

**UNIVERSIDAD RICARDO PALMA**

**ESCUELA DE POSGRADO**

**MAESTRÍA EN CIENCIA DE LOS DATOS**



**TESIS**

Para optar el Grado Académico de Maestro en Ciencia de los Datos

Algoritmos de aprendizaje automático para predecir la curación de malformaciones arteriovenosas cerebrales tratadas con radiocirugía

Autor: Bach. Salas Ojeda, María Isabel

Asesor: Mg. León Aguilar, Alfredo

LIMA- PERÚ

2021

Página de Jurado

Presidente: Dr Oscar Efrain Ramos Ponce

Miembro: Mg. Mirko Jerber Rodríguez Mallma

Miembro: Ludmer Edward Arcaya Arhuata

Asesor: Mg Alfredo León Aguilar

Representante de la EPG: Dra Ofelia Roque Paredes

## Dedicatoria

A la memoria del Dr Erwin Kraenau Espinal, por impulsar esta maestría que tanta satisfacción me brinda.

Dedico esta tesis a mi esposo, a mis hijos cuyo amor y comprensión son mi fortaleza.

A mis padres por su cariño y aliento constante en mi progreso profesional y personal.

## Agradecimiento

A mi asesor Mg. Alfredo León Aguilar, por su paciencia, sugerencias y ayuda en la elaboración de mi tesis

A la Dra Ofelia Roque Paredes, por impulsarme y alentarme en el desarrollo de este trabajo.

Al Mg. Mirko Rodríguez Mallma, por su gran ayuda y aporte académico en el desarrollo de esta tesis.

## Índice de contenidos

INTRODUCCIÓN .....	1
CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA .....	3
1.1 Descripción del Problema .....	3
1.2. Formulación del problema .....	3
1.2.1 Problema general.....	4
1.2.2 Problemas específicos .....	4
1.3. Importancia y justificación del estudio .....	4
1.4 Delimitación del estudio .....	5
1.5 Objetivos de la investigación .....	6
1.5.1 Objetivo General .....	6
Objetivos Específicos.....	6
CAPÍTULO II: MARCO TEÓRICO .....	7
2.1 Marco histórico .....	7
2.2 Investigaciones relacionadas con el tema .....	9
2.2.1. Antecedentes nacionales: .....	9
2.2.2 Antecedentes internacionales: .....	9
2.3 Estructura teórica que sustenta el estudio .....	16
2.3.1 Las malformaciones arteriovenosas cerebrales.....	16
2.3.1.1 Epidemiología, características clínicas y diagnóstico.....	16
2.3.1.2 Tratamiento .....	18
2.3.1.2.1 Técnica microquirúrgica .....	18
2.3.1.2.2 La terapia endovascular.....	18
2.3.1.2.3 La Radiocirugía estereotáxica .....	19
2.3.2 Sistemas de Calificación y clasificación de las MAVs.....	22
2.3.3 Aprendizaje automático y análisis predictivo .....	27
2.3.3.2 Preparación de los datos .....	29
2.3.3.3 Entrenamiento del modelo .....	31
2.3.3.4 Evaluación del modelo.....	37
2.3.3.5 Aprendizaje supervisado: Algoritmos de clasificación.....	39
2.3.4 Modelos predictivos en datos de la salud.....	54
2.3.5 Ética en la aplicación de aprendizaje automático en datos de salud .....	56
2.4 Definición de términos básicos .....	57

2.5 Fundamentos teóricos que sustentan las hipótesis .....	59
2.6 Hipótesis.....	60
2.6.1 Hipótesis general.....	60
2.6.2 Hipótesis específicas .....	60
2.7 Variables (definición y operacionalización de variables): Dimensiones e indicadores .....	61
<b>CAPÍTULO III: MARCO METODOLÓGICO .....</b>	<b>68</b>
3.1 Tipo de investigación .....	68
3.2 Población y muestra .....	68
3.3 Técnicas e instrumentos de recolección de datos.....	69
3.4 Descripción del procedimiento de análisis.....	70
<b>CAPÍTULO IV: RESULTADOS Y ANÁLISIS DE RESULTADOS .....</b>	<b>75</b>
4.1 Resultados .....	75
4.2 Análisis de resultados o discusión de resultados .....	94
<b>CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>97</b>
<b>REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>99</b>
<b>ANEXOS .....</b>	<b>106</b>
Anexo 1 Declaración de autenticidad .....	106
Anexo 2 Autorización de consentimiento para realizar la investigación .....	106
Anexo 3: Matriz de consistencia .....	107

## Listado de tablas y figuras

### Listado de Tablas:

Tabla 1. Escala de Spetzler-Martin y de Lawton-Young para malformaciones arteriovenosas cerebrales	23
Tabla 2. Escala de malformaciones arteriovenosas cerebrales de Radiocirugía de Virginia	25
Tabla 3. Escala de Buffalo para tratamiento endovascular de MAV cerebrales	26
Tabla 4. Escala de Estado Funcional de Karnofsky	27
Tabla 5. Matriz de confusión para la clasificación con dos categorías de respuesta	37
Tabla 6: Características sociodemográficas de los pacientes con malformación arteriovenosa cerebrales tratados con radiocirugía 2005 al 2018	76
Tabla 7. Estadísticas descriptivas de las variables cuantitativas de los pacientes con malformación arteriovenosa cerebrales tratados con radiocirugía 2005 al 2018.	77
Tabla 8. Variables clínicas y de angioarquitectura de las MAV	78
Tabla 9. Localización de la malformación arteriovenosa cerebrales en los pacientes tratados con radiocirugía 2005 al 2018	79
Tabla 10. Correlación entre las variables independientes cuantitativas	81
Tabla 11. Predictores que ayudarán a predecir la curación de MAV cerebrales.	82
Tabla 12. Comparación de los algoritmos en el primer experimento	83
Tabla 13 Comparación de los algoritmos en el segundo experimento	86
Tabla 14. Comparación de los algoritmos en el tercer experimento	88
Tabla 15. Comparación de los algoritmos en el cuarto experimento	91

### Listado de Figuras:

Figura 1. Ejemplos de grados la escala de Spetzler–Martin. Grados 1 hasta 4 de malformaciones arteriovenosas cerebrales (MAV) y opciones de tratamiento.	24
Figura 2. División del conjunto de datos en entrenamiento, validación y prueba para aplicar a aprendizaje automático	33
Figura 3. Proceso de validación cruzada con k-iteraciones $k=10$ .	35
Figura 4. Proceso de validación cruzada anidada, con $k = 10$ .	36
Figura 5. Idea básica de máquina de soporte vectorial para el caso lineal.	45
Figura 6. Clasificador de margen suave	47
Figura 7. Representación del truco de kernel	47
Figura 8. Neurona biológica	49
Figura 9. La sinapsis de una neurona	50
Figura 10. Implementación de una red neuronal	51
Figura 11. Ejemplo de una red feedward con múltiples capas	53

<i>Figura 12.</i> Ejemplo de una red recurrente.....	53
<i>Figura 13.</i> Ruta de trabajo para la aplicación de algoritmos de aprendizaje automático en análisis predictivo. ....	60
<i>Figura 14.</i> Proceso de recolección de los datos.....	70
<i>Figura 15.</i> Pasos principales para aprendizaje automático supervisado.....	70
<i>Figura 16.</i> Selección de variables.....	72
<i>Figura 17.</i> Descripción del procedimiento de análisis.....	74
<i>Figura 18.</i> Distribución de la variable Curación de la MAV.....	75
<i>Figura 19.</i> Localización de la Malformación arteriovenosa cerebral en los pacientes tratados con radiocirugía 2005 al 2018.....	79
<i>Figura 20.</i> Visualización de las variables predictoras y del target: Curación.....	80
<i>Figura 21.</i> Mapa de calor de las correlaciones de las variables cualitativas.....	82
<i>Figura 22.</i> Comparación de los algoritmos en el primer experimento.....	84
<i>Figura 23.</i> Resultados de la validación cruzada en el primer experimento.....	84
<i>Figura 24.</i> Área bajo la curva de los algoritmos: regresión logística, Naive Bayes, Random forest, Máquina de soporte vectorial, y redes neuronales, en el primer experimento.....	85
<i>Figura 25.</i> Comparación de los algoritmos en el segundo experimento.....	86
<i>Figura 26.</i> Resultados de la validación cruzada del segundo experimento.....	87
<i>Figura 27.</i> Área bajo la curva de los algoritmos: regresión logística, Naive Bayes, Random forest, Máquina de soporte vectorial, y redes neuronales, en el segundo experimento.....	88
<i>Figura 28.</i> Comparación de los algoritmos en el tercer experimento.....	89
<i>Figura 29.</i> Resultados de la validación cruzada del tercer experimento.....	89
<i>Figura 30.</i> Área bajo la curva de los algoritmos: regresión logística, Naive Bayes, Random forest, Máquina de soporte vectorial, y redes neuronales, en el tercer experimento.....	90
<i>Figura 31.</i> Comparación de los algoritmos en el cuarto experimento.....	91
<i>Figura 32.</i> Resultados de la validación cruzada del cuarto experimento.....	92
<i>Figura 33.</i> Área bajo la curva de los algoritmos: regresión logística, Naive Bayes, Random forest, Máquina de soporte vectorial, y redes neuronales, en el cuarto experimento.....	93



## RESUMEN

El objetivo de esta tesis es identificar los factores que influyen en la curación de las malformaciones arteriovenosas (MAV) cerebrales en pacientes que fueron tratados mediante radiocirugía estereotáctica y comparar diversos algoritmos de aprendizaje automático para predecir la curación de esta enfermedad.

Las MAVs cerebrales son enfermedades de baja prevalencia, pero de gran impacto por la elevada morbimortalidad que producen como consecuencia de su ruptura por lo que evaluamos 202 pacientes con este diagnóstico intervenidos mediante radiocirugía estereotáctica en el Instituto de Radiocirugía de la Clínica San Pablo entre los años 2005 al 2018, se anotó 45 variables entre socio demográficas y clínicas de los pacientes, así como las características morfológicas y hemodinámicas a través de imágenes de las MAVs cerebrales. Los pacientes que se curaron de las MAVs fueron 167 (83%), durante el periodo de seguimiento de hasta 3 años. El modo de presentación clínica más frecuente fue cefalea (88.12 %) seguido de hemorragia cerebral (76.73%), y convulsiones (55 %). A fin de determinar el mejor modelo predictivo se seleccionaron 23 variables predictores y luego se evaluaron los algoritmos aprendizaje automático de regresión logística, Naive Bayes, Máquina de Soporte Vectorial, Random Forest y Redes Neuronales. Se utilizó, 80% de los datos para el entrenamiento de los algoritmos de aprendizaje automático y 20% para el test de la performance predictiva de los modelos.

El algoritmo que obtuvo las mejores métricas fue el algoritmo de máquina de soporte vectorial donde el accuracy, la sensibilidad y especificidad fue de 98.51%, 100% y 96.97%, respectivamente. Asimismo, el algoritmo random forest también mostro buenos resultados cuyas métricas de accuracy, sensibilidad y especificidad fue de 96.57%, 99.12% y 93.94%, respectivamente. La predicción de la curación de la MAV cerebral a través de los algoritmos de aprendizaje automático puede ayudar a los especialistas a tomar las mejores decisiones respecto al tratamiento de pacientes con este diagnóstico. Por consiguiente, estos algoritmos podrían ser implementados en las guías de tratamiento de MAVs cerebrales para predecir su curación mediante la automatización de los procesos a partir de los datos clínicos y de la técnica de aplicación de la radiocirugía para lograr un flujo de atención eficaz y personalizado en la curación de esta enfermedad.

Palabras claves: aprendizaje automático, malformaciones arteriovenosas cerebrales, algoritmos, predecir, radiocirugía, supervisado.

## ABSTRACT

The objective of this thesis is to identify the factors that influence the cure of cerebral arteriovenous malformations (AVMs) in patients who were treated by stereotactic radiosurgery and to compare various machine learning algorithms to predict the cure of this disease.

Brain AVMs are diseases of low prevalence, but of great impact due to the high morbidity and mortality that they produce as a consequence of their rupture, which is why we evaluated 202 patients with this diagnosis who underwent stereotactic radiosurgery at the Institute of Radiosurgery of the San Pablo Clinic between the years From 2005 to 2018, 45 variables were recorded between the sociodemographic and clinical of the patients, as well as the morphological and hemodynamic characteristics through images of the brain AVMs. The patients who were cured of AVMs were 167 (83%), during the follow-up period of up to 3 years. The most frequent mode of clinical presentation was headache (88.12%) followed by cerebral hemorrhage (76.73%), and seizures (55%). In order to determine the best predictive model, 23 predictor variables were selected and then the logistic regression machine learning algorithms, Naive Bayes, Vector Support Machine, Random Forest and Neural Networks were evaluated. 80% of the data was used for the training of the machine learning algorithms and 20% for the test of the predictive performance of the models.

The algorithm that obtained the best metrics was the vector support machine algorithm where the accuracy, sensitivity and specificity were 98.51%, 100% and 96.97%, respectively. Likewise, the random forest algorithm also showed good results, whose accuracy, sensitivity and specificity metrics were 96.57%, 99.12% and 93.94%, respectively. Predicting the cure for cerebral AVM using machine learning algorithms can help specialists make the best decisions regarding the treatment of patients with this diagnosis. Consequently, these algorithms could be implemented in the treatment guidelines for cerebral AVMs to predict their cure by automating the processes based on clinical data and the technique of applying radiosurgery to achieve an efficient and personalized flow of care. in curing this disease.

Keywords: machine learning, cerebral arteriovenous malformations, algorithms, predicting, radiosurgery, supervised.

## INTRODUCCIÓN

Las malformaciones arteriovenosas (MAV) cerebrales son anomalías congénitas de los vasos sanguíneos que permiten conexiones directas entre las arterias y las venas en una parte del cerebro y es una de las principales causas de hemorragia cerebral y muerte en las personas que lo sufren. También son causa de secuelas neurológicas importantes en los supervivientes. (Spetzler, Kondziolka, Higashida, & Kalani, 2015).

El tratamiento ha representado un desafío por los riesgos que asume el paciente sobre todo cuando la MAV es grande o está ubicada en áreas con elevada probabilidad de secuelas neurológicas o riesgo de vida en lo que los especialistas denominan áreas elocuentes y profundas. Desde que las intervenciones quirúrgicas han sido el método tradicional para tratar de curar las MAVs, estas tienen una mortalidad entre 5.7% y 30.18% y morbilidad es decir riesgo de secuelas neurológicas entre 4.8 y 32% (Drake, 1979). Además, por la dificultad quirúrgica la extirpación de la MAV es incompleta en el 16% persistiendo el riesgo de morbimortalidad por sangrado (Rosell, Rocca, Chavez, Dávila, & Palacios, 2004). A fin de reducir estos riesgos se han desarrollado nuevos métodos complementarios y a veces de primera elección como la terapia endovascular y la radiocirugía estereotáctica (Bradac, Charvat, & Benes, 2013).

La radiocirugía estereotáctica es una técnica moderna, de mínima invasión y consiste en la generación de haces de fotones de alta energía, desde varios ángulos, sobre un blanco intracraneal, en este caso la MAV, con la finalidad de excluirla (curarla) mientras que las áreas cerebrales colindantes son respetadas.

La curación de la MAV con radiocirugía es un proceso de obliteración paulatina de los vasos enfermos que puede demorar de 2 a 3 años, persistiendo en ese tiempo el riesgo de sangrado (Moorthy & Rajshekhar, 2015). El efecto final de la radiocirugía es producir radionecrosis en los blanco-objetivos y con ello la desaparición de la MAV.

No obstante, hay experiencias que refieren que la radiocirugía es una técnica eficaz y válida para el tratamiento de las MAV, está aún en estudio para sus indicaciones y recomendaciones. Se ha reportado una morbimortalidad permanente de 4.2 % (rango de 0-35%) con una curación de 79% (60-100%) de las MAV tratadas con radiocirugía (Van Beijnum & Van der Worp, 2011)

El tratamiento adecuado de las MAV puede ser complejo y depende de múltiples variables clínicas y de características de la lesión haciendo de la categorización y pronóstico de curación un proceso necesario.

Los modelos estadísticos convencionales resultan insuficientes para este propósito, los algoritmos de aprendizaje automático han sido propuestos como una aproximación más adecuada, teniendo la ventaja de incorporar nueva información para mejorar el desempeño de una predicción de resultados o ayudar al pronóstico de los pacientes con MAV cerebral.

La presente tesis está conformada por cuatro capítulos:

Capítulo I, donde se describe y formula el problema de investigación, además se delimita y justifica la investigación y se indica los objetivos de la investigación.

Capítulo II, contiene el marco histórico, investigaciones relacionadas con el tema de investigación, la estructura teórica que sustenta el estudio, las hipótesis y las variables.

Capítulo III, está formado por el marco metodológico, donde se describe el tipo de investigación, la población y muestra, las técnicas e instrumentos de recolección de datos y la descripción y procedimiento de análisis.

Capítulo IV, comprende los resultados y análisis de los resultados, así como las conclusiones, las referencias y anexos.

# CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA

## 1.1 Descripción del Problema

La hemorragia cerebral es una causa importante de enfermedad y muerte por los daños que produce en el sistema nervioso central. Una de las principales causas es la MAV cerebral que es un conglomerado de vasos anormalmente formado por ausencia de capilares y que se ubica en distintas partes del cerebro (Aminoff, Boller, & Swaab, 2017). La MAV cerebral se descubre como hemorragia en un 30 a 60% de pacientes. Otras formas de presentarse son mediante convulsiones en un 28 a 46% y dolor de cabeza en un 10 a 34% (Hernesniemi & Reza, 2008). El riesgo de muerte por ruptura de una MAV es de 10-29% y las secuelas neurológicas a largo plazo están entre 20 a 60% (Badakhshi, 2016). El modo de tratarlos es un desafío para los especialistas por el alto riesgo quirúrgico. La radiocirugía estereotáctica es una técnica novedosa en nuestro país y se realiza en el Instituto de radiocirugía del Perú desde el año 2005, consiste en aplicar radiación de alta energía sobre la malformación arteriovenosa respetando el tejido cerebral colindante. No obstante, también tiene riesgos ya que la curación no es inmediata sino es progresiva con seguimientos hasta dos a tres años. Son varios los factores que contribuyen al éxito del tratamiento y son dependientes de la propia lesión, del paciente, de la técnica de radiocirugía y la cantidad de energía que se entrega en el tratamiento de la radiocirugía (Solomon & Connolly, 2017). El aprendizaje automático es un valioso elemento que nos permitirá seleccionar el mejor algoritmo para un pronóstico efectivo que ayude a los especialistas a tomar las mejores decisiones respecto al tratamiento de estos pacientes. Estos algoritmos se están explorando como herramientas para la predicción del resultado neuroquirúrgico (Senders, y otros, 2017).

## 1.2. Formulación del problema

A pesar que la radiocirugía es una técnica eficaz y válida para el tratamiento de las MAV, está en evolución y tiene también sus riesgos. Se ha reportado una morbimortalidad permanente de 4.2 % (rango de 0-35%) con una curación de 79% (60-100%) de las MAV tratadas con radiocirugía (Van Beijnum & Van der Worp, 2011).

Hace falta el diseño de un sistema de predicción de resultados que pueda modelar el efecto de las variables como factores de pronóstico sobre los resultados de la radiocirugía en los pacientes con malformaciones arteriovenosas para conocer estas

variables influyentes se podría aplicar y comparar técnicas de aprendizaje automático para encontrar el modelo más eficiente.

### 1.2.1 Problema general

En base a lo expuesto anteriormente, se formula la pregunta de esta investigación: ¿Es posible predecir la curación de las malformaciones arteriovenosas cerebrales tratadas con radiocirugía utilizando algoritmos de clasificación de aprendizaje automático?

### 1.2.2 Problemas específicos

Asimismo, se tienen los siguientes problemas específicos:

¿Cuáles son las variables sociodemográficas, clínicas, de angioarquitectura y del procedimiento de la radiocirugía que ayudarán a predecir la curación de los pacientes con diagnóstico de malformación arteriovenosa cerebral tratados con radiocirugía?

¿Es posible que los algoritmos de Naive bayes, regresión logística, máquina de soporte vectorial, random forest y redes neuronales con oversampling tengan mejores resultados que los algoritmos de Naive bayes, regresión logística, máquina de soporte vectorial, random forest y redes neuronales sin oversampling en el pronóstico de la curación de los pacientes con malformación arteriovenosa cerebral tratados con radiocirugía?

## 1.3. Importancia y justificación del estudio

La malformación arteriovenosa (MAV) cerebral es un desafío para la medicina por el elevado riesgo de morbilidad y mortalidad. Las técnicas de tratamiento no están exentas de riesgos y complicaciones. La radiocirugía no obstante ser la menos invasiva, requiere un periodo de tiempo para su efectividad, en el cual la MAV disminuye de tamaño progresivamente hasta su total desaparición (curación). En este periodo mientras la MAV persista aún puede romperse y causar una hemorragia cerebral. Las técnicas de aprendizaje automático podrían identificar las variables más influyentes que favorecen su curación. El desarrollo de esta investigación y sus modelos podrían ayudar en predecir los factores de las MAV que favorecen su curación y disminuir la morbilidad así como mejorar el proceso de la técnica de radiocirugía estereotáctica.

## Teórico

El aprendizaje automático: permite identificar las variables clínicas, físicas, estructurales y sociales de las MAV cerebrales en pacientes que sufren esta enfermedad y son tratados con radiocirugía, para luego extraer información y convertirla en conocimiento mediante la integración de la estadística, la inteligencia artificial y predecir la curación de las MAV. La radiocirugía en el tratamiento de las MAV cerebrales es una técnica valiosa que, no obstante, su modernidad y eficacia en el tratamiento de las MAV cerebrales es necesario identificar variables que predisponen su curación o complicación.

## Práctica

El establecimiento y adopción de un modelo de aprendizaje automático podría predecir la curación de las MAV mediante la utilización de la técnica de radiocirugía estereotáctica y permitiría confirmar su efectividad, tomar medidas correctivas, preventivas y alentar cambios en la planificación de tratamientos futuros.

## Metodológica

Mediante las etapas definidas de la implementación de machine learning se sigue un análisis previo sobre el comportamiento de las MAV cerebrales considerando todas las variables utilizando aprendizaje automático.

Considerando el paso a paso de aprendizaje automático que utiliza técnicas matemáticas, estadísticas validadas fundamentadas estructuradas y ordenadas.

## Social

Reducir la morbimortalidad permite devolver a la sociedad pacientes curados evitando costos sociales de las secuelas invalidantes o el sufrimiento de los casos fatales. Un caso resuelto representa un enorme ahorro social y económico para las familias y el Estado.

### 1.4 Delimitación del estudio

Desde que la malformación arteriovenosa (MAV) cerebral es una enfermedad vascular cerebral con riesgo de morbimortalidad por su ruptura y que requiere curación. La presente investigación permite analizar las variables más influyentes de las MAV



cerebrales para predecir su curación mediante técnicas de aprendizaje automático en pacientes referidos para su tratamiento mediante la técnica quirúrgica de radiocirugía estereotáxica en el Instituto de radiocirugía del Perú de la ciudad de Lima entre los años 2005 al 2018. Los algoritmos de aprendizaje automático supervisado de clasificación usadas en este trabajo son: Naive Bayes regresión logística, máquina de soporte vectorial, random forest y redes neuronales.

## 1.5 Objetivos de la investigación

### 1.5.1 Objetivo General

Predecir la curación de las malformaciones arteriovenosas cerebrales en pacientes tratados con radiocirugía utilizando algoritmos de clasificación de aprendizaje automático.

### Objetivos Específicos

- Identificar las variables sociodemográficas, clínicas, de angioarquitectura y del procedimiento de la radiocirugía de las malformaciones arteriovenosas cerebrales. que ayudarán a predecir la curación de los pacientes con malformación arteriovenosa cerebral.
- Determinar la predicción de los algoritmos de Naive Bayes, regresión logística, random forest, máquina de soporte vectorial y redes neuronales usando todas las variables, las variables más importantes, con oversampling y sin oversampling y elegir el mejor resultado en la predicción de la curación de los pacientes con malformación arteriovenosa cerebral tratados con radiocirugía.

## CAPÍTULO II: MARCO TEÓRICO

### 2.1 Marco histórico

#### Algoritmos de aprendizaje automático

En las últimas décadas con la mejora exponencial mejora del rendimiento del software para el cálculo basado en computadora, los algoritmos de aprendizaje automático se han desarrollado en varios campos científico como patrones de reconocimiento, visión de computadora, ingeniería espacial, finanzas, entretenimiento, ecología, biología computacional y aplicaciones biomédicas. Históricamente, el inicio del aprendizaje automático se remonta al siglo XVII y con el desarrollo por Pascal y Leibniz de máquinas que pueden emular la capacidad humana de agregar y sustraer. Arthur Samuel de IBM acuñó el término "aprendizaje automático" y demostró que las computadoras podían programarse para aprender a jugar a las damas. La inteligencia artificial comenzó en la década de 1950. En esos años, Alan Turing ejecutó la prueba de Turing para probar la capacidad de una máquina de imitar a un humano. En 1956, el término "inteligencia artificial" se utilizó por primera vez por Marvin Minsky del Instituto de Tecnología de Massachusetts, John McCarthy de Stanford University y Allen Newell y Herbert Simon de la Universidad Carnegie-Mellon. Esto fue seguido por el desarrollo del perceptrón por Rosenblatt, una de las primeras arquitecturas de redes neuronales en 1958. Sin embargo, el entusiasmo inicial sobre el perceptrón se vio atenuado por la observación hecha por Minsky de que la capacidad de clasificación del perceptrón se limita a problemas linealmente separables y problemas no lineales no comunes como una lógica XOR simple. En 1975 se logró un gran avance mediante el desarrollo del perceptrón multicapa (MLP) por Werbos. Esto fue seguido por el desarrollo de árboles de decisión por Quinlan en 1986 y el soporte de máquinas de vectores por Cortes y Vapnik. Posteriormente se propusieron algoritmos de aprendizaje automático de ensamble, que combinan múltiples aprendizajes, incluidos Adaboost y bosques aleatorios. Más recientemente, los algoritmos de aprendizaje multicapa distribuidos se han fusionado bajo la noción de aprendizaje profundo. Estos algoritmos pueden aprender buenas representaciones de los datos que facilitan la extracción de información útil al construir clasificadores u otros predictores. Hoy en día, la inteligencia artificial y el aprendizaje automático se utilizan en muchas áreas de investigación y sectores de trabajo. (El Naqa & Murphy, 2015; Çelik, Ö. ,2018).

Un campo de trabajo de estos algoritmos es la salud, así un modelo de aprendizaje automático puede aprender los patrones de comportamiento de las enfermedades y ayudar al profesional de la salud a anticipar futuros eventos en un nivel de experto disponiendo de información más allá de su experiencia individual práctica. (Rajkomar, Dean & Kohane, 2019). Los modelos de clasificación de aprendizaje automático son empleados para análisis genómicos, celulares, moleculares y ayudar en la predicción temprana del riesgo, diagnóstico y tratamientos. Los modelos de clasificación pueden servir de base en la implementación de la salud predictiva y la medicina de precisión para lograr un diagnóstico temprano y una intervención oportuna y personalizada. (Lee, 2018).

De otro lado Yoo y otros (2012), usando la tecnología de minería de datos pueden predecir fraude en los seguros de salud, costos en la salud, pronóstico, diagnóstico y estancia hospitalaria.

#### Malformaciones arteriovenosas cerebrales

Las MAVs cerebrales representan una enfermedad compleja, con un perfil de presentación consistente en hemorragia, convulsiones o dolor de cabeza. La incidencia y la prevalencia anuales de las MAVs es de aproximadamente de 1 caso en 100 000 y 50 casos en 100 000 respectivamente (Aminoff M. y otros, 2017, Lawtony otros, 2015)

Los primeros casos publicados sobre el diagnóstico y tratamiento de las malformaciones arteriovenosas (MAV) cerebrales fueron descritos por Cushing y Dandy a finales de los años 20. La altísima mortalidad asociada a este tipo de patología hizo que en esa época las consideraban intratables. Con el advenimiento de la angiografía, en los años 30, se conoció más acerca de las MAVs. permitiendo la visualización preoperatoria de estas. (Yáñez, A., 2001).

En la década del 40 comienzan a aparecer publicaciones de intervenciones quirúrgicas con resultados aceptables por Olivecrona y Norlén. Posteriormente, Yasagil a comienzos de la década del 70 introduce el microscopio quirúrgico en la manipulación de las diversas estructuras del cerebro. (Yáñez, A.,2001). Actualmente, la técnica microquirúrgica es el estándar en el tratamiento de las MAV, no obstante, el alto riesgo que demanda sobre todo lesiones ubicadas profundamente y en áreas elocuentes. Por estas consideraciones se han desarrollado técnicas alternativas y complementarias como

la terapia endovascular y la radiocirugía estereotáctica. (Yáñez, A. ,2001; Solomon & Connolly 2017).

Asimismo, la radiocirugía es una modalidad de tratamiento que consiste en la aplicación de fotones de alta energía en un blanco objetivo intracraneal, en este caso la MAV con el objetivo de causar radio necrosis y con ello la curación de la MAV, respetando el área colindante a la MAV (Solomon & Connolly, 2017). La radiocirugía es un tratamiento efectivo, particularmente en los pacientes que tienen un alto riesgo para realizarse la cirugía (Friedman, 2011; Pollock & col, 2016). A pesar que la radiocirugía es una técnica eficaz y válida para el tratamiento de las MAV, está en evolución y tiene también sus riesgos. Se ha reportado una morbimortalidad permanente de 4.2 % (rango de 0-35%) con una curación de 79% (60-100%) de las MAV tratadas con radiocirugía (Van Beijnum & Van der Worp, 2011; Lawton, 2015; Aminoff, 2017; Solomon, 2017).

## 2.2 Investigaciones relacionadas con el tema

### 2.2.1. Antecedentes nacionales:

Sánchez (2017), en su tesis “características epidemiológicas y clínicas de los pacientes con malformaciones arteriovenosas cerebrales que recibieron radiocirugía estereotáctica, en el centro integral de radioncoterapia del hospital militar central durante el periodo 2013 – 2016, se revisaron a 20 pacientes y se encontró que el 80% de estos pacientes tenían edades de 10 a 49 años, asimismo 60% eran varones. Además, 90% de estos pacientes presentaron cefalea, 45% mareos. La localización más frecuente de la malformación arteriovenosa cerebral fue en la parte temporal izquierda (50%) seguido por el lado temporal derecho (30%) y 20% ubicados en otras zonas. Se menciona que 60% de estos pacientes recibieron tratamiento adicional a la radiación. Se considera que la radiocirugía estereotáctica es recomendable ya que no deja secuelas y llega a áreas donde es difícil acceder con la cirugía convencional.

### 2.2.2 Antecedentes internacionales:

Larios y otros (2010), mostraron que los resultados postoperatorios de MAV cerebral en una realidad latinoamericana donde las MAVs cerebrales fueron más frecuentes en hombres (57%) que mujeres (43%) más en menores de 35 años de edad (65%). La forma de presentación más frecuente fue con hemorragia intracraneal (41%),

seguido de epilepsia y focalización neurológica progresiva. El 55% de los pacientes fue sometido a un tratamiento invasivo (embolización, cirugía o los dos con radiocirugía). La evolución clínica con estos tratamientos mostró que 248 casos (82%) tuvo una buena evolución, mientras que 37 pacientes (12%) tuvieron evolución con discapacidad severa y 17 casos (6%) fallecieron.

Además, Kandai, Abdullah, & Naing (2010), reportaron pacientes que evolucionaron mal con o sin tratamiento, para lo cual estudiaron factores de las MAV que hacen que algunos pacientes sangran y otros no. De esta forma, hallaron que los sangrados de las MAV eran más frecuentes en pacientes varones jóvenes (65.5%), tamaño pequeño de las MAV ( $p=0.004$ ), localización cerebral profunda ( $p=0.003$ ) y drenaje venoso profundo ( $p=0.006$ ), además todos los pacientes con sangrado tuvieron un aneurisma prenidial incidental, es decir antes de la MAV propiamente dicha.

Asimismo, Parkhutik (2014), en su tesis doctoral: “Radiocirugía estereotáxica en el tratamiento de malformaciones arteriovenosas cerebrales”, realizó un estudio observacional en una cohorte de pacientes con malformaciones arteriovenosas cerebrales en el hospital de Valencia, antes y después de ser tratados con radiocirugía. Se obtuvo datos de 108 pacientes, evaluados en una media de 25 meses antes y 65 meses después del tratamiento. La edad media al momento del diagnóstico fue de 36 años, 55% eran hombres. El 39.8% de las malformaciones fueron hemorrágicas, del resto la mayoría fue diagnosticada tras crisis epilépticas (32.4%) o cefaleas (15.7%). La mayoría (78%) tenía un nido de localización córtico-subcortical, con un volumen mediano de 5.8 cc. Se encontró que estos pacientes tratados con radiocirugía tienen menos hemorragias cerebrales, además la tasa de sangrado de las malformaciones hemorrágicas baja de 3.3% a 1.1% y de las malformaciones no hemorrágicas a 0.3%. Esto demuestra la protección de la radiocirugía contra la hemorragia cerebral. Se reporta que luego de los 5 años se observan importantes lesiones radiológicas en el 20% de estos pacientes. El riesgo global de desarrollar nuevas complicaciones clínicas se cifra en 13.7%. La aparición tardía de estas complicaciones sugiere que los protocolos de seguimiento de estos pacientes han de alargarse en el tiempo. Los factores contribuyentes al desarrollo de complicaciones son el diámetro de nido mayor de 3cm y tratamiento con dos radiocirugías. (Parkhutik Matveera, 2014)

Por otro lado, Asadi, y otros, (2016), realizaron un estudio para identificar los factores que influyen en el resultado en las malformaciones arteriovenosas cerebrales tratadas con embolización endovascular. Evaluaron la viabilidad del uso de técnicas de aprendizaje automático para pronosticar y predecir la mortalidad comparándolas con el análisis estadístico de regresión logística convencional. El estudio fue retrospectivo, se trabajó con 199 pacientes sometidos a tratamiento endovascular de malformaciones durante un período de 22 años en un centro nacional de neurociencia. Se registraron la presentación clínica, las imágenes, los detalles del procedimiento, las complicaciones y los resultados. Los datos se analizaron utilizando técnicas de inteligencia artificial para identificar predictores de resultados y se evaluó mediante el accuracy en la predicción de resultados clínicos en el seguimiento final. Entre las características clínicas de presentación se tiene que 56% tuvieron hemorragia cerebral antes del tratamiento. Luego de la embolización 51 (25.63%) volvieron a sangrar con una mortalidad a 90 días de 10 pacientes (3.7%). El modelo de análisis de regresión logística estándar tenía una precisión de 43% en la predicción del resultado final (mortalidad), con el tipo de complicación del tratamiento identificado como el predictor más importante. En cuanto a los algoritmos de aprendizaje automático empleados, se usó las redes neuronales artificiales y máquinas de soporte vectorial. El algoritmo de red neuronal uso dos capas ocultas: sigmoide y lineal, mostró el mejor accuracy ( $97.5\% \pm 2.5\%$ ) en la predicción del resultado e identificó la presencia o ausencia de fístulas nidales como el factor más importante de hemorragia y muerte subsiguiente post embolización, asimismo se estimó un área bajo la curva de 0.76. Se concluyó que las malformaciones arteriovenosas cerebrales pueden tratarse con éxito mediante técnicas endovasculares o combinarse con cirugía y radiocirugía con un perfil de riesgo aceptable. Además, las técnicas de aprendizaje automático pudieron predecir el resultado final con mayor precisión que la regresión logística convencional y pueden ayudar a individualizar el tratamiento en función de factores clave de predicción.

Asimismo, Panesar, Souza, Yeh, & Fernández-Miranda (2019), publicaron un artículo comparando técnicas de aprendizaje automático con la regresión logística para predecir la mortalidad en pacientes portadores de pequeños y heterogéneos gliomas cerebrales. Con el objetivo de lograr un pronóstico de la mortalidad a 2 años de estos pacientes, se trabajó con una base de datos pequeña y con alta dimensionalidad que consistió en 76 pacientes recolectados desde el 2009 hasta 2017, se estudiaron 21

variables: edad, sexo, diámetro del glioma, tipo de intervención, escala de coma de Glasgow, otros tratamientos, número de cirugías, localización del glioma, grado de clasificación, características moleculares. Todos los modelos fueron corridos 15 veces cada uno para reducir el problema de sobreajuste debido al pequeño tamaño de base de datos. Se comparó el efecto de aplicar algoritmos de clasificación de aprendizaje automático: redes neuronales artificiales, árboles de decisión, máquina de soporte vectorial y regresión logística a una base de datos cruda (todas las variables) versus datos donde solo se incluyeron variables con significancia estadística (selección de atributos: 14 características). En los resultados, donde se trabajó con todas las 21 variables sin procesar se logró un rendimiento de accuracy/ área bajo la curva de: 70.7% / 0.70 (49.9-88.5) para las redes neuronales artificiales, 68% / 0.72 (53.4-90.4) para máquina de soporte vectorial, 66.7% / 0.64 (43.6-85.0) para la regresión logística y 65% / 0.70 (51.6-89.5) para árboles de decisión. En cambio, al trabajar con las 14 variables se obtuvo un rendimiento de accuracy/área bajo la curva de: 73.4% / 0.75 (62.9-87.9) para las redes neuronales artificiales, 73.3% / 0.74 (62.1-87.4) para máquina de soporte vectorial, 69.3%/0.73 (60.0-85.8) para la regresión logística y 65.2% / 0.63 (49.1-76.9) para la técnica de árboles de decisión. Se concluyó que estas técnicas también se pueden aplicar a conjuntos de datos pequeños y altamente dimensionales. Además, se encontró que las técnicas de aprendizaje automático lograron un rendimiento razonable en comparación con estudios similares en la literatura. A pesar de que las bases de datos locales pueden ser repositorios de cáncer pequeños versus grandes, se demuestra que las técnicas de aprendizaje automático aún pueden aplicarse a su análisis; sin embargo, los métodos estadísticos tradicionales son de beneficio similar. El pequeño tamaño de una base de datos local no proporciona una barrera para la implementación de técnicas de aprendizaje automático para fines de pronóstico.

Además, Sami (2013), en su tesis de maestría titulada “Predicción analítica e interpretación de aneurismas cerebrales”, diseñó un marco de predicción tanto para la ruptura y la respuesta al tratamiento para entender los factores de riesgo de aneurisma cerebral. Se emplearon las técnicas de: clasificación basada en reglas de asociación de minería de datos, la regresión logística y el algoritmo clustering de maximización de expectativa para el diseño de tres casos reales de estudio. La precisión de la regresión logística y del algoritmo clustering fueron de 90%, estos son útiles como apoyo en la toma de decisiones en la práctica médica.

Asimismo, Miranda (2016), en su tesis doctoral: “Aplicaciones de principios de ingeniería en la comparación de métodos de clasificación para predecir la respuesta al tratamiento en pacientes con cáncer en la cabeza y el cuello”, usó varios métodos de clasificación como árboles de decisión, máquina de soporte vectorial, regresión logística, Naive Bayes, análisis discriminante lineal, vecinos más cercanos, comparando estos con los modelos de clasificación de ensamble (bagging y boosting) para predecir la pérdida de peso de cinco a más kilogramos y la toxicidad de cinco a más Grays sobre la dosis de terapia de radiación recibida por los pacientes en 90 días post tratamiento. Se obtuvieron datos de 326 pacientes del hospital Johns Hopkins en Baltimore, USA, se trabajó con 295 variables predictoras para predecir la pérdida de peso. Entre las variables predictoras se incluyó: factores del tumor, diagnóstico, tratamiento, localización del cáncer, y una encuesta de calidad de vida entre otros. Para predecir la toxicidad se usó 597 pacientes y 37 variables predictoras entre las cuales se incluyó la toxicidad a varios órganos y tejidos. Los datos fueron colectados desde los años 2006 hasta 2014. La selección de atributos y los modelos fueron validados evaluando la exactitud de la performance predictiva con validación cruzada de diez iteraciones. Se construyeron los modelos usando el proceso de entrenamiento y validación con Matlab. Los árboles de clasificación de ensamble bagging mostraron una exactitud de predicción de 86.1% (toxicidad) y 96.3% (pérdida de peso). En cambio, los árboles de ensamble boosting mostraron una exactitud de 92.3% (toxicidad) y 100% (pérdida de peso). Se concluye que los métodos de ensamble mostraron la exactitud de predicción más altas que las otras técnicas de clasificación.

De esta forma, Oermann y otros (2016), presentaron un estudio titulado “Usando un enfoque de machine learning para predecir resultados de malformaciones arteriovenosas cerebrales después de la radiocirugía”, estudiaron durante un periodo de ocho años de seguimiento, con los objetivos de predecir resultados individuales de pacientes con MAVs después de la radiocirugía, aplicando un enfoque de aprendizaje automático y además analizar la capacidad predictiva de los sistemas clínicos de clasificación ya existentes en estos pacientes. Se evaluó las bases de datos de radiocirugía de pacientes con malformaciones arteriovenosas cerebrales de tres instituciones pertenecientes al Consorcio International Gamma Knife (IGKRF) que comprendieron 1442 pacientes provenientes de la Universidad de Virginia y de la Universidad de Pittsburgh y 60 pacientes de la Universidad de Nueva York cuyos datos



se usaron para la validación. Se estudiaron 23 variables que incluían parámetros clínicos y de tratamiento estándar: sexo, edad, tratamientos previos, hemorragia previa a la malformación, presencia de aneurismas dentro del nido (intranidales) o alrededor (perinidales) de la MAV, dosis marginal, dosis máxima, isodosis, número de isocentros, duración del seguimiento y síntomas clínicos. Las características de la angioarquitectura incluían la ubicación (elocuente versus no elocuente y profunda versus superficial) de diámetro máximo, volumen, número de venas de drenaje y ubicación de las venas de drenaje (solo superficial versus cualquier componente profundo). Las regiones de ubicación se definieron en términos de los lóbulos principales (frontal, temporal, parietal, occipital), ganglios basales, tálamo, cuerpo caloso, tronco encefálico, cerebelo o ínsula. La edad se escaló en escala logarítmica. Los datos fueron particionados en la relación 2:1 y se usó 100 iteraciones. Los modelos predictores que se incluyeron fueron: regresión logística, random forest, descenso de gradiente estocástico y máquina de soporte vectorial. Los modelos clínicos de predicción ya existentes incluyeron a la escala de Spetzler-Martin (SM), los puntajes basado en la escala de radiocirugía modificada (RBAS) y el de la escala de Virginia. El mejor modelo predictor fue el de regresión logística y obtuvo un promedio de área bajo la curva (AUC) de 0,71, en cambio con los sistemas de predicción clínicos se obtuvo un AUC de 0,63 en todos los puntos de tiempo. El modelo de regresión logística tuvo un accuracy de 0,74 en todos los umbrales de todos los tiempos con una especificidad y sensibilidad del 62% y 85% respectivamente. Cuando se realizó la optimización de hiperparámetros todos los clasificadores de aprendizaje automático mostraron buenos resultados. Se concluye que el aprendizaje automático fue capaz de proporcionar las mejores predicciones posibles de los resultados de radiocirugía y podría ser una herramienta poderosa para la planificación del tratamiento y garantizando una correcta atención al paciente.

Sender y col (2017), realizaron un estudio de revisión sistemática donde se evaluó el potencial de predicción en los resultados neuroquirúrgicos usando técnicas de aprendizaje automático. Se identificaron treinta estudios que evaluaron los algoritmos de aprendizaje automático utilizados como modelos de predicción para la supervivencia, la recurrencia, la mejoría de los síntomas y los eventos adversos en pacientes sometidos a cirugía por epilepsia, tumor cerebral, lesiones espinales, enfermedad neurovascular, trastornos del movimiento, lesión cerebral traumática e hidrocefalia. Los algoritmos de aprendizaje automático supervisados más frecuentemente encontrados en esta revisión

fueron: redes neuronales y máquina de soporte vectorial. Otros algoritmos también usados fueron random forest y naive bayes. Los modelos de aprendizaje automático predijeron los resultados después de la neurocirugía con una mediana de accuracy y área bajo la curva de 94.5% y 0.83, respectivamente. En los estudios donde se comparó los modelos de aprendizaje automático con la regresión logística, los modelos de aprendizaje automático tuvieron un rendimiento significativamente mejor. Se concluyó que los modelos de aprendizaje automático tienen un gran potencial para mejorar la predicción del resultado neuroquirúrgico, se menciona además que estos pueden ser una ayuda valiosa para los médicos, pacientes y sus familias en el proceso de toma de decisiones quirúrgicas y médicas. Se recomienda que futuros estudios deberían explorar los obstáculos asociados con la creación, validación y despliegue de modelos de aprendizaje automático en la atención clínica paralela al desarrollo de estos métodos, así como las implicaciones éticas y sociales de su adopción.

Asimismo, Uddin, Khan, Hossain, & Ali Moni, (2019), realizaron una investigación con el objetivo de comparar diferentes tipos de algoritmos supervisados de aprendizaje automático, así como su rendimiento y uso para la predicción del riesgo de enfermedad. Se usaron dos bases de datos, Scopus y Pubmed, seleccionaron 48 artículos para la predicción de enfermedades, donde se usaron diferentes algoritmos supervisados de aprendizaje automático comparándolos. Se eligieron estudios donde se implementaron múltiples métodos de aprendizaje automático en los mismos datos y predicción de la enfermedad. Los algoritmos empleados fueron: máquina de soporte vectorial, Naive Bayes, random forest, regresión logística, redes neuronales, K vecinos más cercanos. Se encontró que el algoritmo máquina de soporte vectorial fue el más frecuentemente empleado (29 artículos), seguido del algoritmo de Naive bayes (23 artículos), random forest (17 artículos). El algoritmo random forest mostró superior accuracy en el 53% de estos estudios seguido por máquina de soporte vectorial con el 41% de los estudios. En los artículos en los que se utilizaron métodos de validación de 10 y 5 iteraciones, se encontró que el algoritmo máquina de soporte vectorial muestra un mejor accuracy, además en los artículos donde no se utilizó ningún método para la validación, las redes neuronales mostraron el mejor accuracy. Los autores concluyen que independientemente de las variaciones en la frecuencia y el rendimiento, los resultados demostraron el potencial de estas familias de algoritmos en la predicción de la enfermedad.

## 2.3 Estructura teórica que sustenta el estudio

### 2.3.1 Las malformaciones arteriovenosas cerebrales

#### 2.3.1.1 Epidemiología, características clínicas y diagnóstico

Las malformaciones arteriovenosas cerebrales (MAV) son lesiones vasculares cerebrales debido a la ausencia de capilares en un sector del cerebro generando un pasaje directo de arterias a venas cerebrales los que causan varices intracraneales, donde su mayor riesgo es la hemorragia cerebral por rotura del mismo (Aminoff, Boller & Swaab, 2017; Spetzler, Kondziolka, Higashida, & Kalani, 2015).

Las MAVs cerebrales tienen un perfil de presentación consistente en hemorragia, convulsiones o dolor de cabeza. La incidencia y la prevalencia anuales de las MAVs es de aproximadamente de 1 caso en 100 000 y de 50 casos en 100 000 respectivamente. Las tasas de mortalidad anuales generales varían de 0.7% a 2.9% en diferentes poblaciones, teniendo los pacientes con MAV un exceso de mortalidad significativa en comparación con la población general (Aminoff, Boller & Swaab, 2017; Lawton y otros, 2015; Spetzler, Kondziolka, Higashida, & Kalani, 2015).

Por ser una enfermedad de relativa rareza, la verdadera incidencia y prevalencia de MAV cerebrales es desconocida a esto se agrega la falta de estudios epidemiológicos en poblaciones a gran escala.

El riesgo anual de hemorragia por una malformación arteriovenosa cerebral es aproximadamente del 3%, pero depende de la clínica y las características anatómicas de la malformación, el riesgo puede ser tan bajo como 1% o tan alto como 33%. El riesgo de hemorragia cerebral aumenta en 3.2 si el paciente ha tenido episodios previos de sangrado o si la malformación se localiza en el interior del cerebro o el tallo cerebral (razón de riesgo, 2.4; IC del 95%, 1.4 a 3.4) o es caracterizado exclusivamente por un drenaje venoso profundo (razón de riesgo 2,4; IC del 95%, 1,1 a 3,8). Los pacientes con ninguno de estos factores de riesgo tienen un riesgo muy bajo de hemorragia cerebral (<1% anual), los pacientes con uno de estos factores tienen un riesgo bajo (3 a 5% anual), pacientes con dos los factores tienen un riesgo medio (8 a 15% anual), y los pacientes con los tres factores tienen un riesgo alto (> 30% anual). (Solomon & Connolly, 2017).

La causa exacta de la MAV sigue siendo desconocida. A pesar de su origen de desarrollo, generalmente no son lesiones hereditarias. (Spetzler, Kondziolka, Higashida, & Kalani, 2015).

Los factores de riesgo significativos para la hemorragia incluyen ruptura previa, localización profunda, drenaje venoso profundo, aneurismas asociados y embarazo. La edad del paciente, sexo y los MAV pequeños no se consideran factores de riesgo significativos para la hemorragia MAV. Además del riesgo de hemorragia y de convulsiones, la historia natural de una MAV también abarca la carga psicológica diaria que un paciente debe soportar sabiendo que posee un MAV no tratada. Todos los factores anteriores deben tenerse en cuenta al determinar el tratamiento óptimo de una MAV. (Aminoff, Boller & Swaab, 2017).

Florestean & Ciurea (2012), refieren que las MAVs pueden ser fistulosas, plexiformes o mixtas. Las fistulosas se caracterizan por la unión directa entre una arteria normal y una vena, mientras que las plexiformes se definen como una arteria conectada a una red indiferenciada de vasos inmaduros (nunca a una red capilar) la cual se conecta a su vez con una vena. Las MAVs cerebrales habitualmente tienen tres componentes: las arterias aferentes, el nido y las venas de drenaje. Sus características generales incluyen la presencia de una o varias conexiones directas arteriovenosas, las cuales generan una comunicación directa “shunt” de alto flujo a través de pequeñas arterias sin capa muscular, y la ausencia de una red de capilares. Las MAVs pueden presentarse con hemorragia cerebral, crisis comiciales, cefalea o déficit focal, siendo la hemorragia la forma más frecuente. Se diagnostica una MAV tras realizarse una prueba de imagen cerebral bajo la sintomatología expuesta previamente o al hallarla de manera incidental tras realizar un estudio de tomografía computarizada (TC) o resonancia magnética (RM) cerebral por otro motivo.

Para el diagnóstico de las MAV cerebrales se utilizan tecnología imagenológica entre ellas tenemos: la tomografía computarizada (TC), la resonancia magnética cerebral (RM) y angiografías cerebrales digitales.; las que nos permiten ver las características anatómicas estructurales y la dinámica de flujo de las MAV. Nos permiten determinar la ubicación anatómica exacta, tamaño del nido, los alimentadores arteriales y el drenaje venoso. (Aminoff, Boller & Swaab, 2017) y con ello la toma de decisión para un tratamiento adecuado.

### 2.3.1.2 Tratamiento

Una vez diagnosticada se necesita un tratamiento cuya finalidad es la oclusión total de la MAV. Los tratamientos parciales no son efectivos porque no excluyen el riesgo de sangrado permaneciendo una alta morbimortalidad. (Samblas, Sallabanda, Martínez, & Calvo, 2012). El tratamiento de la malformación arteriovenosa cerebral es desafiante y se ha recurrido a la cirugía directa, la embolización y la radiocirugía.

#### 2.3.1.2.1 Técnica microquirúrgica

La resección quirúrgica es la primera opción terapéutica, la cirugía de una MAV pequeña, con drenaje venoso superficial localizado en área motora, no está libre de riesgos. Las cirugías de las MAV localizadas en el tronco, tálamo y en zonas elocuentes del cerebro es comprometida y riesgosa (Samblas, Sallabanda, Martínez, & Calvo, 2012)

La escala de graduación de Spetzler-Martin ha sido utilizado para determinar el riesgo de déficits neurológico postoperatoria o muerte como resultado del tratamiento quirúrgico de una malformación arteriovenosa del cerebro. Los resultados de malformaciones sin ruptura generalmente son mejores que las que han sangrado y causan déficits neurológicos antes de la cirugía. Además, los resultados publicados reflejan la experiencia de centros especializados que operan en grandes cantidades de pacientes y no pueden reflejar los riesgos en los centros con menos volúmenes de cirugías. (Solomon & Connolly 2017).

#### 2.3.1.2.2 La terapia endovascular

El tratamiento endovascular de malformaciones arteriovenosas se logra mediante la inserción de un catéter guía generalmente por la arteria femoral y a su través la navegación intracraneal de un microcatéter hasta la MAV, donde se intenta ocluir la MAV con agentes como como N-butil-2-cianoacrilato o un copolímero etileno-alcohol vinílico no adhesivo. La embolización como modalidad única de tratamiento es efectiva en el 10 a 15 % de pacientes, siendo su utilidad principal un complemento valioso a la microcirugía para reducir la morbilidad quirúrgica al ocluir los vasos de alimentación arterial profunda, minimizando la pérdida de sangre y disección extensa durante la operación. Sin embargo, la embolización preoperatoria ha resultado en déficits neurológicos persistentes en el 2.5% de los pacientes. La embolización también es útil

para reducir el tamaño de las MAV grandes a un tamaño que permita ofrecer el tratamiento complementario con radiocirugía estereotáctica. La embolización pre radioquirúrgica también puede ocluir aneurismas en la alimentación de arterias y aneurismas que ocurren dentro del nidus de la malformación arteriovenosa. Para reducir el riesgo de sangrado recurrente durante el intervalo entre la radiocirugía y la desaparición de la malformación arteriovenosa. (Solomon & Connolly 2017).

#### 2.3.1.2.3 La Radiocirugía estereotáctica

La radiocirugía estereotáctica es un tratamiento bien estudiado para malformaciones arteriovenosas cerebrales y consiste en la entrega de radiación focalizada de alta dosis a la malformación arteriovenosa, induciendo la esclerosis gradual de los vasos sanguíneos y la trombosis de la lesión y con ello su curación u obliteración. La obliteración exitosa de la malformación arteriovenosa se predice en función del tamaño de la lesión y la dosis de radiación administrada a los márgenes de la malformación. (la "dosis de marginal"). Las lesiones que responden favorablemente a la radiocirugía estereotáctica son las pequeñas malformaciones (<4 cm) con grado 1 o 2 en la escala de Virginia, que son tratadas con una dosis de radiación de 18 Gy o más. Con este tratamiento, la tasa de obliteración de la lesión, evaluada mediante resonancia magnética (MRI) o angiografía, se aproxima al 80%. Las lesiones más grandes (grados 3, 4 o 5) se tratan con dosis de radiación marginal más bajas y se curan con menor frecuencia (48%), y el tratamiento se asocia con un riesgo considerable de necrosis inducida por radiación del cerebro adyacente (3%). Las ventajas de la radiocirugía comparada con los otros tratamientos es que es no invasivo y tiene mínimo riesgo de complicaciones agudas además se realiza como un procedimiento ambulatorio que no requiere tiempo de recuperación para el paciente. Una de las mayores limitaciones de la radiocirugía estereotáctica para una MAV es el retraso sustancial en la obliteración de la lesión, la cura no es inmediata y lleva entre 2 a 4 años en promedio para conseguir la obliteración total. Durante el intervalo entre el tratamiento radioquirúrgico y la obliteración de la MAV, el riesgo de hemorragia permanece. La mayoría de los datos sugieren que el riesgo de sangrado durante este período sólo es ligeramente menor que el riesgo durante el período previo al tratamiento. (Friedman & Bova, 2011; Solomon & Connolly 2017).

El tratamiento con radiocirugía tiene como objetivos: a) La obliteración de la malformación arteriovenosa. b) La disminución del riesgo de sangrado, hasta su

eliminación. c) La mejoría de la sintomatología clínica y la prevención de la misma. d) Mantener y mejorar la calidad de vida del paciente. (Samblas, Sallabanda, Martínez, & Calvo, 2012)

La eficacia del tratamiento de radiocirugía en las MAV está comprobada y en la actualidad se utiliza como tratamiento sólo o en combinación ya sea con la terapia endovascular o la cirugía. La tasa de cierre completo varía de 54 a 90%, esto depende de la dosis, del volumen y ocurre en un periodo comprendido de 1 a 5 años (Pollock, y otros, 2016; Samblas, Sallabanda, Martínez, & Calvo, 2012)

Las complicaciones de la radiocirugía se relacionan con el volumen de la MAV, con el edema que se produce por falta de vascularización hasta la reestructuración de esta y con la dosis administrada. El aumento de la dosis del margen radioquirúrgico de hasta aproximadamente 25 Gy, mejora la probabilidad de obliteración MAV, sin embargo, el riesgo de efectos adversos de radiación también aumenta concomitantemente con esa dosis, relacionando con un daño cerebral. (Aminoff, Boller, & Swaab, 2017; Samblas, Sallabanda, Martínez, & Calvo, 2012)

La radiocirugía tiene modalidades para entregar los fotones al blanco seleccionado el cual se hace mediante un acelerador lineal o mediante un sistema Gamma Knife, mientras que el primero obtiene los fotones a partir de rayos X, el segundo los obtiene a partir de fuentes de cobalto 60, no habiendo diferencias en cuanto a la efectividad de sus tratamientos.

Javalkar, y otros (2009) determinaron que la efectividad de la radiocirugía puede obedecer a muchos factores y en lo que más se ha incidido es en un balance entre la dosis de radiación y la demarcación de la lesión.

En general la efectividad de la radiocirugía en la curación de las MAV se da con alta probabilidad con la aplicación de radiación en un rango de 18 a 23 Gy y hasta 10 cm cúbicos en volumen (3 cm de diámetro), con mínimos efectos colaterales (Flickinger, y otros, 1999).

Puesto que las MAV son de diferentes tamaños y difieren en su forma y ubicación en el cerebro, además otros factores no relacionados como la experiencia, el uso de equipos, el software utilizado, dosis de radiación, selección de blancos y de modo general se reporta que las tasas de cura de la MAV con radiocirugía varían de 51

a 92%. El volumen pequeño de MAV o la alta dosis marginal son los predictores más comunes para la oclusión o cura de la MAV (Kwon, y otros, 2000).

Las fallas en el tratamiento se deben usualmente a dos factores: mala demarcación del nido de la MAV, por embolización previa, desplazamiento de la MAV por el hematoma, baja dosis por ubicación de la MAV en áreas elocuentes o tamaño demasiado grande de las MAV (Flickinger, y otros, 1999).

Se ha reportado que las complicaciones por hemorragia cerebral por el tratamiento con radiocirugía de las MAV son más en las estructuras profundas, localización en los ganglios basales, lóbulo occipital o cerebelo (Flickinger, Kondziolka, Maitz, & Lunsford, 1998)

Para MAV con alto grado de riesgo por estar ubicadas en zonas elocuentes o muy profundas la radiocirugía mostró una efectividad de curación en el 44 % con un seguimiento a 5 años. Los factores que favorecen la curación fueron pacientes no embolizados, localización superficial, alta dosis de prescripción. La tasa anual de hemorragia post radiocirugía fue de 3% y necrosis inducida por radiación de 12% (más en las MAV que se habían roto). El deterioro clínico temporal o permanente fue en 9 y 10% respectivamente (Ding, y otros, 2014).

No obstante que la radiocirugía para el tratamiento de malformaciones arteriovenosas cerebrales es una alternativa moderna y de mínima invasión con resultados favorables no está exenta de complicaciones. Es una técnica que ha madurado con el tiempo desde su creación y basados en reportes se han estandarizando las dosis a aplicar, mejoría de software y mejoramiento de equipos de diagnóstico y tratamiento. Las hemorragias y falta de curación u obliteración de los pacientes con MAV cerebrales tratados con radiocirugía se debe a una falta de estandarización del manejo y consideración de las múltiples dimensiones que pueden influir en los resultados finales del tratamiento ya sea de curación, falta de cierre de la MAV en el periodo de seguimiento, persistencia de convulsiones o su complicación con hemorragias cerebrales. Entre estas dimensiones, se puede mencionar a: las dimensiones clínicas: tratamientos anteriores, forma de presentación, sangrados, convulsiones, dolor de cabeza. Las Dimensiones Físicas del hardware y software de radiocirugía, como: tamaño del blanco, conformidad de conos con la lesión, dosis de radiación, número de isocentros, número de arcos de radiación, volúmenes irradiados. Las dimensiones



estructurales de la MAV como ubicación cerebral, tamaño, forma y número de vasos aferentes, eferentes, velocidad de flujo, presencia de aneurismas, fístulas y estenosis venosas. Las Dimensiones epidemiológicas como edad, sexo, actividad, procedencia, nivel socioeconómico, grado de instrucción, otras enfermedades, ocupación, tipo de seguro, vivienda. Dimensiones del seguimiento y control, tiempo de obliteración de la MAV, imágenes de diagnóstico como resonancia magnética, tomografía cerebral y angiografía digital cerebral.

### 2.3.2 Sistemas de Calificación y clasificación de las MAVs

Los sistemas de clasificación son un método importante para describir las malformaciones arteriovenosas y evitar complicaciones quirúrgicas se han identificado factores que determinan los riesgos operativos y han desarrollado criterios que orientan esta decisión. Estos sistemas de clasificación se han convertido en algo más que predictores de riesgos quirúrgicos; son descripciones abreviadas de la patología en sí y una parte del lenguaje de las malformaciones arteriovenosas cerebrales. El sistema de clasificación de Spetzler-Martin es el predominante porque es simple, aplicable y preciso. Fue originalmente desarrollado para predecir la respuesta de tratamiento de microcirugía, pero puede ser usado para predecir la respuesta al tratamiento de radiocirugía. La escala tiene cinco grados e incorpora puntajes para tres características de la malformación arteriovenosa cerebral: tamaño del diámetro, elocuencia (función crítica del parénquima cerebral circundante) y patrones de drenaje venoso. El más bajo grado indica más bajo riesgo para el tratamiento. (Lawton, y otros, 2015)

En este sentido, el artículo publicado por Spetzler y Martin significó un gran aporte a la uniformidad de los distintos reportes acerca del tema. Este autor clasifica a las MAV en 5 grados de acuerdo a sólo tres aspectos: su tamaño, su localización (área elocuente o no) y a la presencia o no de drenaje venoso profundo. Su simpleza y la alta correlación pronóstica han hecho que esta clasificación se siga usando y que sea un estándar dentro de la especialidad. (Yáñez, A., 2001).

El sistema Spetzler-Martin no incluye todos los factores importantes para la selección quirúrgica y los resultados. Factores como la edad, la presentación hemorrágica y la compacidad que también son importantes. Malformaciones arteriovenosas compactas tienen bordes definidos que separan la malformación arteriovenosa cerebral del tejido cerebral adyacente, mientras que las malformaciones

difusas tienen bordes irregulares y están entremezclados con la materia cerebral. El sistema de clasificación de Lawton – Young incorpora estos otros factores. Se asignan puntos para el ABC de malformaciones arteriovenosas: edad, sangrado y compacidad. La suma de los grados Spetzler – Martin y Lawton – Young, proporciona mayor precisión predictiva y establece límites para la operabilidad de la malformación arteriovenosa. (Ver Tabla 1) (Lawton, y otros, 2015; Solomon & Connolly, 2017)

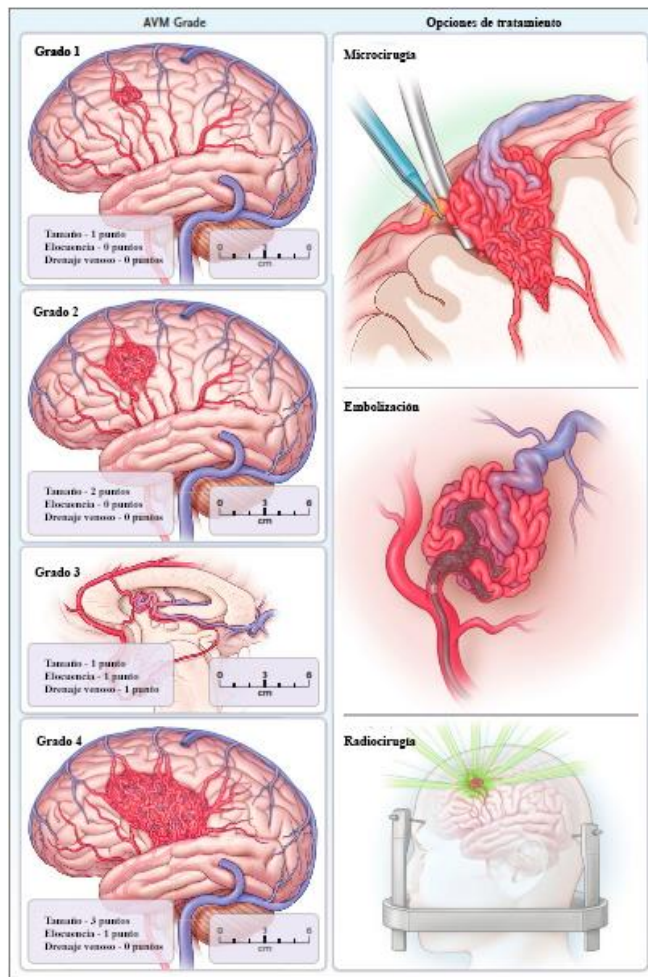
Tabla 1. *Escala de Spetzler-Martin y de Lawton-Young para malformaciones arteriovenosas cerebrales*

<b>Parámetros</b>	<b>Escala de Spetzler-Martin (Score total =5)</b>	<b>Puntos</b>
<b>Tamaño</b>	<3 cm	1
	3-6 cm	2
	>6cm	3
<b>Drenaje venoso</b>	Drenaje superficial en vena cortical y senos de convexidades	0
	Drenaje profundo en la vena de Galeno	1
<b>Elocuencia</b>	No	0
	Si (Localizado en la corteza sensoriomotor, áreas del lenguaje, corteza visual, hipotalamo, cápsula interna, tronco cerebral, pedúnculo cerebelar o núcleo cerebelar profundo).	1
	<b>Escala de Lawton- Young (Score total =5)</b>	
<b>Edad</b>	< 20 años	1
	20-40 años	2
	> 40 años	3
<b>Sangrado</b>	Si	0
	No	1
<b>Compacidad</b>	Si	0
	No	1

Fuente (Lawton, y otros, 2015; Solomon & Connolly, 2017)

Las lesiones de grado 1 son malformaciones superficiales pequeñas que no afectan a las regiones elocuentes del cerebro. Grado 2 las lesiones afectan tanta región elocuente en el cerebro o son mayores de 3 cm en diámetro. La microcirugía, embolización y la radiocirugía estereotáctica son todas las opciones de tratamiento apropiadas para el grado 1 y malformaciones de grado 2. La malformación de grado 3 es pequeña, pero profundamente ubicado en el tejido cerebral elocuente, con drenaje en el

sistema venoso profundo. La radiocirugía es la mejor opción de tratamiento para lesiones no rotas en esta categoría. Las lesiones de grado 4 son todas mayores de 3 cm de diámetro e involucran regiones elocuentes del cerebro. Tales lesiones son generalmente tratadas conservadoramente, ya que todas las opciones de tratamiento son peligrosas y a menudo fracasan en lograr una cura. (Solomon & Connolly, 2017) (Ver Figura 1)



*Figura 1.* Ejemplos de grados la escala de Spetzler–Martin. Grados 1 hasta 4 de malformaciones arteriovenosas cerebrales (MAV) y opciones de tratamiento.

Fuente: (Solomon & Connolly, 2017)

Para predecir resultados más precisos después de la radiocirugía se utiliza la escala de radiocirugía de Virginia (VRAS). Esta escala fue desarrollada en la Universidad de Virginia basada en 20 años de experiencia institucional en más de 1400 pacientes con malformaciones arteriovenosas cerebrales y tratados con radiocirugía, esta

escala asigna uno de los cinco grados sobre la base de puntos para el volumen ( $<2 \text{ cm}^3$  [0 puntos], 2 a  $4 \text{ cm}^3$  [1 punto],  $>4 \text{ cm}^3$  [2 puntos]), ubicación elocuenta (1 punto) y antecedentes de hemorragia cerebral (1 punto). Las malformaciones arteriovenosas de grado 1 tienen 0 puntos y las malformaciones de grado 5 tienen 4 puntos. (Tabla 2) (Aminoff, Boller, & Swaab, 2017)

Tabla 2. *Escala de malformaciones arteriovenosas cerebrales de Radiocirugía de Virginia*

<b>Variable</b>	<b>Puntos</b>
<b>Volumen de la malformación arteriovenosa cerebral (<math>\text{cm}^3</math>)</b>	
<2	0
de 2 – 4	1
>4	2
<b>Malformación situada en área elocuenta</b>	1
<b>Antecedente de hemorragia</b>	1

Fuente (Aminoff, Boller, & Swaab, 2017)

Asimismo, Dumont y otros (2015), propusieron la escala Buffalo para calificar los posibles resultados obtenidos cuando el tipo de tratamiento para las MAV cerebrales es el endovascular. Se determina la calificación propuesta del puntaje de Buffalo contabilizando el número de pedículo arterial, diámetro del pedículo arterial y elocuencia de la corteza cercana (ver tabla 3).

En consecuencia, el poder de esta nueva escala de calificación está en ayudar a los cirujanos a estratificar mejor las estrategias de tratamiento a los pacientes, entre quirúrgicos y endovasculares. El puntaje Buffalo representa una nueva herramienta potencial de utilidad para guiar el manejo endovascular de MAV intracraneales. (Aminoff, Boller, & Swaab, 2017)

Tabla 3. *Escala de Buffalo para tratamiento endovascular de MAV cerebrales*

<b>Características</b>	<b>Puntos</b>
<b>Número de pedículos arteriales</b>	
1 o 2	1
3 o 4	2
5 o más	3
<b>Diámetro de pedículos arteriales</b>	
Mayoría de > 1mm	0
Mayoría de ≤ 1mm	1
<b>Localización del nido</b>	
No elocuente	0
Elocuente	1

Fuente (Dumont, y otros, 2015)

Además, se debe agregar que existen escalas que caracterizan la condición funcional del paciente. Así se tiene la escala de performance de Karnofsky (KPS) que evalúa la capacidad funcional y fue la primera que se desarrolló en pacientes con cáncer para obtener información sobre el pronóstico de supervivencia, ayuda a tomar decisiones sobre usos y riesgos de terapias, y contribuye a definir el nivel de asistencia que un paciente necesita. La escala de Karnofsky (KPS) evalúa diferentes dimensiones de un individuo como nivel de actividad, capacidad de trabajo, nivel de cuidados requeridos, gravedad y progresión de la enfermedad y necesidad de hospitalización. Esta escala va de 100 (salud normal) a 0 (muerte), descendiendo en valores de 10 en 10 (Tabla 4). (Pérez-Cruz & Acevedo, 2014; Reponen, Tuominen, & Korja, 2014).

Tabla 4. *Escala de Estado Funcional de Karnofsky*

<b>Definición</b>	<b>%</b>	<b>Criterio</b>
Capaz de realizar actividad normal y de trabajar. No requiere cuidados especiales.	100	Normal, no hay molestias. No hay evidencia de enfermedad
	90	Capaz de realizar actividad normal. Signos y síntomas menores de enfermedad
	80	Actividad normal con esfuerzo. Algunos signos y síntomas de enfermedad.
Incapaz de trabajar. Capaz de vivir en casa y de satisfacer la mayoría de las necesidades personales. Requiere una cantidad variable de asistencia	70	Cuida de si mismo. Incapaz de realizar actividad normal o de realizar trabajo activo
	60	Requiere de asistencia ocasional, pero es capaz de satisfacer la mayoría de sus necesidades
	50	Requiere asistencia considerable y cuidados médicos frecuentes
Incapaz de cuidarse por si mismo. Requiere cuidados institucionales u hospitalarios. La enfermedad puede estar progresando rápido.	40	Discapacitado. Requiere cuidados especiales y asistencia
	30	Gravemente discapacitado. Hospitalización esta indicada aunque muerte no se ve inminente
	20	Muy enfermo, hospitalización es necesaria. Tratamientos de soporte activo son necesarios
	10	Moribundo, proceso fatal progresando rápidamente
	0	Fallecido

Fuente (Pérez-Cruz & Acevedo, 2014)

Predecir el resultado de un paciente específico con un tratamiento particular es fundamental en la práctica médica.

### 2.3.3 Aprendizaje automático y análisis predictivo

#### 2.3.3.1 Definiciones Generales

El aprendizaje automático, es un área de la inteligencia artificial que se ocupa de identificar patrones a partir de los datos cuyo fin es dotar a una máquina y a través de algoritmos utiliza datos de entrada para lograr una tarea deseada sin ser literalmente programado para producir un resultado particular, tiene la capacidad de entrenar y aprender en base a esos datos, logrando al final del entrenamiento un modelo que permite la generalización (El Napa, Li, & Murphy, 2015; Shobha & Rangaswamy, 2018; Veliz Capuñay, 2018; Mechelli & Vieira, 2019).

El aprendizaje automático se clasifica en términos generales como aprendizaje supervisado, no supervisado, semi supervisado y de refuerzo. Un modelo de aprendizaje supervisado tiene dos tareas principales: la clasificación y la regresión. La clasificación se trata de predecir una etiqueta de clase nominal, mientras que la regresión se trata de predecir el valor numérico para la etiqueta de clase. Matemáticamente, construir un modelo de regresión es tratar de identificar la relación entre la etiqueta de clase y los predictores de entrada. Los predictores también se llaman atributos. En términos estadísticos, los predictores se llaman variables independientes, mientras que la etiqueta de clase se llama variable dependiente. Un modelo de regresión es una representación de esta relación entre variables dependientes e independientes. Una vez que esto se aprende durante la fase de entrenamiento, cualquier dato nuevo se conecta a la curva de relación para encontrar la predicción (Shobha & Rangaswamy, 2018; Mechelli & Vieira, 2019).

Asimismo, en el aprendizaje no supervisado, no existe la variable dependiente. El proceso de aprendizaje se logra usando las relaciones entre un conjunto de variables. Con este tipo de aprendizaje se construyen modelos descriptivos, en cambio con el aprendizaje supervisado se construyen modelos predictivos. Tenemos en el aprendizaje no supervisado: clustering, análisis de componentes principales, conglomerados, reglas de asociación (Çelik & Altunaydin, 2018; Veliz Capuñay, 2018; Mechelli & Vieira, 2019).

En el aprendizaje semi supervisado, las etiquetas o la variable dependiente solo están disponibles para una parte de los datos. El aprendizaje semi supervisado aborda este problema al permitir que el modelo integre los datos no etiquetados disponibles en su aprendizaje supervisado. Este enfoque es útil cuando no es práctico o es demasiado costoso acceder o medir la variable objetivo para todos los participantes. Por ejemplo, los estudios longitudinales que investigan la progresión de la enfermedad requieren varios años de seguimiento para obtener una etiqueta confiable de la enfermedad. En tales casos, el aprendizaje semi supervisado puede usarse para modelar los datos etiquetados existentes, es decir, pacientes que han completado su participación en el estudio, así como datos no etiquetados, es decir, pacientes para quienes aún no se ha establecido una etiqueta confiable. Este enfoque maximiza la cantidad de datos de entrada en el contexto de estudios con recursos y tiempo finitos (Mechelli & Vieira, 2019).

Por otra parte, en el aprendizaje por refuerzo, el objetivo es construir un sistema que pueda aprender interactuando con el entorno. En este tipo de aprendizaje, el comportamiento del algoritmo se forma a través de una secuencia de recompensas y penalizaciones, que dependen de si sus decisiones hacia un objetivo definido son correctas o incorrectas, según lo definido por el investigador. A diferencia del aprendizaje supervisado, donde el algoritmo usa los ejemplos dados para el comportamiento del modelo, en el aprendizaje por refuerzo, el algoritmo puede comportarse libremente, es decir, en base a prueba y error. Una computadora jugando al ajedrez con el humano, aprendiendo a reconocer palabras habladas y aprendiendo a clasificar nuevas estructuras astronómicas son algunos ejemplos de aprendizaje de refuerzo (Shobha & Rangaswamy, 2018; Mechelli & Vieira, 2019).

Así que, el aprendizaje automático abre nuevas posibilidades en la medicina personalizada, al permitir el desarrollo de nuevas herramientas que podrían utilizarse para informar la toma de decisiones de diagnóstico y tratamiento en la práctica clínica diaria (Mechelli & Vieira, 2019).

En esta tesis se usó, el aprendizaje supervisado de clasificación para predecir la etiqueta (variable dependiente): curación de las malformaciones arteriovenosas cerebrales tratadas con radiocirugía. De manera que se describirán las etapas de un sistema de aprendizaje automático supervisado: preprocesamiento de los datos, división aleatoria de los datos en entrenamiento y prueba, aprendizaje y selección de modelos en datos de entrenamiento, predicción de la respuesta de interés en datos de test y evaluación de los modelos seleccionados. (Ver la figura 13)

#### 2.3.3.2 Preparación de los datos

Los datos sin procesar rara vez se presentan en la forma que son necesarias para el rendimiento óptimo de un algoritmo de aprendizaje. Por lo que, el preprocesamiento de los datos es uno de los pasos más importante en cualquier aplicación de aprendizaje automático, siendo necesario considerar características como la dimensión del conjunto de datos, presencia de datos perdidos, dispersos, algunas de las características seleccionadas pueden estar altamente correlacionadas y, por lo tanto, redundantes hasta cierto punto, siendo útil para esto la reducción de dimensionalidad, conllevando a que se requiera menos espacio de almacenamiento y el algoritmo de aprendizaje puede ejecutarse mucho más rápido (Kuhn & Johnson, 2013; Raschka, 2015)



Así tenemos, el tratamiento de datos faltantes la cual podría ser una ocurrencia aleatoria de no traer cualquier información nueva al proceso de clasificación, pero puede haber circunstancias en que la falta es información importante en sí mismo que vale la pena preservar para la clasificación. El impacto de los valores perdidos depende en gran medida del algoritmo utilizado. Mientras que algunos algoritmos toleran datos faltantes sin procesamiento previo (por ejemplo, algunos métodos basados en árboles de decisión, hay otros que requieren que todos los valores estén disponibles para todas las variables. Por lo tanto, se han propuesto varios tratamientos para enmendar esta cuestión. (Consejo, Melcer, & Rozema, 2018). Para ello se tiene que tener en cuenta: que la distribución de la variable no se altere al introducir los datos que reemplazaran a los faltantes y que los datos que se imputan sean consistentes con los valores de las otras variables. La imputación o reemplazo se realiza: asignando un valor basado en los valores de la misma variable, se puede usar la media, la mediana y si son categóricos se usa la moda. También se puede predecir en términos de una o más variables que no tengan datos faltantes, se puede usar la regresión lineal o el modelo del vecino más cercano, también se puede reemplazar el valor perdido, generando  $k$  estimaciones, otra manera es completando con información anterior disponible y también duplicando un valor ya existente (Veliz Capuñay, 2018).

Además, el preprocesamiento de los datos incluye las transformaciones de los datos originales dando lugar a la mejora del desempeño de muchos algoritmos que necesitan predictores medidos en la misma escala y es mediante la normalización y la estandarización que se llega a esto. En la normalización, se redimensiona las variables para tener una escala entre 0 y 1, también se puede normalizar usando la forma de la escala decimal, la cual cambia el punto decimal de modo que el valor absoluto máximo sea siempre menos de 1 y en la estandarización se tendrá valores con media 0 y desviación estándar 1 (Kuhn & Johnson, 2013; Raschka, 2015; García, Luengo, & Herrera, 2015). Asimismo, existen transformaciones como log, raíz cuadrada del valor inverso que contribuyen a corregir la asimetría de los datos. Además, la transformación Box Cox puede ser estimada a partir de los datos de entrenamiento (Kuhn & Johnson, 2013).

La presencia de predictores altamente correlacionados puede contribuir a que el modelo de aprendizaje automático sea inestable, para evitar esto se procede a remover algunos predictores o también se utiliza la técnica de reducción de la dimensionalidad

que genera un nuevo conjunto de atributos tratando de no perder mucha información y mantener o mejorar el rendimiento del modelo. Así tenemos el análisis de componentes principales que explica la estructura de varianza y covarianza de un conjunto de variables con base en sus combinaciones lineales (Kuhn & Johnson, 2013; Mechelli & Vieira, 2019). Una extensión del análisis de componentes principales es el análisis de componentes independientes, este trabaja con estadísticas de orden superior para maximizar la independencia estadística de los componentes estimados. (Mechelli & Vieira, 2019)

Por otra parte, es importante diferenciar las variables de tipo nominales y las ordinales. Para las variables ordinales se pueden convertir los valores tipo texto a enteros y para las variables nominales se puede convertir la variable original en un conjunto de variables ficticias (dummy), se construye un nuevo atributo para cada valor posible, representadas por valores binarios 0 y 1 para cada categoría la variable nominal. La decisión de incluir todas las variables dummy puede depender de la selección del modelo, si el modelo incluye un intercepto habría problemas numéricos al agregar todas las variables ficticias, pero si el modelo no presenta intercepto, entonces usar el conjunto completo de variables ficticias más bien puede contribuir a la interpretación del modelo (Kuhn & Johnson, 2013; Raschka, 2015)

### 2.3.3.3 Entrenamiento del modelo

En el aprendizaje supervisado, el entrenamiento del modelo se refiere al proceso en el que un algoritmo de aprendizaje automático encuentra una función  $f$  que mapea mejor algunas características de entrada  $X$  y la variable objetivo  $y$ .

$$y=f(X)$$

Sin embargo, es importante que el objetivo fundamental del aprendizaje automático sea encontrar una función  $f$  que sea capaz de alto rendimiento cuando se aplica a nuevos datos. Por lo tanto, la prueba final de un algoritmo de aprendizaje automático es ver qué tan bien la función aprendida se generaliza a los datos que no se utilizan para el entrenamiento. (Mechelli & Vieira, 2019)

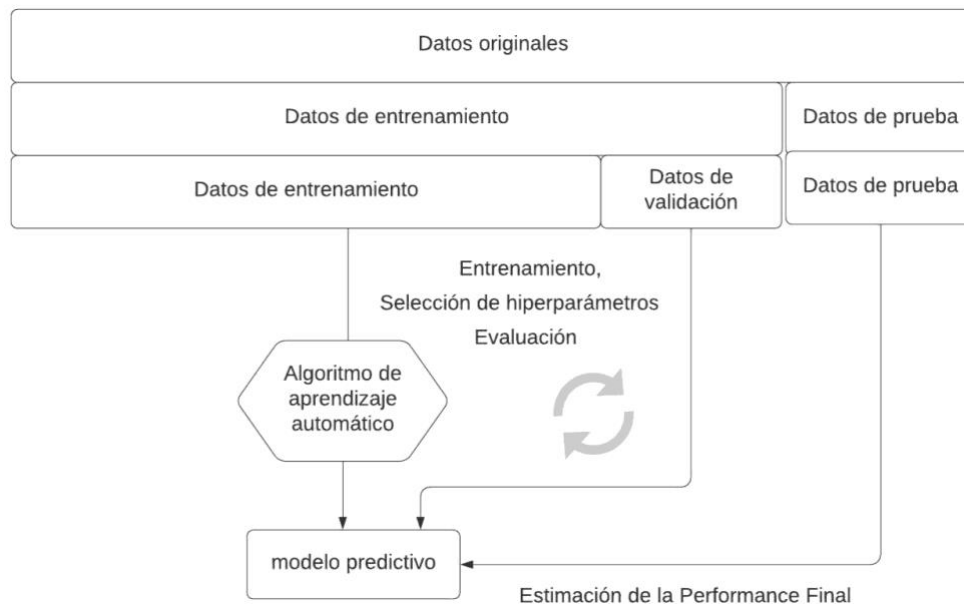
Para Veliz Capuñay (2018), en el aprendizaje supervisado se tiene la presencia de una variable dependiente  $Y$  que debe ser explicada con las variables independientes  $X_1, X_2, \dots, X_p$ . Aquí, las variables independientes y dependientes se observan de manera

conjunta en cada objeto. La variable  $Y$  se llama variable respuesta u objetivo, la cual actúa en el aprendizaje como un supervisor del trabajo que se realiza. en el contexto de los modelos supervisados, cuando se quiere predecir una variable  $Y$  si esta variable es categórica el modelo se llama de clasificación. Si la variable  $Y$  es continua, el modelo se llama de regresión. Con la clasificación, se trata de construir un modelo para ubicar cada elemento de la población en una de las categorías de  $Y$  a partir de las características determinadas por las variables independientes, las que pueden ser categóricas o numéricas. Así, un problema de clasificación se refiere, por ejemplo, a estimar los parámetros del modelo que permite ubicar a cada uno de los clientes que solicita un préstamo a un banco como buen pagador o mal pagador a partir del conocimiento de su sueldo, su edad, sus años de estudio y del valor del préstamo solicitado; también puede referirse a predecir si un paciente tiene tal o cual dolencia en base a características corporales como la presión arterial la temperatura corporal, el colesterol, etc. Estos modelos se llaman clasificadores y son determinados por aplicaciones del espacio de los atributos (determinados por las variables independientes) al espacio de las clases (definidas por la variable dependiente). Entre los modelos de clasificación, se encuentran: el modelo de los  $K$ -vecinos más cercanos, el clasificador ingenuo de Bayes, el análisis discriminante, el análisis de regresión logística, el modelo de árboles de clasificación, el modelo de redes neuronales para la clasificación, el modelo de clasificación de máquinas de soporte vectorial.

La forma más sencilla de evaluar la generalización de un modelo es dividir aleatoriamente todo el conjunto de datos en dos subconjuntos: el conjunto de entrenamiento y el conjunto de prueba. Primero, el algoritmo de aprendizaje automático se ajusta utilizando el conjunto de entrenamiento; este paso implica aprender la función  $f$ . Luego se evalúa el modelo utilizando el conjunto de prueba. Este paso consiste en evaluar el rendimiento de la función aprendida en datos no vistos, que el modelo no utilizó para el entrenamiento (Mechelli & Vieira, 2019).

Los datos de entrenamiento se usan para ajustar los modelos y los datos de validación para seleccionar un modelo en función de su rendimiento predictivo. Por lo que, los datos de prueba se utilizarán para evaluar el error de generalización del modelo seleccionado (Hastie, Tibshirani, & Friedman, 2009).

Al dividir los datos en un solo conjunto de entrenamiento y prueba, el modelo es muy vulnerable a las variaciones fortuitas en los conjuntos de entrenamiento y prueba, para evitar esto se usa el enfoque alternativo para estimar el rendimiento del modelo dividiendo los datos disponibles en conjuntos de entrenamiento / prueba varias veces, cada vez con un grupo diferente de observaciones en cada conjunto. Por lo que, en lugar de tener solo una medida de rendimiento, se tiene varias, una para cada división de entrenamiento / prueba o iteración. Una vez que se completan todas las iteraciones, tenemos una mejor estimación del rendimiento general del algoritmo de aprendizaje automático mediante el cálculo del rendimiento promedio en todas las iteraciones Este proceso también proporciona una medida de variabilidad y estabilidad del rendimiento del modelo. Este método puede ejecutarse utilizando el procedimiento de validación cruzada (Mechelli & Vieira, 2019) Figura 2



*Figura 2.* División del conjunto de datos en entrenamiento, validación y prueba para aplicar a aprendizaje automático (Raschka, 2015)

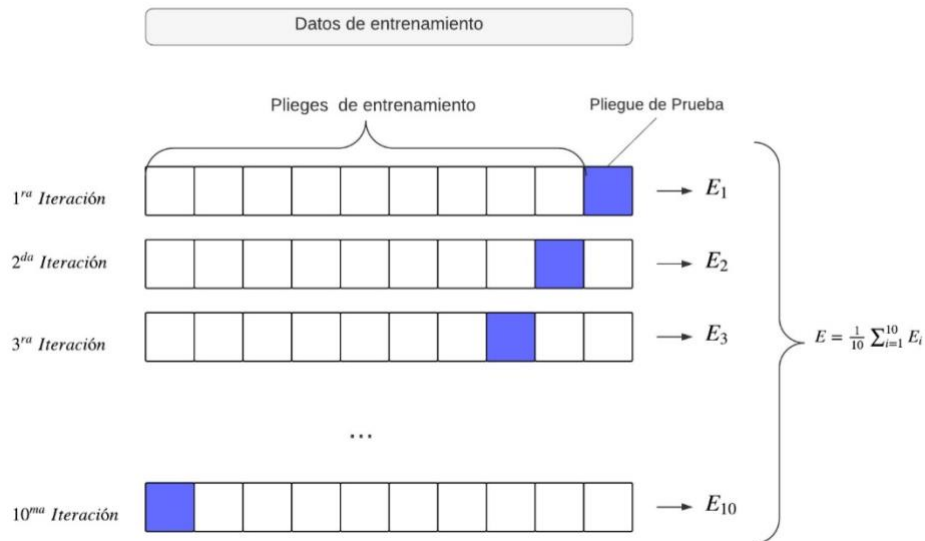
Entre las técnicas de remuestreo, la validación cruzada k-iteraciones representa uno de los más usados en el aprendizaje automático. Esta técnica consiste en la división aleatoria del conjunto de entrenamiento original en k partes de aproximadamente el mismo tamaño, en el que representarán datos de entrenamiento para ajustar el modelo predictivo y la otra parte se reservará para la estimación de su rendimiento. El proceso se repite hasta que todas las partes hayan participado tanto en el entrenamiento como en

la validación del modelo, lo que resulta en  $k$  estimaciones de rendimiento que se resumirá, generalmente, calculando la media y el error estándar (Kuhn & Johnson, 2013; Mechelli & Vieira, 2019). Otro método común implica que el número de particiones es igual al número de observaciones en el conjunto del entrenamiento original,  $K = N$ , esto se conoce como validación cruzada dejando uno fuera (leave-one-out) (Kuhn & Johnson, 2013; Raschka, 2015).

No hay una regla general para elegir  $K$ , uno de los esquemas comúnmente usados es dividir los datos en 5 o 10 partes. En cada iteración, se usan nueve grupos para el entrenamiento y el grupo restante se usa para las pruebas; esto se repite 10 veces, cada vez con un grupo diferente como en el conjunto de prueba. Para la elección de  $K$  es importante considerar como el número de iteraciones afecta el sesgo y la varianza del proceso de prueba. Existe una compensación entre el sesgo y la varianza asociada con la elección de  $K$ : cuanto más cerca esté  $K$  de  $N$  (tamaño de la muestra), menor será el sesgo y mayor la varianza. Por otro lado, cuanto más cerca esté  $K$  de 2, mayor será el sesgo y menor la varianza. Por ello, se sugiere un valor de  $K$  de 5 o 10, ya que se ha demostrado empíricamente que producen estimaciones de tasa de error de prueba que no sufren ni un sesgo excesivamente alto ni una varianza muy alta (Kuhn & Johnson, 2013; Mechelli & Vieira, 2019)

La validación cruzada dejando uno fuera (leave-one-out) es uno de los métodos más utilizados en la literatura sobre trastornos cerebrales. No es factible usar este método para datos complejos, como datos de neuroimágenes de alta dimensión, datos multimodales etc. por la gran variación y por ser muy costoso (Mechelli & Vieira, 2019).

En la Figura 3, abajo, se describe el proceso de validación cruzada de  $K$ -Iteraciones, con  $k= 10$ . Los datos de entrenamiento se dividieron en 10 partes, y en cada iteración, se usaron 9 partes para entrenar el modelo con diferentes hiperparámetros, y uno será usado como el conjunto de prueba para la evaluación del modelo. Al final de este procedimiento, se utilizan los rendimientos estimados de cada modelo,  $E_i$ , para calcular el rendimiento promedio (Raschka, 2015).



**Figura 3.** Proceso de validación cruzada con k-iteraciones k=10 (Raschka, 2015)

Si el objetivo es seleccionar uno entre diferentes algoritmos de aprendizaje automático se recomienda usar la validación cruzada anidada, consiguiendo una estimación insesgada del error de predicción. (Raschka, 2015)

En la validación cruzada anidada se usa una validación cruzada adicional, llamada validación cruzada interna. Las varias iteraciones de esta validación cruzada interna se utilizan para entrenar y probar diferentes hiperparámetros. El hiperparámetro de mejor rendimiento en la validación cruzada interna se usa para entrenar un modelo en todo el conjunto de entrenamiento definido por la validación cruzada externa. Finalmente, este modelo se usa para predecir la variable objetivo en el conjunto de prueba de la validación cruzada externa. El rendimiento final se estima promediando el rendimiento en el conjunto de pruebas en todas las iteraciones de la validación cruzada externas (Mechelli & Vieira, 2019).

En la Figura 4, se muestra el proceso de validación cruzada anidada, con k = 10 para la validación cruzada externa y para la validación cruzada interna. En la validación cruzada externa, los datos originales se dividieron en 10 partes y para cada iteración, se consideraron 9 partes para el entrenamiento y una para prueba. En la validación cruzada interna, realizada en cada iteración de la validación cruzada externa, el conjunto de entrenamiento también se dividió en 10 partes y para las iteraciones 1 a 10, se usaron 9

partes para entrenar modelos con diferentes hiperparámetros y uno para estimar su rendimiento predictivo. (Raschka, 2015)

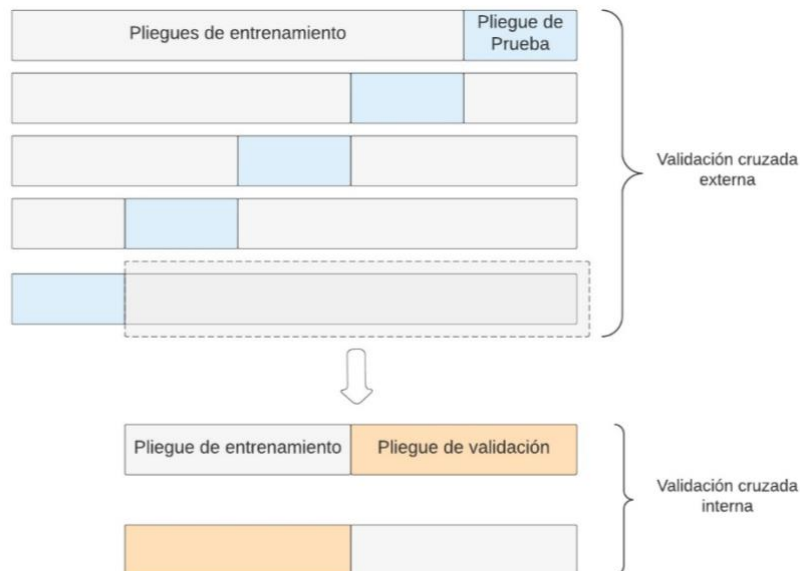


Figura 4. Proceso de validación cruzada anidada, con  $k = 10$  (Raschka, 2015)

Hay algunos casos en los que el conjunto de datos puede estar sesgado hacia un valor particular. Por ejemplo, en un conjunto de datos de una enfermedad rara, puede haber un gran número de predicciones negativas que positivas. En tales escenarios, la validación de  $k$ -iteraciones no funciona bien, y se usa una variación de la validación de  $k$ -iteraciones, donde hay una representación igual de cada tipo de etiqueta de clase, y también se asegura de que cada iteración tenga el mismo porcentaje de representación. Este tipo de validación se denomina validación cruzada estratificada de  $k$ -iteraciones. (Shobha & Rangaswamy, 2018).

Existen varios métodos de sobremuestreo para manejar datos desbalanceados, así tenemos: el algoritmo de técnica de sobremuestreo de minoría sintética (SMOTE) que produce datos artificiales basados en las similitudes de espacio de características entre puntos de datos de minorías existentes, además el algoritmo de muestreo sintético adaptativo (ADASYN) utiliza un método sistemático para crear diferentes cantidades de observaciones sintéticas de forma adaptativa según sus distribuciones. (Kang & Tian, 2018)

En la mayoría de los algoritmos de clasificación cuando una categoría respuesta es mayoritaria dan resultados sesgados hacia la clase mayoritaria, este problema es especialmente común con las bases de datos clínicas, donde la prevalencia de la enfermedad suele ser bastante baja, por lo que es importante en datos médicos abordar adecuadamente este desbalance. (Srinivasan Perumbakkam & Kohl, 2018).

#### 2.3.3.4 Evaluación del modelo

Luego que el algoritmo ha sido entrenado y luego aplicado al conjunto de prueba, lo que sigue es evaluar el desempeño de la clasificación. Para cuantificar existen métricas que permiten evaluar este rendimiento, tanto para la clasificación como para la regresión. A continuación, se revisará las métricas usadas para evaluar el modelo de clasificación.

Un método común para describir el rendimiento de un modelo de clasificación es la matriz de confusión. Esta es una simple tabulación cruzada de las clases observadas y predichas para los datos. La tabla 5 muestra un ejemplo cuando el resultado tiene dos clases. Las celdas diagonales principales (a y d) indican los casos en los que las clases se predicen correctamente, mientras que las celdas fuera de la diagonal (b y c) representan los errores de clasificación (Kuhn & Johnson, 2013; Mechelli & Vieira, 2019)

Tabla 5. *Matriz de confusión para la clasificación con dos categorías de respuesta*

		RESPUESTA OBSERVADA		
		Presente	Ausente	
RESPUESTA PREDICHA	Presente	Verdadero Positivo (VP)	Falso Positivo (FP)	VP+FP
	Ausente	Falso negativo (FN)	Verdadero Negativo (VN)	FN+VN
		VP+FN	FP+VN	

La exactitud (accuracy) es la proporción de predicciones correctas, esto corresponde a la división del número de predicciones correctas sobre el número de



predicciones totales,  $(VP+VN)/(VP+FP+VN+FN)$ . Es una de las métricas más comúnmente usadas, sin embargo, una de sus limitaciones de esta métrica es que no toma en cuenta cuando hay desbalance entre las clases de respuesta. Por lo tanto, una métrica más apropiada es la exactitud balanceada, que se obtiene calculando el promedio entre las precisiones para cada clase (Mechelli & Vieira, 2019).

$$\text{Exactitud Balanceada} = \frac{\frac{VP}{VP+FN} + \frac{VN}{VN+FP}}{2}$$

Cuando el interés es evaluar el modelo predictivo de acuerdo con la clase de respuesta de interés, esto puede ser explicado por la sensibilidad y la especificidad. La sensibilidad llamada tasa de verdaderos positivos o recall es la proporción de positivos correctamente identificados como positivos  $(VP) / (VP+FN)$ . Lo opuesto de la sensibilidad es la tasa de precisión que se calcula:  $(VP) / (VP+FP)$ . Tanto la sensibilidad y la precisión no proporcionan información sobre el número de verdaderos negativos (Shobha & Rangaswamy, 2018; Mechelli & Vieira, 2019). Para identificar a los verdaderos negativos, se usa la especificidad, que es la proporción de negativos que son correctamente identificados como negativos  $(VN) / (VN+FP)$  (Mechelli & Vieira, 2019).

Además, para la clasificación de respuesta binaria, se tiene a Score F (o medida F) es una métrica de exactitud de una prueba. Toma en consideración la precisión y la sensibilidad de la prueba para calcular su puntaje. La medida F es el promedio armónico de la precisión y la sensibilidad. Cuando la medida F alcanza su mejor valor a 1 (se dice que la precisión y la sensibilidad son perfectas) y su peor puntuación es 0 (Shobha & Rangaswamy, 2018).

$$F = \frac{2}{\left(\frac{1}{\text{sensibilidad}}\right) + \left(\frac{1}{\text{precisión}}\right)}$$

Asimismo, existen gráficos para comprenderlos con mayor detalle y para seleccionar modelos para la clasificación en función de su performance con respecto a las tasas de falsos positivos y verdaderos positivos, que se calculan al cambiar el umbral de decisión del clasificador. Así, la curva ROC (receiver operating characteristic) es útil para visualizar la eficacia de los modelos de aprendizaje automático, examina el

equilibrio entre la detección de verdaderos positivos, mientras se evitan los falsos positivos. Las curvas se definen en un gráfico con la proporción de positivos verdaderos en el eje vertical y la proporción de falsos positivos en el eje horizontal. Debido a que estos valores son equivalentes a la sensibilidad y  $(1 - \text{especificidad})$ , respectivamente, el diagrama también se conoce como gráfico de sensibilidad / especificidad. El desempeño de un clasificador en todos los puntos de corte posibles es dado por el área bajo la curva, esta varía de 0.5 (para un clasificador sin valor predictivo), 0.7-0.8 se considera aceptable, 0.8-0.9 se considera excelente, más de 0.9 se considera sobresaliente y 1.0 (para un clasificador perfecto) (Lantz, 2015).

#### 2.3.3.5 Aprendizaje supervisado: Algoritmos de clasificación

##### **La regresión Logística**

La regresión logística es uno de los instrumentos más usados para el análisis de datos en la clínica y en epidemiología. En la regresión logística: la variable dependiente o respuesta es discreta dicotómica ( $Y=0$ ,  $Y=1$ ); las variables predictoras (independientes) pueden ser cuantitativas o cualitativas y la ecuación del modelo no es una función lineal de partida, sino exponencial; si bien, por una sencilla transformación logarítmica (logit), puede finalmente presentarse como una función lineal. La respuesta tiene sólo dos posibles resultados; presencia o ausencia del evento con probabilidades  $\pi$  y  $1-\pi$ , respectivamente. Por ejemplo, interesa estimar la probabilidad de que un paciente presente (o no) una enfermedad, dado su perfil de factores de riesgo (o variables explicativas), el modelo será de utilidad puesto que, muchas veces, el perfil de variables explicativas puede estar formado por características medidas en escala numérica (cuantitativas) o categórica (cualitativas), constituyendo una mayor aproximación a la realidad biológica ya que la influencia de un factor de riesgo sobre la probabilidad de una enfermedad se asemeja a una curva del tipo “sigmoidea”. (Hosmer, Lemeshow, & Sturdivant, 2013)

Para la formulación del modelo de regresión logística, se tiene una variable dicotómica,  $Y$ , que puede tomar valores  $Y=1$  (ocurrió el evento) e  $Y=0$  (no ocurrió el evento). Además la ocurrencia del evento está explicado por un conjunto de variables  $\{X_1, X_2, \dots, X_p\}$ . (Kuhn & Johnson, 2013)

$$E(Y / X_1 = x_1, \dots, X_p = x_p) = \pi = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}$$

Aplicando transformaciones para linealizar la ecuación se obtiene

$$\text{Log}\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

## El clasificador ingenuo de Bayes

Este es un clasificador estadístico basado en el teorema de Bayes eso hace la suposición ingenua de que cada variable contribuye independientemente a la probabilidad de un caso perteneciente a una clase particular. En la práctica, cumplir con esta suposición es raro. Independientemente, el clasificador aún puede lograr buenos resultados debido a su simplicidad y resistencia a ruidos, variables faltantes o irrelevantes. (Consejo, Melcer, & Rozema, 2018)

Un clasificador probabilístico para predecir una de las categorías de Y:  $c_1, c_2, \dots, c_k$  en términos de las variables  $X_1, X_2, \dots, X_p$  es el clasificador de Bayes. Este clasificador, apoyándose en el teorema de Bayes, asigna cada caso determinado por  $x_0 = (x_{01}, x_{02}, \dots, x_{0p})$  a la clase  $c^*$  de Y si la probabilidad condicional:

$P(Y = c^* | X_1 = x_{01}, X_2 = x_{02}, \dots, X_p = x_{0p})$  es máxima.

En general y según el teorema de Bayes:

$$P(Y = c_k | X_1 = x_{01}, \dots, X_p = x_{0p}) = \frac{P(X_1 = x_{01}, \dots, X_p = x_{0p}) \cdot P(Y = c_k)}{\sum_{i=1}^k P(X_1 = x_{01}, \dots, X_p = x_{0p}) \cdot P(Y = c_i)}$$

Sin embargo, se menciona que se presentan dificultades de este cálculo si no se cuenta con las distribuciones conjuntas de X e Y y las condiciones de independencia de las variables predictoras no se relajan.

Si se supone la independencia de las variables predictoras en cada una de las categorías de Y se obtienen las probabilidades condicionales (supuesto de independencia condicional).

Como antes, el caso  $x_0$  se asigna a la categoría que le corresponde mayor probabilidad condicional. Así, se obtiene el clasificador ingenuo de Bayes. Notándose que, como los denominadores de las relaciones que determinan las probabilidades condicionales son iguales, serán suficientes los numeradores para comparar las probabilidades. (Veliz Capuñay, 2018)

Para los predictores categóricos, la distribución de probabilidad se puede determinar con las frecuencias observadas en los datos del conjunto de entrenamiento. Sin embargo, utilizando el supuesto de independencia, es probable que esta probabilidad se sobreestime. La probabilidad previa le permite al modelador inclinar la probabilidad final hacia una o más clases. Por ejemplo, cuando se modela un evento raro, es común muestrear selectivamente los datos para que la distribución de la clase en el conjunto de entrenamiento sea más equilibrada. Sin embargo, el modelador puede desear especificar que el evento es realmente raro asignándole una probabilidad previa baja. Si no se da explícitamente ningún previo, la convención es usar las proporciones observadas del conjunto de entrenamiento para estimar el previo. A pesar de todo, el clasificador de Bayes muestra ventajas, como el hecho de ser fácil y rápido, incluso para grandes conjuntos de entrenamiento. Por ejemplo, cuando los predictores son todos categóricos, todo lo que se requiere son tablas de búsqueda simples con las distribuciones de frecuencia del conjunto de entrenamiento. Además, la regla de Bayes, funciona de manera competitiva en muchos casos. Esta es esencialmente una declaración de probabilidad. Se crean las probabilidades de clase y la clase predicha es la asociada con la clase más grande de probabilidad (Kuhn & Johnson, 2013)

### **Técnicas de clasificación basado en ensamble**

La combinación de clasificadores múltiples, comúnmente denominado como clasificador de ensamble, han demostrado la capacidad de mejorar la precisión de la clasificación en muchos dominios de aplicaciones. Como resultado, esta área ha atraído una gran cantidad de investigaciones en los últimos años.

Un clasificador de ensamble es un grupo de clasificadores cuyas decisiones individuales se fusionan de alguna manera para proporcionar, como resultado, una decisión de consenso. El objetivo general de este enfoque es combinar salidas de varios

modelos, también conocidos como clasificadores base, para generar un único resultado agregada que supera a cualquiera de los clasificadores básicos de forma aislada. La primera etapa en el proceso de generar un conjunto clasificador es la generación de una colección de clasificadores básicos. Un enfoque es aplicar N diferentes métodos de aprendizaje, con un solo conjunto de datos de entrenamiento, para obtener N diferentes modelos de clasificación. Un enfoque alternativo es crear N porciones diferentes de datos del conjunto de datos de entrenamiento y emplea un único algoritmo de aprendizaje con cada parte. Durante la fase de aprendizaje, es importante adoptar un enfoque que permita la creación de diversos clasificadores. También se ha demostrado que la selección adecuada de clasificadores básicos, en lugar de simplemente combinarlos a todos en un solo ensamble, puede afectar la clasificación general final. La selección del clasificador base se puede realizar de una manera estática o dinámica. En el enfoque estático, el mismo subconjunto de base de los clasificadores que se seleccionan se aplican para todas las muestras de prueba. En el enfoque dinámico, la selección es realizado para cada nueva instancia individualmente. Después de obtener una colección de clasificadores básicos, el siguiente paso es combinar sus resultados para obtener la decisión final. Durante esta fase, los principales problemas a considerar están relacionados con los tipos de información se combinarán y qué método de combinación se va a aplicar (Jurek, BI, Wu, & Nugent, 2013)

Los métodos de ensamble también son conocidos como métodos agregados, estos se forman agregando modelos que apenas superan en precisión a la regla ingenua. Los modelos que se agregan cubren diferentes partes de los datos de entrada y al complementarse entre ellos reducen las deficiencias que los modelos básicos puedan presentar. Estos métodos se aplican a diversos clasificadores, en particular, a los árboles de decisión. El clasificador que se obtiene como agregación de los clasificadores básicos tiene la forma:

$$F_C(X) = \sum_{m=1}^M \alpha_m \hat{f}_m(X),$$

Donde  $\hat{f}_m$  representa un estimador básico.

Entre los procedimientos para obtener conjuntos de clasificadores, tenemos: método bagging, el método de random forest o bosques aleatorios y el método de adaboost. (Veliz Capuñay, 2018)

### **Random Forest o método de bosque aleatorio**

Los bosques aleatorios son una combinación de predictores de árboles, de modo que cada árbol depende de los valores de un vector aleatorio muestreado de forma independiente y con la misma distribución para todos los árboles en el bosque. El error de generalización para los bosques converge hasta un límite a medida que aumenta el número de árboles en el bosque. La generalización del error de un bosque de clasificadores de árboles depende de la fuerza de los árboles individuales en el bosque y la correlación entre ellos. El uso de una selección aleatoria de características para dividir cada nodo produce tasas de error que se comparan favorablemente con Adaboost. Las estimaciones internas supervisan el error, la fuerza y la correlación, y se utilizan para mostrar la respuesta al aumento del número de características utilizadas en la división. Las estimaciones internas también se utilizan para medir la importancia variable. Estas ideas también son aplicables a la regresión. Por lo que, estos clasificadores son una herramienta efectiva en la predicción. Agregando la aleatoriedad los convierte en clasificadores y regresores precisos. Además, la forma como trabajan estos clasificadores en cuanto a su fuerza de los predictores individuales y sus correlaciones dan la idea de la capacidad de predicción del bosque aleatorio. Los bosques dan resultados competitivos con boosting y bagging adaptativo, pero no cambian progresivamente el conjunto de entrenamiento. Su precisión indica que actúan para reducir el sesgo (Breiman, 2001).

Los bosques aleatorios son una combinación de técnicas asociada con los árboles de clasificación y bagging. Omitiendo un detalle, un clasificador de bosque aleatorio es simplemente árboles en bolsas: en  $L$  muestras bootstrap de los datos, los árboles son entrenados (crecen para estar en forma, sin poda) y luego los árboles botan a la clase que corresponden. El detalle omitido es que, en cada paso del crecimiento del árbol, una decisión de entrenamiento aleatorio es hecha: una pequeña muestra aleatoria de las  $m$  características (generalmente aproximadamente  $\sqrt{m}$  de ellos) se consideran para la división en lugar de todos  $m$ . El propósito de esto es reducir la correlación entre pares de árboles en el bosque (Knox, 2018).

Knox (2018), indica que los bosques aleatorios son demostrablemente resistentes al sobreajuste en el sentido de que aumentar el número de árboles  $L$  sin límite no hace que el bosque aleatorio se sobreajuste, a pesar de que el número de parámetros necesarios para describir el bosque va a infinito. Sea  $P_{(X, Y)}$  la distribución de probabilidad de la que se extraen los datos: se supone que los datos de entrenamiento y todos los datos futuros se extraen de esta distribución. Sea  $(X, Y)$  denotar un dato aleatorio extraído de  $P_{(X, Y)}$ . Se puede demostrar que bajo 0-1 pierde el riesgo de un bosque aleatorio con árboles  $T_1 \dots, T_L$  es:

$$R(T_1, \dots, T_L) = P_{(X, Y)} (\text{lárboles votan por la clase } Y) \\ < \max_{c \neq Y} |\text{árboles votan por la clase } c|$$

Según Véliz (2018), al aplicar la metodología Bagging, existe la posibilidad de que los árboles que se construyan sean similares sobre todo si existe un predictor que, por su gran importancia, es elegido al iniciar la construcción de cada árbol. Si esto sucede, los árboles que se construyen serán similares con predicciones altamente correlacionadas y sin la reducción de la variabilidad en los resultados finales. Para evitar esta dificultad, se introdujo la metodología llamada bosque aleatorio, que consiste en seleccionar de manera aleatoria  $m$  de las variables predictoras en cada paso del método bagging. De esta manera, al usar solo una parte de los datos y una parte de las variables, se entrega, en cada etapa, una visión parcial del problema reduciendo la correlación de las predicciones. Si se dispone de  $p$  variables predictoras, el número  $m$  de variables a elegir es  $p/3$  para el caso de los árboles de regresión y  $\sqrt{p}$  para la clasificación. El clasificador final que se obtiene es el que asigna la categoría con mayor votación en el caso de la clasificación; y el promedio de las predicciones, en el caso de la regresión.

### **Máquina de soporte vectorial**

La máquina de soporte vectorial es una técnica de aprendizaje automático para resolver la discriminación y problemas de regresión, en sus comienzos se propuso para datos binarios. Hoy en día, se utiliza en varias áreas de investigación, como el reconocimiento facial, reconocimiento de voz, diagnóstico médico y para datos de clasificación del clima. De hecho, la máquina de soporte vectorial consiste en construir

el hiperplano marginal máximo que separa de manera óptima dos clases de un conjunto de datos  $D = \{(x, y), \dots, (x_m, y_m)\}$  donde  $x_i \in \mathbb{R}^n$   $y_i \in \{-1, 1\}$  son la etiqueta de clase de  $x_i$  (Figura 5). Sin embargo, muchos hiperplanos pueden separar las dos clases. Por lo tanto, SVM usa una fase de entrenamiento para encontrar el hiperplano óptimo llamado hiperplano óptimo separado, la cual cumple  $f(x) = w^T x - b$ , donde  $w$  es un vector de  $m$  dimensiones y  $b$  es un término de "sesgo" que separa dos clases y maximiza el margen entre las dos clases. Por lo tanto, siguiendo la figura 5, el hiperplano óptimo  $H$  verifica que los hiperplanos  $H1$  y  $H2$  son paralelos a través de él y pasan a través de los puntos más cercanos a  $H$  que se denominan vectores de soporte (SVs). Por lo tanto, los SVM eligen el hiperplano óptimo que maximiza el margen entre las dos clases, que es la distancia entre  $H1$  y  $H2$ . El hiperplano con el margen más grande está definido por  $\frac{2}{\|w\|}$

(Yahyaoui & Yahyaoui, 2018)

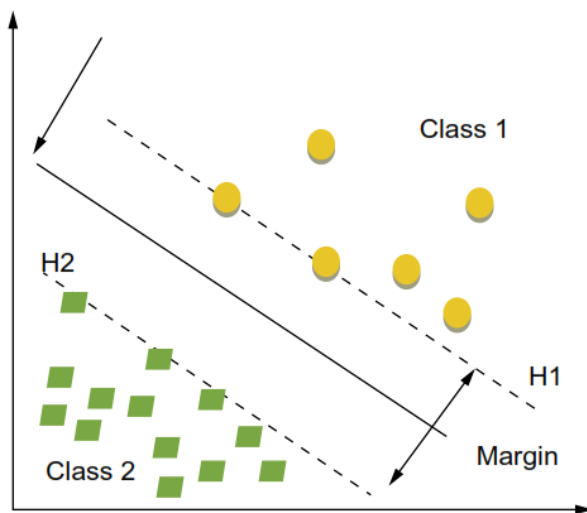


Figura 5. Idea básica de máquina de soporte vectorial para el caso lineal. (Yahyaoui & Yahyaoui, 2018)



De acuerdo con Véliz (2018), las máquinas de vectores de soporte (SVM), se presentan como modelos que tienen un marco matemático bien fundamentado, que generalizan satisfactoriamente, que utilizan fronteras lineales para separar las clases y sobre todo, que hacen uso tan solo de una parte de los datos, evitando el sobreajuste. Las máquinas de vectores de soporte desde su creación se utilizan en un gran número y variedad de aplicaciones como, por ejemplo, en el reconocimiento de señales y en la clasificación de textos. Estos modelos fueron introducidos con la finalidad de resolver problemas de clasificación binaria, sin embargo, en la actualidad se han extendido a problemas de clasificación para más de dos clases y para la regresión, así como para la formación de agrupamiento. Para la clasificación binaria, se consideran los siguientes casos:

Caso 1: existe un hiperplano que separa completamente a los dos conjuntos que determinan las clases (las clases son perfectamente separables linealmente)

Caso 2: existe un solapamiento entre las clases (estas clases se llama de margen blando)

Caso 3: los conjuntos no son separables linealmente; sin embargo, es posible establecer transformaciones de los espacios originales a espacios de mayor dimensión en donde será posible encontrar hiperplanos que separan satisfactoriamente a las clases. (Veliz Capuñay, 2018)

Los datos en tiempo real sobre los que se aplican los algoritmos de aprendizaje automático tienen ruido. Es prácticamente imposible separar los datos con un hiperplano. Una forma de manejar esto es relajar la restricción de tener el margen máximo para la línea que separa las clases. Esta forma de relajar la restricción se conoce como clasificador de margen suave. La relajación permite que algún punto viole la línea de separación y esté del otro lado. Las variables flojas son algunos coeficientes adicionales que se introducen para hacer espacio en cada dimensión. La complejidad del modelo aumenta cuando se utilizan más parámetros para ajustar el modelo a los datos proporcionados. El objetivo de un clasificador de margen suave es maximizar el margen, al tiempo que minimiza la suma de holguras, es decir, reducir el número de valores atípicos en cada grupo. La figura 6, muestra un espacio vectorial de un clasificador de margen suave (Shobha & Rangaswamy, 2018).

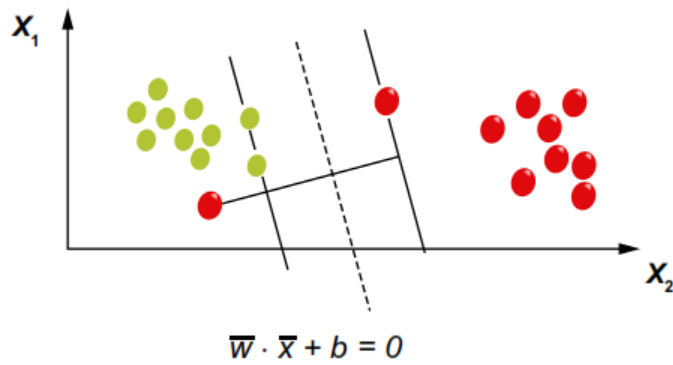


Figura 6. Clasificador de margen suave (Shobha & Rangaswamy, 2018).

En realidad, existen muchos conjuntos de datos no se pueden separar linealmente. La linealidad de una máquina de soporte vectorial puede extenderse a una máquina de soporte vectorial no lineal mediante la aplicación de funciones de kernel conocidos o trucos de kernel. En la Figura 7, se muestra un caso unidimensional. El conjunto de datos no se puede separar linealmente. Este problema se resuelve mapeando los datos en un espacio dimensional superior y luego usando un clasificador lineal en el recién proyectado espacio dimensional superior. La figura 7 muestra la transformación de datos de entrada sin procesar en un espacio de características de alta dimensión a través de una función de "similitud" con la matriz del kernel correspondiente. Esta forma de proyectar datos en un espacio dimensional superior se denomina "truco kernel". El truco kernel hace explotar las matemáticas para proyectar los datos inseparables linealmente en datos separables linealmente (Mechelli & Vieira, 2019).

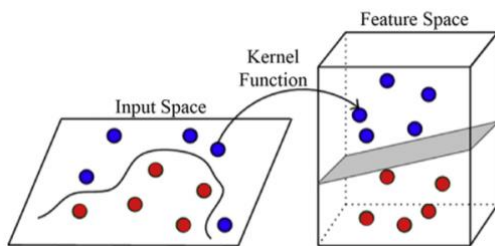


Figura 7. Representación del truco de kernel (Mechelli & Vieira, 2019).

## Redes neuronales

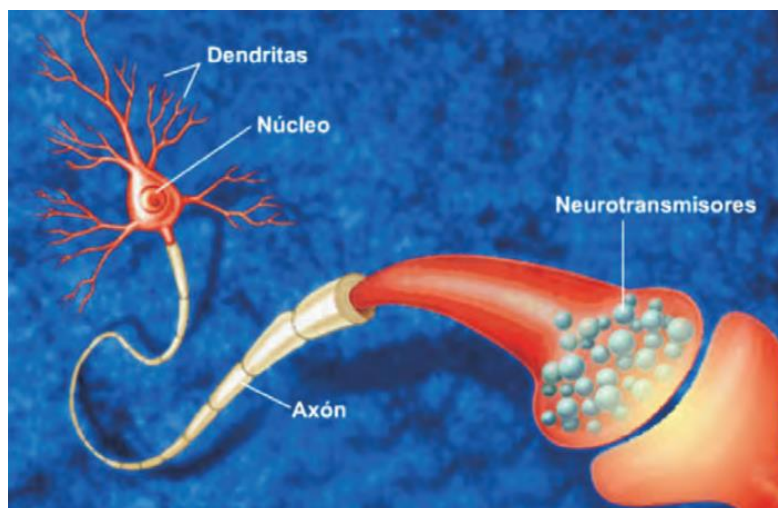
Las redes neuronales artificiales son modelos computacionales inspirados en el sistema nervioso de los seres vivos, sobretodo de la corteza cerebral. Tienen la capacidad de adquirir y mantener conocimiento (basado en información) y pueden definirse como un conjunto de unidades de procesamiento, representadas por neuronas artificiales, interconectadas por muchas interconexiones (sinapsis artificiales), implementadas por vectores y matrices de pesos sinápticos (Sidey-Gibbons & Sidey-Gibbon, 2019; Nunes da Silva, Hernane, Andreda, Bartocci, & Franco, 2017)).

Las redes neuronales no permiten la explicación de los resultados debido a que emplean un sistema matemático complejo y, en términos de las variables predictoras, estos modelos son, considerados como “cajas negras”, las redes neuronales aprenden por sí mismas qué variables de entrada son las más importante. Por lo que es imposible saber cómo las variables contribuyen con la salida (Veliz Capuñay, 2018; Kubben, Dumontier, & Dekker, 2019).

Asimismo, las redes neuronales artificiales pueden emplearse en varios problemas relacionados con la ingeniería y las ciencias. Así en medicina las redes neuronales se usan para clasificar y predecir el cáncer, también constituye en un sistema de apoyo a la decisión en el diagnóstico de enfermedades como las cardíacas (Nunes da Silva, Hernane, Andreda, Bartocci, & Franco, 2017).

La célula fundamental del sistema nervioso central es la neurona, y su función se reduce a conducir impulsos (estímulos eléctricos originados por reacciones físico-químicas) bajo ciertas condiciones de operación. Este componente biológico se puede dividir en tres partes principales: dendritas, cuerpo celular (también conocido como "soma") y axón. Las dendritas están compuestas por varias extensiones delgadas que forman el árbol dendrítico). El propósito fundamental de las dendritas es adquirir, continuamente, estímulos de varias otras neuronas (conectores) o del ambiente externo, como es el caso de algunas neuronas en contacto con el ambiente (también llamadas neuronas sensoriales). El cuerpo celular es responsable de procesar toda la información que proviene de las dendritas, para producir un potencial de activación que indica si la neurona puede desencadenar un impulso eléctrico a lo largo de su axón. También es en el cuerpo celular donde se encuentran los principales orgánulos citoplasmáticos (núcleo, mitocondrias, centriolos, lisosomas, etc.) de Se puede encontrar la neurona. El axón está

compuesto por una única extensión cuya misión es guiar los impulsos eléctricos a otras neuronas conectadas, o a neuronas conectadas directamente al tejido muscular (neuronas eferentes). La terminación del axón también se compone de ramas llamadas terminales sinápticas. El procesamiento de la información realizado por el cerebro humano se lleva a cabo mediante componentes de procesamiento biológico, que funcionan en paralelo, para producir funciones adecuadas, como el pensamiento y el aprendizaje (Figura 8). Las sinapsis son las conexiones que permiten la transferencia de los impulsos axónicos eléctricos de una neurona particular a las dendritas de otras neuronas, como se ilustra en la figura 9. Es importante tener en cuenta que no hay contacto físico entre las neuronas que forman la unión sináptica, por lo que los elementos neurotransmisores liberados en la unión se encargan de ponderar la transmisión de una neurona a otra. La funcionalidad de una neurona depende de su ponderación sináptica, que también es dinámica y depende de la química (Nunes da Silva, Hernane, Andreda, Bartocci, & Franco, 2017)



*Figura 8.* Neurona biológica  
(Caicedo & López, 2017)



*Figura 9.* La sinapsis de una neurona  
(Caicedo & López, 2017)

De manera semejante, los componentes computacionales o unidades de procesamiento, llamadas neuronas artificiales, son modelos simplificados de neuronas biológicas. Las estructuras de redes neuronales artificiales se desarrollaron a partir de modelos conocidos de sistemas nerviosos biológicos y del propio cerebro humano. Estos modelos se inspiraron en el análisis de cómo una membrana celular de una neurona genera y propaga impulsos eléctricos. En ese modelo, cada neurona de una red se puede implementar como se muestra en la figura 10. Las múltiples señales de entrada que provienen del entorno externo (aplicación) están representadas por el conjunto  $\{x_1, x_2, x_3, \dots, x_n\}$ , análogo a los impulsos eléctricos externos reunidos por las dendritas en la neurona biológica. El pesaje realizado por las uniones sinápticas de la red se implementa en la neurona artificial como un conjunto de pesos sinápticos  $\{w\}$ . Análogamente, la relevancia de cada una de las  $\{x\}$  entradas de neuronas se calcula multiplicándose por su correspondiente peso sináptico  $\{w_i\}$ , ponderando así toda la información externa que llega a la neurona. Por lo tanto, es posible verificar que la salida del cuerpo celular artificial, denotado por  $u$ , es la suma ponderada de sus entradas. (Nunes da Silva, Hernane, Andreda, Bartocci, & Franco, 2017)

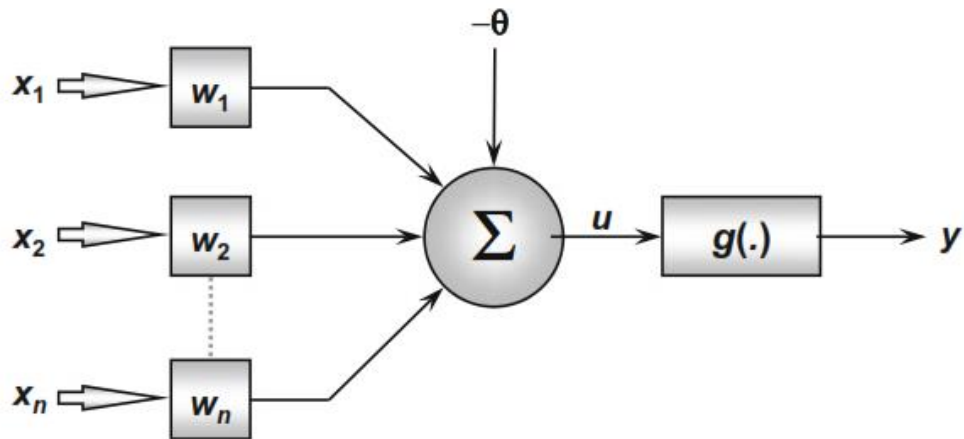


Figura 10. Implementación de una red neuronal  
(Nunes da Silva, Hernane, Andreda, Bartocci, & Franco, 2017)

Véliz (2018), menciona que el caso más simple de red neuronal artificial lo constituye el perceptrón, la red que tiene solamente la capa de entrada y la capa de salida con un único nodo o neurona  $j$  al que llegan  $m$  señales que corresponden a  $m$  variables predictoras  $X_1, X_2, \dots, X_m$ ,

La información que llega a la neurona de salida a través de las neuronas de entrada se resume usando ponderaciones llamadas pesos sinápticos, mediante la función llamada potencial.

Si para las variables predictoras y para la variable dependiente  $Y$ , los datos que corresponden a  $n$  instancias o ejemplos son:

$$(x_{11}, x_{12}, \dots, x_{1m}, y_1), (x_{21}, x_{22}, \dots, x_{2m}, y_2), \dots, (x_{n1}, x_{n2}, \dots, x_{nm}, y_n)$$

La función potencial en el nodo  $s$  de salida y para cada instancia  $I$  se define como:

$$n = \sum_{k=1}^m w_k x_{ik}$$

En donde  $w_k$  es el peso que le corresponde al valor  $x_{ik}$  de la variable  $X_k$ .

Habitualmente, a la suma se agrega un parámetro  $\theta$ , y se obtiene:

$$n = \sum_{k=1}^m (w_k x_{ik}) + \theta$$

Al valor de  $\theta$  se le llama sesgo o umbral.

Si se acuerda escribir  $x_{i0} = 1$  y  $w_0 = \theta$ , la función potencial podrá escribirse como:

$$n = \sum_{k=0}^m w_k \cdot x_{ik}$$

Para cada ejemplo o instancia  $i$ , la salida  $\hat{y}_i$  de la neurona se determina aplicando al potencial la función  $f$  llamada de activación o de transferencia.

$$\hat{y}_i = f(n) = f\left(\sum_{k=0}^m w_k \cdot x_{ik}\right)$$

La función de activación simula el estado cerebral en donde las señales simultáneas que llegan a la neurona a través de las dendritas es activada o inhibida por ciertos procesos bioquímicos, determinando de esta manera la transmisión o no de la señal de salidas a través del axón.

Si en el perceptrón no se considera el umbral y la función de activación es la función identidad  $f(n)=n$ , se tiene el modelo asociador lineal. Si en el perceptrón se considera el umbral y la función de activación es la identidad  $f(n)=n$ , se tiene la red neuronal llamada Adeline. Es el caso de la regresión lineal. Otras funciones que también se usan son: el regulador lineal, la tangente hiperbólica, la función gaussiana (Veliz Capuñay, 2018)

Por otro lado, la arquitectura de una red neuronal artificial se entiende como la forma como se organizan las neuronas en su interior y está estrechamente ligada al algoritmo de aprendizaje usado para entrenar la red. Dependiendo del número de capas, se definen las redes como monocapa y multicapa; y si se toma como elemento de clasificación la forma como fluye la información, se define las redes como Feedforward y Recurrentes. Así, en las redes feedforward, la información fluye en un único sentido

desde las neuronas de entrada a la capa o capas de procesamiento, para los casos de redes monocapa y multicapa, respectivamente; hasta llegar a la capa de salida de la red neuronal. La información fluye desde la neurona *i-ésima* de entrada hacia la *j-ésima* de salida (Caicedo & López, 2017) Figura N°11.

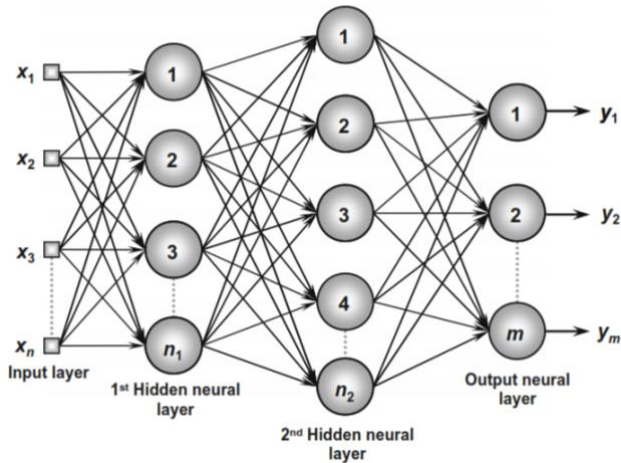


Figura 11. Ejemplo de una red feedward con múltiples capas (Nunes da Silva, Hernane, Andreda, Bartocci, & Franco, 2017)

En cambio, en las redes recurrentes, la información no siempre fluye en un sentido, puesto que puede realimentar hacia capas anteriores a través de conexiones sinápticas. Este tipo de red neuronal puede ser monocapa o multicapa (Figura 12).

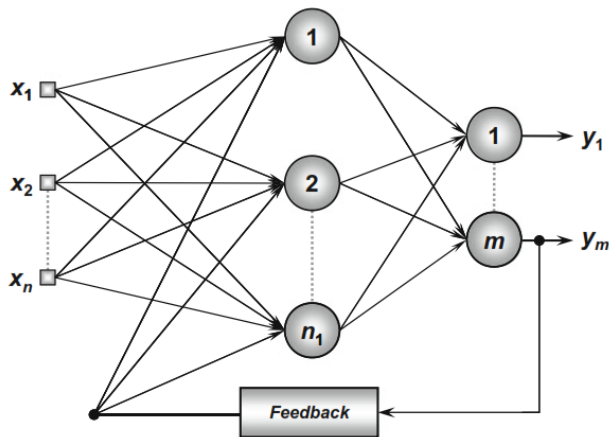


Figura 12. Ejemplo de una red recurrente (Nunes da Silva, Hernane, Andreda, Bartocci, & Franco, 2017)



### 2.3.4 Modelos predictivos en datos de la salud

El aprendizaje automático mejora la capacidad de los profesionales de la salud para establecer un pronóstico. En cuanto a la predicción de resultados binarios, se pueden distinguir dos tipos de herramientas de predicción, primero una herramienta que se puede utilizar para predecir la probabilidad en un individuo de la presencia de enfermedad en el momento de la predicción (es decir, un modelo de predicción de diagnóstico) y otra herramienta que se puede usar para predecir la probabilidad de la ocurrencia futura de un evento (es decir, un modelo de predicción pronóstica). Aunque la aplicación puede diferir sustancialmente, los métodos que se emplean para desarrollar dichos modelos son similares. (Kubben, Dumontier, & Dekker, 2019). En esta tesis, el aprendizaje automático la predicción fue desarrollada para realizar un modelo de predicción pronóstica.

En el desarrollo de herramientas para la predicción individual, es importante resaltar que no es el interés descifrar asociaciones causales entre predictores y el resultado. El interés es seleccionar el mejor conjunto de predictores e incluirlos en el modelo de tal manera que las predicciones del modelo sean lo más precisas posible. Fenómenos epidemiológicos como la confusión (es decir, el sesgo es introducido en la estimación de coeficientes debido a una variable asociada tanto con el predictor y el resultado, pero no está controlado) y la mediación (es decir, la presencia de una variable intermedia que explica la asociación entre el predictor y el resultado) estas no son relevantes en el contexto del modelo de predicción. Términos de interacción, que son variables que moderan la asociación entre un predictor y el resultado puede ser útil para aumentar el rendimiento predictivo de un modelo si las asociaciones entre las variables predictoras y el resultado difieren entre los subgrupos, pero no se utilizan para ayudar a la interpretación causal (Kubben, Dumontier, & Dekker, 2019). De esta manera, Mechelli & Vieira, (2019), mencionan que las variables confusoras pueden afectar negativamente el desarrollo del modelo de aprendizaje automático por lo que recomiendan que durante la preparación de los datos es necesario verificar la presencia de cualquier posible fuente de sesgo en el conjunto de datos utilizando pruebas estadísticas (como, por ejemplo, test de homogeneidad) y el balanceo de los datos.

En ese sentido, Alahmari (2020), investigó el desbalance de clases para una aplicación médica relacionada con el cribado del trastorno del espectro autista para identificar el método ideal de remuestreo de datos que puede estabilizar el rendimiento

de la clasificación. Utilizando diferentes técnicas de sobremuestreo y submuestreo en un conjunto de datos reales desbalanceados: Smote (Synthetic Minority Oversampling Technique) y el submuestro aleatorio. Los resultados producidos por múltiples clasificadores en los conjuntos de datos considerados mostraron superioridad en términos de especificidad, sensibilidad y precisión, entre otros, al adoptar técnicas de sobremuestreo en la fase de preprocesamiento.

En efecto, los modelos de aprendizaje automático se están explorando como herramientas para la predicción del resultado neuroquirúrgico con un amplio rango de campos que abarcan epilepsia, tumor cerebral, columna vertebral, sistema neurovascular, enfermedad de Parkinson, lesión traumática cerebral y pacientes con hidrocefalia. Además, algunos modelos de aprendizaje automático incluso han demostrado que superan los índices pronósticos a los clásicos modelos estadísticos, de desempeño similar o mejor que los expertos clínicos bajo ciertas condiciones (Senders, y otros, 2017). Asimismo, la aplicación del aprendizaje automático a los datos neurobiológicos podría ayudar a desarrollar herramientas para predecir la progresión de la enfermedad o la respuesta al tratamiento a nivel individual. Esto sería muy importante para los médicos, ya que podría ayudar a tomar decisiones clínicas más efectivas en las primeras etapas del trastorno, lo que daría como resultado menos ensayos ineficaces y tasas de remisión más altas. Además, tales herramientas también tendrían implicaciones considerables para el costo económico de la atención médica, por ejemplo, al reducir la necesidad de hospitalización (Mechelli & Vieira, 2019).

Además, el conjunto de datos pequeños no representa una barrera para realizar estudios de aprendizaje automático para fines de predicción en la medicina, según: Panesar, Souza, Yeh, & Fernández-Miranda, 2019, Asadi, y otros, 2016; Miranda Hernández, 2016.

Además, Elshawi, Al-Mallah, & Sakr (2019) en su estudio de interpretabilidad de aprendizaje automático basado en modelos para predecir la hipertensión arterial, desarrollaron técnicas de la interpretabilidad para ayudar al personal clínico a comprender mejor y confiar más en los resultados del aprendizaje automático para las predicciones. El trabajar con datos complejos y heterogéneos de pacientes, además de centrarse en el rendimiento de la métrica de los algoritmos de aprendizaje automático hacen que sean descritos como modelos de caja negra, como es el caso de las redes

neuronales, máquina de soporte vectorial y bosques aleatorios. Así, Du, Liu, & Hu (2020) describen técnicas para interpretar el comportamiento del modelo de aprendizaje automático.

### 2.3.5 Ética en la aplicación de aprendizaje automático en datos de salud

El alcance del poder del análisis predictivo y el aprendizaje automático en el área de la salud es un desafío para los marcos éticos actuales y abre una gama de dilemas éticos complejos y sin respuesta. Algunos de estos desafíos éticos son discernibles y requieren una atención más cuidadosa. Este es el caso por ejemplo sobre las implicaciones éticas de los prejuicios humanos que existen en los datos que luego se utilizan para entrenar en el aprendizaje automático para tomar decisiones clínicas a escala. En otros dominios, los desafíos éticos son más difíciles y requieren la implicación de un nuevo conjunto de consultas en el ámbito de la bioética. Hoy en día, los dispositivos son capaces de cosechar casi cada dato de nuestro comportamiento humano. Por lo que, estos puntos de datos podrían utilizarse para crear una huella digital de cada individuo y luego ser usado para predecir resultados de su comportamiento. Pero más aún, podría usarse para "empujar" incesantemente al individuo hacia comportamientos prescritos. Esto plantea varias preguntas éticas que superan el ámbito actual de juramento Hipocrático y la ubicación del análisis predictivo como un "tercer participante" en la relación entre un paciente y los profesionales de la salud. Ciertos defensores están llamando a investigadores e instituciones a adoptar un nuevo código de ética para la inteligencia artificial (Saber, Somai, Rajah, Scalzo, & Liebeskind, 2019).

Asimismo, la fijación de datos y la digitalización de la medicina, que también tiene lugar en la neurología y psiquiatría, requiere el almacenamiento y procesamiento de grandes cantidades de datos, que a su vez pueden estar sujetos a vulnerabilidades inherentes con respecto a la seguridad de los datos y la privacidad de las personas y grupos sociales particulares. Además, la creciente capacidad de los métodos avanzados y adaptativos de aprendizaje automático para inferir estados mentales a partir de datos de neuroimagen puede crear tensiones éticas y desafíos legales con respecto a la privacidad mental y posiblemente requerir una extensión de los instrumentos legales y derechos existentes y derechos. Se debe considerar, la transparencia e interpretabilidad esperada de los métodos de aprendizaje automático debe ser proporcional a la relación entre riesgos y beneficios. Paralelamente a esto, el problema omnipresente de

estructuras de datos sesgadas y ontologías que se usan para entrenar algoritmos de aprendizaje automático debería ser más ampliamente reconocido y se debería promover la investigación interdisciplinaria sobre estos. Estas y muchas otras tensiones éticas apuntan al hecho de que la investigación proactiva e interdisciplinaria, así como el discurso y la deliberación de múltiples partes interesadas son necesarios para garantizar el desarrollo y la implementación efectiva y responsable de estas tecnologías emergentes (Mechelli & Vieira, 2019)

## 2.4 Definición de términos básicos

### a. Relacionados con las malformaciones arteriovenosas y la radiocirugía:

- *Fotones de alta energía.*: Es una radiación de alta energía generado por rayos X (acelerador lineal) o una fuente radiactiva de cobalto (Linac) (Badakhshi, 2016)
- *Grays*: Es la unidad de energía de radiación que emite gamma Knife o Linac (Badakhshi, 2016).
- *Hemorragia cerebral*: Una acumulación de sangre, secundario a una rotura vascular, se produce en el interior del cerebro.
- *Malformación arteriovenosa cerebral (MAV)*: Son anomalías vasculares congénitas compuestas por un número anormal de vasos sanguíneos anormalmente construidos donde no existe una red capilar intermedia para amortiguar la presión. Tanto la construcción anormal de los vasos sanguíneos como el flujo sanguíneo anormal conducen a un riesgo de ruptura y hemorragia intracraneal, dolor de cabeza o convulsiones. (Badakhshi, 2016).
- *Radiocirugía*: Es una técnica quirúrgica mínimamente invasiva que consiste en la aplicación de fotones de alta energía generados por un acelerador lineal o un gamma knife, sobre un blanco intracraneal produciendo la necrosis de ese blanco mientras que el tejido circundante no es afectado Ese blanco puede ser malformación arteriovenosa o un tumor (Badakhshi, 2016)
- *Obliteración*: Es la exclusión de la malformación arteriovenosa cerebral (MAV) mediante la observación de un examen angiográfico cerebral durante el periodo de seguimiento después de la radiocirugía. También se le llama curación o cierre de la malformación arteriovenosa cerebral

b. Relacionados a los algoritmos de aprendizaje automático:

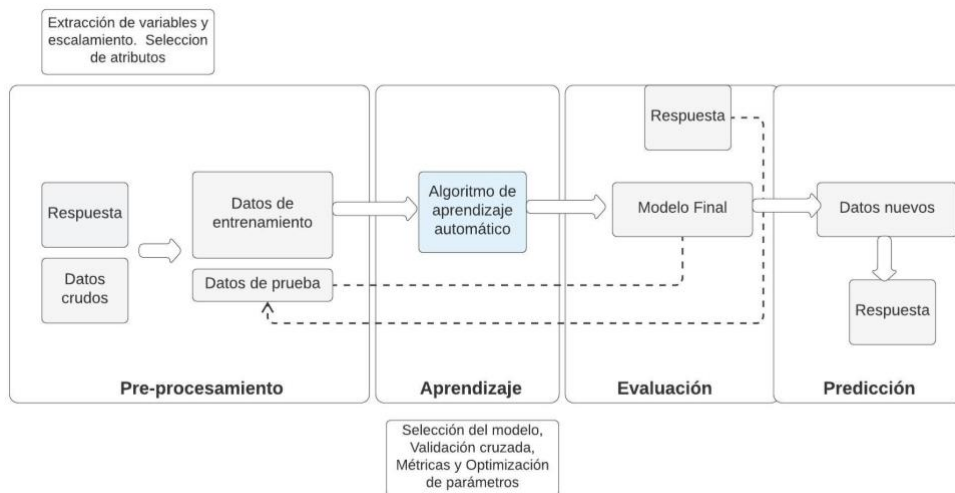
- *AUC de la Curva ROC*: Es el área bajo la curva que cuanto más alejada este de la diagonal principal mejor es el método de diagnóstico.
- *Aprendizaje automático*: Es una rama de la inteligencia artificial cuyo fin es dotar a una máquina, a través de algoritmos, la capacidad de entrenar y aprender en base a datos, sin ser explícitamente programada, imitando la capacidad que tienen las personas de aprender mediante ejemplos sin recurrir a fórmulas ni reglas entre las variables y posibilitando al término del entrenamiento un modelo que permite la generalización; es decir, la obtención de resultados en nuevas situaciones no conocidas durante el aprendizaje. (Veliz Capuñay, 2018)
- *Minería de datos*: Proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos. Siendo su tarea fundamental de encontrar modelos inteligibles a partir de los datos (Hernandez, Ramirez Quintana, & Ferri Ramirez, 2005)
- *Modelación estadística*: Consiste en explicar el comportamiento de una variable a partir del conocimiento de otras (Hernandez, Ramirez Quintana, & Ferri Ramirez, 2005).
- *Naive Bayes*: Modelo simple de clasificación, su estructura de red es fija y solo necesita aprender de los parámetros (probabilidades) (Hernandez, Ramirez Quintana, & Ferri Ramirez, 2005)
- *Random forest*: También conocido como bosque aleatorio, es un algoritmo de aprendizaje de máquina supervisado, combinación de métodos de clasificación basado en teoría de aprendizaje estadístico. Es una técnica mejorada de Bagging que mejora la precisión en la clasificación mediante la incorporación de aleatoriedad en la construcción de cada clasificador individual.
- *Redes neuronales*: es una técnica que se comporta de forma parecida al cerebro humano aprendiendo de la experiencia y el pasado y aplicando tal conocimiento a la resolución de problemas nuevos. Están formados en una serie de niveles o capas (al menos tres: entrada, procesamiento u oculta y salida) compuestas por nodos o “neuronas”. Siendo muy usados en distintos sectores: industria, ejército,

comunicaciones, medicina, robótica, marketing, banca y finanzas, etc. (García, y otros, 2018)

- *Accuracy (Exactitud)*: es la proporción de predicciones correctas, esto corresponde a la división del número de predicciones correctas sobre el número de predicciones totales,  $(VP+VN) / (VP+FP+VN+FN)$ . Es una de las métricas más usadas.
- *Accuracy balanceada*, se obtiene calculando el promedio entre las precisiones para cada clase. Esta métrica toma en cuenta cuando hay desbalance entre las clases de respuesta.
- *Sensibilidad (recall)*: llamada tasa de verdaderos positivos es la proporción de positivos correctamente identificados como positivos  $(VP) / (VP+FN)$ .
- *Especificidad*: es la tasa de verdaderos negativos es la proporción de negativos que son correctamente identificados como negativos  $(VN) / (VN+FP)$  (Mechelli & Vieira, 2019).
- *Features*: Son las variables de entrada para un modelo de aprendizaje automático. Estas pueden ser procesadas antes que ellas entren al modelo.

## 2.5 Fundamentos teóricos que sustentan las hipótesis

En la figura 13 se muestra la ruta de trabajo empleada para el análisis predictivo de los algoritmos de aprendizaje automático, este fundamento teórico permitirá encontrar las variables que ayudarán a predecir la curación de las MAVs tratadas con radiocirugía en esta tesis, esas variables serán identificadas mediante la selección de atributos, esto corresponde a la etapa de preprocesamiento. Asimismo, para encontrar el algoritmo que predice mejor la curación de la Mavs cerebral, los datos seguirán el proceso de preprocesamiento, aprendizaje, evaluación y predicción.



*Figura 13.* Ruta de trabajo para la aplicación de algoritmos de aprendizaje automático en análisis predictivo. (Raschka, 2015)

## 2.6 Hipótesis

### 2.6.1 Hipótesis general

Los algoritmos de clasificación de aprendizaje automático predicen de forma adecuada la curación de las malformaciones arteriovenosas cerebrales en pacientes tratados con radiocirugía.

### 2.6.2 Hipótesis específicas

Las variables sociodemográficas, clínicas, de angioarquitectura y del procedimiento de la radiocirugía son un grupo de variables que ayudarán a predecir la curación de los pacientes con diagnóstico de malformación arteriovenosa cerebral tratados con radiocirugía.

Los algoritmos Naive bayes, regresión logística, máquina de soporte vectorial, random forest y redes neuronales con oversampling tiene mejores resultados que los algoritmos de regresión logística, Naive Bayes, random forest, redes neuronales sin oversampling en el pronóstico de la curación de los pacientes con malformación arteriovenosa cerebral tratados con radiocirugía.

2.7 Variables (definición y operacionalización de variables): Dimensiones e indicadores

**Variable dependiente:** Curación de las malformaciones arteriovenosas. (Variable cualitativa nominal): 1: Curación de la MAV 0: No curación de la MAV

**Variables Independientes:**

Características sociodemográficas: Género, edad, tipo de ocupación, lugar de procedencia, grado de instrucción, tipo de seguro de salud.

Antecedentes clínicos: presencia de hemorragia, tipo de hemorragia, presencia de cefalea, presencia de convulsiones, presencia de encefalomalacia, presencia de déficit, puntaje Karnofsky, puntaje en la escala de coma de Glasgow.

Características del tratamiento previo a la radiocirugía: cirugía previa, embolización, tipo de agente embolizante, cirugía y embolización, puntaje de la escala de Escala de Spetzler Martin.

Características de angioarquitectura de la MAV: presencia de aneurisma venoso, presencia de aneurisma arterial, presencia de dolicoectasia, número de vasos aferentes, profundidad de la MAV, diámetro de la MAV, lado de ubicación de la MAV, tipo de malformación, tipo de drenaje venoso, elocuencia, velocidad de flujo, estenosis venosa, localización de la malformación, volumen de la malformación, puntaje de la escala endovascular de Buffalo, puntaje de la escal endovascular de Virginia, tiempo de obliteración de la malformación.

Características físicas del hardware y software de radiocirugía para el tratamiento de la MAV: dosis de radiación prescrita, número de isocentros, número de sesiones, porcentaje de isodosis.

Algoritmos de aprendizaje automático: Naive Bayes, regresión logística, random forest, máquina de soporte vectorial redes neuronales. Indicadores: accuracy, accuracy balanceado, AUC.



## OPERACIONALIZACIÓN DE LAS VARIABLES

Variable	Definición operacional	Indicador	Tipo	Escala de medición	Valores
Curación de la MAV	Es la exclusión de la malformación arteriovenosa cerebral (MAV) mediante la observación de un examen angiográfico cerebral durante el periodo de seguimiento de 3 años después de la radiocirugía	Demostrado mediante una angiografía o angiotomografía cerebral	Cualitativa	Nominal	1. Curación de la MAV 2. No curación de la MAV
Género	Condición orgsnica que define hombre o mujer	Sexo de acuerdo al DNI	Cualitativa	Nominal	1. Masculino 2. Femenino
Edad	tiempo cronológico de vida	Años cumplidos	Cuantitativa	razón	4 a 75 años
Lugar de procedencia	Ubicación de su residencia en el ultimo año.	Según registro en la historia clínica	Cualitativa	Nominal	1. Lima y Callao 2. Provincias 3. Extranjero
Tipo de ocupación	Actividad o trabajo	Labor que desempeña	Cualitativa	Nominal	1. Profesional universitario- técnico 2. Obrero 3. Ama de casa 4. Policía o fuerzas armadas 5. Estudiante universitario-no universitario 6. Preescolar-Escolar 7. Desocupado-subempleado 8. Comerciante-independiente
Grado de Instrucción	Nivel educativo	Según registro en la historia clínica	Cualitativa	Ordinal	1. Preescolar 2. Primaria 3. Secundaria 4. Superior
Tipo de seguro de salud	Institución aseguradora que cubre sus gastos en salud	Según registro en la historia clínica	Cualitativa	Nominal	1. Privado 2. Essalud 3. SIS 4. Particular 5. Militar, Policial, naval

Variable	Definición operacional	Indicador	Tipo	Escala de medición	Valores
Hemorragia	Presencia de sangre a raíz de una ruptura de la MAV cerebral.	Según registro en la historia clínica	Cualitativa	Nominal	1. Si 2. No
Tipo de hemorragia	El tipo de sangrado que acompaña a la malformación.	Según registro en la historia clínica y lectura de imágenes.	Cualitativa	Nominal	1. Parenquimal 2. Ventricular 3. Parenquimal y ventricular 4. Ninguno
Cefalea	Dolor de cabeza persistente	Según registro en la historia clínica	Cualitativa	Nominal	1. Si 2. No
Convulsiones	Presencia de convulsiones al momento del diagnóstico de la MAV.	Según registro en la historia clínica	Cualitativa	Nominal	1. Si 2. No
Encefalomalasia	Reblandecimiento cerebral en un área donde anteriormente hubo un infarto o hemorragia cerebral.	Según registro en la historia clínica y lectura de imágenes.	Cualitativa	Nominal	1. Si 2. No
Presencia de otras enfermedades	Presencia de enfermedades sistémicas o degenerativas	Según registro en la historia clínica	Cualitativa	Nominal	1. Si 2. No
Deficit	Funciones limitadas en los sentidos.	Según registro en la historia clínica	Cualitativa	Nominal	1. Motor 2. Sensorial 3. Cognitivo 4. Ninguno
Cirugía	Operación a cráneo abierto	Según registro en la historia clínica y lectura de imágenes.	Cualitativa	Nominal	1. Presenta 2. No presenta

Variable	Definición operacional	Indicador	Tipo	Escala de medición	Valores
Embolizacion	Colocación de material embolizante para ocluir una MAV	Según registro en la historia clínica y lectura de imágenes.	Cualitativa	Nominal	1. Si 2. No
Agente embolizante	Tipo de material embolizante usado para la embolización.	Lectura de imágenes.	Cualitativa	Nominal	1. Onix 2. Histoacryl 3. Ninguno
Cirugía y embolización	Al momento de ser tratado con radiocirugía el paciente ya tenía intervención quirúrgica y de embolización.	Según registro en la historia clínica y lectura de imágenes.	Cualitativa	Nominal	1. Solo cirugía 2. Solo embolización 3. Cirugía y embolización 4. Ninguno
Localización de la MAV	Lugar anatómico donde se ubica la malformación arteriovenosa cerebral.	Lectura de imágenes.	Cualitativa	Nominal	1. Lóbulo frontal 2. Lóbulo temporal 3. Lóbulo parietal 4. Lóbulo occipital 5. Cuerpo calloso 6. Insula 7. Ganglios basales 8. Cerebelo 9. Ventricular 10. Vermis 11. Fronto mesial 12. Fronto parietal 13. Fronto temporal 14. Mesencefalo 15. Occipital mesial 16. Parietal mesial 17. Parieto occipital 18. Protuberancia 19. Temporal mesial 20. Temporal occipital 21. Temporal parietal 22. Tronco cerebral
Aneurisma venoso	Presencia de aneurisma venoso acompañando a la malformación arteriovenosa cerebral.	Lectura de imágenes.	Cualitativa	Nominal	1. Si 2. No
Aneurisma arterial	Presencia de aneurisma arterial acompañando a la malformación arteriovenosa cerebral.	Lectura de imágenes.	Cualitativa	Nominal	1. Si 2. No
Dolicoestasia	Elongación, dilatación y tortuosidad de las venas de drenaje de la MAV.	Lectura de imágenes.	Cualitativa	Nominal	1. Si 2. No

Variable	Definición operacional	Indicador	Tipo	Escala de medición	Valores
Número de vasos aferentes	Número de arterias que nutren la MAV.	Lectura de imágenes.	Cuantitativa	Discreta	Conteo
Profundidad de la MAV	Ubicación de la MAV dentro de la estructura cerebral.	Lectura de imágenes.	Cualitativa	Nominal	1. Cortical 2. Subcortical 3. Cortico subcortical 4. profundo 5. ventricular
Diámetro	Medición del diametro mayor de la MAV.	Lectura de imágenes.	Cuantitativa	Razón	0.5 a 8cm
Lado de la MAV	Lado del cerebro donde se encuentra la malformación.	Lectura de imágenes.	Cualitativa	Nominal	1. Derecho 2. Izquierdo 3. Medial
Tipo de MAV	Forma como la malformación se expande en el área cerebral.	Lectura de imágenes.	Cualitativa	Nominal	1. Compacto 2. Difuso 3. Disperso-mixto
Tipo de drenaje venoso	Tipo de desague de sangre venosa de la MAV	Lectura de imágenes.	Cualitativa	Nominal	1. Superficial 2. Profundo 3. Mixto
Elocuencia	La malformación se encuentra en una zona que compromete las funciones vitales y/o funcionales del paciente.	Lectura de imágenes.	Cualitativa	Nominal	1. Si 2. No
Drenaje	Es el vaso o los vasos venosos que evacúan la sangre de la MAV.	Lectura de imágenes.	Cualitativa	Nominal	1. Circulación venosa superficial 2. Circulación venosa profunda
Velocidad del flujo	Velocidad de flujo de la sangre en la MAV	Lectura de imágenes.	Cualitativa	ordinal	1. Lento 2. Moderado 3. Alto
Estenosis venosa	Estrechamiento de la luz de un vaso venoso a la salida del drenaje de una malformación arteriovenoso cerebral.	Lectura de imágenes.	Cualitativa	Nominal	1. Si 2. No

Variable	Definición operacional	Indicador	Tipo	Escala de medición	Valores
Volumen	Volumen de la masa compone la MAV.	Lectura de imágenes.	Cuantitativa	Razón	0.05 a 75 cm <sup>3</sup>
Tiempo de obliteración	Obliteración o cura de la MAV en el tiempo	Según registro en la historia clínica .	Cuantitativa	Razón	6 a 36 meses
Número de sesiones	Número de radiocirugías necesarios para estabilizar la malformación arterio venosa cerebral.	Tomado de la historia clínica.	Cuantitativa	Discreta	Conteo
Resonancia magnetica	Se evaluo la malformacion mediante resonancia magnetica nuclear.	Resonancia magnética	Cualitativa	Nominal	1. Si 2. No
Tomografía cerebral	Se evaluo la malformacion mediante tomografía cerebral.	Tomografía cerebral.	Cualitativa	Nominal	1. Si 2. No
Angiografía digital cerebral	Se evaluo la malformacion mediante angiografía digital cerebral.	Angiografía cerebral	Cualitativa	Nominal	1. Si 2. No
Número de isocentros	Cobertura y calculo de conos y tamaños necesarios para cubrir la MAV a tratar	Mediante software SIMUPLAN	Cuantitativa	Discreta	Conteo
Dosis de radiación	radiación de tratamiento aplicada durante la radiocirugia estereotactica a la MAV	Aplicación de dosis a la MAV generados por un acelerador lineal	Cuantitativa	Razón	1 a 50 Greys
Isodosis	Porcentaje de isodosis	Tomado de los registros de física	Cuantitativa	Razón	40 a 80 %

<b>Variable</b>	<b>Definición operacional</b>	<b>Dimensión</b>	<b>Indicador</b>	<b>Tipo</b>	<b>Escala de medición</b>	<b>Valores</b>
Puntaje de Karnofsky	Estado de salud general y las actividades diarias que un paciente puede llevar a cabo	Incapacidad leve Incapacidad moderada Incapacidad severa	Escala de Performance de Karnofsky	Cuantitativo	Intervalo	0 a 100%
Estado de conciencia	Estados de conciencia desde el coma hasta el estado de alerta	Ocular, Verbal, motora	Escala de coma de Glasgow	Cuantitativa	Razón	3 a 15
Clasificación de la MAV	Clasificación de los pacientes con MAV, teniendo en cuenta el tamaño, localización y drenaje venoso	Tamaño, elocuencia drenaje venoso	Escala de Spetzler Martin	Cuantitativa	Razón	0 a 5
Grado de radiocirugía	Posible resultado al tratamiento on radiocirugía	Volumen, localización elocuente, antecedente de hemorragia.	Escala de radiocirugía de Virginia	Cuantitativa	Intervalo	0 a 4
Tratamiento endovascular	Califica los posibles resultados obtenidos cuando el tipo de tratamiento para las MAV cerebrales es el endovascular	Número de pedículos arteriales, diámetro de pedículos arteriales localización del nido.	Escala endovascular de Buffalo	Cuantitativa	Razón	1 a 5
Algoritmos de aprendizaje automático	Conjunto de códigos que predicen la curación de las malformaciones arteriovenosas cerebrales	Algoritmos: Naive Bayes, regresión logística, random forest, máquina de soporte vectorial y redes neuronales.	Accuracy, Accuracy balanceado, sensibilidad, especificidad, AUC.			

## CAPÍTULO III: MARCO METODOLÓGICO

### 3.1 Tipo de investigación

La presente investigación por el enfoque es de tipo cuantitativo, por el nivel de investigación es explicativa porque permite identificar las variables que predicen la curación de las malformaciones arteriovenosas mediante los algoritmos de clasificación de aprendizaje automático. Por el diseño, es no experimental porque no manipula las variables independientes. Por el método de investigación es hipotético-deductivo porque a partir del comportamiento de datos históricos de las malformaciones en cuanto a su curación es posible tomar características que puedan predecir la curación individual de estos pacientes.

### 3.2 Población y muestra

La población estuvo formada por los pacientes portadores de malformaciones arteriovenosa cerebrales que fueron intervenidos con radiocirugía estereotáctica en el Instituto de Radiocirugía del Perú desde el año 2005 al 2018.

Se trabajo con toda la población de pacientes con el diagnóstico de malformación cerebral arteriovenosa a quienes se realizó tratamiento mediante radiocirugía estereotáctica.

Se calculó la precisión del tamaño de muestra, según los siguientes datos: para una población probable de 202 casos de MAV tratados en el periodo de estudio, con una frecuencia esperada de curaciones de la MAV de 85%, obteniéndose una precisión de 5%, aceptándose como aceptables precisiones entre 5 y 10%.

$$n = \frac{Z_{\alpha/2}^2 * p * q}{d^2}$$

Donde:

p: Proporción esperada de pacientes con MAV cerebral curados con radiocirugía

q: Proporción de pacientes de MAV cerebral no curados con radiocirugía

$Z_{\alpha/2}$ : Nivel de confianza al 95%: 1.96

d: Nivel de precisión: 0.05.

Asimismo, para desarrollar los algoritmos de aprendizaje automático supervisado se seleccionó una muestra por conveniencia del 80% de los datos para el entrenamiento y 20% para el test de la performance predictiva de los modelos. La selección de la muestra fue aleatoria simple.

### 3.3 Técnicas e instrumentos de recolección de datos

Los datos utilizados en esta tesis, provinieron de pacientes con MAVs cerebrales tratados con radiocirugía en el Instituto de Radiocirugía del Perú desde el año 2005 al 2018. La información fue recolectada de las historias clínicas, de las imágenes de resonancia magnética nuclear, imágenes de angiogramas o angiografías por sustracción digital cerebral, informes de física médica.

Se consideró como paciente curado de la malformación arteriovenosa cerebral, luego del tratamiento de radiocirugía, al observar a través de la angiogramas o angiografías por sustracción digital la obliteración total de la malformación en un tiempo de hasta 3 años (36 meses).

El proceso de recolección de datos para esta investigación consistió en los siguientes pasos:

Previo permiso de las autoridades respectivas del Instituto de radiocirugía del Perú, se identificó al paciente con malformación arteriovenosa que recibió tratamiento de radiocirugía en los archivos digitales generales del Instituto desde 2005 al 2018 y se elaboró una lista de estos pacientes.

Los datos clínicos de cada paciente se recogieron de las historias clínicas respectivas.

Los datos de las variables de radiación de física médica usados para el tratamiento de radiocirugía fueron tomados de los archivos de física médica.

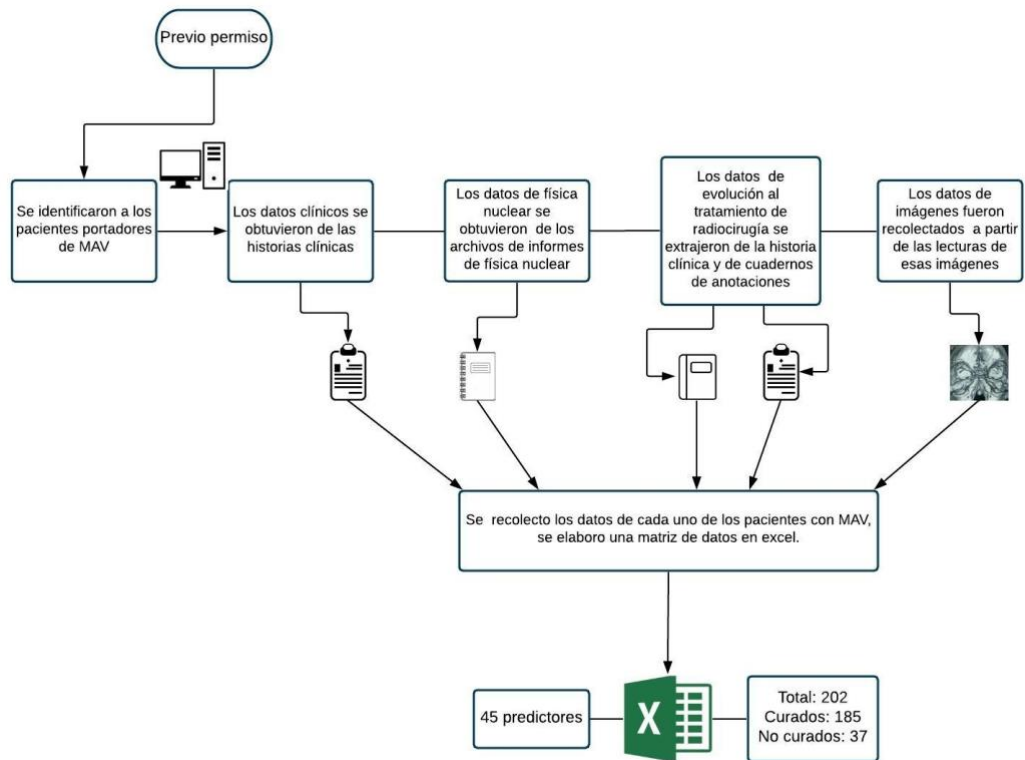
Los datos de evolución al tratamiento de la radiocirugía fueron extraídos de la historia clínica respectiva.

Los datos de las imágenes de la malformación arteriovenosa cerebral fueron recolectados de la lectura de imágenes de resonancia magnética nuclear, angiogramas o angiografías por sustracción digital cerebral de cada uno de los pacientes.



Con la información recolectada se confeccionó una matriz de datos en Excel. Para identificar a los pacientes se etiquetó como clase 1 los pacientes curados y clase 0 los no curados.

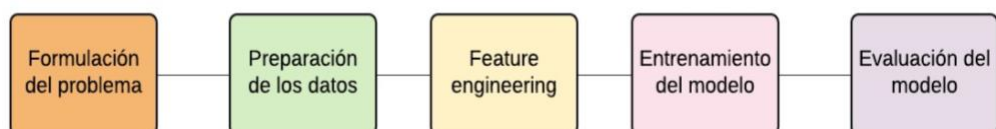
El proceso de recolección de los datos seguido en esta tesis se muestra en la figura 14.



*Figura 14.* Proceso de recolección de los datos  
Fuente: Elaboración propia

### 3.4 Descripción del procedimiento de análisis

Los datos fueron analizados utilizando técnicas de aprendizaje automático supervisado de clasificación. Para implementar el modelo de aprendizaje automático se siguió las etapas mostradas en la figura 15.



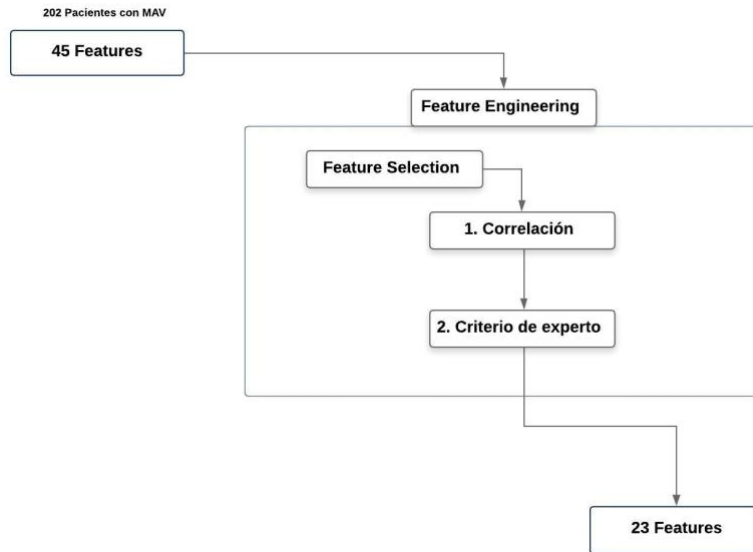
*Figura 15.* Pasos principales para aprendizaje automático supervisado  
Fuente. Modificado del libro de Mechelli & Vieira, 2019.

Para la fase de preparación de los datos, feature engineering, entrenamiento y evaluación del modelo se utilizó el lenguaje de programación Python que es un software libre, además se usaron las librerías scikit learn, pandas, matplotlib, scipy y numpy.

El target (la curación) de la MAV cerebral, presentó categorías desbalanceadas (n= 167; 35). Los resultados obtenidos en esta tesis fueron organizados de acuerdo con las etapas descritas en la Figura 15. Así, en la etapa del preprocesamiento, no fue necesario aplicar técnicas de imputación porque se pudo acceder a todos los valores de las variables predictoras, estas estuvieron debidamente registradas, siendo que la MAV cerebral es una patología poco frecuente. Se realizó el escalamiento de las variables predictoras.

Para identificar las variables que ayudarán a predecir la curación de los pacientes con diagnóstico de MAV cerebral tratados con radiocirugía se realizó el proceso de feature engineering. Se realizaron los procedimientos que se describen a continuación.

La figura 16, muestra el proceso de feature engineering, para esto se seleccionaron los atributos mediante correlación, criterio de experto. Se usó el coeficiente de correlación de Pearson para las variables cuantitativas y el coeficiente de correlación de Cramer para las variables categóricas. Adicionalmente, para evaluar las variables confusoras se aplicó el test de homogeneidad. Se logró una selección de atributos más representativos, reduciéndose estos en 23 atributos: edad, diámetro de la mav, localización de la mav, dosis, cirugía previa, embolización, cirugía y embolización, hemorragia, tipo de hemorragia, cefalea, convulsiones, encefalomalacia, aneurisma venoso, dolicoectasia, profundidad de la mav, tipo de mav, número de vasos eferentes, tipo de drenaje venoso, velocidad de flujo, estenosis venosa, puntaje en la escala de Spetzler Martin, puntaje en la escala de Virginia y número de isocentros.



*Figura 16.* Selección de variables

Fuente: Elaboración propia

Asimismo, para comparar los resultados de los algoritmos y verificar cual de los algoritmos obtiene el mejor pronóstico para la curación de las malformaciones arteriovenosas tratadas con radiocirugía, se desarrollaron procedimientos que se describen a continuación.

Se utilizó, 80% de los datos para el entrenamiento de los algoritmos de aprendizaje automático y 20% para el test de la performance predictiva de los modelos.

Siguiendo la ruta de trabajo de la figura 17, que se muestra abajo, se realizó 4 experimentos:

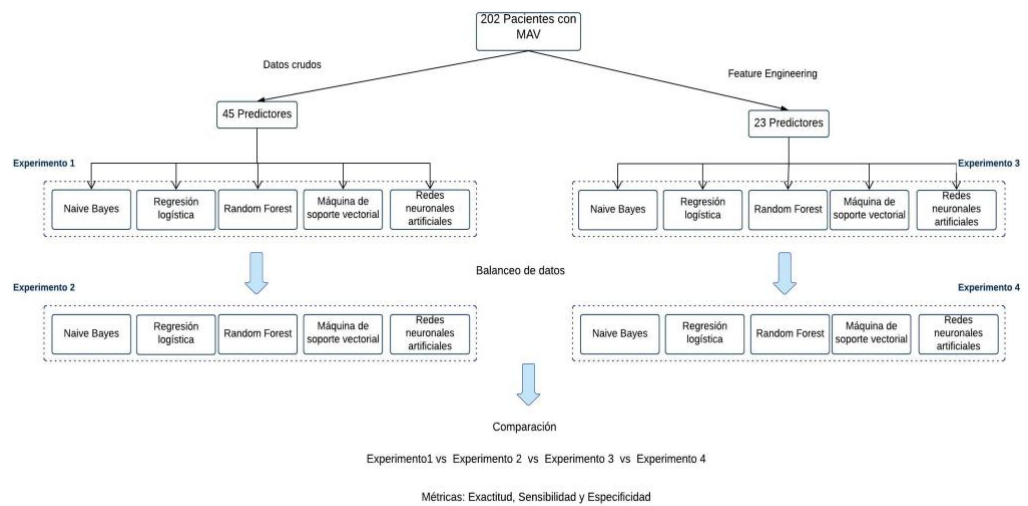
1. Experimento 1: Se trabajó con todas las variables predictoras (sin balanceo).
2. Experimento 2: Se trabajó con todas las variables predictoras (con balanceo)
3. Experimento 3: Se trabajó con los 23 atributos obtenidos mediante feature engineering (sin balanceo).
4. Experimento 4: Se trabajó con los 23 atributos obtenidos mediante feature engineering (con balanceo).

El balanceo del target se hizo mediante la técnica de SMOTE (Synthetic Minority Over-sampling Technique) más Tomek links.

En nuestro estudio no fue necesario aplicar técnicas de imputación para datos faltantes porque se tuvo acceso a cada uno de esos datos, lo cual garantiza la validez de nuestro conjunto de datos para este trabajo.

Para el análisis predictivo se usaron los siguientes algoritmos: regresión logística, naive Bayes, Máquina de soporte vectorial, random forest y redes neuronales, que, según los autores, Senders, y otros, (2017), Uddin, Khan, Hossain, & Ali Moni, (2019) son los modelos que más comúnmente se utilizan en el área médica.

Para maximizar el aprovechamiento de los datos realizamos varias iteraciones, se hicieron 10 iteraciones en cada algoritmo. La evaluación de estos algoritmos se realizó utilizando métricas de clasificación: accuracy, accuracy balanceado, precisión, recall y el área bajo la curva. Para todas las métricas se tomaron los promedios de las 10 iteraciones ejecutadas para cada modelo. La validación de los modelos se realizó utilizando la técnica de resampling llamada validación cruzada. Finalmente, se comparó los indicadores de cada modelo y se eligió el modelo que mejores valores de las métricas obtuvieron para la predicción de la curación de la MAV cerebral. El desempeño del indicador AUC (área bajo la curva) se considero sin valor predictivo si el área es 0.5, aceptable si el área es 0.7-0,80, excelente si el área es 0.8-0.9 y sobresaliente si es mayor de 0.90. Se uso la prueba Kruskal Wallis para evaluar las diferencias significativas entre el área bajo la curva de los algoritmos con un nivel de significancia de 0.05. El detalle del procedimiento descrito se puede visualizar en la figura 17.



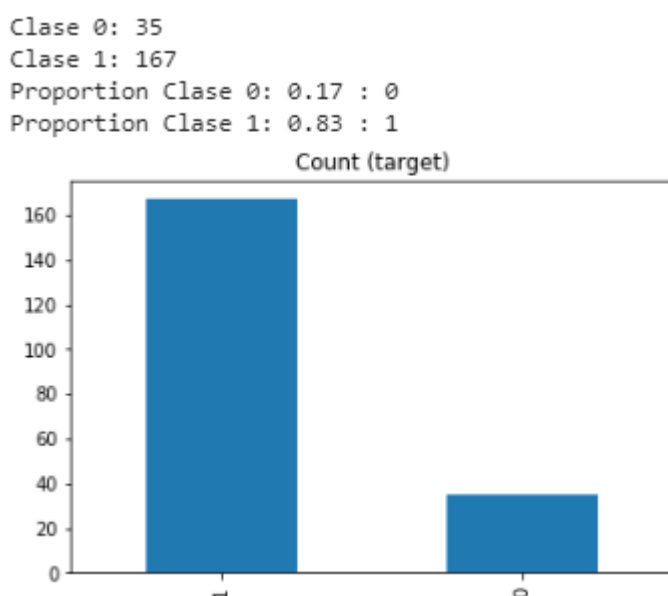
*Figura 17.* Descripción del procedimiento de análisis  
Fuente: Elaboración propia

Se realizó el análisis descriptivo de los predictores mediante tablas de frecuencias para las variables categóricas y para las cuantitativas se usó medidas de resumen.

## CAPÍTULO IV: RESULTADOS Y ANÁLISIS DE RESULTADOS

### 4.1 Resultados

Para un mejor entendimiento y a fin de que se conozcan los pacientes que conforman nuestra base de datos, presentaremos en primer lugar tablas descriptivas de las características sociodemográficas y clínicas de los 202 pacientes con MAV cerebral tratados con radiocirugía estereotáctica entre los años 2005 al 2018. En la figura 18 se observa, que la radiocirugía fue efectiva para la curación en 167 (82.2%) de 202 pacientes con MAV cerebral.



*Figura 18.* Distribución de la variable Curación de la MAV

La tabla 6, muestra las características sociodemográficas de nuestra población; la mayoría tenían entre 18 a 59 años de edad (70.49%), procedían de Lima y Callao (80.69%), eran preescolares o escolares (27.72%) y 52.97 % solo tenían secundaria y un 42.08 % tenían seguro integral de salud (SIS) del Ministerio de Salud.

Tabla 6: *Características sociodemográficas de los pacientes con malformación arteriovenosa cerebrales tratados con radiocirugía 2005 al 2018*

<b>Datos Sociodemográficos</b>	<b>Total</b>	<b>%</b>
<b>Sexo</b>		
Masculino	107	52.97
Femenino	95	47.0
<b>Edad</b>		
Niños ( 1-11 años)	30	14.85
Adolescentes (12 - 17 años)	24	11.88
Jóvenes (18 - 29 años)	66	32.37
Adultos (30 - 59 años)	77	38.12
Adultos mayores (60 a más años)	5	2.48
<b>Lugar de procedencia</b>		
Lima y Callao	163	80.69
Provincias	37	18.32
Extranjero	2	0.99
<b>Tipo de Ocupación</b>		
Profesional Universitario/No universitario	35	17.33
Obrero	24	11.88
Ama de casa	34	16.83
Fuerzas armadas y policiales	9	4.46
Estudiante Universitario/ No universitario	20	9.9
Preescolar/escolar	56	27.72
Desocupado/subempleado	10	4.95
Comerciante/Independiente	14	6.93
<b>Grado de instrucción</b>		
Preescolar	5	2.48
Primaria	33	16.34
Secundaria	107	52.97
Superior	57	28.22
<b>Tipo de seguro de salud</b>		
Privado	20	9.9
Essalud	29	14.36
SIS	85	42.08
Particular	49	24.26
Fuerzas armadas y policiales	19	9.41
<b>Total</b>	<b>202</b>	<b>100.0</b>

Fuente: Elaboración propia

La tabla 7, muestra las características clínicas donde se observa que el tiempo promedio desde la aplicación de la radiocirugía hasta su curación (obliteración) en los pacientes que se curaron fue de 22.07 meses, la dosis prescrita en promedio fue de 17.86 Greys con diámetro promedio de las MAV de 2.144 centímetros cúbicos y el número de

isocentros aplicados en promedio es de 1.35. La radiocirugía fue realizada en promedio en sola 1 sesión.

Tabla 7. *Estadísticas descriptivas de las variables cuantitativas de los pacientes con malformación arteriovenosa cerebrales tratados con radiocirugía 2005 al 2018.*

Variable	Media	Desv.Est.	Mínimo	Mediana	Máximo
Edad	27.63	14.9	4	25	68
Karnofsky	82.78	9.374	40	80	100
Escala de coma de Glasgow	14.79	0.464	12	15	15
Spetzler-Martin	2.545	0.8869	1	3	5
Escalas de Buffalo	2.426	0.9758	0	2	5
Escala de Virginia	2.079	0.8248	0	2	4
Diametro de la malformación (cm)	2.144	0.8863	0.5	2.1	6
Volumen de la malformación (cc)	6.297	8.327	0.063	4	75
Número de vasos aferentes	2.51	0.9155	1	2	6
Número de sesiones	1.356	1.556	1	1	10
Número de Isocentros	1.352	0.5556	1	1	4
Dosis prescrita (Greys)	17.86	4.441	10	17	40
% Isodosis	69.31	14.37	50	80	90
Tiempo de obliteración de la MAV (meses)	22.07	6.474	6	24	36

Fuente: Elaboración propia

La tabla 8 muestra los tratamientos previos de los pacientes antes la radiocirugía estereotáctica. De los 202 pacientes, 31 recibieron tratamiento quirúrgico previo, 22 fueron tratados con embolización y además se intervinieron quirúrgicamente de manera convencional, 40 recibieron terapia endovascular mediante embolización. La mayoría de las embolizaciones fueron realizadas con ónix. Asimismo, se muestra variables del cuadro clínico de ingreso al estudio, de los 202 pacientes 155 tuvieron hemorragia cerebral, 76 desarrollaron encefalomalasia., 178 tuvieron cefalea y 112 convulsiones. Un 55% presentaron antes del procedimiento algún tipo de déficit, ya sea motor, sensitivo o cognitivo.

La tabla 8, además muestra las características anatómicas de la angioarquitectura de la MAV es decir características propias de la lesión. La mayoría se ubicaron en el lado izquierdo del cerebro (49.5%), con una ubicación profunda en el (46.52%). La presencia de una varice dolicoectásica estaba en el 69.31% de los pacientes. La mayoría de las MAV tratadas tienen un flujo de intensidad moderada (47.03 %) y un 56.4% tiene drenaje profundo y 25.74 % tenía estenosis venosa a la salida de la MAV. Un 28.71 % tenían aneurismas venosos y solo 3 casos tenían aneurismas arteriales. 145 de las MAV tratadas era de tipo compacto.



Tabla 8. *Variables clínicas y de angioarquitectura de las MAV*

Variables	Categorías	Frecuencia por categoría	Valores
Cirugía previa	2	1=31 ; 2= 171	1: Si ; 2: No
Embolización	2	1 =49; 2 =153	1: Si ; 2: No
Tipo de agente embolizante	3	1 =26; 2 =24 ; 3 =152	1: Onix ; 2: Histoacryl; 3: Ninguno
Cirugía y embolización	4	1=22; 2=40; 3=9; 4= 131	1: Solo cirugía ; 2: Solo embolización; 3: Cirugía y embolización; 4:Ninguno
Hemorragia	2	1=155; 2= 47	1: Si ; 2: No
Tipo de hemorragia	4	1=91; 2=13; 3=29; 4= 69	1: Parenquimal ; 2: Ventricular; 3: Parenquimal y ventricular; 4:Ninguno
Cefalea	2	1=178 ; 2= 24	1: Si ; 2: No
Convulsiones	2	1=112 ; 2= 90	1: Si ; 2: No
Encefalomalasia	2	1=76 ; 2= 126	1: Si ; 2: No
Deficit	4	1=53; 2=26; 3=32; 4= 91	1: Motor ; 2: Sensorial; 3: Cognitivo; 4:Ninguno
Aneurisma venoso	2	1=58 ; 2= 144	1: Si ; 2: No
aneurisma arterial	2	1=3 ; 2= 199	1: Si ; 2: No
Dolicoestasia	2	1=140 ; 2= 62	1: Si ; 2: No
Profundidad de la MAV	5	1=9; 2=48; 3=47; 4= 96; 5=2	1: Cortical ; 2: Subcortical; 3: Cortico subcortical; 4:Profundo; 5: Ventricular
Lado de la MAV	3	1=85; 2=100; 3=17	1: Derecho ; 2: Izquierda; 3: Medial
Tipo de MAV	3	1=145; 2= 39; 3=18	1: Compacto ; 2: Difuso; 3: Disperso-mixto
Tipo de drenaje venoso	3	1=75; 2= 114; 3=13	1: Superficial ; 2: Profundo; 3:Mixto
Elocuencia	2	1=63 ; 2= 139	1: Si ; 2: No
Drenaje	2	1=70; 2=132	1: Circulación venosa superficial; 2: Circulación venosa profunda
Velocidad de flujo	3	1=80; 2= 95; 3=27	1: Lento ; 2: Moderado; 3:Alto
Estenosis venosa	2	1=52 ; 2= 150	1: Si ; 2: No

Fuente: Elaboración propia

La tabla 9 y la figura 19 muestran la ubicación anatómica de las MAVs dentro del cerebro, la mayoría se ubican a nivel de los ganglios basales con 16.83 % de los casos, seguidos de los ubicados en el lóbulo frontal (9.9%).

Tabla 9. Localización de la malformación arteriovenosa cerebrales en los pacientes tratados con radiocirugía 2005 al 2018

Localización de la Malformación	Total	%
Lóbulo frontal	20	9.90
Lóbulo temporal	11	5.45
Lóbulo parietal	11	5.45
Lóbulo occipital	7	3.47
Cuerpo calloso	12	5.94
Insula	14	6.93
Ganglios basales	34	16.83
Cerebelo	13	6.44
Ventricular	10	4.95
Vermis	3	1.49
Fronto mesial	6	2.97
Fronto parietal	9	4.46
Fronto temporal	3	1.49
Mesencéfalo	7	3.47
Occipital mesial	3	1.49
Parietal mesial	2	0.99
Parieto occipital	14	6.93
Protuberancia	3	1.49
Temporal mesial	14	6.93
Temporal occipital	3	1.49
Temporal parietal	1	0.49
Tronco cerebral	2	0.99
<b>Total</b>	<b>202</b>	<b>100.0</b>

Fuente: Elaboración propia

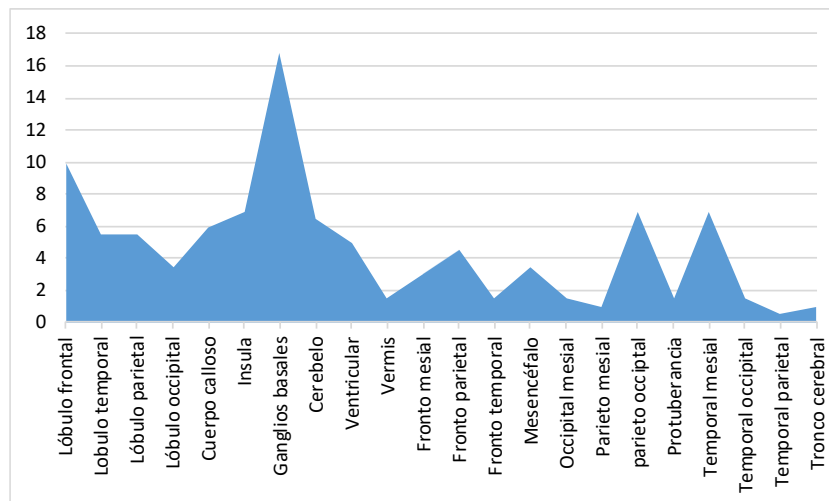


Figura 19. Localización de la Malformación arteriovenosa cerebral en los pacientes tratados con radiocirugía 2005 al 2018

Fuente: Elaboración propia

En la figura 20 se puede visualizar todas las variables predictoras y del target a curación de la malformación arteriovenosa cerebral.

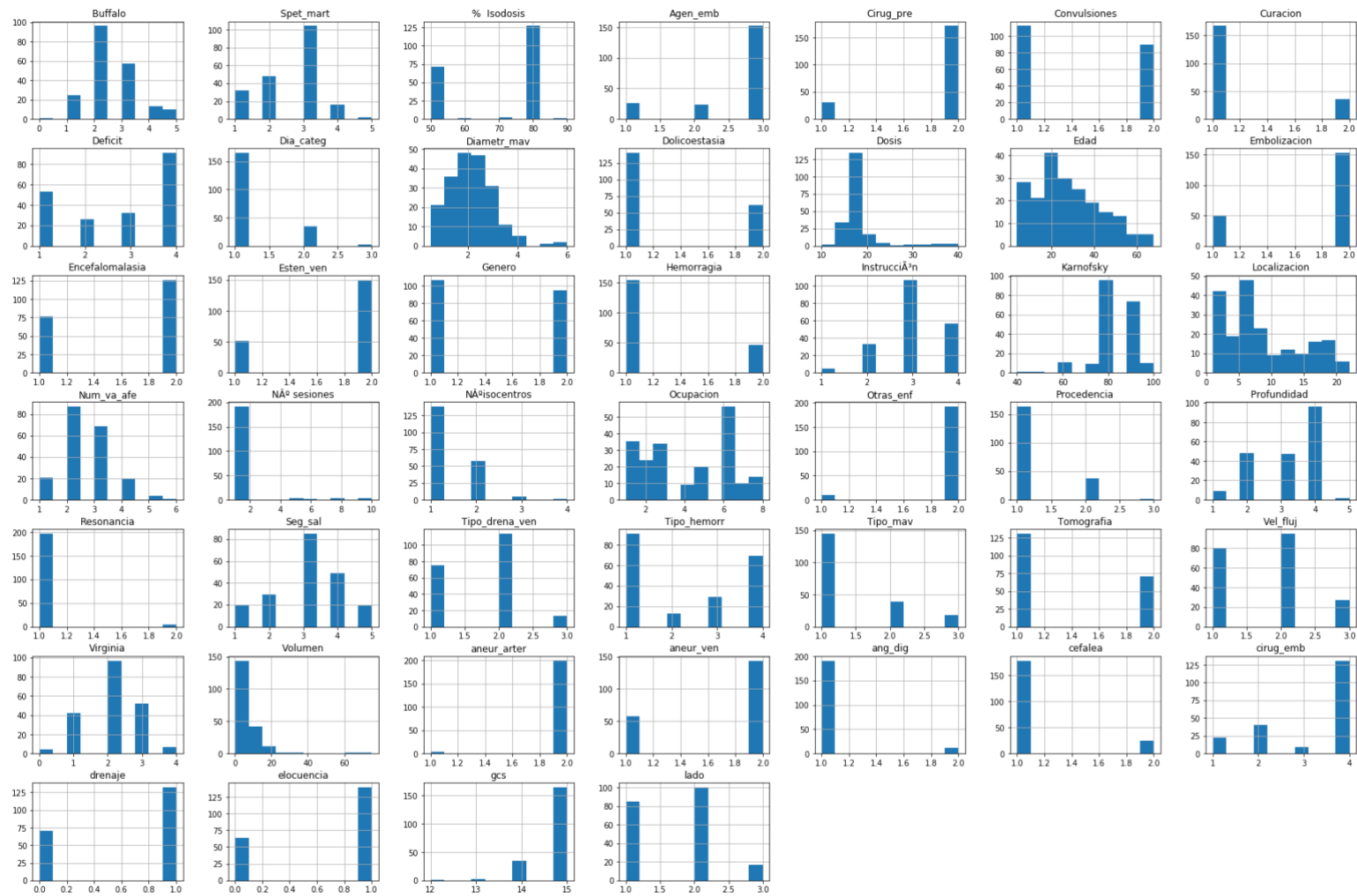


Figura 20. Visualización de las variables predictoras y del target: Curación

Fuente: elaboración propia

Con el objetivo de identificar las variables sociodemográficas, clínicas, de angioarquitectura y del procedimiento de la radiocirugía de las malformaciones arteriovenosas cerebrales que ayudarán a predecir la curación de los pacientes con malformación arteriovenosa cerebral, se realizó el proceso de de feature engineering, los resultados se muestran a continuación.

La tabla 10 muestra las correlaciones entre las variables predictoras cuantitativas, se retiraron las siguientes variables: volumen, número de sesiones, isodosis, agente embolizante y drenaje; teniendo en cuenta las correlaciones altas y además el criterio de experto:

Tabla 10. *Correlación entre las variables independientes cuantitativas*

	edad	gcs	karnofsky	diametr_mav	num_va_afe	volumen	spet_mart	buffalo	virginia	nro_ses	nro_isocent	dosis	isodosis
edad	1.000000	-0.219905	-0.068336	0.057894	-0.003423	-0.029008	0.052643	-0.088640	-0.023502	-0.087414	-0.003388	-0.112657	-0.010037
gcs	-0.219905	1.000000	0.396526	-0.025301	0.054159	0.006320	-0.062010	-0.023284	-0.164768	0.048051	0.053315	0.084847	-0.066497
karnofsky	-0.068336	0.396526	1.000000	0.000043	-0.085123	-0.018230	-0.151696	-0.063549	-0.195253	-0.051165	0.138336	-0.009713	-0.141502
diametr_mav	0.057894	-0.025301	0.000043	1.000000	0.412947	0.786159	0.295334	0.286212	0.449742	0.500934	0.296005	0.359094	-0.260637
num_va_afe	-0.003423	0.054159	-0.085123	0.412947	1.000000	0.373532	0.290244	0.432863	0.288352	0.293231	0.199988	0.200069	-0.199658
volumen	-0.029008	0.006320	-0.018230	0.786159	0.373532	1.000000	0.329463	0.298056	0.335062	0.648043	0.309155	0.447341	-0.220079
spet_mart	0.052643	-0.062010	-0.151696	0.295334	0.290244	0.329463	1.000000	0.414884	0.559669	0.276939	0.205335	0.215515	-0.095162
buffalo	-0.088640	-0.023284	-0.063549	0.286212	0.432863	0.298056	0.414884	1.000000	0.396792	0.201079	0.089679	0.188425	0.006956
virginia	-0.023502	-0.164768	-0.195253	0.449742	0.288352	0.335062	0.559669	0.396792	1.000000	0.253215	0.199512	0.171929	-0.142269
nro_ses	-0.087414	0.048051	-0.051165	0.500934	0.293231	0.648043	0.276939	0.201079	0.253215	1.000000	0.182450	0.898658	-0.164733
nro_isocent	-0.003388	0.053315	0.138336	0.296005	0.199988	0.309155	0.205335	0.089679	0.199512	0.182450	1.000000	-0.025547	-0.810618
dosis	-0.112657	0.084847	-0.009713	0.359094	0.200069	0.447341	0.215515	0.188425	0.171929	0.898658	-0.025547	1.000000	0.065483
isodosis	-0.010037	-0.066497	-0.141502	-0.260637	-0.199658	-0.220079	-0.095162	0.006956	-0.142269	-0.164733	-0.810618	0.065483	1.000000

La figura 21, muestra un mapa de calor con las correlaciones de las variables cualitativas, se retiro las variables con correlación alta y las consideradas siguiendo el criterio de experto.

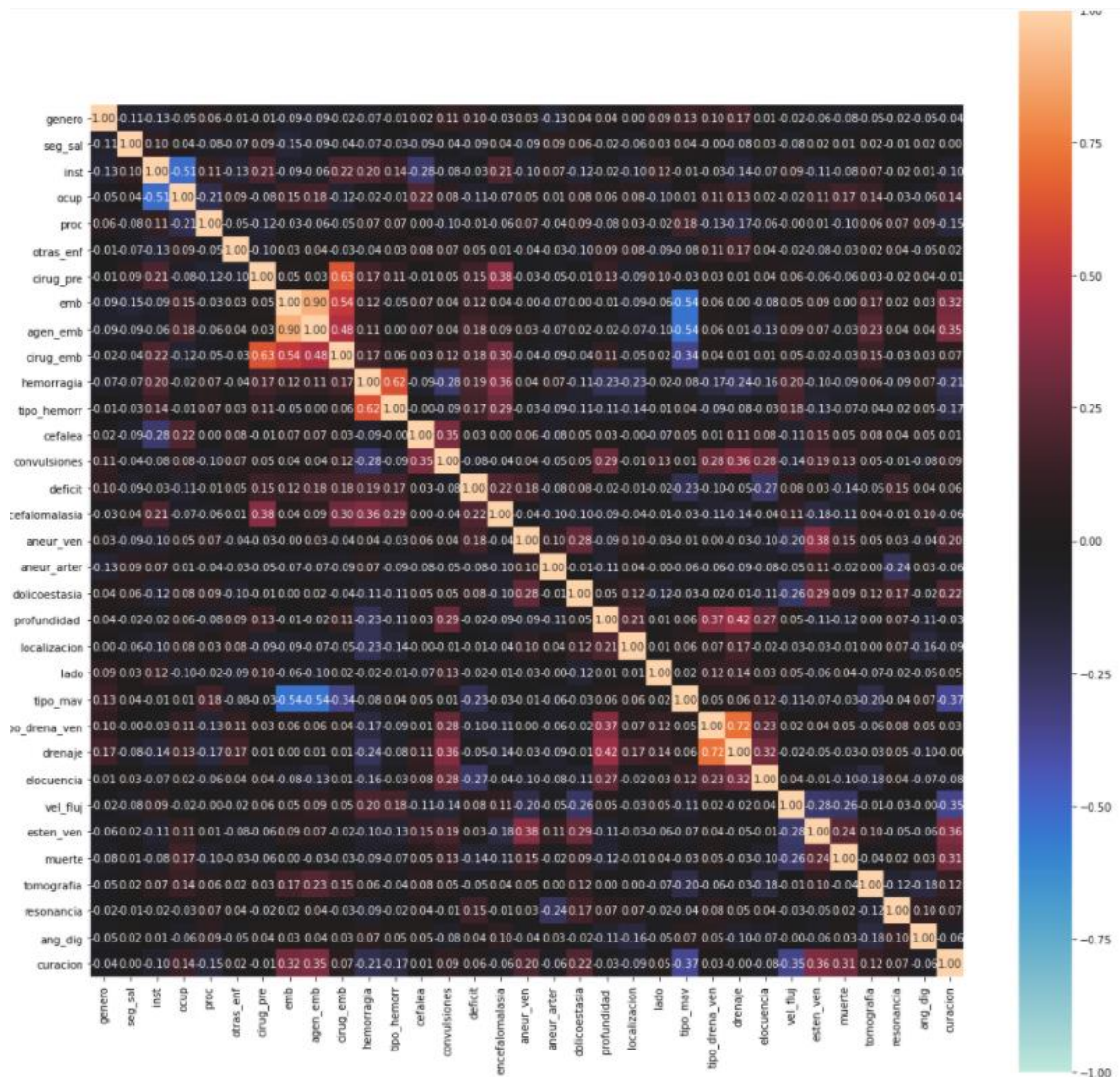


Figura 21. Mapa de calor de las correlaciones de las variables cualitativas

La tabla 11, muestra las 23 variables que finalmente ayudaran a predecir la curación de las malformaciones arteriales tratadas con radiocirugía.

Tabla 11. *Predictores que ayudarán a predecir la curación de MAV cerebrales.*

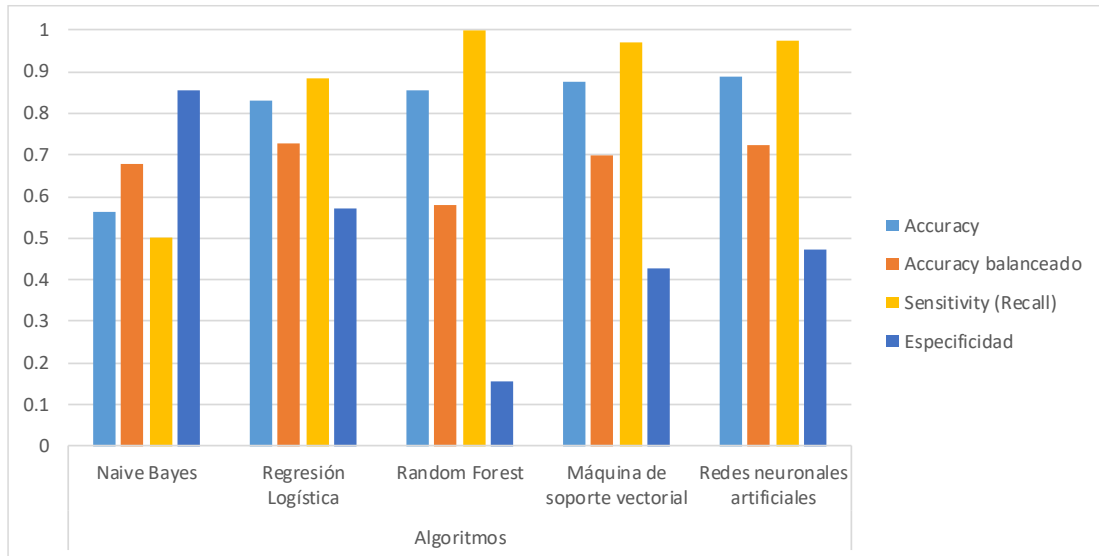
Sociodemográficos	Edad
Clínicos	Hemorragia, tipo de hemorragia, cefalea, convulsiones, encefalomalasia
Tratamiento previo a la radiocirugía	Cirugía previa, embolización, cirugía y embolización, puntaje de Spetzler Martin
Angioarquitectura de la MAV cerebral	Diámetro, localización, aneurisma venosos, dolicoestasia, profundidad, tipo de hemorragia, número de vasos aferentes, tipo de drenaje venoso, velocidad de flujo, estenosis venosa, puntaje de Virginia.
Física médica	Dosis, número de isocentros

Con el objetivo de la tesis de determinar y evaluar la predicción de los algoritmos de Naive Bayes, regresión logística, random forest, máquina de soporte vectorial y redes neuronales usando todas las variables, las variables más importantes, con oversampling y sin oversampling y elegir el mejor resultado en la predicción de la curación de los pacientes con malformación arteriovenosa cerebral tratados con radiocirugía, se realizaron cuatro experimentos (ver sección 3.4), en cada uno de los experimentos se compararon las métricas de los algoritmos. Esto se describe a continuación.

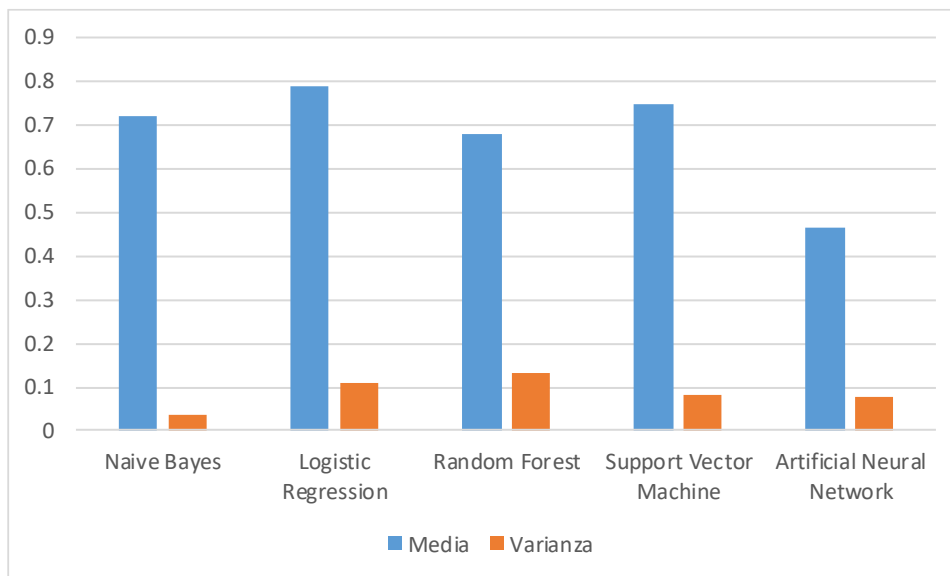
En la tabla 12 y figura 22, 23, muestran las métricas de los diferentes algoritmos de clasificación obtenidos en el primer experimento. El algoritmo con mejor accuracy balanceado es el de regresión logística (72.69%), al validar la métrica con validación cruzada se obtuvo 78.75% y la varianza fue de 0.11. Asimismo, el algoritmo con mejor métrica de sensibilidad es random forest (100%) y el algoritmo con mejor especificidad es el de naive bayes (85.71%)

Tabla 12. *Comparación de los algoritmos en el primer experimento*

Métricas de rendimiento de los algoritmos de clasificación	Algoritmos				
	Naive Bayes	Regresión Logística	Random Forest	Máquina de soporte vectorial	Redes neuronales artificiales
Accuracy	0.56097561	0.829268293	0.856097561	0.87804878	0.887804878
Accuracy balanceado	0.678571429	<b>0.726890756</b>	0.578571429	0.699579832	0.722478992
Prevalence	0.829268293	0.829268293	0.829268293	0.829268293	0.829268293
Sensitivity (Recall)	0.5	<b>0.882352941</b>	1	0.970588235	0.973529412
Especificidad	0.857142857	<b>0.571428571</b>	0.157142857	0.428571429	0.471428571
Precision (PPV)	0.944444444	0.909090909	0.852179487	0.891891892	0.899869059
NPV	0.260869565	0.5	1	0.75	0.801666667
F1	0.653846154	0.895522388	0.920177712	0.929577465	0.934987014
AUC	0.678571429	0.726890756	0.572571429	0.699579832	0.712478992
<b>Validación cruzada</b>					
Score (mean)	0.720868347	<b>0.78745862</b>	0.67965368	0.74763178	0.462837573
Score (variance)	0.036267571	<b>0.108657919</b>	0.131376545	0.080893684	0.077105218



*Figura 22.* Comparación de los algoritmos en el primer experimento  
Fuente: Elaboración propia



*Figura 23.* Resultados de la validación cruzada en el primer experimento  
Fuente: Elaboración propia

En cuanto el área bajo la curva, el algoritmo de regresión logística presentó la mejor métrica, AUC de 72.69%. Ver Figura 24

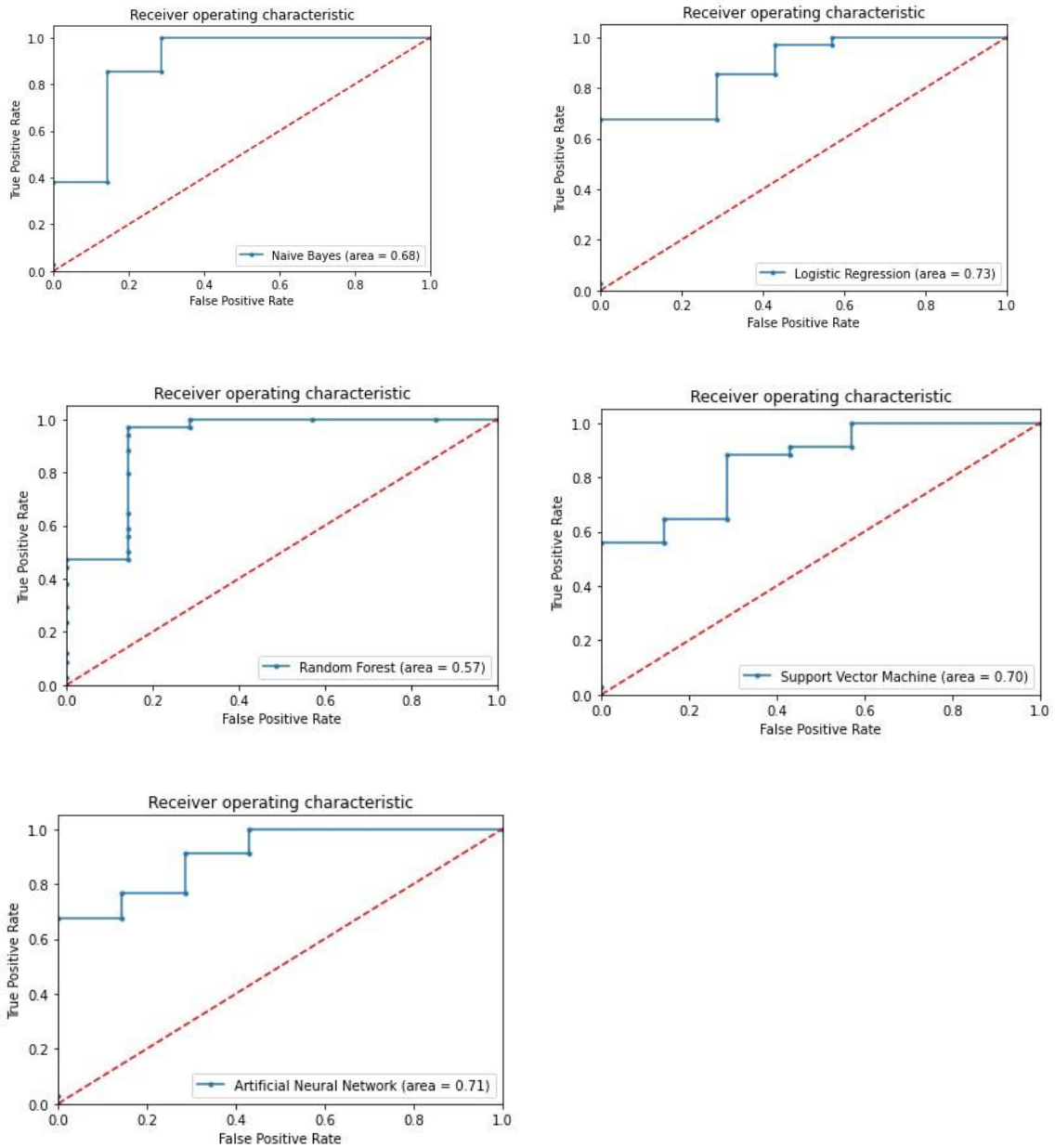


Figura 24. Área bajo la curva de los algoritmos: regresión logística, Naive Bayes, Random forest, Máquina de soporte vectorial, y redes neuronales, en el primer experimento.

Mediante la prueba de Kruskal Wallis se comprobó que existe diferencias significativas al comparar el área bajo la curva de cada grupo de algoritmos ( $\lambda^2=38.114$ ,  $p=0.000$ ).

En la tabla 13 y figura 25, 26 se muestran las métricas de los diferentes algoritmos de clasificación obtenidos en el segundo experimento. El algoritmo con



mejor accuracy es el de máquina de soporte vectorial (98.5%), al validar la métrica con validación cruzada se obtuvo 97.91% y la varianza fue de 0.022. El algoritmo con mejor métrica de sensibilidad es maquina de soporte vectorial (100%) y el algoritmo con mejor especificidad es el de máquina de soporte vectorial (96.97%).

Tabla 13 Comparación de los algoritmos en el segundo experimento

Métricas de rendimiento de los algoritmos de clasificación	Algoritmos				
	Naive Bayes	Regresión Logística	Random Forest	Máquina de soporte vectorial	Redes neuronales artificiales
Accuracy	0.701492537	0.955223881	<b>0.965671642</b>	<b>0.985074627</b>	0.964179104
Accuracy balanceado	0.704991087	0.95543672	0.965285205	0.984848485	0.964705882
Prevalence	0.507462687	0.507462687	0.507462687	0.507462687	0.507462687
Sensitivity (Recall)	0.470588235	0.941176471	<b>0.991176471</b>	<b>1</b>	0.929411765
Especificidad	0.93939393	0.969696969	<b>0.939393939</b>	<b>0.96969696</b>	1
Precision (PPV)	0.888888889	0.96969697	0.943968254	0.971428571	1
NPV	0.63265306	0.94117647	0.990625	1	0.93386194
F1	0.615384615	0.955223881	0.966956522	0.985507246	0.962872124
AUC	0.704991087	0.95543672	0.945285205	0.984848485	0.982705882
<b>Validación cruzada</b>					
Score (mean)	0.742559928	0.946132972	<b>0.9461782</b>	<b>0.979059249</b>	0.451108096
Score (variance)	0.044409041	0.036036112	<b>0.042793814</b>	<b>0.022326345</b>	0.199065531

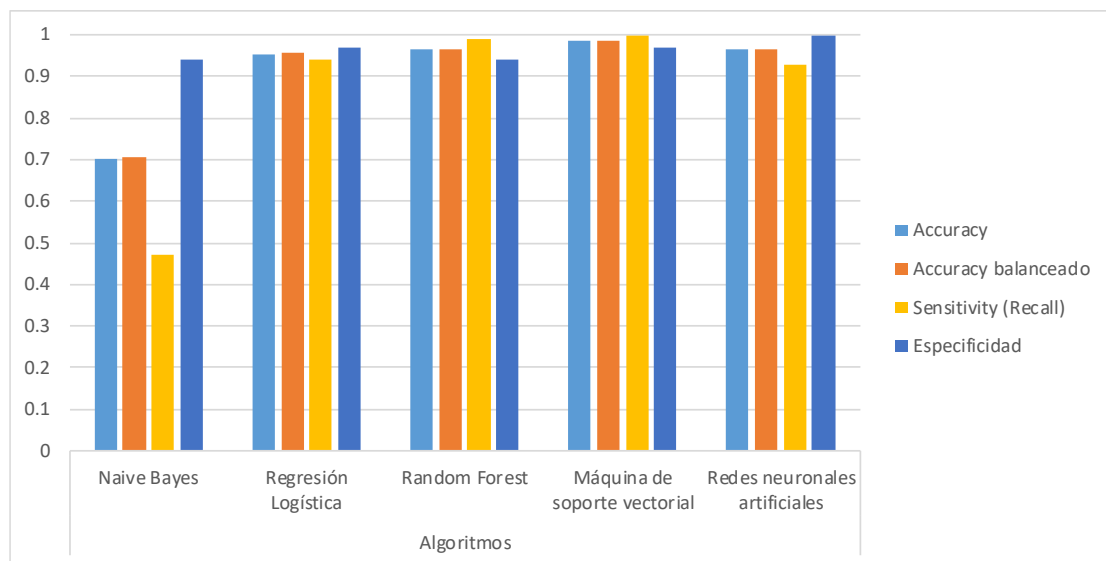


Figura 25 Comparación de los algoritmos en el segundo experimento  
Fuente: Elaboración propia

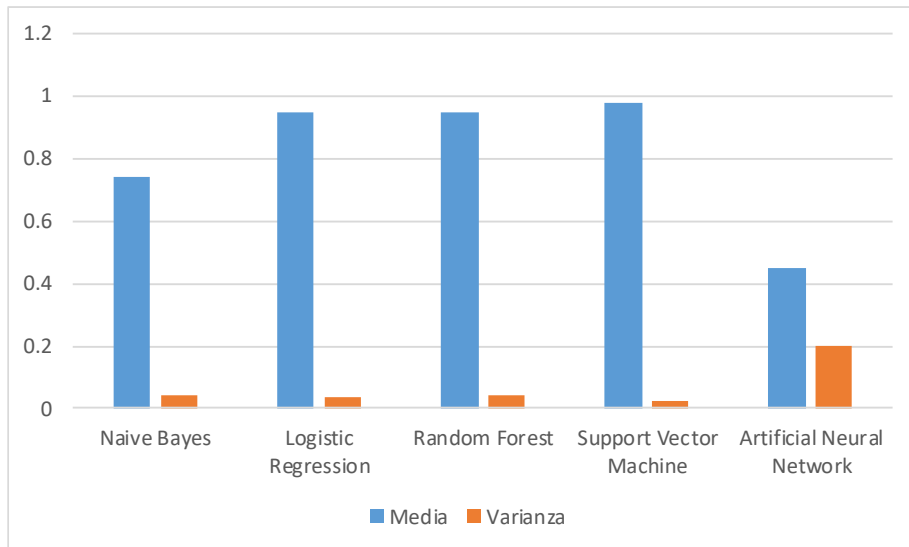
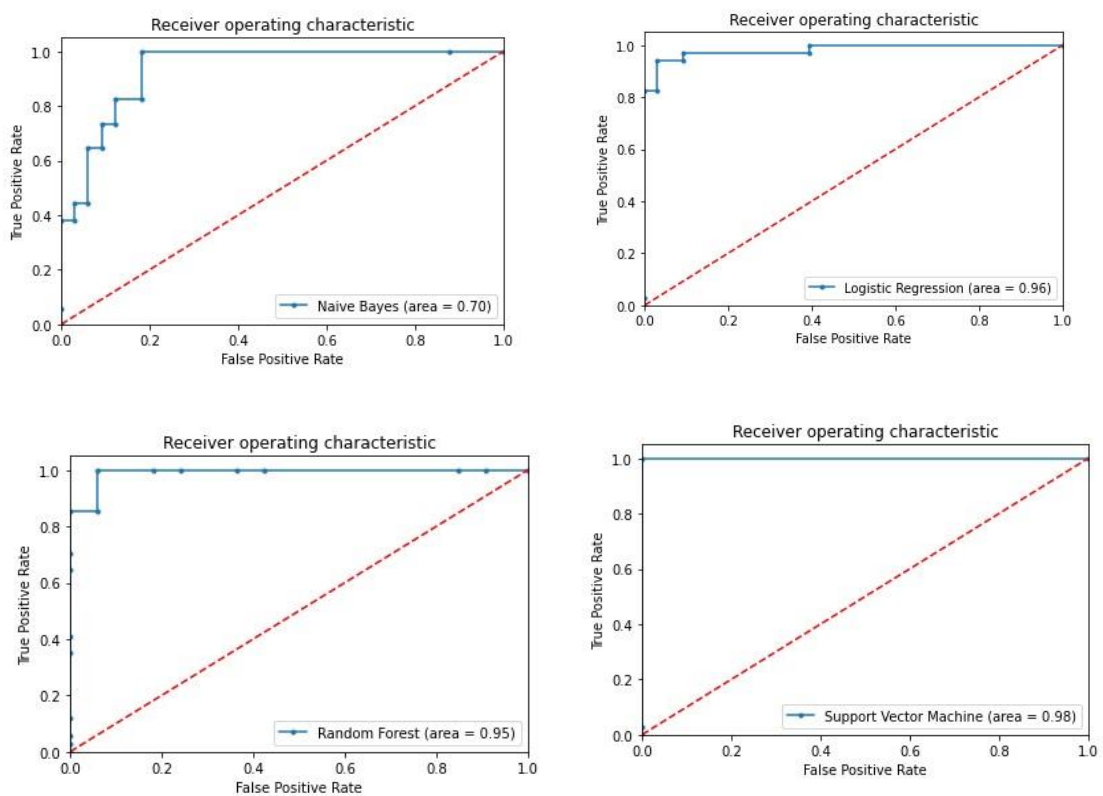


Figura 26. Resultados de la validación cruzada del segundo experimento.  
Fuente: Elaboración propia

De la misma manera, en cuanto al área bajo la curva en el experimento 2, el algoritmo maquina de soporte vectorial tiene la mejor área 98.48%. Ver figura 27



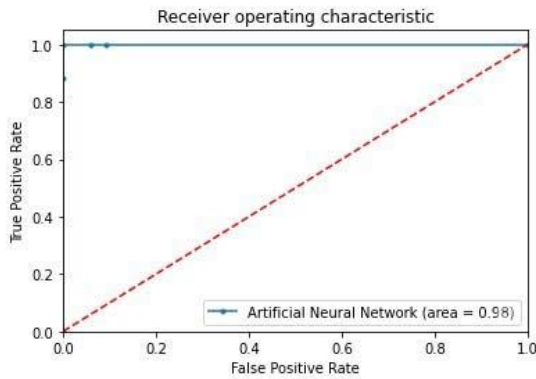


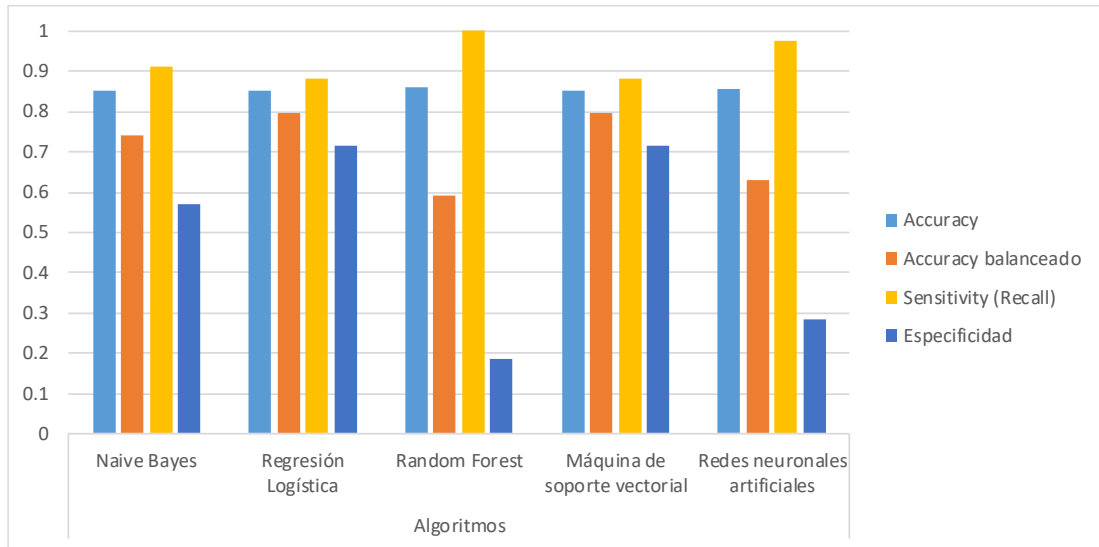
Figura 27. Área bajo la curva de los algoritmos: regresión logística, Naive Bayes, Random forest, Máquina de soporte vectorial, y redes neuronales, en el segundo experimento.

Mediante la prueba de Kruskal Wallis se comprobó que existe diferencias significativas al comparar el área bajo la curva de cada grupo de algoritmos ( $\lambda^2=36.038$ ,  $p=0.000$ ).

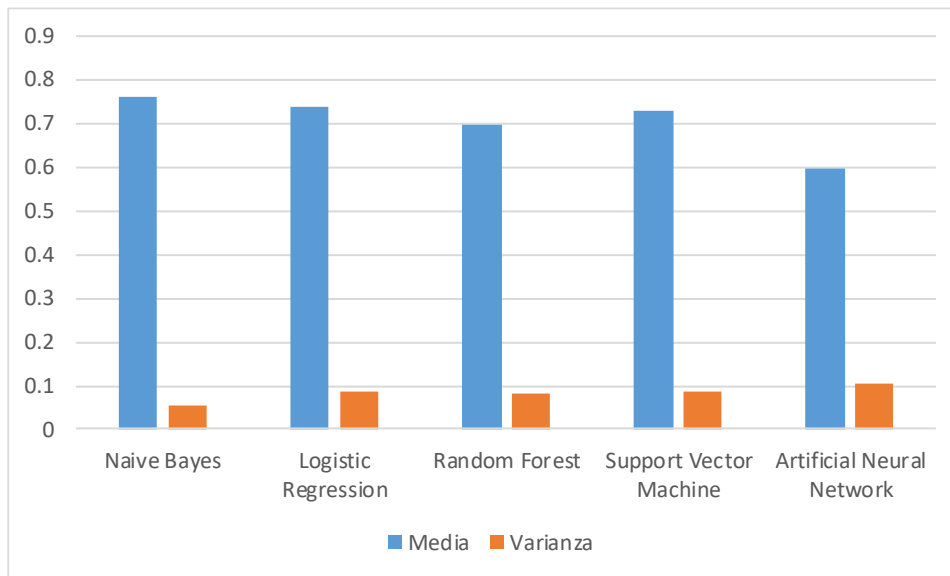
La tabla 14 y figura 28, 29 muestra las métricas de los diferentes algoritmos de clasificación obtenidos en el tercer experimento, se trabajo con los 23 atributos seleccionados. El algoritmo con mejor accuracy balanceado es el de regresión logística (79.83%) al validar la métrica con validación cruzada se obtuvo 74.09% y la varianza fue de 0.087. El algoritmo con mejor métrica de sensibilidad es naive bayes (91.18%) y el algoritmo con mejor especificidad es el de regresión logística (71.43%).

Tabla 14. Comparación de los algoritmos en el tercer experimento

Métricas de rendimiento de los algoritmos de clasificación	Algoritmos				
	Naive Bayes	Regresión Logística	Random Forest	Máquina de soporte vectorial	Redes neuronales artificiales
Accuracy	0.853658537	0.853658537	0.86097561	0.853658537	0.858536585
Accuracy balanceado	0.741596639	<b>0.798319328</b>	0.592857143	<b>0.798319328</b>	0.631092437
Prevalence	0.829268293	0.829268293	0.829268293	0.829268293	0.829268293
Sensitivity (Recall)	0.911764706	<b>0.882352941</b>	1	<b>0.882352941</b>	0.976470588
Especificidad	0.571428571	<b>0.714285714</b>	0.185714286	<b>0.714285714</b>	0.285714286
Precision (PPV)	0.911764706	0.9375	0.856538462	0.9375	0.870533558
NPV	0.571428571	0.555555556	1	0.555555556	0.601666667
F1	0.911764706	0.909090909	0.922695298	0.909090909	0.919842303
AUC	0.741596639	0.798319328	0.572857143	0.798319328	0.711092437
<b>Validación cruzada</b>					
Score (mean)	0.763216196	<b>0.740883626</b>	0.696969697	<b>0.728762414</b>	0.596781622
Score (variance)	0.052144589	<b>0.087465814</b>	0.080349399	<b>0.087075434</b>	0.106044667



*Figura 28.* Comparación de los algoritmos en el tercer experimento  
Fuente: Elaboración propia



*Figura 29.* Resultados de la validación cruzada del tercer experimento.  
Fuente: Elaboración propia

De igual forma, en cuanto al área bajo la curva del experimento 3, el algoritmo con la mejor área fue el de máquina de soporte vectorial y el de regresión logística ambos con una predicción de 0.7983%. Ver figura 30

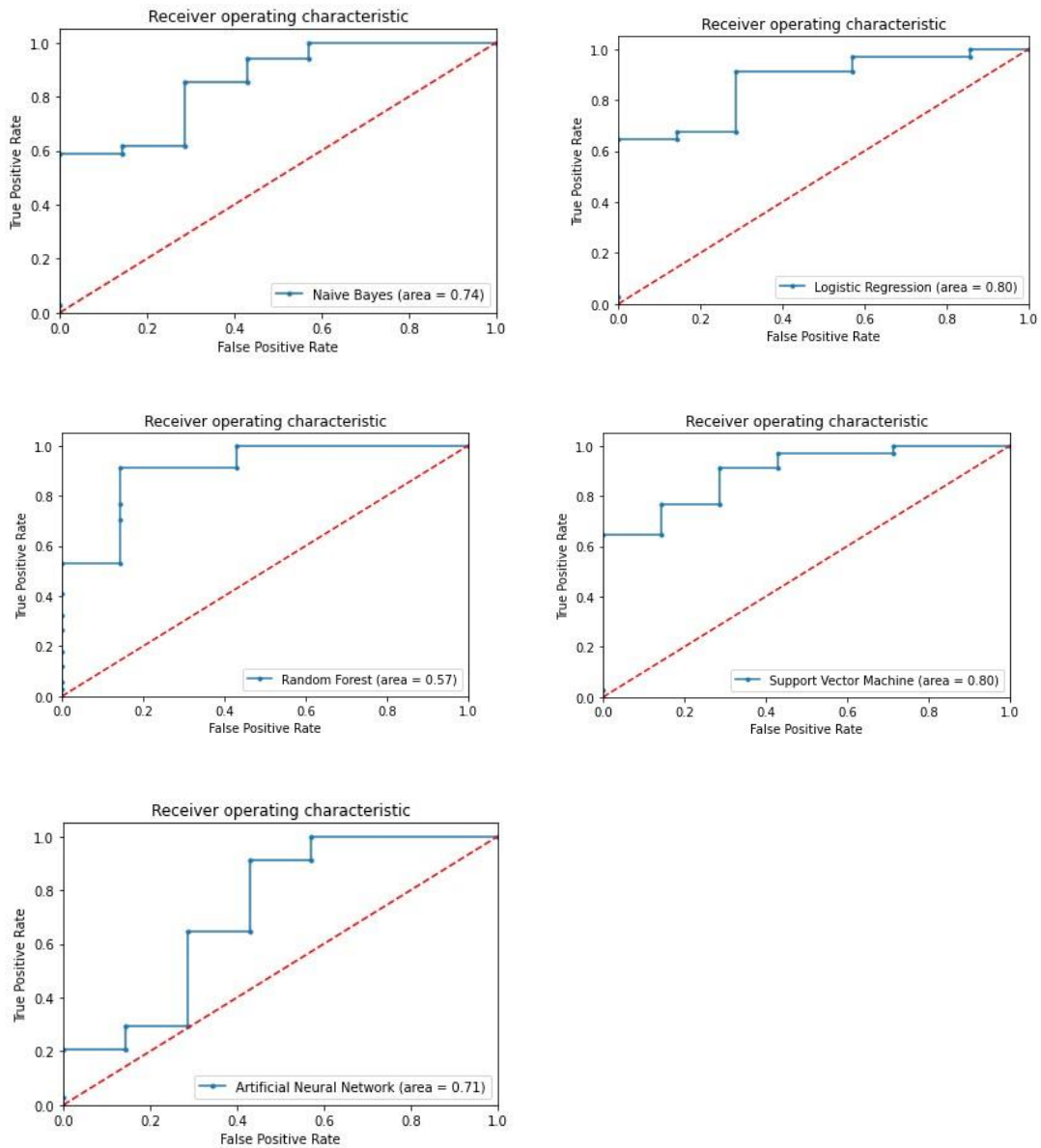


Figura 30. Área bajo la curva de los algoritmos: regresión logística, Naive Bayes, Random forest, Máquina de soporte vectorial, y redes neuronales, en el tercer experimento.

Mediante la prueba de Kruskal Wallis se comprobó que existe diferencias significativas al comparar el área bajo la curva de cada grupo de algoritmos ( $\lambda^2=44.706$ ,  $p=0.000$ ).

La tabla 15 y figura 31, 32, se muestran las métricas de los diferentes algoritmos de clasificación obtenidos en el cuarto experimento. El algoritmo con mejor accuracy es el algoritmo de máquina de soporte vectorial (89.55%) al validar la métrica con validación cruzada se obtuvo 89.53% y la varianza fue de 0.0364. El algoritmo con

mejor métrica de sensibilidad es random forest (92.35%) y el algoritmo con mejor especificidad es el de máquina de soporte vectorial (96.97%).

Tabla 15. Comparación de los algoritmos en el cuarto experimento

Métricas de rendimiento de los algoritmos de clasificación	Algoritmos				
	Naive Bayes	Regresión Logística	Random Forest	Máquina de soporte vectorial	Redes neuronales artificiales
	Accuracy	0.731343284	0.880597015	<b>0.901492537</b>	<b>0.895522388</b>
Accuracy balanceado	0.730392157	0.881461676	0.901158645	0.896613191	0.818048128
Prevalence	0.507462687	0.507462687	0.507462687	0.507462687	0.507462687
Sensitivity (Recall)	0.794117647	0.823529412	<b>0.923529412</b>	<b>0.823529412</b>	0.808823529
Especificidad	0.666666667	0.939393939	<b>0.878787879</b>	<b>0.96969697</b>	0.827272727
Precision (PPV)	0.710526316	0.933333333	0.886984127	0.965517241	0.829997876
NPV	0.75862069	0.837837838	0.917943548	0.842105263	0.810162204
F1	0.75	0.875	0.90484472	0.888888889	0.817846151
AUC	0.730392157	0.881461676	0.911158645	0.896613191	0.848048128
<b>Validación cruzada</b>					
Score (mean)	0.775531434	0.901221167	0.9461782	<b>0.895251018</b>	0.607869742
Score (variance)	0.06016357	0.038400593	0.063006916	<b>0.036452634</b>	0.075312015

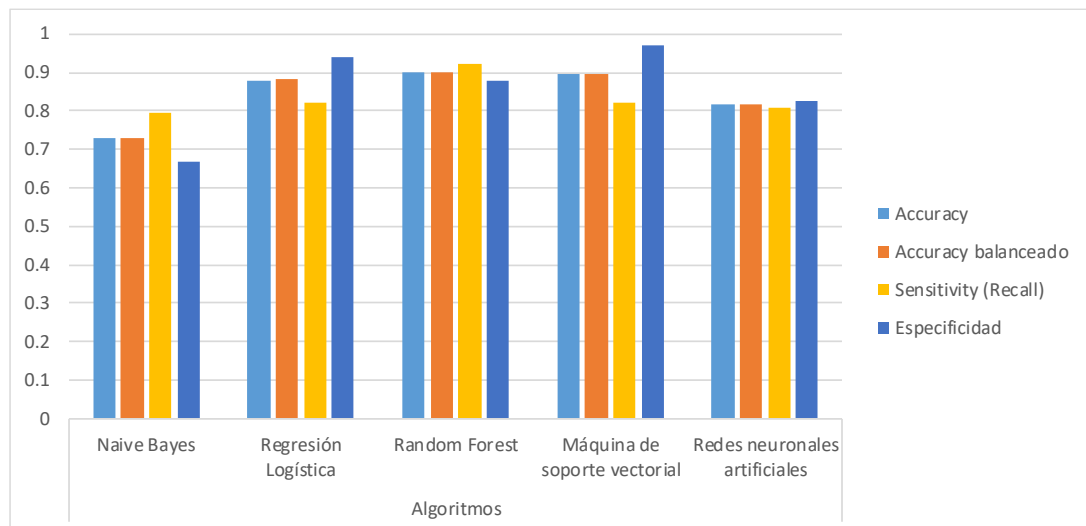
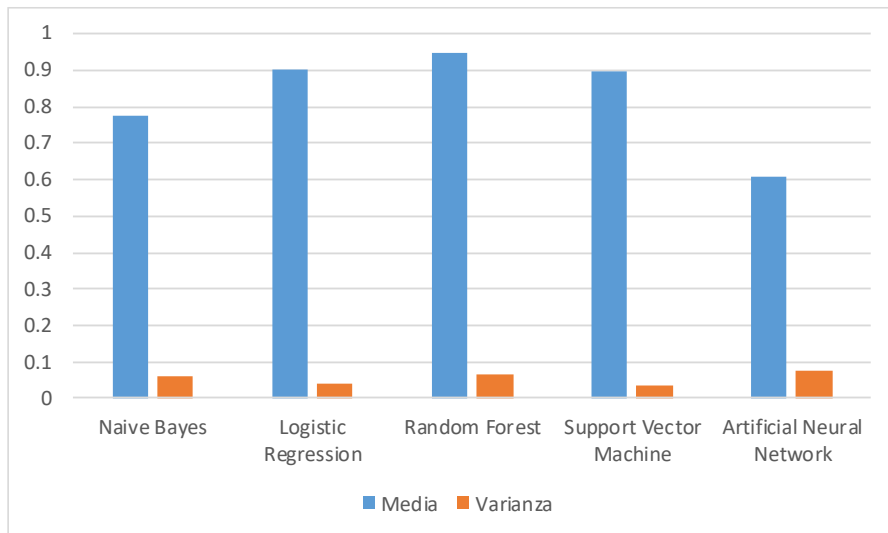


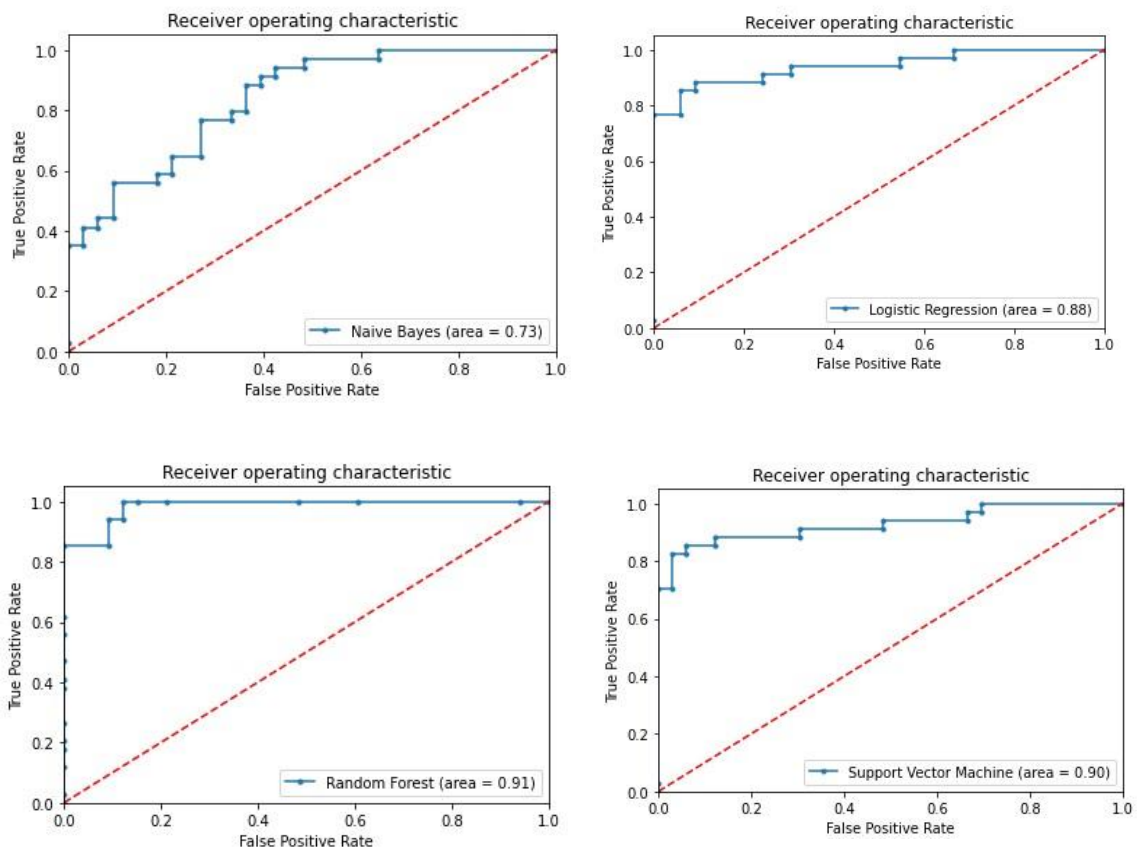
Figura 31. Comparación de los algoritmos en el cuarto experimento

Fuente: Elaboración propia



*Figura 32.* Resultados de la validación cruzada del cuarto experimento.  
Fuente: Elaboración propia

En cuanto al área bajo la curva del experimento, el algoritmo de random forest y el de máquina de soporte vectorial tuvieron la mejor área con 91% y 90% respectivamente. Ver Figura 33.



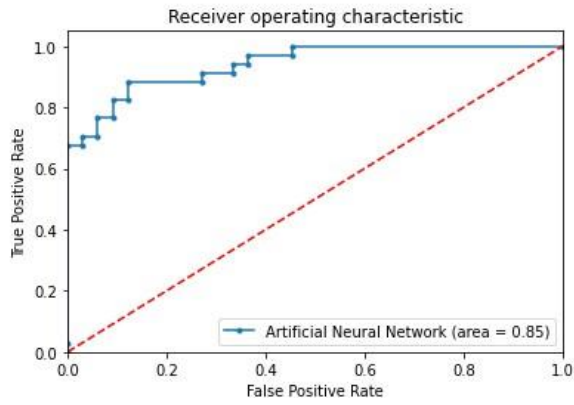


Figura 33. Área bajo la curva de los algoritmos: regresión logística, Naive Bayes, Random forest, Máquina de soporte vectorial, y redes neuronales, en el cuarto experimento.

Mediante la prueba de Kruskal Wallis se comprobó que existe diferencias significativas al comparar el área bajo la curva de cada grupo de algoritmos ( $\lambda^2=45.495$ ,  $p=0.000$ ).



## 4.2 Análisis de resultados o discusión de resultados

Este trabajo demuestra un enfoque de la utilidad de técnicas de aprendizaje automático para clasificación y predecir resultados de la radiocirugía estereotáctica en la curación de las MAV cerebrales. Debido a la edad relativamente joven en que esta enfermedad es diagnosticada y a los efectos catastróficos que siguen a su ruptura, hay un genuino interés en predecir los resultados de curación de los tratamientos a las MAVs tratadas con radiocirugía.

En esta tesis, se encontró que las edades de la mayoría de pacientes con malformaciones arteriovenosas estaban comprendidas en el rango de edades entre 12 a 60 años (82%) y que la gran parte eran del sexo masculino (53%). Esto coincide con estudios tanto nacionales como internacionales que mencionan que gran porcentaje de pacientes portadores de malformaciones arteriovenosas cerebrales son de edades comprendidas entre 12 a 60 años de edad y del sexo masculino (Sánchez, 2017; Larios, y otros, 2010). Por otro lado, en esta tesis se encontró que la localización más frecuente fue en los ganglios basales (16.83%) lo que no concuerda con el estudio de Sánchez (2017), donde encontró que la malformación arteriovenosa cerebral más frecuente fue en la parte frontal izquierda. Además, de la localización resalta la ubicación profunda de la malformación arteriovenosa cerebral (47.52%). Esto se podría explicar porque la gran mayoría de pacientes con malformaciones arteriovenosas cerebrales que son tratados con radiocirugía tiene localizado sus malformaciones en zonas de difícil acceso para realizarse la cirugía convencional abierta.

A pesar de que existen escasas fuentes teóricas en relación al empleo de la técnica de aprendizaje automático en el tratamiento de las malformaciones arteriovenosas cerebrales con radiocirugía para poder contrastar los resultados más ampliamente, la cual es una limitación, nuestro estudio pretende demostrar que la aplicación del aprendizaje automático obtiene resultados que pueden utilizarse en el área médica. Por lo cual nuestra investigación realizó cuatro experimentos:

En el experimento 1 se observó que el mejor modelo, considerando el accuracy balanceado, es el de regresión logística (72.69%) seguido de redes neuronales artificiales (72.25%) sin embargo en la validación cruzada los resultados para redes neuronales obtienen valores demasiado bajos como para poder utilizarlo de forma confiable (promedio accuracy 46%). Considerando la sensibilidad, el mejor modelo es

el de random forest seguido de redes neuronales, sin embargo, la especificidad en ambos modelos son demasiados bajos lo cual descartaría un uso óptimo de estos modelos. La especificidad es mejor en el modelo naive bayes, sin embargo, este modelo no tiene valores adecuados para la sensibilidad y el accuracy balanceado. Finalmente, en este primer experimento ninguno de los modelos analizados tiene métricas óptimas para el accuracy, sensibilidad y especificidad a la vez. En cuanto el área bajo la curva, el algoritmo de regresión logística presentó la mejor métrica, AUC de 72.69%, la cual se considera como predicción aceptable.

Asimismo, en el experimento 2, se encontró que el mejor modelo, teniendo en cuenta el accuracy es el de máquina soporte vectorial (98.51%) seguido del algoritmo de random forest (96.57%), además en la validación cruzada también se confirmó que el modelo de máquina de soporte vectorial consiguió un accuracy alto (97.91%) a la vez que mostró estabilidad (varianza de 0.022). En cuanto a la sensibilidad el mejor modelo es el de máquina de soporte vectorial (100%) seguido de random forest (99%), asimismo, teniendo en cuenta la especificidad, el modelo de máquina de soporte vectorial presentó el mejor valor (96.97%) y el de random forest presentó un valor aceptable (93.94%) aunque el modelo de regresión logística presentó el mejor valor de especificidad (96.97%), pero este modelo no tiene valores adecuados para la sensibilidad y el accuracy. Finalmente, entre los modelos analizados en este segundo experimento el modelo de máquina de soporte vectorial es el que presentó las métricas óptimas tanto para el accuracy, sensibilidad y especificidad a la vez, seguido del modelo de random forest. Asimismo, en cuando al área bajo la curva, el algoritmo máquina de soporte vectorial es considerado que tuvo una predicción sobresaliente con 98.48%

En el experimento 3, se observó que los mejores modelos, considerando el accuracy balanceado, es el de regresión logística y máquina de soporte vectorial (79.83%) pero en la validación cruzada los resultados para la regresión logística obtienen valores bajos (promedio accuracy 74.09% y varianza 0.087). Considerando a la sensibilidad, el mejor modelo es el de redes neuronales (97.65%) seguido del modelo de regresión logística (88.24%). Asimismo, la especificidad es mejor en el modelo de regresión logística (71.43%) aunque sus valores son bajos. Sin embargo, el accuracy, accuracy balanceado, la especificidad, sensibilidad, en el modelo de redes neuronales y en la regresión logística son demasiados bajos, lo cual descartaría su uso. Finalmente, en este tercer experimento ninguno de los modelos analizados tiene métricas óptimas

para el accuracy, sensibilidad y especificidad a la vez. En cuando al comparar las predicciones de los algoritmos teniendo en cuenta el área bajo la curva el algoritmo de regresión logística tuvo una predicción aceptable mostrando un área de 79.83%.

En el experimento 4, se encontró que el mejor modelo, teniendo en cuenta el accuracy es el random forest (90.15%) seguido del algoritmo de máquina de soporte vectorial (89.55%), en la validación cruzada se confirmó que el modelo de máquina de soporte vectorial consiguió los valores óptimos (promedio de accuracy 89.53% y varianza de 0.036). En cuanto a la sensibilidad el mejor modelo es el de random forest (92.35%), asimismo, teniendo en cuenta la especificidad, el modelo de máquina de soporte vectorial presentó el mejor valor (96.97%). En este cuarto experimento, ninguno de los modelos analizados tiene métricas óptimas para el accuracy, sensibilidad y especificidad a la vez. Comparando las predicciones de los algoritmos en cuanto al área bajo la curva del experimento, el algoritmo de random forest y el de máquina de soporte vectorial tuvieron la mejor área con 90.11% y 89.66% respectivamente

Finalmente, de los cuatro experimentos evaluados en esta investigación, el segundo experimento obtuvo la mejor métrica para la predicción con el algoritmo máquina de soporte vectorial. (accuracy de 98.51%) y área bajo la curva de 98.48%, considerándose como una predicción sobresaliente, sin embargo, este resultado no coincide con el estudio de Oermann, y otros (2016), que encontraron que el mejor modelo fue el modelo de regresión logística con accuracy de 74%, así como los resultados son también diferentes con el estudio de Asadi, y otros, (2016) donde se encontró que el mejor modelo fueron el de redes neuronales ( $97.5\% \pm 2.5\%$ ), aunque en este estudio la predicción fue de la mortalidad de pacientes con malformaciones arteriovenosas cerebrales.

Además, se puede considerar que el resultado tiene mejores métricas en el segundo experimento porque se efectuó el balanceo de las clases y según Srinivasan Perumbakkam & Kohl (2018), este desbalance de las clases es un problema común con las bases de datos clínicos más aún cuándo la prevalencia de la enfermedad es bastante baja, por lo que es importante en datos médicos tratar este desbalance. Aunque no se encontraron estudios de malformaciones arteriovenosas cerebrales donde se aborda el desbalanceo de las clases podemos mencionar que se analizó estudios aplicados a la medicina donde se estudio este tema, los resultados mostraron superioridad en términos

de especificidad, sensibilidad y precisión, entre otros. (Alahmari, 2020; Mechelli & Vieira, 2019; Razzaghi, Safro, Ewing, Sadrfaidpour, & Scott, 2019).

## CONCLUSIONES Y RECOMENDACIONES

### Conclusiones:

1. Los predictores que ayudan a predecir la curación de las MAVs son: las variables sociodemográficas (edad), clínicas (hemorragia, tipo de hemorragia, cefalea, convulsiones, encefalomalasia), tratamiento previo a la radiocirugía (cirugía previa, embolización, cirugía y embolización, puntaje Spetzler Martin), angioarquitectura (diámetro, localización, aneurisma venoso, dolicoectasia, profundidad, tipo de hemorragia, número de vasos eferentes, tipo de drenaje venosos, velocidad de flujo, estenosis venosa, puntaje de Virginia), física (dosis, número de isocentros).

2. Entre los algoritmos de aprendizaje automático comparados, el que obtuvo las mejores métricas es el algoritmo de máquina de soporte vectorial y pertenece al segundo experimento en cuanto al accuracy, la sensibilidad y especificidad fue de 98.51%, 100% y 96.97.

En este experimento se usó los 45 predictores y se balanceó las clases con oversampling. Además, presenta un área bajo la curva de 98.48%, la cual indica que es una predicción sobresaliente para la predicción de pacientes curados tratados con radiocirugía, esta área es significativo ( $p = 0.000$ ).

3. Aplicar técnicas de oversampling mejora las métricas de los modelos de aprendizaje automático, debido a que puede estabilizar la performance de la clasificación. Esta mejora en la predicción se aprecia considerablemente en esta tesis (experimento 2 y 4). Así en el experimento 4, el algoritmo con mejor accuracy fue el de máquina de soporte vectorial (89.55%) y el área bajo la curva fue el de los algoritmos de random forest y el de máquina de soporte vectorial tuvieron la mejor área con 91% y 90% respectivamente, siendo esto significativo ( $p = 0.000$ ).

4. Las predicciones de la curación de la malformación a través de los algoritmos de aprendizaje automático permiten obtener resultados que podrían tomarse en cuenta para tomar mejores decisiones en el tratamiento de esta enfermedad.

5. La aplicación de los métodos de aprendizaje automático a datos médicos constituye un nuevo paradigma de investigación mediante el preprocesamiento de datos, selección de características que ayuden a generar percepciones clínicas e idealmente acciones y decisiones que pueden ser implementados en la práctica clínica.
6. A pesar de tener pocos datos, como en este estudio, los algoritmos de aprendizaje automático también son aplicables al área médica.

#### Recomendaciones:

1. Aunque las malformaciones arteriovenosas cerebrales es una enfermedad poco frecuente, se hace necesario tener estudios donde se incluyan otras variables que ayudan a la predicción de su curación para poder hacer comparaciones y obtener resultados más confiables.
2. Es importante analizar cuidadosamente cada variable que se retiran en el proceso de feature engineering debido a que puede tener una implicancia negativa en los resultados finales que se obtienen si es que “no se utiliza criterios adecuados” en la selección de los predictores, esto es especialmente importante cuando el conjunto de datos es pequeño.
3. Actualmente, algunos modelos de aprendizaje automático son considerados “caja negra” sería importante realizar estudios con técnicas de interpretación y de explicabilidad del comportamiento de los algoritmos de aprendizaje automático para identificar los predictores mas importantes que explican la curación de las MAVs que ayudarían a optimizar las decisiones clínicas de forma precisa.
4. Se recomienda realizar estudios donde se comparen técnicas de balanceo para posteriormente evaluar la predicción de la curación de las MAVs.
5. A pesar de tener pocos datos en este estudio de medicina se recomienda aplicar estos modelos de aprendizaje automático y compararlos con métodos tradicionales, lo cual ayuda a tomar decisiones en el área de la medicina.
6. Se recomienda que los algoritmos de aprendizaje automático sean implementados en las guías de tratamiento de MAVs cerebrales para predecir la curación de las MAVs cerebrales mediante la automatización de los procesos a partir de los datos clínicos y de la técnica de aplicación de la radiocirugía para lograr un flujo de atención eficaz y personalizado en la curación de esta enfermedad.

## REFERENCIAS BIBLIOGRÁFICAS

- Alahmari, F. (2020). A comparison of resampling Techniques for medical Data Using Machine Learning. *Journal of Information & Knowledge Management*, 19(1).
- Aminoff, M. J., Boller, F., & Swaab, D. F. (2017). *Handbook of Clinical Neurology* (Vol. 143). United Kingdom: Elsevier.
- Asadi, H., Kok, F., Looby, S., Brennan, P., O'Hare, A., & Thornton, J. (2016). Outcomes and Complications following Endovascular Treatment of Brain Arteriovenous Malformations-A Prognostication Attempt using Artificial Intelligence. *World Neurosurgery*.
- Badakhshi, H. (2016). *Image- Guided Stereotactic Radiosurgery*. Switzerland: Springer International Publishing.
- Berkhin, P. (2006). A survey of clustering data mining techniques. En P. Berkhin. Grouping multidimensional data.
- Bradac, O., Charvat, F., & Benes, V. (2013). Treatment for brain arteriovenous malformation in the 1998-2011 period and review of the literature. *Acta Neurochir*, 155:199-209.
- Breiman, L. (2001). Random Forest. *Machine learning*, 45, 5-32.
- Caicedo, E. F., & López, J. A. (2017). *Una aproximación práctica a las redes neuronales artificiales*. Programa editorial Universidad del Valle, edición digital.
- Çelik, Ö., & Altunaydin, S. S. (2018). A research on Machine Learning Methods and Its Applications. *Journal of Educational Technology & Online Training*, 1(3), 25-40.
- Consejo, A., Melcer, T., & Rozema, J. (2018). Introduction to Machine Learning for Ophthalmologists. *Seminars in Ophthalmology*.
- Ding, D., Yen, C. P., Starke, R. M., Xu, Z., Sun, X., & Sheehan, J. P. (2014). Outcomes following single-session radiosurgery for high-grade intracranial arteriovenous malformations. *British journal of neurosurgery*, 28(5), 666-674.

- Drake, C. (1979). Cerebral Arteriovenous Malformations: Considerations for and Experience with Surgical Treatment in 166 cases. *Neurosurgery*(26 (CN-suppl-1)), 145-208.
- Du, M., Liu, N., & Hu, X. (2020). Techniques for interpretable Machine Learning. *Communications of the ACM*, 63(1).
- Dumont, T. M., Kan, P., Snyder, K. V., Hopkins, N. L., Siddiqui, A. H., & Levy, E. I. (2015). A proposed grading system for endovascular treatment of cerebral arteriovenous malformations: Buffalo score. *Surgical Neurology International*, 6-3.
- El Napa, I., Li, R., & Murphy, M. J. (2015). *Machine Learning in Radiation Oncology*. Switzerland: Springer.
- Elshawi, R., Al-Mallah, M., & Sakr, S. (2019). On the interpretability of machine learning based model for predicting hypertension. *BMC Medical Informatics and Decision Making*, 19(146).
- Flickinger, J. C., Kondziolka, D., Lunsford, L. D., Pollock, B. E., Yamamoto, M., Gorman, D. A., & McDermott, M. W. (1999). A multi-institutional analysis of complication outcomes after arteriovenous malformation radiosurgery. *International Journal of Radiation Oncology\*Biophysics*, 44(1), 67-74.
- Flickinger, J. C., Kondziolka, D., Maitz, A. H., & Lunsford, L. D. (1998). Analysis of neurological sequelae from radiosurgery of arteriovenous malformations: how location affects outcome. *International Journal of Radiation\*Biophysics*, 40(2), 273-278.
- Florestean, N., & Ciurea, A. (2012). Anatomical and clinical aspects of cerebral arteriovenous malformations. *Romanian Journal of Neurology*, XI(4).
- Friedman, W. A., & Bova, F. J. (2011). Radiosurgery for arteriovenous malformations. *Neurological research*, 803-819.
- García, J., Molina, J. M., Berlanga, A., Patricio, M. A., Bustamante, A. L., & Padilla, W. R. (2018). *Ciencia de Datos. Técnicas analíticas y aprendizaje estadístico*. Ciudad de México, México: Alfa y omega.

- García, S., Luengo, J., & Herrera, F. (2015). *Data Preprocessing in Data mining. Intelligent Systems Reference Library 72*. Springer.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of Statistical Learning. Data Mining, Inference and Prediction*. Springer.
- Hernández, J., Ramirez Quintana, J., & Ferri Ramirez, C. (2005). *Introducción a la minería de datos*. Madrid: Pearson.
- Hernesniemi, J., & Reza, D. (2008). Natural history of brain arteriovenous malformations: a long-term follow-up study of risk of hemorrhage in 238 patients. *Neurosurgery*, 63(5), 823-29.
- Hosmer, D., Lemeshow, S., & Sturdivant, R. (2013). *Applied Logistic Regression*. Wiley.
- Javalkar, V., Pillai, P., Vannemreddy, P., Caldito, G., Ampil, F., & Nanda, A. (2009). Gamma knife radiosurgery for arteriovenous malformations located in eloquent regions of the brain. *Neurology India*, 57(5), 617.
- Jurek, A., BI, Y., Wu, S., & Nugent, C. (2013). A survey of commonly used ensemble-based classification techniques. *The knowledge Engineering Review*, 29(5), 551-581.
- Kandai, S., Abdullah, M., & Naing, N. (2010). Angioarchitecture of brain arteriovenous malformations and the risk of bleeding: an analysis of pacientes in Northeastern Malasya. *The Malaysian journal of medical sciences: MJMS*, 17(1), 44-48.
- Kang, M., & Tian, J. (2018). Machine Learning: Data Pre-processing. En M. G. Pecht, & M. Kang, *Prognostics and health Management of electronics* . John Wiley & Sons.
- Knox, S. (2018). *Machine Learning: a concise introduction*. Wiley.
- Kubben, P., Dumontier, M., & Dekker, A. (2019). *Fundamentals of Clinical Data Science*. Switzerland: Springer.
- Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. New York, USA: Springer.



- Kwon, Y., Jeon, S. R., Kim, J. H., Lee, J. K., Ra, D. S., Lee, D. J., & Kwun, B. D. (2000). Analysis of the causes of treatment failure in gamma knife radiosurgery for intracranial arteriovenous malformations. *Journal of neurosurgery*, *93*, 104-106.
- Lantz, B. (2015). *Machine learning with R*. Birminhan, UK: Packt Publishing.
- Larios, C., Mimenza, A. A., Zenteno, M., Murillo-Bonilla, L., Sanchez-Hernandez, J., Garcia-Ramos, G., & Cantu-Brito, C. (Septiembre-Octubre de 2010). Comportamiento de las malformaciones arteriovenosas cerebrales: estudio de una cohorte de 302 pacientes. *Revista Mexicana de Neurociencia*, *11*(5), 330-337.
- Lawton, M. T., Rutledge, C., Kim, H., Stapf, C., Whitehead, K. J., Li, D. Y., . . . Kondziolka, D. (2015). Brain arteriovenous malformations. *Nature Reviews Disease Primers*. doi:doi.org/10.1038/nrdp.2015.47
- Lee, E. K. (2018). Machine learning for early detection and treatment outcome prediction. En N. Kong, & S. Zhang, *Decision Analytics and Optimizacion in diseases Prevention and Treatment*. USA: Wiley.
- Mechelli, A., & Vieira, S. (2019). *Machine Learning. Methods and Applications to Brain Disorders*. London, United Kingdom: Elsevier.
- Miranda Hernandez, A. (August de 2016). Applications of Engineering Principles with a Comparison of Machine Learning Classification Methods to Predict Treatment Outcomes in Head and Neck Cancer Patients. *Thesis for degree of Doctor of Engineering*. George Washington University.
- Mohr, J., Kejda-Scharler, J., & Pile-Spellman, j. (2013). Diagnosis and Treatment of Aeteriovenous Malformations. *Curr Neurolo Neurosci Rep*, *13*, 324.
- Moorthy, R., & Rajshekhar, V. (2015). Stereotactic radiosurgery for intracranial arteriovenous malformations: A review. *Neurol India*, *63*, 841-51.
- Nunes da Silva, I., Hernane, D., Andreda, R., Bartocci, L., & Franco, S. (2017). *Artificial Neural Networks. A practical course*. Springer.

- Oermann, E., Rubinsteyn, A., Ding, D., Mascitelli, J., Starke, R., Bederson, J., & Kondziolka, D. (2016). Using a Machine Learning Approach to Predict Outcomes after Radiosurgery for Central Arteriovenous Malformations. *Scientific reports*, 6.
- Panesar, S. S., Souza, R. N., Yeh, F.-C., & Fernández-Miranda, J. C. (2019). Machine Learning Versus Logistic Regression Methods for 2-Year Mortality Prognostication in a small, Heterogeneous Glioma Database. *World Neurosurgery*.
- Parkhutik Matveera, V. (2014). Radiocirugía estereotáxica en el tratamiento de malformaciones arteriovenosas cerebrales. *Tesis Doctoral*. Barcelona: Universidad Autónoma de Barcelona. Facultad de Medicina. Obtenido de <https://tdx.cat/bitstream/handle/10803/283352/vpm1de1.pdf?sequence=1&isAllowed=y>
- Pérez-Cruz, P. E., & Acevedo, F. C. (2014). Escalas de estado funcional (o performance status). *Gastroenterología Latinoamericana*, 25(3), 219-226.
- Pollock, B. E., Storlie, C. B., Link, M. J., Stafford, S. L., Garces, Y. I., & Foote, R. L. (8 de April de 2016). Comparative and analysis of arteriovenous malformation grading scales in predicting outcomes after stereotactic radiosurgery. *Journal Neurosurgery*.
- Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *The New England Journal of Medicine*, 380(14), 1347-1358.
- Raschka, S. (2015). *Python Machine Learning. Unlockdeeper insights into machine learning with this vital guide to cutting-edge predictive analytics*. Birmingham, United Kingdom: Packt Publishing Ltd.
- Razzaghi, T., Safro, I., Ewing, J., Sadrfaidpour, E., & Scott, J. (2019). Predictive models for bariatric surgery risks with imbalanced medical datasets. *Annals of Operations Research*.
- Reponen, E., Tuominen, H., & Korja, M. (2014). Evidence for the use of preoperative risk Assessment scores inelective cranial neurosurgery:A systematic review of the literature. *Anesthesia & Analgesia*, 119(2), 420-432.

- Rosell, A., Rocca, U., Chavez, L., Dávila, A., & Palacios, F. (2004). Tratamiento quirúrgico de las malformaciones vasculares arteriovenosas. *Revista de Neuropsiquiatría*, 64(4).
- Saber, H., Somai, M., Rajah, G., Scalzo, F., & Liebeskind, D. (2019). Predictive analytics and machine learning in stroke and endovascular. *Neurological Research*.
- Samblas, J., Sallabanda, K., Martínez, R., & Calvo, F. (2012). *Radiocirugía. Fundamentos, avances tecnológicos, indicaciones y resultados*. Madrid, España: Arán.
- Sami, K. (2013). Predictive Analytics and Interpretation of Brain Aneurysm. *Thesis the degree of masters*. (B. U. City, Ed.)
- Sánchez, E. D. (2017). Características epidemiológicas y clínicas de los pacientes con malformaciones arteriovenosas cerebrales wque recibieron radiocirugía estereotáxica en el centro integral de radioncología del hospital Militar Central durante el período 2013-2016. *Tesis*. Perú: Universidad Privada San Juan Bautista. Obtenido de <http://repositorio.upsjb.edu.pe/bitstream/handle/upsjb/538/T-TPMC-Edinson%20David%20S%C3%A1nchez%20Ferreira.pdf?sequence=1&isAllowed=y>
- Senders, J. T., Staples, P. C., Karhade, A. V., Zaki, M. M., Gormley, W. B., Broekman, M. L., . . . Arnaout, O. (2017). Machine Learning and Neurosurgical outcomes Prediction: A Systematic Review. *World Neurosurgery*.
- Shobha, G., & Rangaswamy, S. (2018). Machine Learning. En V. N. Gudivada, & C. R. Rao, *Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications*. North-Holland: Elsevier.
- Sidey-Gibbons, J., & Sidey-Gibbon, C. (2019). Machine learning in medicine: a practical introduction. *BMC Medical Research Methodology*.
- Solomon, R. A., & Connolly, S. (2017). Arteriovenous Malformations of the Brain. *The New England Journal of Medicine*, 1859-66.

- Spetzler, R., Kondziolka, D., Higashida, R., & Kalani, Y. (2015). *Comprehensive Management of Arteriovenous Malformations of the brain and Spine*. United Kingdom: Cambridge Medicine.
- Srinivasan Perumbakkam, S., & Kohl, M. (2018). Statistical Learning in Precision Medicine. En H.-P. Deigner, & M. Kohl, *Precision Medicine. Tools and Quantitative Approaches*. Academic Press.
- Subbulakshmi, C., & Deepa, S. (2015). Medical dataset classification: a machine learning paradigm integrating particle swarm optimization with extreme learning machine classifier. *The Scientific World Journal*.
- Uddin, S., Khan, A., Hossain, E., & Ali Moni, M. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*, 19(281).
- Van Beijnum, H., & Van der Worp. (2011). Treatment of brain arteriovenous malformations: a systematic review and meta-analysis. *JAMA*, 306(18), 2011-2019.
- Véliz Capuñay, C. (2018). *Aprendizaje automático*. Lima: PUCP.
- Viñuela, F. (2004). Neurroradiología intervencionista en malformaciones arteriovenosas cerebrales y dures Cap 32. En J. Mercader, & F. Viñuela, *Neurroradiología Diagnóstica y Terapéutica*. Elsevier Masson.
- Yahyaoui, A., & Yahyaoui, I. (2018). Machine learning Techniques for data classification. En I. Yahyaoui, *Advances in Renewable energies and power Technologies*. Elsevier.
- Yañez, A. (2001). Malformaciones Arteriovenosas encefálicas. *Revista chilena de Neuropsiquiatria*, 53-54.
- Yoo, I., Alafairet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R., Chang, J., & Hua, L. (2012). Data mining in healthcare and biomedicine: A survey of the literature. *Journal of medical Systems*, 36(4), 2431-3448.

## ANEXOS

Anexo 1 Declaración de autenticidad

Anexo 2 Autorización de consentimiento para realizar la investigación

Anexo 3: Matriz de consistencia

Problema	Objetivo	Hipótesis	VARIABLE	INDICADOR
<b>General</b>	<b>General</b>	<b>General</b>		
¿Es posible predecir la curación de las malformaciones arteriovenosas cerebrales tratadas con radiocirugía utilizando algoritmos de clasificación de aprendizaje automático?	Predecir la curación de las malformaciones arteriovenosas cerebrales en pacientes tratados con radiocirugía utilizando algoritmos de clasificación de aprendizaje automático.	Los algoritmos de clasificación de aprendizaje automático predicen de forma adecuada la curación de las malformaciones arteriovenosas cerebrales en pacientes tratados con radiocirugía.	<b>Dependiente:</b> La curación de la malformaciones arteriovenosas cerebrales.	Curado y No Curado
<b>Específico</b>	<b>Específico</b>	<b>Específico</b>		
¿Cuáles son las variables sociodemográficas, clínicas, de angioarquitectura y del procedimiento de la radiocirugía que ayudarán a predecir la curación de los pacientes con diagnóstico de malformación arteriovenosa cerebral tratados con radiocirugía?	Identificar las variables sociodemográficas, clínicas, de angioarquitectura y del procedimiento de la radiocirugía de las malformaciones arteriovenosas cerebrales. que ayudarán a predecir la curación de los pacientes con malformación arteriovenosa cerebral.	Las variables sociodemográficas, clínicas, de angioarquitectura y del procedimiento de la radiocirugía son un grupo de variables que ayudarán a predecir la curación de los pacientes con diagnóstico de malformación arteriovenosa cerebral tratados con radiocirugía.	<b>Independientes:</b> Algoritmos de clasificación de aprendizaje automático .	Indicadores: área bajo la curva (AUC), Accuracy, Sensibilidad, Especificidad.
¿Es posible que los algoritmos de Naive bayes, regresión logística, máquina de soporte vectorial, random forest y redes neuronales con oversampling tengan mejores resultados que los algoritmos de Naive bayes, regresión logística, máquina de soporte vectorial, random forest y redes neuronales sin oversampling en el pronóstico de la curación de los pacientes con malformación arteriovenosa cerebral tratados con radiocirugía?	Determinar la predicción de los algoritmos de Naive Bayes, regresión logística, random forest, máquina de soporte vectorial y redes neuronales usando todas las variables, las variables más importantes, con oversampling y sin oversampling y elegir el mejor resultado en la predicción de la curación de los pacientes con malformación arteriovenosa cerebral tratados con radiocirugía.	Los algoritmos Naive bayes, regresión logística, máquina de soporte vectorial, random forest y redes neuronales con oversampling tiene mejores resultados que los algoritmos de regresión logística, Naive Bayes, random forest, redes neuronales sin oversampling en el pronóstico de la curación de los pacientes con malformación arteriovenosa cerebral tratados con radiocirugía.	Variables sociodemográficas, clínicas, de angioarquitectura y del procedimiento de la radiocirugía	