, $\mathrm{DWSS_n}$ y $\mathrm{TAWSS_n}$ aceptan la hipótesis de distribución normal. En el grupo de aneurismas rotos, esto se cumple para H, AR, SR, UI y α_A. Para el total de aneurismas, únicamente α_A cumple con la normalidad de datos. Así, la mayor parte de los atributos no sigue una distribución normal.

El análisis exploratorio de datos de los atributos morfológicos y hemodinámicos permite vislumbrar que ninguno de los parámetros exhibe un valor umbral que haga posible separar completamente los aneurismas no rotos de los rotos. Lo anterior es consistente con la naturaleza poco predecible de los aneurismas cerebrales en relación a la ruptura. A pesar de este resultado, en varios de los atributos morfológicos se observa una diferencia, clara en mayor o menor grado dependiendo del caso, entre la distribución de ambos grupos, ya sea en forma, modas o rangos, que es consistente con la literatura y la tesis doctoral de Amigo [4]. Por su parte, existen atributos morfológicos cuyas diferencias entre los grupos de aneurismas no rotos y rotos no concuerdan con la literatura, sino que resultan contrarias, como son el caso del índice de no esfericidad y el índice de ondulación. A su vez, existe el caso del ángulo del aneurisma, el cual no resulta ser estadísticamente significativo de acuerdo a Amigo, como lo señala su p-valor mostrado en la Tabla 3.2, mientras que hay estudios que sugieren lo contrario [4, 33, 157]. Esto puede deberse a que la cantidad de aneurismas considerados es baja, lo cual tiende a aumentar las diferencias entre los datos aquí trabajados y los de otros estudios. A su vez, ciertos estudios abarcan incluso menos aneurismas cerebrales, lo cual afecta en la misma dirección. En contraste, las distribuciones de atributos hemodinámicos analizados son consistentes con la literatura y la significancia estadística de Amigo, en cuanto a la comparación entre aneurismas no rotos y rotos [4].

Matriz de Correlación

Para los resultados obtenidos mediante la matriz de correlación de Spearman, mostrada en la Figura 6.7, el análisis se concentra en las correlaciones existentes entre los atributos morfológicos y hemodinámicos, junto con la correlación entre el estado de ruptura y los atributos originales. Al observar los coeficientes de correlación entre los atributos morfológicos y hemodinámicos se puede notar que todos se encuentran por debajo de 0,4 en valor absoluto. Lo anterior significa que los datos trabajados no exhiben ninguna correlación fuerte entre un par de atributos morfológico-hemodinámico.

La correlación de mayor intensidad se da entre AR y $SWSS_n$, con un valor de -0,36. Le siguen en valor absoluto la correlación entre AR con $DWSS_n$ y $TAWSS_n$, con valores de -0,34 y -0,33, respectivamente. Como se ve en la matriz de p-valor de la correlación de Spearman, exhibida en la Figura 6.8, estos tres coeficientes de correlación de mayor intensidad y negativos son estadísticamente significativos. Dichos resultados indican una tendencia leve a una dependencia monotónica entre el WSS normalizado y AR, donde el aumento de una variable se relaciona con la disminución de la otra. $SWSS_n$ y $TAWSS_n$ poseen de las correlaciones más intensas encontradas con H, siendo -0,32 para ambos, además de -0,33 y -0,3 con BNF, respectivamente. A partir de los p-valores mostrados en la matriz de la Tabla 6.8, se tiene que estas cuatro correlaciones son estadísticamente significativas. Estos resultados sugieren levemente que a medida que el aneurisma cerebral aumenta en tamaño y en desproporción geométrica, el WSS normalizado tiende a disminuir, de forma monotónica. Otra correlación que se incluye entre aquellas de mayor intensidad, dentro de las halladas, ocurre entre H y OSI, con un valor de 0,31. Esto señala que hay una leve tendencia al aumento de la oscilación del vector WSS, al aumentar el tamaño del aneurisma cerebral. Estos resultados proponen débilmente una correlación monotónica entre ciertos atributos morfológicos y hemodinámicos. Sin embargo, no permiten validar el hecho de que la morfología determina instantáneamente la hemodinámica, y esta determina la morfología futura [83].

La última fila de la matriz de correlación de Spearman de la Figura 6.7 exhibe los coeficientes de correlación entre el estado de ruptura y los 15 atributos originales. Tomando en cuenta que este coeficiente de correlación no es el ideal para atributos binarios y continuos, ninguno de los coeficientes supera o iguala el valor de 0,5 en valor absoluto, implicando que la correlación de Spearman no evidencia dependencias monotónicas fuertes entre los atributos y el estado de ruptura.

El atributo con el cual se encuentra una mayor correlación corresponde a SR, con un valor de 0,49. Seguidamente, se encuentran $DWSS_n$ junto con RRT_n, con valores de -0,35 y 0,35 respectivamente. Posteriormente, se tiene $SWSS_n$, con un coeficiente de -0,33, y α_F con 0,32. Tanto AR como BNF poseen una correlación de 0,31. Observando el p-valor de cada coeficiente de correlación anteriormente señalado, en la matriz de la Tabla 6.8, se tiene que todos son estadísticamente significativos. Considerando que el estado no roto posee el valor 0 y en caso contrario el valor 1, estos resultados indican que los aneurismas cerebrales rotos se asocian débilmente con una mayor SR, RRT_n, α_F, AR y BNF, alineándose con la tendencia reportada en la literatura [33, 76, 111] y siendo consistente con la significancia estadística de los atributos individuales y su AUC, como se muestra en las Tablas 3.2 y 3.3 respectivamente, calculadas por Amigo [4].

Por su parte, $DWSS_n$ y $SWSS_n$, al tener coeficientes de correlación negativos, tienden a disminuir para los aneurismas cerebrales rotos, de acuerdo a lo encontrado. Lo anterior concuerda con lo reportado por Zhou *et al.* en su revisión sistemática y meta-análsis [158]. Por otro lado, α_A y NSI resultan ser los menos correlacionados, con valores de 0,04 y -0,16 respectivamente. Esto se condice con la significancia estadística de los atributos individuales y su AUC, mostradas en las Tablas 3.2 y 3.3, respectivamente, encontradas por Amigo [4].

7.1.2. Atributos Adicionales

Gráficos y Normalidad

A partir del histograma normalizado de la Figura 6.9(a) y el diagrama de caja y bigote de la Figura 6.9(b) de la edad, se tiene que este parámetro distribuye dentro de un rango razonable, siendo entre 30 y 80 años aproximadamente, sin exhibir la presencia de valores perdidos. Este rango de edades concuerda con lo reportado acerca de la prevalencia de aneurismas cerebrales según la edad de los pacientes [54, 147]. La moda de la edad en el grupo no roto es levemente menor que la del grupo roto, lo cual a su vez, resulta consistente con el aumento del riesgo de sufrir ruptura conforme la edad es mayor, según la literatura [6]. Se observa en el histograma normalizado de la Figura 6.9(a), que entre las edades 40 a 50 años, la cantidad de aneurismas no rotos es cercana al doble de la cantidad de aneurismas rotos. Por su parte, entre las edades 50 y 60 años, la cantidad de aneurismas registrados es prácticamente la misma en ambos grupos. En cuanto al rango entre 60 y 70 años, la cantidad de aneurismas rotos supera la de aneurismas no rotos, siendo aproximadamente 1,5 veces mayor. Lo anterior sigue la tendencia de los estudios considerados sobre la edad como factor de riesgo para la ruptura [149, 6]. El p-valor obtenido para los grupos de aneurismas no rotos, rotos y el total, mostrado en la Tabla 6.2, valida la distribución normal de la edad en los tres grupos.

En cuanto a los atributos categóricos, del gráfico de barras del sexo de los paciente mostrado en la Figura 6.10(a), los aneurismas cerebrales en mujeres superan en más del doble a los casos en hombres, acorde a lo señalado por la literatura [147, 115]. Si bien en mujeres existen más aneurismas no rotos y en hombres predominan los rotos, la diferencia entre ambos grupos, en las dos categorías, es muy leve, lo cual sugiere que este atributo no permitiría discriminar el estado de ruptura. Para el tipo de aneurisma, como se muestra en el gráfico de barras de la Figura 6.10(b), se tiene que la mayor parte de los aneurismas cerebrales son del tipo lateral. Para este tipo de lesiones, predominan los aneurismas no rotos. Para el tipo lateral bifurcación, la cantidad de aneurismas rotos es levemente superior. En los aneurismas cerebrales terminales, los rotos corresponden aproximadamente al doble de los no rotos. Se considera que estos resultados indican el tipo de aneurisma como un parámetro poco promisorio para discriminar el estado de ruptura.

Para el tipo de circulación, según se observa en el gráfico de barras de la Figura 6.10(c), la mayor cantidad de aneurismas se da en la circulación posterior, seguida de la anterior, carótida y media, en forma decreciente. El orden decreciente de la cantidad de aneurismas cerebrales en la circulación anterior, carótida y media, es consistente con lo reportado, sin embargo, el hecho de que la circulación posterior aparezca como aquella con mayor frecuencia de esta patología, se opone a la literatura [122]. Aún con lo anterior, la circulación posterior es donde ocurre un mayor predominio de aneurismas rotos, en comparación con las demás categorías, lo cual concuerda con lo indicado por Wermer *et al.* en su estudio [149]. Para la circulación anterior, los aneurismas rotos predominan por una leve diferencia. En la circulación carótida, proporcionalmente, se tiene la mayor diferencia entre aneurismas rotos y no rotos, siendo los últimos mayoría. En cuanto a la circulación media, la cantidad de aneurismas no rotos es ligeramente mayor. De las 4 posibles circulaciones, existen dos que tienden a discriminar de forma más marcada el estado de ruptura, sin lograr una separación absoluta.

La ubicación según el diagnóstico, cuyo gráfico de barras se muestra en la Figura 6.10(d), muestra que existen 4 de las 9 categorías que se asocian con un único estado de ruptura: arteria basilar o BA (no roto), arteria comunicante anterior o ACOM (roto), arteria pericallosa o PA (roto), y arteria cerebral posterior o PCA (roto). Si bien este hecho ayuda a discriminar de mejor forma el estado de ruptura, la desventaja de estas 4 categorías es que cada una posee únicamente entre 1 y 3 aneurismas asociados. La arteria carótida interna o ICA agrupa la mayor cantidad de aneurismas y exhibe un predominio considerable de los casos no rotos, además de ser la diferencia más marcada en proporción, entre las categorías. Le sigue en cantidad la arteria comunicante posterior, donde el número de aneurismas rotos supera el doble de los no rotos, siendo acorde a los estudios realizados [149, 127]. En tercer lugar, respecto a la cantidad de lesiones abarcadas, se encuentra la arteria cerebral media o MCA, en cuyo caso los aneurismas no rotos poseen un ligero predominio. A partir de esto, la ubicación según el diagnóstico se vislumbra como un atributo potencialmente útil para distinguir el estado de ruptura de la patología.

La ubicación según las categorías de Koivisto, como se observan en el gráfico de barras de la Figura 6.10(e), posee una sola categoría que engloba únicamente aneurismas rotos: ACA. Sin embargo, la categoría ACA encierra la menor cantidad de aneurismas. El resto de las categorías en orden decreciente en número de aneurismas son ICA , MCA y VBA. En el caso de la VBA, hay igual cantidad de aneurismas no rotos y rotos. Por su parte, tanto la ICA como la MCA exhiben un ligero predominio del estado no roto. A pesar de tener una categoría con una sola clase, este atributo no se considera promisorio para separar el estado de ruptura de los aneurismas cerebrales. Finalmente, la multiplicidad de aneurismas, como se observa en el gráfico de barras de la Figura 6.10(f), aparece como un fenómeno más infrecuente que frecuente. A su vez, existe una predominancia considerable de casos no rotos al tenerse múltiples aneurismas. En caso de no haber múltiples aneurismas, se da una mayor cantidad de aneurismas rotos. En vista de estos resultados gráficos, la multiplicidad de aneurismas podría ser útil en cierto grado como discriminante del estado de ruptura.

Tablas de Contingencia y Significancia Estadística

Las Tablas A.1-A.6 del Apéndice A, donde se muestran las tablas de contingencia de los 6 atributos adicionales categóricos, permiten visualizar tanto las frecuencias observadas como esperadas. En particular, las frecuencias esperadas permiten establecer una prueba de significancia estadística apropiada para los atributos categóricos. Las Tablas A.4 y A.5 poseen celdas con frecuencias esperadas menores a 5. Debido a esto, se lleva a cabo la prueba exacta de Fisher para determinar la significancia estadística de los atributos categóricos. Los resultados de esta prueba se observan en la Tabla 6.3, junto el resultado de la prueba de la suma de rangos de Wilcoxon para la edad. A partir del p-valor obtenido para cada atributo adicional, se tiene que la ubicación según el diagnóstico y la multiplicidad de aneurismas son atributos estadísticamente significativos en relación a diferenciar el estado de ruptura. La significancia estadística asociada a la ubicación es consistente con lo reportado en la literatura [157]. El sexo del paciente discrepa con lo encontrado por Anderson *et al.* al no resultar significativo en este caso [149].

7.2. Clasificación

7.2.1. Conjuntos de Atributos

Los resultados de clasificación en los distintos conjuntos de atributos definidos, como se observan en la Tabla 6.4, señalan que el algoritmo GB alcanza la mayor exactitud promedio, con un valor de 68,2 %, seguido por LDA, con el cual se logra un 66,3 % en promedio, y en tercer lugar RF, obteniendo un 65,4 %. Por su parte, con el algoritmo SVM se obtiene la peor exactitud promedio, siendo de 50,8 %. Al analizar el promedio de exactitud lograda en los conjuntos, se observa que en los de atributos originales y significativos se alcanza una exactitud promedio de 66,8 % (redondeando al primer decimal). Lo anterior valida la utilidad de la selección de atributos realizada a partir de su significancia estadística, mostrada en la Tabla 3.2, ya que en promedio se alcanza la misma exactitud utilizando únicamente 4 de los 15 atributos originales. Observando la exactitud promedio en los conjuntos de atributos morfológicos y hemodinámicos, de 61,4 % y 55,5 % respectivamente, se tiene que ninguno de los dos supera el promedio alcanzado en el conjunto original. Esto apoyaría el planteo de que el desarrollo y la ruptura de los aneurismas cerebrales tiene relación con una interacción compleja entre la morfología y la hemodinámica, al menos de forma preliminar en virtud de los atributos a disposición [4]. A su vez, este resultado señala la mayor calidad del grupo de atributos morfológicos tenidos, frente al de atributos hemodinámicos, para discriminar el estado de ruptura vía machine learning. Esto podría explicarse a priori por el hecho de contar con una mayor cantidad de atributos morfológicos que hemodinámicos. Por otro lado, al observar los coeficientes de correlación de Spearman mostrados en la matriz de correlación de la Figura 6.7, se tiene que, junto con ser menos en cantidad, se alcanzan valores elevados de correlación entre varios pares de atributos hemodinámicos, siendo así más redundantes, en comparación con los morfológicos, lo cual podría explicar su menor efectividad para distinguir el estado de ruptura.

La exactitud promedio obtenida en el conjunto de atributos morfológicos-4 es de 66,9 %, superando lo logrado en los atributos morfológicos sin la aplicación de selección de atributos, acorde a lo esperado. Por su parte, la exactitud promedio resultante en el conjunto de atributos hemodinámicos-4 es de 56,2 %, superando ligeramente el promedio obtenido en los atributos hemodinámicos sin la aplicación de selección de atributos. El hecho de que la selección de atributos realizada otorgue un mayor aumento de exactitud promedio en los atributos morfológicos comparado con los hemodinámicos, puede explicarse debido a que en el primer caso, se reduce de 9 a 4 atributos, y en el segundo caso, de 6 a 4. Esto quiere decir que la reducción de los datos provocada por la selección de atributos es mucho menor en proporción, en el conjunto de atributos hemodinámicos-4.

En cuanto al conjunto de atributos aumentados, se logra una exactitud promedio de 61,2 %, superando el promedio en el conjunto de atributos hemodinámicos y prácticamente alcanzando el del conjunto de atributos morfológicos. Respecto al conjunto de atributos originales, no se logra equiparar su exactitud promedio en este caso. Aplicando selección de atributos, se logra elevar la exactitud promedio, resultando en 67,7 % en el conjunto de atributos aumentados*. Este valor corresponde al promedio más alto logrado entre los 8 conjuntos evaluados, seguido por lo obtenido en el conjunto de atributos morfológicos-4. Junto con esto, la mayor exactitud promedio alcanzada en términos globales, se logra con el algoritmo GB en el conjunto de atributos adicionales, con un valor de 78,9 %, al cual le sigue el mismo algoritmo en el conjunto de atributos adicionales*, llegando a 78,8 %. Esto indica que la inclusión de los atributos adicionales facilita la distinción del estado de ruptura de los aneurismas cerebrales mediante machine learning. A su vez, se valida la efectividad de la selección de atributos dentro del grupo de atributos adicionales.

7.2.2. Transformaciones de Atributos

Los resultados de clasificación al evaluar el conjunto de atributos originales con la aplicación de 8 transformaciones de datos diferentes, como se observa en la Tabla 6.5, revelan que el algoritmo GB logra la mayor exactitud promedio, con un valor de 76,4 %. A este le siguen los algoritmos DT y LDA, alcanzando en promedio 70,6 % y 70,3 %, respectivamente. En contraparte, el algoritmo SVM exhibe el peor desempeño promedio entre el conjunto de atributos originales y sus 8 transformaciones, llegando a 59,9 % de exactitud. Del análisis de la exactitud promedio alcanzada por conjunto o transformación, se tiene que con excepción de la normalización, todas las transformaciones elevan el desempeño promedio entre los algoritmos logrado en los atributos originales. La transformación con la cual se logran mejores resultados corresponde a la estandarización, con una exactitud promedio de 69,9 %. Le sigue la transformación cuantil-uniforme, donde se alcanza 68,7 % y la transformación mín-máx, con 68,5 %. A nivel global, el desempeño más elevado se obtiene con el algoritmo GB, llegando a 83,0 % con la transformación cuantil-normal.

Se observa que en general, los modelos de los algoritmos LDA, NB, y aquellos en base a árboles de decisión, es decir, DT, RF, AB y GB, no varían su exactitud alcanzada para la estandarización, y las transformaciones máx-abs, mín-máx, y robusta. En particular, RF únicamente varía su exactitud con la normalización y la transformación cuantil-uniforme. Por su parte, algoritmos como KNN y SVM, los cuales trabajan a través de distancias entre datos, varían su desempeño en función de las transformaciones utilizadas. Este comportamiento resulta ser acorde a lo esperado [11].

El hecho de que en términos generales la aplicación de transformaciones de datos eleve la exactitud alcanzada es consistente, dado que, como se ve en los diagramas de caja y bigote de las Figuras 6.4, 6.5 y 6.6, existe una cantidad no despreciable de valores perdidos entre los distintos atributos morfológicos y hemodinámicos. Así, las transformaciones de carácter robusto corrigen dicho problema. Por su parte, para ciertos algoritmos como KNN y SVM que se valen de distancias, es conveniente tener los atributos en la misma escala, lo cual se logra con las transformaciones. Por ejemplo, los atributos que representan ángulos se encuentran en escalas distintas a las de otros atributos morfológicos y hemodinámicos. De igual forma, dado que en su mayoría los valores de los atributos no siguen una distribución normal, es esperable que el desempeño mejore al aplicar transformaciones que tienden a hacer más normales los datos. En el caso del algoritmo SVM, se ve que la exactitud lograda en los atributos originales posee un cambio importante de 46,6 % a 70,5 % al aplicar estandarización, lo cual hace posible vislumbrar la utilidad de las transformaciones de datos para el desempeño de ciertos algoritmos.

7.2.3. Búsqueda Exhaustiva

Los resultados de clasificación al utilizar el método de búsqueda exhaustiva, mostrados en la Tabla 6.6, indican un aumento generalizado de la exactitud alcanzada. Al comparar con los resultados de clasificación de la etapa previa, utilizando transformaciones, vistos en la Tabla 6.5, se tiene que la búsqueda exhaustiva eleva la exactitud promedio de los 9 algoritmos empleados, así como la exactitud promedio lograda en el conjunto de atributos originales y sus 8 transformaciones. Lo anterior demuestra la efectividad de la técnica de selección de atributos con el objetivo de mejorar el desempeño en clasificación vía machine learning. Nuevamente, el algoritmo GB obtiene la exactitud promedio más alta, siendo de 86,8 %. Le siguen los algoritmos DT y RF, con exactitudes promedio de 83,7 % y 82,8 %, respectivamente. De igual forma, el algoritmo SVM obtiene la exactitud promedio más baja, con un valor de 74,5 %. Respecto a las transformaciones, la exactitud promedio más elevada se alcanza utilizando la transformación robusta, con un 82,8 %, seguida de la estandarización con 82,6 % y la transformación cuantil-uniforme con 82,2 %. En términos globales, el mejor desempeño se logra con el algoritmo GB, con el empleo de las transformaciones cuantil-normal y cuantil-uniforme, obteniendo 88,8 % de exactitud.

Los resultados explícitos de los mejores modelos obtenidos mediante búsqueda exhaustiva, por algoritmo, exhibidos en las Tablas 6.7 y 6.8 permiten identificar ciertas características propias de los algoritmos utilizados en este problema. Se observa que el algoritmo LDA requiere el uso de 11 de los 15 atributos originales, para alcanzar su mejor desempeño. A este le sigue GB, requiriendo 9 atributos. Por el contrario, los algoritmos que requieren de menos atributos para lograr su máximo desempeño en cuanto a exactitud son NB, SVM y RF, con únicamente 3 de los 15 atributos. En este caso, NB y RF logran la mayor exactitud con el mínimo de atributos, con un valor de 83 %. Junto con esto, el algoritmo RF alcanza su exactitud más elevada en el conjunto de atributos originales y 7 de sus transformaciones, siendo el que más posibilidades de transformaciones abarca. Cabe decir que únicamente para las transformaciones cuantil-normal y cuantil-uniforme se gatilla un cambio (en cantidad) en el subconjunto de atributos seleccionado que maximiza la exactitud, siendo igual en los otros 5 casos, con la cantidad mínima de 3 atributos. A este algoritmo le siguen los algoritmos LDA y NB, los cuales alcanzan su mayor exactitud en el conjunto de atributos originales y 4 de sus transformaciones, siendo estas las mismas para ambos.

En términos generales, se tiene una variación baja, por algoritmo, entre los subconjuntos hallados vía búsqueda exhaustiva para los mejores modelos. Más aún, para un par específico algoritmo-transformación, el subconjunto seleccionado es único en la mayoría de los casos. Las excepciones se dan para LR, sin aplicar transformaciones, con 2 subconjuntos posibles, SVM con la transformación Box-Cox, teniendo 2 subconjuntos posibles, RF con la transformación cuantil-normal, exhibiendo 3 subconjuntos posibles, y la transformación cuantil-uniforme, con 2 subconjuntos posibles, y AB junto con la transformación cuantil-uniforme, teniendo 2 subconjuntos posibles. Por otro lado, se advierte que todos los subconjuntos de atributos asociados a los modelos de mayor exactitud por algoritmo, poseen una combinación de atributos morfológicos y hemodinámicos, lo cual apoya la necesidad y dependencia entre ambas características para el estudio y la discriminación del estado de ruptura de los aneurismas cerebrales [4].

7.2.4. Reglas de Asociación

De acuerdo a los resultados de la Tabla 6.9, se tiene que los 4 atributos que más aparecen en los subconjuntos de atributos hallados por búsqueda exhaustiva, de los modelos que logran o superan una exactitud promedio de 80 %, son en primer lugar α_F, con 91,9 % de soporte, seguido por BNF con 81,1 %, SWSS_n con 75,7 % y SR con 56,8 %. El hecho de que estos atributos sean los más frecuentes en los mejores modelos, destaca su importancia a la hora de discriminar el estado de ruptura. Si bien el soporte obtenido no genera un ordenamiento jerárquico equivalente de estos atributos al gatillado por su significancia estadística, mostrada en la Tabla 3.2, el resultado es consistente con las correlaciones de Spearman de la Figura 6.7, la literatura y los resultados de Amigo [76, 15, 153, 111, 4]. Lo anterior sugiere que la combinación de ciertos atributos con otros, altera sus valores individuales como predictores del estado de ruptura dentro de un problema de clasificación binario abordado con machine learning. Como ejemplo, se tiene que los atributos α_A y NSI, con soportes de 29,7 % y 27,0 % respectivamente, son más frecuentes en los mejores subconjuntos que TAWSS_n, que posee un soporte de 18,9 %, aún cuando según la significancia estadística y el análisis ROC de las Tablas 3.2 y 3.3, junto a la matriz de correlación de la Figura 6.7, se debiese tener el resultado opuesto.

Los resultados de la Tabla 6.10 indican que los parámetros OSI, RRT_n y TAWSS_n aparecen por lo menos 3 veces más en conjunto con otros atributos, en su mayoría morfológicos, de lo que correspondería bajo el supuesto de independencia entre ambos. Los atributos OSI, RRT_n y TAWSS_n son hemodinámicos, por lo cual las reglas de asociación encontradas apoyan una dependencia entre morfología y hemodinámica para lograr una mejor distinción del estado de ruptura mediante machine learning, acorde a lo señalado por Amigo [4]. Estos resultados sugieren que OSI, RRT_n y TAWSS_n requieren de la presencia de otros atributos para aportar a esta tarea de clasificación, considerando los datos trabajados. Es interesante notar que dentro de los atributos que aparecen relacionados mediante las reglas de asociación, se encuentran desde los más relevantes como SR y α_F hasta los más irrelevantes como NSI y AFI. Es por ello que la utilidad de cada atributo no se puede analizar de forma puramente independiente, sino en relación con los demás.

7.2.5. Optimización de Hiperparámetros

A partir de las Tablas 6.11 a la 6.18, se observa que los únicos modelos de máxima exactitud por algoritmo, para los cuales se logra una variación de los hiperparámetros iniciales mediante optimización, son LR y RF. En el caso de LR, se disminuye la cantidad de iteraciones máxima de 100 a 50, y se activa la partida en caliente, lo cual refiere al uso de la solución anterior como inicialización para un nuevo ajuste [71]. Para RF, la cantidad de estimadores o árboles individuales aumenta de 100 a 200. A su vez, la cantidad máxima de hojas o nodos terminales pasa de no tener restricción, a limitarse a 3. A pesar de ello, como se observa en los resultados de la Tabla 6.19, la exactitud alcanzada no varía para ninguno de los modelos de máxima exactitud por algoritmo. Únicamente se aprecia una variación en el valor entregado al utilizar el comando para llevar a cabo la optimización de hiperparámetros, pero esto se debe a un componente aleatorio. A pesar de que las grillas de búsqueda para la optimización de hiperparámetros podrían aumentarse para ampliar el espacio de búsqueda, este resultado indica que para las opciones probadas, los hiperparámetros iniciales ya se encontraban en su mayoría optimizados, al menos en cuanto a exactitud de clasificación se refiere. Junto con lo anterior, el resultado sugiere que el desempeño de los modelos depende mucho más del trabajo realizado en los atributos (selección, transformaciones) que de la optimización de hiperparámetros.

7.2.6. Evaluación de Modelos Finales

Validación Cruzada

Los resultados de la evaluación final mediante validación cruzada, mostrados en la Tabla 6.20, señalan que el mejor desempeño en cuanto a exactitud promedio se logra con el algoritmo GB, alcanzando un valor de 88,8 % ± 8,5 % en el conjunto de prueba. En este caso se combina la mayor exactitud promedio junto con la menor desviación estándar. En cuanto a precisión, el algoritmo DT logra el mayor resultado promedio, con un valor de 78,0 % ± 39,4 %. Para el algoritmo AB no se tienen los resultados de precisión, dado que el componente aleatorio de la validación cruzada no siempre genera la existencia de ambas clases en los conjuntos de entrenamiento y prueba, por lo cual en ciertas iteraciones dicha métrica puede quedar mal definida.

Se observa que en su mayoría, el desempeño promedio es más alto en el conjunto de prueba que en el de entrenamiento, tanto para la exactitud como la precisión. La única excepción se da con el algoritmo LR, el cual exhibe una exactitud promedio de 80,1 % en el conjunto de prueba, superada por 81,6 % en el conjunto de entrenamiento. Esto significa que con excepción del algoritmo LR, todos los modelos de máxima exactitud por algoritmo sufren de sobreajuste en menor o mayor grado. Por su parte, la desviación estándar para ambas métricas siempre es menor en el conjunto de entrenamiento. En el caso de la precisión, la desviación estándar en el conjunto de prueba es mucho mayor, comparando con la exactitud. Esto puede deberse a que aleatoriamente el proceso de validación cruzada genera conjuntos de prueba con alta variación del balance entre clases, otorgando una precisión con alta variación, mientras que el balance de clases en los conjuntos de entrenamiento generados debe poseer una variación mucho más baja. Lo anterior es factible dada la baja cantidad de instancias.

Se observa que en cuanto a exactitud y precisión, todos los modelos en base a algoritmos de árboles, es decir, DT, RF, AB y GB, alcanzan métricas de 100 % ± 0,0 % en el conjunto de entrenamiento, pero su desempeño disminuye en el conjunto de prueba. Así, estos algoritmos sufren de sobrejauste, lo cual es un problema usual de los árboles de decisión. A pesar de ello, GB y DT alcanzan las mayores exactitudes tanto en promedio como en desviación estándar, con valores de 88,8 % ± 8,5 % y 87,3 % ± 10,0 % respectivamente. Con el fin de reducir dicho problema podrían aplicarse técnicas de poda.

Validación Cruzada Estratificada

Los resultados de la evaluación final mediante validación cruzada estratificada, mostrados en la Tabla 6.21, señalan el desempeño según varias métricas. En cuanto a exactitud, RF se posiciona como el mejor algoritmo, logrando 82,6 % ± 13,9 % en el conjunto de prueba, seguido de GB con 82,3 % ± 14,4 %. Esto indica que la exactitud varía y en particular disminuye respecto de la evaluación sin estratificación, como se ve en la Tabla 6.20, para estos dos algoritmos. Lo anterior es consistente dado que los árboles de decisión, base de ambos algoritmos, pueden cambiar considerablemente dependiendo del conjunto de entrenamiento. Por su parte, AB exhibe la peor exactitud, con un valor de 71,5 % ± 17,8 %.

Respecto a la precisión de los modelos finales, el mejor resultado se alcanza con el algoritmo LR, obteniendo 83,2 % ± 18,4 % en el conjunto de prueba. A este le sigue RF con 82,0 % ± 17,1 %. Nuevamente, AB logra el desempeño más bajo, con 74,9 % ± 21,8 %. Por su parte, en términos de sensibilidad, el algoritmo que logra el mejor modelo es LR, con un valor de 85,0 % ± 15,3 %, seguido de RF con 85,0 % ± 21,3 %, con el cual únicamente se genera un aumento de la desviación estándar de la métrica. A su vez, AB alcanza el menor valor de 67,5 % ± 27,2 %. Combinando ambas métricas anteriores, el valor F1 más elevado se logra con el algoritmo LR, llegando a 82,8 % ± 14,1 %, seguido de RF con 81,9 % ± 16,2 %, en el conjunto de prueba. Una vez más, AB exhibe el desempeño más bajo, con 68,1 % ± 22,1 %. Por último, para la AUC de la curva ROC, el algoritmo cuyo modelo logra el valor medio más alto es GB, con un 90,2 % ± 16,1 %, al cual le sigue LDA con 83,5 % ± 16,4 %. En este caso, el desempeño más bajo es el de LR, con 78,5 % ± 16,7 %.

Nuevamente se tiene que todos los modelos en base a algoritmos de árboles de decisión sufren de sobreajuste, ya que logran un desempeño perfecto en el conjunto de entrenamiento en las 5 métricas de evaluación, mientras que el nivel alcanzado en el conjunto de prueba es menor. Por otro lado, en términos generales se considera que la evaluación en virtud de las 5 métricas es satisfactoria, dado que el valor promedio logrado en cada una no varía de forma considerable de una métrica a otra, por algoritmo. Lo anterior es consistente con la validación cruzada estratificada, la cual conserva la proporción entre clases, que se encuentra balanceada para el total de instancias.

Resumen

El modelo de mayor exactitud logrado mediante el algoritmo GB, con un valor de 88,8 % ± 8,5 % en validación cruzada con 10 particiones, logra superar los resultados obtenidos con el mismo algoritmo en el estudio similar desarrollado y publicado por Niemann *et al.* en 2018 [93]. Por otro lado, este modelo se ve ligeramente superado por aquel desarrollado por Aranda y Valencia en el 2018, con una exactitud de 92,86 % en el conjunto de prueba [9]. Además, la exactitud obtenida por Bisbal *et al.*, de 95,5 %, usando SVM, aventaja a su vez lo logrado en el presente trabajo [16]. A pesar de esto, el mejor resultado aquí alcanzado se encuentra en torno a los valores reportados en la literatura, ubicándose satisfactoriamente próximo a los resultados más elevados dentro de lo revisado.

En relación a la evaluación mediante la AUC de la curva ROC, el modelo basado en GB alcanza un 90,2 % ± 16,1 % mediante validación cruzada estratificada. Al respecto, Niemann *et al.* alcanzan un máximo comparativamente inferior de 70 % ± 2 % en su trabajo [93]. Por su parte, Liu *et al.* superan levemente la AUC obtenida, alcanzando un valor de 92,8 % en el conjunto de prueba [78]. Junto con esto, Aranda y Valencia consiguen una AUC máxima de 94,4 % con el algoritmo AB, superando lo obtenido en el presente estudio por un margen pequeño [10]. Por otro lado, Silva *et al.* alcanzan una AUC máxima de 81 %, siendo sobrepasada por el resultado aquí expuesto [127]. A esto se suma el resultado obtenido con el modelo mixto de Amigo, visto en la Tabla 3.4, el cual llega a una AUC de 82 % [4]. Considerando lo anterior, se tiene que en términos de AUC, el desempeño logrado se posiciona dentro del rango reportado en la literatura, acercándose a los mayores valores alcanzados entre los estudios considerados.

En relación al desempeño cuantificado a través de otras métricas, como la sensibilidad y el valor F1, reportados por Suzuki *et al.*, en este estudio se alcanzan valores considerablemente superiores, llegando a una sensibilidad de 85,0 % ± 15,3 % [133]. A pesar de ello, se requiere de una revisión de la literatura más amplia como para poder contrastar de forma más objetiva los resultados en métricas distintas a la exactitud y AUC, vía machine learning.

La evaluación de desempeño por algoritmo se condice parcialmente con la literatura. Por una parte, ciertos estudios alcanzan su máxima exactitud y/o superan el máximo aquí obtenido con SVM, mientras que en este caso, dicho algoritmo presenta un desempeño pobre en términos generales [16, 9]. Por su lado, el algoritmo AB exhibe el peor desempeño en la evaluación realizada mediante validación cruzada estratificada, en contraste a lo obtenido por Aranda y Valencia, siendo el algoritmo de mejor desempeño considerando la AUC alcanzada, entre otros algoritmos probados [10]. Por el contrario, GB aparece como uno de los mejores algoritmos a partir de este estudio, lo cual concuerda con lo reportado por Niemann *et al.*, quienes también probaron una serie de algoritmos de forma simultánea [93]. Lo anterior sugiere que, si bien los algoritmos presentan diferencias en su desempeño, hasta cierto punto bien establecidas o siguiendo tendencias generales, dicho desempeño se encuentra en función de diversos factores, tales como hiperparámetros, atributos, y forma de evaluación, entre otros. Así, no es posible predecir con exactitud qué algoritmo obtendrá un desempeño mejor en cada problema particular. Por esta razón, es conveniente probar varias opciones de algoritmos con el fin de aumentar las posibilidades de alcanzar valores mayores en las métricas de evaluación.

Por último, en cuanto a los resultados médicos, Sailer *et al.* en su revisión sistemática y meta-análisis del 2013, señalan una sensibilidad agrupada de 95 % para la angiografía por resonancia magnética [118]. Por su parte, Yang *et al.* reportan en 2017 exactitudes máximas mediante angiografía por tomografía computarizada, de 97,2 % y 97 % en aneurismas cerebrales pequeños rotos y no rotos, respectivamente. Asimismo, reportan sensibilidades máximas de 97,4 % y 90,3 % para los casos rotos y no rotos, respectivamente [155]. En virtud de estos valores, se tiene que los resultados de exactitud y sensibilidad obtenidos mediante clasificación binaria con machine learning, tanto en el presente trabajo como en los estudios previos, se encuentran ligeramente por debajo de lo logrado con las técnicas de diagnóstico actuales en medicina. Lo anterior, junto con la mejora continua y rápida de las técnicas de machine learning, CFD, y neuroimagenología, plantea el desafío de proseguir con el desarrollo de esta área de investigación, la cual cuenta con perspectivas promisorias de equiparar y eventualmente mejorar el diagnóstico actual de la práctica clínica en un futuro cercano.

Capítulo 8

Conclusiones

Mediante el trabajo realizado se generan varios modelos de machine learning que permiten clasificar el estado de ruptura de aneurismas cerebrales basándose en la morfología y la hemodinámica. Dichos modelos poseen distintos desempeños en función de las métricas de evaluación empleadas. Específicamente, se consigue llevar a cabo el entrenamiento de los modelos a partir de atributos morfológicos y hemodinámicos para un problema de clasificación binaria, teniendo como posibles etiquetas los estados no roto y roto. Usando técnicas tales como selección de atributos, transformación de atributos e incorporación de distintos algoritmos, se logra alcanzar una exactitud de clasificación mediante validación cruzada máxima de 88,8 % ± 8,5 %, aproximándose al valor objetivo de 90 %. A su vez, se obtiene una sensibilidad máxima de 85,0 % ± 15,3 % y una AUC de la curva ROC máxima de 90,2 % ± 16,1 %. Estos desempeños se posicionan levemente por debajo de los valores máximos reportados en la práctica clínica y mediante machine learning para este problema de clasificación, por lo cual el trabajo se considera satisfactorio, más aún, tomando en cuenta que la cantidad de instancias disponibles es baja.

Los resultados obtenidos ponen de manifiesto la utilidad de la caracterización morfológica y hemodinámica de los aneurismas cerebrales, mediante los atributos o parámetros utilizados, para discriminar su estado de ruptura. Las técnicas desarrolladas validan la relevancia individual de ciertos atributos de manera consiste con lo reportado previamente. A pesar de que no se consigue evidenciar una dependencia explícita entre atributos morfológicos y hemodinámicos, lo encontrado sugiere el uso combinado de ambos tipos de parámetros para generar modelos de clasificación de mayor exactitud. Junto con esto, se resalta la importancia de cada atributo en relación con los otros, además de su importancia individual, lo cual apunta hacia la complejidad de la ruptura de este tipo de lesiones, que involucra la interacción morfológica y hemodinámica. Por otro lado, se vislumbra que la combinación de una cantidad reducida del total de atributos a disposición permite alcanzar exactitudes elevadas. Esto, junto con la estadística, ayuda a focalizar la obtención de ciertos parámetros particulares, apuntando a una generación de datos más eficiente.

A partir de lo presentado se pueden implementar distintos focos de trabajo y modificaciones con el fin de explorar la generación de modelos de mayor desempeño a la hora de clasificar el estado de ruptura de aneurismas cerebrales. Las posibilidades a considerar son múltiples, incluyendo: generación de más instancias y/o atributos, uso de otras transformaciones o modificación de las ya utilizadas (si aplica) e incorporación de otros algoritmos de machine learning para clasificación, entre otras. Dentro de estas, se identifica el potencial de atributos adicionales a los morfológicos y hemodinámicos, incorporándolos a estos últimos, para discriminar el estado de ruptura. Por esta razón, se requiere de una mayor investigación en este tema, con el fin de alcanzar mejores resultados y ampliar el conocimiento.

Las técnicas provenientes de la inteligencia artificial, como el machine learning, se encuentran en avance para la resolución de diversas tareas de aspecto práctico. Este trabajo constituye una prueba de la utilidad del machine learning para abordar la clasificación del estado de ruptura de aneurismas cerebrales. Los resultados reportados tanto aquí como en la literatura son promisorios, señalando a la aparición de máquinas inteligentes que asesoren e incluso mejoren el diagnóstico actual en la práctica clínica, dentro del futuro cercano. Dentro de los diversos desafíos para lograrlo, se encuentra la rapidez requerida en la generación de datos morfológicos y hemodinámicos de un paciente particular, como parte del proceso clínico. La tendencia creciente de la ciencia y tecnología, abarcando las simulaciones computacionales, el machine learning, y la neuroimagenología, auguran la superación de este y otros desafíos, además del mejor manejo clínico y comprensión de los aneurismas cerebrales.

Bibliografía

[1] Tammam Abboud, Jihad Rustom, Maxim Bester, Patrick Czorlich, Eik Vittorazzi, Hans O. Pinnschmidt, Manfred Westphal, and Jan Regelsberger. Morphology of ruptured and unruptured intracranial aneurysms. *World Neurosurgery*, 99:610 – 617, 2017.

[2] Norman Ajiboye, Nohra Chalouhi, Robert Starke, Mario Zanaty, and Rodney Bell. Unruptured cerebral aneurysms: Evaluation and management. *The Scientific World Journal*, 2015:1–10, 07 2015.

[3] M al Yamany and IB Ross. Giant fusiform aneurysm of the middle cerebral artery: successful hunterian ligation without distal bypass. *British journal of neurosurgery*, 12(6):572—575, December 1998.

[4] Nicolás Amigo. *Caracterización morfológica y estudio de la hemodinámica de aneurismas cerebrales humanos mediante simulaciones computacionales*. Tesis doctoral, Universidad de Chile, 2018.

[5] Nicolás Amigo and Alvaro Valencia. Determining significant morphological and hemodynamic parameters to assess the rupture risk of cerebral aneurysms. *Journal of Medical and Biological Engineering*, 39, 04 2018.

[6] C. Anderson, G. Hankey, K. Jamrozik, and D. Dunbabin. Epidemiology of aneurysmal subarachnoid hemorrhage in australia and new zealand: Incidence and case fatality from the australasian cooperative research on subarachnoid hemorrhage study (across). *Stroke*, 31(8):1843–1850, 2000. cited By 151.

[7] J.D. Anderson. Governing equations of fluid dynamics. In John F. Wendt, editor, *Computational Fluid Dynamics*, pages 15–51. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[8] Aplimed. Aneurisma cerebral [en línea]. [Recuperado el 14 de julio de 2019 de `http://www.aplimed.com/aneurisma-cerebral.php#.XSt7S-tKjIU`].

[9] Alfredo Aranda and Alvaro Valencia. Study on cerebral aneurysms: Rupture risk prediction using geometrical parameters and wall shear stress with cfd and machine learning tools. *Machine Learning and Applications: An International Journal*, 5, 12 2018.

[10] Alfredo Aranda and Alvaro Valencia. Computational study on the rupture risk in real cerebral aneurysms with geometrical and fluid-mechanical parameters using fsi simulations and machine learning algorithms. *Journal of Mechanics in Medicine and Biology*, page 1950014, 02 2019.

[11] Sudharsan Asaithambi. Why, how and when to scale your features [en línea], 2017. [Recuperado el 28 de agosto de 2019 de `https://medium.com/greyatom/why-how-and-when-to-scale-your-features-4b30ab09db5e`].

[12] V. Kishore Ayyadevara. Logistic regression. In *Pro Machine Learning Algorithms : A Hands-On Approach to Implementing Algorithms in Python and R*, pages 49–69. Apress, Berkeley, CA, 2018.

[13] Mihalj Bakator and Dragica Radosav. Deep learning and medical diagnosis: A review of literature. *Multimodal Technologies and Interaction*, 2:47, 08 2018.

[14] Viv Bewick, Liz Cheek, and Jonathan Ball. Statistics review 8: Qualitative data - tests of association. *Critical Care*, 8(1):46, 2003.

[15] Pervinder Bhogal, Muhammad Almatter, Victoria Hellstern, Oliver Ganslandt, H Bäzner, Hans Henkes, and MAguilar Pérez. Difference in aneurysm characteristics between ruptured and unruptured aneurysms in patients with multiple intracranial aneurysms. *Surgical Neurology International*, 9:1, 01 2018.

[16] Jesus Bisbal, Gerhard Engelbrecht, Maria-Cruz Villa-Uriol, and Alejandro Frangi. Prediction of cerebral aneurysm rupture using hemodynamic, morphologic and clinical features: A data mining approach. In *Database and Expert Systems Applications: 22nd International Conference, DEXA 2011, Toulouse, France, August 29 - September 2, 2011, Proceedings, Part II*, volume 6861, pages 59–73, 08 2011.

[17] G. Boulouis, C. Rodriguez-Régent, E.C. Rasolonjatovo, W. Ben Hassen, D. Trystram, M. Edjlali-Goujon, J.-F. Meder, C. Oppenheim, and O. Naggara. Unruptured intracranial aneurysms: An updated review of current concepts for risk factors, detection and management. *Revue Neurologique*, 173(9):542 – 551, 2017. INTERNATIONAL SFN / SFNV MEETING 2017.

[18] Max Bramer. *Principles of Data Mining*. Springer London, London, 2016.

[19] Tom Brijs, Koen Vanhoof, and Geert WETS. Defining interestingness for association rule. *International Journal Information Theories and Applications*, 10, 01 2003.

[20] Jonathan L. Brisman, Joon K. Song, and David W. Newell. Cerebral aneurysms. *New England Journal of Medicine*, 355(9):928–939, 2006. PMID: 16943405.

[21] Jason Brownlee. A gentle introduction to the gradient boosting algorithm for machine learning [en línea], 2016. [Recuperado el 1 de agosto de `https://machinelearningmastery.com/gentle-introduction-gradient-boosting-algorithm-machine-learning/`].

[22] Jason Brownlee. A gentle introduction to the bootstrap method [en línea], 2018. [Recuperado el 31 de julio de 2019 de https://machinelearningmastery.com/a-gentle-introduction-to-the-bootstrap-method/].

[23] Jason Brownlee. How to calculate correlation between variables in python [en línea], 2018. [Recuperado el 25 de julio de 2019 de https://machinelearningmastery.com/how-to-use-correlation-to-understand-the-relationship-between-variables/].

[24] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

[25] Jun Soo Byun, Sun Young Choi, and Taewon Seo. The numerical study of the hemodynamic characteristics in the patient-specific intracranial aneurysms before and after surgery. *Computational and Mathematical Methods in Medicine*, 2016:1–12, 05 2016.

[26] Ian Campbell, Jared Ries, Saurabh S Dhawan, Arshed Quyyumi, W Taylor, and John Oshinski. Effect of inlet velocity profiles on patient-specific computational fluid dynamics simulations of the carotid bifurcation. *Journal of biomechanical engineering*, 134:051001, 05 2012.

[27] Frank Ceballos. Scikit-learn decision trees explained [en línea]. [Recuperado el 31 de julio de 2019 de https://towardsdatascience.com/scikit-learn-decision-trees-explained-803f3812290d].

[28] J.R. Cebral, F. Mut, J. Weir, and C.M. Putman. Association of hemodynamic characteristics and cerebral aneurysm rupture. *American Journal of Neuroradiology*, 32(2):264–270, 2011.

[29] Nohra Chalouhi, Brian L. Hoh, and David Hasan. Review of cerebral aneurysm formation, growth, and rupture. *Stroke*, 44(12):3613–3622, 2013.

[30] BJ Chung, F Mut, CM Putman, F Hamzei-Sichani, W Brinjikji, D Kallmes, CM Jimenez, and JR Cebral. Identification of hostile hemodynamics and geometries of cerebral aneurysms: A case-control study. *AJNR. American journal of neuroradiology*, 39(10):1860—1866, October 2018.

[31] Alessandro Cianfoni, Emanuele Pravatà, Roberto De Blasi, Costa Silvia Tschuor, and Giuseppe Bonaldi. Clinical presentation of cerebral aneurysms. *European Journal of Radiology*, 82(10):1618 – 1622, 2013.

[32] T. David and R.G. Brown. Chapter 7 - models of cerebrovascular perfusion. In Sid M. Becker and Andrey V. Kuznetsov, editors, *Transport in Biological Media*, pages 253 – 273. Elsevier, Boston, 2013.

[33] Sujan Dhar, Markus Tremmel, J Mocco, Minsuok Kim, Junichi Yamamoto, Adnan H. Siddiqui, L. Nelson Hopkins, and Hui Meng. MORPHOLOGY PARAMETERS FOR INTRACRANIAL ANEURYSM RUPTURE RISK ASSESSMENT. *Neurosurgery*, 63(2):185–197, 08 2008.

[34] Zhihui Duan, Yuanhui Li, Sheng Guan, Congmin Ma, Yuezhen Han, Xiangyang Ren, Liping Wei, Wenbo Li, Jiyu Lou, and Zhiyuan Yang. Morphological parameters and anatomical locations associated with rupture status of small intracranial aneurysms. *Scientific Reports*, 8, 12 2018.

[35] Christopher S. Eddleman, Christopher C. Getch, Bernard R. Bendok, and H. Hunt Batjer. Chapter 13 - intracranial aneurysms. In Richard G. Ellenbogen, Saleem I. Abdulrauf, and Laligam N. Sekhar, editors, *Principles of Neurological Surgery (Third Edition)*, pages 209 – 228. W.B. Saunders, Philadelphia, third edition edition, 2012.

[36] Øyvind Evju and Kent-Andre Mardal. On the assumption of laminar flow in physiological flows: Cerebral aneurysms as an illustrative example. In Alfio Quarteroni, editor, *Modeling the Heart and the Circulatory System*, pages 177–195. Springer International Publishing, Cham, 2015.

[37] Usama Fayyad and Ramasamy Uthurusamy. Data mining and knowledge discovery in databases. *Commun. ACM*, 39(11):24–26, November 1996.

[38] David L. Felten, M. Kerry O'Banion, and Mary Summo Maida. 7 - vasculature. In David L. Felten, M. Kerry O'Banion, and Mary Summo Maida, editors, *Netter's Atlas of Neuroscience (Third Edition)*, pages 93 – 124. Elsevier, Philadelphia, third edition edition, 2016.

[39] Vernard S. Fennell, Nikolay L. Martirosyan, Sheri K. Palejwala, G. Michael Lemole, and Travis M. Dumont. Morbidity and mortality of patients with endovascularly treated intracerebral aneurysms: does physician specialty matter? *Journal of Neurosurgery JNS*, 124(1):13 – 17, 2016.

[40] Carolyn U. Fisher and Jenn Stroud Rossmann. Effect of non-newtonian behavior on hemodynamics of cerebral aneurysms. *Journal of biomechanical engineering*, 131 9:091004, 2009.

[41] Phil B Fontanarosa. Recognition of subarachnoid hemorrhage. *Annals of Emergency Medicine*, 18(11):1199 – 1205, 1989.

[42] Joe Niekro Foundation. Types of cerebral aneurysms [en línea]. [Recuperado el 14 de julio de 2019 de `https://www.joeniekrofoundation.com/understanding/types-of-cerebral-aneurysms/`].

[43] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232, 2001.

[44] Juhana Frösen, Riikka Tulamo, Anders Paetau, Elisa Laaksamo, Miikka Korja, Aki Laakso, Mika Niemelä, and Juha Hernesniemi. Saccular intracranial aneurysm: pathology and mechanisms. *Acta Neuropathologica*, 123(6):773–786, Jun 2012.

[45] A.J. Geers, I. Larrabide, H.G. Morales, and A.F. Frangi. Approximating hemodynamics of cerebral aneurysms with steady flow simulations. *Journal of Biomechanics*, 47(1):178 – 185, 2014.

[46] Prince Grover. Gradient boosting from scratch [en línea], 2017. [Recuperado el 1 de agosto de `https://medium.com/mlreview/gradient-boosting-from-scratch-1e317ae4587d`].

[47] Fabrizio Gómez. Estudio numérico de la turbulencia en aneurismas cerebrales. Memoria de título, Universidad de Chile, 2010.

[48] Harp. Algorithms in harp-daal [en línea]. [Recuperado el 31 de julio de `https://dsc-spidal.github.io/harp/docs/harpdaal/algorithms/`].

[49] Michael P. Hartung, Thomas M. Grist, and Christopher J. François. Magnetic resonance angiography: current status and future directions. *Journal of Cardiovascular Magnetic Resonance*, 13(1):19, Mar 2011.

[50] Tomoki Hashimoto, Hui Meng, and William L Young. Intracranial aneurysms: Links among inflammation, hemodynamics and vascular remodeling. *Neurological research*, 28:372–80, 07 2006.

[51] J.W. Hop, G.J.E. Rinkel, A. Algra, and J. Van Gijn. Case-fatality rates and functional outcome after subarachnoid hemorrhage: A systematic review. *Stroke*, 28(3):660–664, 1997. cited By 731.

[52] Howard H. Hu. Chapter 10 - computational fluid dynamics. In Pijush K. Kundu, Ira M. Cohen, and David R. Dowling, editors, *Fluid Mechanics (Fifth Edition)*, pages 421 – 472. Academic Press, Boston, fifth edition edition, 2012.

[53] Yohsuke Imai, Kodai Sato, Takuji Ishikawa, and Takami Yamaguchi. Inflow into saccular cerebral aneurysms at arterial bends. *Annals of Biomedical Engineering*, 36(9):1489, Jun 2008.

[54] Yohichi Imaizumi, Tohru Mizutani, Katsuyoshi Shimizu, Yosuke Sato, and Junichi Taguchi. Detection rates and sites of unruptured intracranial aneurysms according to sex and age: an analysis of mr angiography–based brain examinations of 4070 healthy japanese adults. *Journal of Neurosurgery JNS*, 130(2):573 – 578, 2018.

[55] Timothy John Ingall, Kjell Asplund, Markku S Mähönen, and Ruth Bonita. A multinational comparison of subarachnoid hemorrhage epidemiology in the who monica stroke study. *Stroke*, 31 5:1054–61, 2000.

[56] Jørgen Gjernes Isaksen, Yuri Bazilevs, Trond Kvamsdal, Yongjie Zhang, Jon H. Kaspersen, Knut Waterloo, Bertil Romner, and Tor Ingebrigtsen. Determination of wall tension in cerebral artery aneurysms by numerical simulation. *Stroke*, 39(12):3172–3178, 2008.

[57] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *Resampling Methods*, pages 175–201. Springer New York, New York, NY, 2013.

[58] Ivo Jansen, Joppe Schneiders, W V Potters, Pim Ooij, René Berg, Ed Vanbavel, Henk Marquering, and Charles Majoie. Generalized versus patient-specific inflow boundary conditions in computational fluid dynamics simulations of cerebral aneurysmal hemodynamics. *AJNR. American journal of neuroradiology*, 35, 03 2014.

[59] Sreenivas Jayanti. Equations governing fluid motion. In *Computational Fluid Dynamics for Engineers and Scientists*, pages 17–60. Springer Netherlands, Dordrecht, 2018.

[60] Sina Jelodar, Mohammad Shirani, Amin Tavallayi, Mehdi Ketabchi, and Maysam Alimohamadi. Bilateral fusiform aneurysms of the internal carotid arteries presenting with subarachnoid hemorrhage. *Journal of Stroke and Cerebrovascular Diseases*, 26(6):e111 – e113, 2017.

[61] Woowon Jeong and Kyehan Rhee. Hemodynamics of cerebral aneurysms: Computational analyses of aneurysm progress and treatment. *Computational and mathematical methods in medicine*, 2012:782801, 02 2012.

[62] Pengjun Jiang, Qingyuan Liu, Jun Wu, Xin Chen, Maogui Li, Zhengsong Li, Shuzhe Yang, Rui Guo, Bin Gao, Yong Cao, and Shuo Wang. A novel scoring system for rupture risk stratification of intracranial aneurysms: A hemodynamic and morphological study. *Frontiers in Neuroscience*, 12:596, 2018.

[63] Yo-El S. Ju and Todd J Schwedt. Abrupt-onset severe headaches. *Seminars in neurology*, 30 2:192–200, 2010.

[64] Abhishek Kar. Introduction to linear discriminant analysis [en línea], 2017. [Recuperado el 29 de julio de `https://analyticsdefined.com/introduction-linear-discriminant-analysis/`].

[65] Sai Nikhilesh Kasturi. Underfitting and overfitting in machine learning and how to deal with it !!! [en línea], 2019. [Recuperado el 5 de septiembre de `https://towardsdatascience.com/underfitting-and-overfitting-in-machine-learning-and-how-to-deal-with-it-6fe4a8a49dbf`].

[66] Timo Koivisto, Ritva Vanninen, Heleena Hurskainen, Tapani Saari, Juha Hernesniemi, and Matti Vapalahti. Outcomes of early endovascular versus surgical treatment of ruptured cerebral aneurysms. *Stroke*, 31(10):2369–2377, 2000.

[67] Marc Kotowski, Olivier Naggara, Tim E Darsaut, Suzanne Nolet, Guylaine Gevry, Evgueni Kouznetsov, and Jean Raymond. Safety and occlusion rates of surgical treatment of unruptured intracranial aneurysms: a systematic review and meta-analysis of the literature from 1990 to 2011. *Journal of Neurology, Neurosurgery & Psychiatry*, 84(1):42–48, 2013.

[68] Masaomi Koyanagi, Akira Ishii, Hirotoshi Imamura, Tetsu Satow, Kazumichi Yoshida, Hitoshi Hasegawa, Takayuki Kikuchi, Yohei Takenobu, Mitsushige Ando, Jun C. Takahashi, Ichiro Nakahara, Nobuyuki Sakai, and Susumu Miyamoto. Long-term outcomes of coil embolization of unruptured intracranial aneurysms. *Journal of Neurosurgery JNS*, 129(6):1492 – 1498, 2018.

[69] D Krex, H K Schackert, and G Schackert. Genesis of cerebral aneurysms - an update. *Acta neurochirurgica*, 143:429–48; discussion 448, 02 2001.

[70] Miroslav Kubat. *An Introduction to Machine Learning*. Springer USSpringer International Publishing, Cham, 2017.

[71] Scikit learn developers. Scikit-learn user guide (release 0.21.2) [en línea], 2019. [Recuperado el 28 de julio de `https://scikit-learn.org/stable/documentation.html`].

[72] C.J. Lee, Y. Zhang, H. Takao, Y. Murayama, and Y. Qian. A fluid–structure interaction study using patient-specific ruptured and unruptured aneurysm: The effect of aneurysm morphology, hypertension and elasticity. *Journal of Biomechanics*, 46(14):2402 – 2410, 2013.

[73] Gwang-Jin Lee, Ki-Seong Eom, Cheol Lee, Dae-Won Kim, and Sung-Don Kang. Rupture of very small intracranial aneurysms: Incidence and clinical characteristics. *Journal of Cerebrovascular and Endovascular Neurosurgery*, 17:217, 11 2015.

[74] Michael R Levitt, Michael C Barbour, Sabine Rolland du Roscoat, Christian Geindreau, Venkat K Chivukula, Patrick M McGah, John D Nerva, Ryan P Morton, Louis J Kim, and Alberto Aliseda. Computational fluid dynamics of cerebral aneurysm coiling using high-resolution and high-energy synchrotron x-ray microtomography: comparison with the homogeneous porous medium approach. *Journal of NeuroInterventional Surgery*, 9(8):00–00, 2017.

[75] Ming-Hua Li, Shi-Wen Chen, Yong-Dong Li, Yuan-Chang Chen, Yingsheng Cheng, Ding-Jun Hu, Hua-Qiao Tan, Qian Wu, daze xinnan, Zhen-Kui Sun, Xiao-Er Wei, Jia-Yin Zhang, Rui-Hua Qiao, Wen-Hong Zong, Yin Zhang, Wei Lou, Zhi-Yuan Chen, Yu Zhu, De-Rong Peng, and Wei-Ping Jia. Prevalence of unruptured cerebral aneurysms in chinese adults aged 35 to 75 years: A cross-sectional study. *Annals of internal medicine*, 159:514–521, 10 2013.

[76] Ning Lin, Allen Ho, Nareerat Charoenvimolphan, Kai U Frerichs, Arthur L Day, and Rose Du. Analysis of morphological parameters to differentiate rupture status in anterior communicating artery aneurysms. *PloS one*, 8:e79635, 11 2013.

[77] F.H.H. Linn, G.J.E. Rinkel, A. Algra, and J. van Gijn. Incidence of subarachnoid hemorrhage. *Stroke*, 27(4):625–629, 1996.

[78] Jinjin Liu, Yongchun Chen, Li Lan, Boli Lin, Weijian Chen, Meihao Wang, Rui Li, Yunjun Yang, Bing Zhao, Zilong Hu, and Yuxia Duan. Prediction of rupture risk in anterior communicating artery aneurysms with a feed-forward artificial neural network. *European Radiology*, 28(8):3268–3275, Aug 2018.

[79] Daniel M Sforza, Christopher M Putman, and Juan Raul Cebral. Hemodynamics of cerebral aneurysms. *Annual review of fluid mechanics*, 41:91–107, 01 2009.

[80] Oded Maimon and Lior Rokach. *Introduction to Knowledge Discovery and Data Mining*, pages 1–15. Springer US, Boston, MA, 2010.

[81] A. Mantha, C. Karmonik, G. Benndorf, C. Strother, and R. Metcalfe. Hemodynamics in a cerebral artery before and after the formation of an aneurysm. *American Journal of Neuroradiology*, 27(5):1113–1118, 2006.

[82] Manik Mehra, Gabriela Spilberg, Matthew J. Gounis, and Ajay K. Wakhloo. Intracranial aneurysms: Clinical assessment and treatment options. In Tim McGloughlin, editor, *Biomechanics and Mechanobiology of Aneurysms*, pages 331–372. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[83] H. Meng, V.M. Tutino, J. Xiang, and A. Siddiqui. High wss or low wss? complex interactions of hemodynamics with intracranial aneurysm initiation, growth, and rupture: Toward a unifying hypothesis. *American Journal of Neuroradiology*, 35(7):1254–1262, 2014.

[84] Julia Mikhal, Cornelis H. Slump, and Bernard J. Geurts. Simulation of pulsatile flow in cerebral aneurysms: From medical images to flow and forces. In Yasuo Murai, editor, *Aneurysm*, chapter 10. IntechOpen, Rijeka, 2012.

[85] Diego Miranda. Estudio numérico de diferente modelos de pared en aneurismas cerebrales. Memoria de título, Universidad de Chile, 2017.

[86] Yoichi Miura, Fujimaro Ishida, Yasuyuki Umeda, Hiroshi Tanemura, Hidenori Suzuki, Satoshi Matsushima, Shinichi Shimosaka, and Waro Taki. Low wall shear stress is independently associated with the rupture status of middle cerebral artery aneurysms. *Stroke*, 44(2):519–521, 2013.

[87] Hernan Morales, Ignacio Larrabide, Arjan Geers, Martha L Aguilar, and Alejandro Frangi. Newtonian and non-newtonian blood flow in coiled cerebral aneurysms. *Journal of biomechanics*, 46, 07 2013.

[88] Pablo Munarriz, Pedro A. Gómez, Igor Paredes, Ana Maria Castaño Leon, Santiago Cepeda, and Alfonso Lagares. Basic principles of hemodynamics and cerebral aneurysms. *World Neurosurgery*, 88, 01 2016.

[89] Yuichi Murayama, Hiroyuki Takao, Toshihiro Ishibashi, Takayuki Saguchi, Masaki Ebara, Ichiro Yuki, Hideki Arakawa, Koreaki Irie, Mitsuyoshi Urashima, and Andrew J. Molyneux. Risk analysis of unruptured intracranial aneurysms. *Stroke*, 47(2):365–371, 2016.

[90] Fernando Mut, Rainald Löhner, Aichi Chien, Satoshi Tateshima, Fernando Viñuela, Christopher Putman, and Juan R. Cebral. Computational hemodynamics framework for the analysis of cerebral aneurysms. *International Journal for Numerical Methods in Biomedical Engineering*, 27(6):822–839, 2011.

[91] Avinash Navlani. Knn classification using scikit-learn [en línea], 2018. [Recuperado el 29 de julio de `https://www.datacamp.com/community/tutorials/k-nearest-neighbor-classification-scikit-learn`].

[92] Avinash Navlani. Support vector machines with scikit-learn [en línea], 2018. [Recuperado el 31 de julio de `https://www.datacamp.com/community/tutorials/svm-classification-scikit-learn-python`].

[93] U. Niemann, P. Berg, A. Niemann, O. Beuing, B. Preim, M. Spiliopoulou, and S. Saalfeld. Rupture status classification of intracranial aneurysms using morphological parameters. In *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, pages 48–53, June 2018.

[94] Akira Nishimoto, Kazuyoshi Ueta, Hideaki Onbe, Katsutoshi Kitamura, Tsuyoshi Omae, Fumio Goto, Genju Ohneda, Hiroo Chigasaki, Masanobu Tsuru, and Jiro Suzuki. Nationwide co-operative study of intracranial aneurysm surgery in japan. *Stroke*, 16 1:48–52, 1985.

[95] National Institute of Biomedical Imaging and Bioengineering [en línea]. Computed tomography (ct). [Recuperado el 15 de julio de 2019 de `https://www.nibib.nih.gov/science-education/science-topics/computed-tomography-ct`].

[96] National Institute of Neurological Disorders and Stroke [en línea]. Cerebral aneurysms fact sheet. [Recuperado el 15 de julio de 2019 de `https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Fact-Sheets/Cerebral-Aneurysms-Fact-Sheet`].

[97] Columbia University Department of Neurology. Cerebral aneurysm [en línea]. [Recuperado el 27 de junio de 2019 de `http://www.columbianeurology.org/neurology/staywell/document.php?id=35871`].

[98] California Institute of Neuroscience. Flow diversion for aneurysm [en línea]. [Recuperado el 15 de julio de 2019 de `https://www.cineuro.org/specialties/neuro-intervention/flow-diversion-for-aneurysm/`].

[99] Health Encyclopedia-University of Rochester Medical Center. Endovascular coiling [en línea]. [Recuperado el 15 de julio de 2019 de `https://www.urmc.rochester.edu/encyclopedia/content.aspx?contenttypeid=92&contentid=P08768`].

[100] University Medical Center of the Johannes Gutenberg University Mainz. Conventional angiography [en línea]. [Recuperado el 15 de julio de 2019 de https://www.unimedizin-mainz.de/neuroradiologie/information-for-patients/conventional-angiography.html?L=1].

[101] The International Study of Unruptured Intracranial Aneurysms Investigators. Unruptured intracranial aneurysms — risk of rupture and risks of surgical intervention. *New England Journal of Medicine*, 339(24):1725–1733, 1998. PMID: 9867550.

[102] David L. Olson. *Association Rules*, pages 61–69. Springer Singapore, Singapore, 2017.

[103] Rafik Ouared, Ignacio Larrabide, Olivier Brina, Pierre Bouillot, Gorislav Erceg, Hasan Yilmaz, Karl-Olof Lovblad, and Vitor Mendes Pereira. Computational fluid dynamics analysis of flow reduction induced by flow-diverting stents in intracranial aneurysms: a patient-unspecific hemodynamics change perspective. *Journal of NeuroInterventional Surgery*, 8(12):1288–1293, 2016.

[104] Lee Chang Young Kim Ealmaan Son Eun Ik Park Seong Ho, Yim Man Bin. Intracranial fusiform aneurysms: It's pathogenesis, clinical characteristics and managements. *J Korean Neurosurg Soc*, 44(3):116–123, 2008.

[105] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[106] Michel Piotin, Alessandra Biondi, Nader Sourour, Charbel Mounayer, Maciej Jaworski, Salvatore Mangiafico, Tommy Andersson, Michael Söderman, Pierre Goffette, René Anxionnat, and Raphaël Blanc. The luna aneurysm embolization system for intracranial aneurysm treatment: short-term, mid-term and long-term clinical and angiographic results. *Journal of NeuroInterventional Surgery*, 10(12):e34–e34, 2018.

[107] Dinesh J. Prajapati, Sanjay Garg, and N.C. Chauhan. Interesting association rule mining with consistent and inconsistent rule detection from big sales data in distributed environment. *Future Computing and Informatics Journal*, 2(1):19 – 30, 2017.

[108] Y. Qian, H. Takao, M. Umezu, and Y. Murayama. Risk analysis of unruptured aneurysms using computational fluid dynamics technology: Preliminary results. *American Journal of Neuroradiology*, 32(10):1948–1955, 2011.

[109] Zhuang Qiang Long Jianwu Yang Fan Zhang Hongqi Qin Hao, Yang Qixia. Morphological and hemodynamic parameters for middle cerebral artery bifurcation aneurysm rupture risk assessment. *J Korean Neurosurg Soc*, 60(5):504–510, 2017.

[110] Juan R Cebral, Marcelo Castro, Sunil Appanaboyina, Christopher M Putman, Daniel Millán, and Alejandro Frangi. Efficient pipeline for image-based patient-specific analysis of cerebral aneurysm hemodynamics: Technique and sensitivity. *IEEE transactions on medical imaging*, 24:457–467, 05 2005.

[111] Maryam Rahman, Janel Smietana, Erik Hauck, Brian Hoh, Nick Hopkins, Adnan Siddiqui, Elad I. Levy, Hui Meng, and J. Mocco. Size ratio correlates with intracranial aneurysm rupture status. *Stroke*, 41(5):916–920, 2010.

[112] Bastian E. Rapp. Chapter 29 - computational fluid dynamics. In Bastian E. Rapp, editor, *Microfluidics: Modelling, Mechanics and Mathematics*, Micro and Nano Technologies, pages 609 – 622. Elsevier, Oxford, 2017.

[113] E. C. Raps, J. D. Rogers, S. L. Galetta, R. A. Solomon, L. Lennihan, L. M. Klebanoff, and M. E. Fink. The Clinical Spectrum of Unruptured Intracranial Aneurysms. *JAMA Neurology*, 50(3):265–268, 03 1993.

[114] Payam Refaeilzadeh, Lei Tang, and Huan Liu. Cross-validation. In LING LIU and M. TAMER ÖZSU, editors, *Encyclopedia of Database Systems*, pages 532–538. Springer US, Boston, MA, 2009.

[115] Gabriel J. E. Rinkel, Mamuka Djibuti, Ale Algra, and J. van Gijn. Prevalence and risk of rupture of intracranial aneurysms. *Stroke*, 29(1):251–256, 1998.

[116] Jarl Rosenørn, Vagn Eskesen, Kaare Schmidt, Jens Ole Espersen, Jens Haase, Aage Harmsen, Ole Hein, Verner Knudsen, Steen Midholm, and Erik Marcussen. Clinical features and outcome in 1076 patients with ruptured intracranial saccular aneurysms: A prospective consecutive study. *British journal of neurosurgery*, 1:33–45, 02 1987.

[117] Shewaferaw S Shibeshi and William E Collins. The rheology of blood flow in a branched arterial system. *Applied rheology (Lappersdorf, Germany : Online)*, 15:398–405, 02 2005.

[118] Anna M.H. Sailer, Bart A.J.M. Wagemans, Patricia J. Nelemans, Rick de Graaf, and Willem H. van Zwam. Diagnosing intracranial aneurysms with mr angiography. *Stroke*, 45(1):119–126, 2014.

[119] Remi Sakia. The box-cox transformation technique: A review. *The Statistician*, 41, 01 1992.

[120] Claude Sammut and Geoffrey I. Webb. *Encyclopedia of Machine Learning and Data Mining*. Springer US, Boston, MA, 2017.

[121] Ali Sarrami-Foroushani, Maria-Cruz Villa-Uriol, Mohsen Nasr Esfahany, Stuart C. Coley, Luigi Yuri Di Marco, Alejandro F. Frangi, and Alberto Marzo. Modeling of the acute effects of primary hypertension and hypotension on the hemodynamics of intracranial aneurysms. *Annals of Biomedical Engineering*, 43:207–221, 2014.

[122] Luis E. Savastano, Ankur Bhambri, David Andrew Wilkinson, and Aditya S. Pandey. Chapter 2 - biology of cerebral aneurysm formation, growth, and rupture. In Andrew J. Ringer, editor, *Intracranial Aneurysms*, pages 17 – 32. Academic Press, 2018.

[123] Wouter I. Schievink. Intracranial aneurysms. *New England Journal of Medicine*, 336(1):28–40, 1997. PMID: 8970938.

[124] Shamiul Sarkar Shishir, Md. Abdul Karim Miah, A.K.M. Sadrul Islam, and A.B.M. Toufique Hasan. Blood flow dynamics in cerebral aneurysm - a cfd simulation. *Procedia Engineering*, 105:919 – 927, 2015. The 6th BSME International Conference on Thermal Engineering.

[125] Gangadhar Shobha and Shanta Rangaswamy. Chapter 8 - machine learning. In Venkat N. Gudivada and C.R. Rao, editors, *Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications*, volume 38 of *Handbook of Statistics*, pages 197 – 228. Elsevier, 2018.

[126] Galen R. Shorack. *Distribution and Quantile Functions*, pages 107–126. Springer International Publishing, Cham, 2017.

[127] Michael A. Silva, Jay Patel, Vasileios Kavouridis, Troy Gallerani, Andrew Beers, Ken Chang, Katharina V. Hoebel, James Brown, Alfred P. See, William B. Gormley, Mohammad Ali Aziz-Sultan, Jayashree Kalpathy-Cramer, Omar Arnaout, and Nirav J. Patel. Machine learning models can detect aneurysm rupture and identify clinical features associated with rupture. *World Neurosurgery*, 2019.

[128] Torbjørn Øygard Skodvin, Øyvind Evju, Angelika Sorteberg, and Jørgen Gjernes Isaksen. Prerupture Intracranial Aneurysm Morphology in Predicting Risk of Rupture: A Matched Case-Control Study. *Neurosurgery*, 84(1):132–140, 02 2018.

[129] A. Soriano, N. Pipitone, and C. Salvarani. Chapter 21 - behçet's disease. In Udi Nussinovitch, editor, *The Heart in Rheumatic, Autoimmune and Inflammatory Diseases*, pages 505 – 526. Academic Press, 2017.

[130] Birgitta Stegmayr, Marie Eriksson, and Kjell Asplund. Declining mortality from subarachnoid hemorrhage. *Stroke*, 35(9):2059–2063, 2004.

[131] Masashi Sugiyama. Chapter 1 - statistical machine learning. In Masashi Sugiyama, editor, *Introduction to Statistical Machine Learning*, pages 3 – 8. Morgan Kaufmann, Boston, 2016.

[132] S. Sumathi and S.N. Sivanandam. *Introduction to Data Mining Principles*, pages 1–20. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[133] Masaaki Suzuki, Toshiyuki Haruhara, Hiroyuki Takao, Takashi Suzuki, Soichiro Fujimura, Toshihiro Ishibashi, Makoto Yamamoto, Yuichi Murayama, and Hayato Ohwada. Classification model for cerebral aneurysm rupture prediction using medical and blood-flow-simulation data. In *ICAART*, pages 895–899, 01 2019.

[134] Yoshiteru Tada, Kosuke Wada, Kenji Shimada, Hiroshi Makino, Elena I. Liang, Shoko Murakami, Mari Kudo, Keiko T. Kitazato, Shinji Nagahiro, and Tomoki Hashimoto. Roles of hypertension in the rupture of intracranial aneurysms. *Stroke*, 45(2):579–586, 2014.

[135] Katsuhiro Tanaka, Fujimaro Ishida, Kimito Kawamura, Hideki Yamamoto, Daiki Horikawa, Tomoyuki Kishimoto, Masanori Tsuji, Hiroshi Tanemura, and Shinichi Shimosaka. Hemodynamic assessment of cerebral aneurysms using computational fluid dynamics (cfd) involving the establishment of non-newtonian fluid properties. *Journal of Neuroendovascular Therapy*, 03 2018.

[136] Chris Tong and Duvvuru Sriram. Chapter 1 - introduction. In CHRISTOPHER TONG and DUVVURU SRIRAM, editors, *Artificial Intelligence in Engineering Design*, pages 1 – 53. Academic Press, San Diego, 1992.

[137] Víctor Uc-Cetina. Ensamble de clasificadores usando adaboost [en línea], 2018. [Recuperado el 31 de julio de `https://medium.com/soldai/ensamble-de-clasificadores-usando-adaboost-50bc2ca47640`].

[138] Ádám Ugron and György Paál. On the boundary conditions of cerebral aneurysm simulations. *Periodica Polytechnica Mechanical Engineering*, 58:37–45, 01 2014.

[139] José Unpingco. Machine learning. In *Python for Probability, Statistics, and Machine Learning*, pages 197–273. Springer International Publishing, Cham, 2016.

[140] Linda D. Urden, Kathleen M. Stacy, and Mary E. Lough. *Priorities in Critical Care Nursing*. Elsevier Mosby, St. Louis, MO, 2012.

[141] Pablo Valdivieso. Estudio numérico cfd de condiciones de borde en modelos de aneurismas cerebrales. Memoria de título, Universidad de Chile, 2017.

[142] Alvaro Valencia, Darren Ledermann, Rodrigo Rivera, Eduardo Bravo, and Marcelo Galvez. Blood flow dynamics and fluid–structure interaction in patient-specific bifurcating cerebral aneurysms. *International Journal for Numerical Methods in Fluids*, 58(10):1081–1100, 2008.

[143] Alvaro Valencia and Francisco Torres. Effects of hypertension and pressure gradient in a human cerebral aneurysm using fluid structure interaction simulations. *Journal of Mechanics in Medicine and Biology*, 17:1750018, 02 2017.

[144] Jan van Gijn, Richard S Kerr, and Gabriel JE Rinkel. Subarachnoid haemorrhage. *The Lancet*, 369(9558):306 – 318, 2007.

[145] Irene Vignon-Clementel, Carlos Figueroa, Kenneth Jansen, and C.A. Taylor. Outflow boundary conditions for 3d simulations of non-periodic blood flow and pressure fields in deformable arteries. *Computer methods in biomechanics and biomedical engineering*, 13:625–40, 02 2010.

[146] J. Pablo Villablanca, Gary R. Duckwiler, Reza Jahan, Satoshi Tateshima, Neil A. Martin, John Frazee, Nestor R. Gonzalez, James Sayre, and Fernando V. Vinuela. Natural history of asymptomatic unruptured cerebral aneurysms evaluated at ct angiography: Growth and rupture incidence and correlation with epidemiologic risk factors. *Radiology*, 269(1):258–265, 2013. PMID: 23821755.

[147] Monique HM Vlak, Ale Algra, Raya Brandenburg, and Gabriël JE Rinkel. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. *The Lancet Neurology*, 10(7):626 – 636, 2011.

[148] Mike Walden. Arteries [en línea]. [Recuperado el 15 de julio de 2019 de `https://www.teachpe.com/anatomy/arteries.php`].

[149] Marieke J.H. Wermer, Irene C. van der Schaaf, Ale Algra, and Gabriël J.E. Rinkel. Risk of rupture of unruptured intracranial aneurysms in relation to patient and aneurysm characteristics. *Stroke*, 38(4):1404–1410, 2007.

[150] David O Wiebers. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment. *The Lancet*, 362(9378):103 – 110, 2003.

[151] Lindsy N. Williams and Robert D. Jr Brown. Management of unruptured intracranial aneurysms. *Neurology. Clinical practice*, 3(2):99–108, Apr 2013.

[152] George K.C. Wong and W.S. Poon. Current status of computational fluid dynamics for cerebral aneurysms: The clinician's perspective. *Journal of Clinical Neuroscience*, 18(10):1285 – 1288, 2011.

[153] Jianping Xiang, Sabareesh K Natarajan, Markus Tremmel, Ding Ma, J Mocco, L Nelson Hopkins, Adnan Siddiqui, Elad I Levy, and Hui Meng. Hemodynamic-morphologic discriminants for intracranial aneurysm rupture. *Stroke; a journal of cerebral circulation*, 42:144–52, 03 2011.

[154] Yazhou Yan, Deyuan Zhu, Haishuang Tang, and Qinghai Huang. Safety and efficacy of flow diverter treatment for aneurysm in small cerebral vessels: A systematic review and meta-analysis. *World Neurosurgery*, 115:54 – 64, 2018.

[155] Zhen Lu Yang, Qian Qian Ni, U. Joseph Schoepf, Carlo N. De Cecco, Han Lin, Taylor M. Duguay, Chang Sheng Zhou, Yan E. Zhao, Guang Ming Lu, and Long Jiang Zhang. Small intracranial aneurysms: Diagnostic accuracy of ct angiography. *Radiology*, 285(3):941–952, 2017. PMID: 28654338.

[156] Mair Zamir. Mathematical description of fluid flow. In *Hemo-Dynamics*, pages 13–41. Springer International Publishing, Cham, 2016.

[157] Yongtao Zheng, Feng Xu, Jinma Ren, Qiang Xu, Yingjun Liu, Yanlong Tian, and Bing Leng. Assessment of intracranial aneurysm rupture based on morphology parameters and anatomical locations. *Journal of NeuroInterventional Surgery*, 8(12):1240–1246, 2016.

[158] Geng Zhou, Yueqi Zhu, Yanling Yin, Ming Su, and Minghua Li. Association of wall shear stress with intracranial aneurysm rupture: Systematic review and meta-analysis. *Scientific Reports*, 7, 12 2017.

Apéndice A: Tablas de Contingencia

En este apartado se presentan las tablas de contingencia de los 6 atributos adicionales categóricos, donde se exhiben las frecuencias observadas y esperadas.

Tabla A.1: Tabla de contingencia con las frecuencias observadas y esperadas[6] para el sexo del paciente.

Sexo	No Roto	Roto
F	31 (29,7)	26 (27,3)
M	5 (6,3)	7 (5,7)

Fuente: Elaboración propia.

Tabla A.2: Tabla de contingencia con las frecuencias observadas y esperadas para el tipo de aneurisma.

Tipo de Aneurisma	No Roto	Roto
Lateral	28 (23,8)	19 (23,2)
Lateral Bifurcación	4 (5,6)	7 (5,4)
Terminal	4 (6,6)	9 (6,4)

Fuente: Elaboración propia.

[6]Las frecuencias esperadas se encuentran entre paréntesis.

Tabla A.3: Tabla de contingencia con las frecuencias observadas y esperadas para la circulación.

Circulación	No Roto	Roto
Anterior	9 (9,6)	10 (9,4)
Carótida	10 (6,6)	3 (6,4)
Media	6 (5,6)	5 (5,4)
Posterior	11 (14,2)	17 (13,8)

Fuente: Elaboración propia.

Tabla A.4: Tabla de contingencia con las frecuencias observadas y esperadas para la ubicación según el diagnóstico.

Ubicación (Diagnóstico)	No Roto	Roto
ACOM	0 (1)	2 (1)
AChA	2 (1,5)	1 (1,5)
BA	3 (1,5)	0 (1,5)
ICA	16 (11,2)	6 (10,8)
MCA	8 (7,1)	6 (6,9)
PA	0 (1)	2 (1)
PCA	0 (0,5)	1 (0,5)
PCOM	6 (10,1)	14 (9,9)
PICA	1 (2)	3 (2)

Fuente: Elaboración propia.

Tabla A.5: Tabla de contingencia con las frecuencias observadas y esperadas para la ubicación según Koivisto.

Ubicación (Koivisto)	No Roto	Roto
ACA	0 (2)	4 (2)
ICA	24 (22,8)	21 (22,2)
MCA	8 (7,1)	6 (6,9)
VBA	4 (4,1)	4 (3,9)

Fuente: Elaboración propia.

Tabla A.6: Tabla de contingencia con las frecuencias observadas y esperadas para la multiplicidad de aneurismas.

Multiplicidad	No Roto	Roto
No	17 (22,8)	28 (22,2)
Sí	19 (13,2)	7 (12,8)

Fuente: Elaboración propia.

Apéndice B: Grillas de Búsqueda

En este apartado se detallan las grillas de búsqueda evaluadas en la optimización de hiperparámetros, para cada uno de los mejores modelos por algoritmo hallados mediante búsqueda exhaustiva.

Tabla B.1: Grillas de búsqueda utilizadas en la optimización de hiperparámetros (1/2).

Algoritmo	Hiperparámetro	Grilla
LR	C	0,1; 0,5; 1
	Intercepto de ajuste	Verdadero, Falso
	Método de solución	Liblinear, SAG, SAGA
	Iteraciones máximas	50, 100, 150, 200
	Partida en caliente	Verdadero, Falso
	Tolerancia	0,00001; 0,0001; 0,001
LDA	Método de solución	SVD, LSQR, EIGEN
KNN	Cantidad de vecinos	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
	Peso	Uniforme, Distancia
DT	Criterio	Gini, Entropía
	División	Mejor, Aleatoria
	Profundidad máxima	Ninguna, 2, 3, 4, 5, 6
	Instancias mínimas para división	2, 3, 4
	Instancias mínimas para hoja	1, 2, 3
	Cantidad máxima de hojas	Ninguna, 2, 3, 4, 5
	Disminución mínima de impureza	0; 0,1; 0,2
NB	–	–
SVM	C	0,001; 0,01; 0,1; 1, 10
	Gamma	Auto; Escala; 0,001; 0,01; 0,1; 1
	Tolerancia	0,00001; 0,0001; 0,001

Fuente: Elaboración propia.

Tabla B.2: Grillas de búsqueda utilizadas en la optimización de hiperparámetros (2/2).

Algoritmo	Hiperparámetro	Grilla
RF	Cantidad de estimadores	50, 100, 150, 200
	Profundidad máxima	Ninguna, 2, 3, 4, 5, 6
	Instancias mínimas para división	2, 3, 4
	Instancias mínimas para hoja	1, 2, 3
	Cantidad máxima de hojas	Ninguna, 2, 3, 4, 5
	Disminución mínima de impureza	0; 0,1; 0,2
AB	Cantidad de estimadores	10, 30, 50, 100, 150, 200
	Tasa de aprendizaje	0,1; 0,3; 0,5; 0,8; 1
	Algoritmo	SAMME, SAMME.R
GB	Cantidad de estimadores	50, 100, 150, 200
	Tasa de aprendizaje	0,1; 0,5; 1
	Submuestra	0,5; 0,8; 1
	Profundidad máxima	Ninguna, 2, 3, 4, 5, 6
	Instancias mínimas para división	2, 3, 4
	Instancias mínimas para hoja	1, 2, 3
	Cantidad máxima de hojas	Ninguna, 2, 3, 4, 5
	Disminución mínima de impureza	0; 0,1; 0,2

Fuente: Elaboración propia.