



**UNIVERSIDAD
DE ANTIOQUIA**

**DESARROLLO DE UN MODELO PREDICTIVO
CON INTELIGENCIA ARTIFICIAL PARA
ESTABLECER CLASIFICACIÓN ASA A
PACIENTES EN UNA CONSULTA
PREANESTÉSICA**

Autor

Erika Tatiana Arias Zuluaga

Universidad de Antioquia

Facultad de Ingeniería

Medellín, Colombia

2020



Desarrollo de un modelo predictivo con inteligencia artificial para establecer clasificación
ASA a pacientes en una consulta preanestésica.

Erika Tatiana Arias Zuluaga

Tesis o trabajo de investigación presentada(o) como requisito parcial para optar al
título de:
Bioingeniero

Asesora:
Jenny Kateryne Aristizábal Nieto
MSc Ingeniería Biomédica - Bioingeniera

Universidad de Antioquia
Facultad de Ingeniería.
Medellín, Colombia
2020.

AGRADECIMIENTOS

A mi madre por todo su apoyo durante el proceso de formación que finaliza en este proyecto.
A mis abuelos por todo el amor que me han brindado.

A la docente Jenny Aristizábal Nieto por ser mi guía no solo en este trabajo de grado, si no a través de toda mi formación superior, es de quién más adquiriré el conocimiento en mi pregrado.

Finalmente, a todos los participantes en el proyecto que aportan sus ideas y realimentan para llevar a cabo los objetivos.

RESUMEN

Dentro de las necesidades presentes en el Hospital Universitario San Vicente Fundación en el área de anestesiología se hace necesario el desarrollo de un modelo de inteligencia artificial que facilite la clasificación del riesgo preoperatorio de algún paciente que esté programado para alguna intervención quirúrgica, además con el objetivo de usar esta clasificación para realizar una base de datos global dentro del hospital, en la cual sea verídico la clasificación otorgada por el programa y que sea una base de apoyo para el personal médico, y poder implementar un protocolo de consultas preanestésicas a distancia, se realizó un análisis de los datos y la construcción de una base de datos propia, con el fin de entrenar dos modelos de inteligencia artificial (Clasificadores de Bayes y árboles de clasificación), para clasificar con base en los parámetros recolectados y establecidos en una consulta preanestésica, el riesgo que plantea la anestesia para los distintos estados del paciente, a través de la escala ASA (American Society of Anesthesiologists).

Los modelos implementados para las bases de datos diseñadas se desarrollaron en el entorno de desarrollo de Python. y los resultados obtenidos del clasificador bayesiano y del árbol de clasificación, demuestran que aún se deben ajustar ciertos parámetros y anexar mayor información, también se debe hacer un ajuste a partir de un análisis con una nueva base de datos para asegurar que sea factible la implementación de dichos modelos para realizar consultas preanestésicas a distancia a los pacientes del hospital y estandarizar un programa de clasificación ASA para todo el personal de anestesiología.

CONTENIDO

ÍNDICE DE TABLAS	7
ÍNDICE DE FIGURAS	8
1. INTRODUCCIÓN	9
2. OBJETIVOS.	11
2.1 Objetivo General	11
2.2 Objetivos Específicos.....	11
3. MARCO TEÓRICO	12
3.1 Evaluación Preanestésica	12
3.1.1 Factores de Evaluación.....	12
3.1.2 Escalas de clasificación de riesgo preoperatorio.....	13
3.2 Base de Datos	15
3.2.1 Archivos .CSV	15
3.3 Inteligencia Artificial	15
3.3.1 Preprocesamiento de los datos.	16
3.3.2 Aprendizaje automático supervisado.....	18
3.3.3 Clasificador bayesiano.	18
3.3.4 Árboles de Decisión.	21
3.3.5 Evaluación del modelo.	26
3.3.6 Scikit-Learn.....	30
4. METODOLOGÍA	31
4.1 Construcción Base de Datos.....	33
4.1.1 Datos consulta preanestésica.	33
4.1.2 Recopilación y creación de datos	33
4.1.3 Clasificación de la base de datos.....	33
4.1.4 Organización y división de la base de datos.....	33
4.2 Modelos de Inteligencia Artificial.....	34
4.2.1 Revisión de modelos de inteligencia artificial.....	34
4.2.2 Preprocesamiento de datos	34
4.2.3 Entrenamiento de los modelos.	34
4.2.4 Validación de los modelos	34
4.2.5 Evaluación de los modelos	34
5. RESULTADOS.....	35

5.1 Base de Datos	35
5.1.1 Datos consulta preanestésica.	35
5.1.2 Recopilación y creación de datos	36
5.1.3 Clasificación de la base de datos.....	36
5.1.4 Organización y división de la base de datos.....	36
5.2 Modelos de Inteligencia Artificial.....	39
5.2.1 Preprocesamiento de datos y entrenamiento de modelos.	39
5.2.2 Validación de los modelos	56
5.2.3 Evaluación de los Modelos.....	59
6.ANÁLISIS DE RESULTADOS	62
6.1 Base de Datos.....	62
6.1.1 Clasificación base de datos.....	62
6.2 Modelo de Inteligencia Artificial	62
6.2.1 Preprocesamiento de datos y entrenamientos de los modelos	62
6.2.2 Validación de los modelos	66
6.2.3 Evaluación de los modelos	67
7.CONCLUSIONES	69
BIBLIOGRAFÍA.....	71

ÍNDICE DE TABLAS

Tabla 1. Clasificación ASA según estado del paciente.	14
Tabla 2. Datos recolectados y elegidos para la realización de la base de datos referente a una consulta preanestésica	35
Tabla 3. Distribución de la clasificación de los 50 pacientes creados inicialmente.	40
Tabla 4. Recuento de la cantidad de pacientes existentes por cada clase para 3 de las 14 variables categóricas.....	53
Tabla 5. Recopilación de todos las variables eliminados de la base de datos inicial para la anamnesis	54
Tabla 6. Variables altamente correlacionadas de la anamnesis con su porcentaje y las variables seleccionadas para la eliminación	55
Tabla 7. Matriz de confusión con el método de normalización estándar	56
Tabla 8. Matriz de confusión con el método de normalización de límites definidos.	57
Tabla 9. Métricas calculadas al modelo bayesiano con el método de normalización estándar	57
Tabla 10. Métricas calculadas al modelo bayesiano con el método de normalización de límites definidos.....	57
Tabla 11. Matriz de confusión para el modelo de árbol de decisión	58
Tabla 12. Métricas calculadas para el modelo del árbol del clasificación.....	59
Tabla 13. Matriz de confusión de la clasificación bayesiana para las muestras reales. A) Datos sin normalizar; B) Datos normalizados.....	59
Tabla 14. Reporte de los porcentajes de las métricas para la evaluación del modelo de Bayes en las 10 muestras reales	60
Tabla 15. Matriz de confusión del clasificador Bayesiano, para las 60 muestras de entrenamiento. A) Datos sin normalizar; B) Datos normalizados.....	60
Tabla 16. Reporte de los porcentajes de las métricas para la evaluación del modelo de Bayes en las 60 muestras de entrenamiento	61
Tabla 17. Matriz de confusión resultante del clasificador de árbol para los 10 datos reales.....	61
Tabla 18. Matriz de confusión resultante del clasificador de árbol para las 60 muestras del set de entrenamiento	61
Tabla 19. Reporte de las métricas obtenidas del clasificador de árbol para las 10 muestras reales ..	61
Tabla 20. Reporte de las métricas obtenidas del clasificador de árbol para las 60 muestras del set de entrenamiento	61

ÍNDICE DE FIGURAS

Figura 1. Tipos y técnicas de aprendizaje de máquina	16
Figura 2 Representación gráfica de un CBS, con la variable clase, C, y 4 atributos.....	21
Figura 3. Estructura general de un árbol de decisión.	24
Figura 4. Relación de la entropía con las clases	25
Figura 5. Matriz de confusión.	27
Figura 6. Curva ROC, con AUC=1	29
Figura 7. Curva ROC, con AUC=0.5	29
Figura 8. Curva ROC, con AUC=0	29
Figura 9. Metodología implementada en el proyecto.....	32
Figura 10. Composición de la base de datos	38
Figura 11. Número de clasificaciones ASA otorgadas por cada uno de los médicos anestesiólogos	39
Figura 12. Número de clasificaciones promediadas y usadas para el entrenamiento de los dos modelos de inteligencia artificial.....	39
Figura 13. Porcentajes de las clases de ASA.....	40
Figura 14. Distribución de cada variables perteneciente a la exploración física de la consulta preanestésica, A) Sexo (0-H,1-M); B) ASA; C) Edad; D) IMC (índice de masa corporal); E) Frecuencia cardiaca; F) Presión arterial, G) Saturación de oxígeno.....	42
Figura 15. distribución de las variables de la exploración física según el tipo de ASA, A) Sexo; B) Edad; C) Presión arterial; D) IMC (índice de masa corporal); E) Frecuencia cardiaca; F) Saturación de oxígeno.	43
Figura 16. Metodología del preprocesamiento de los datos de la exploración física para el modelo de Naive Bayes.....	44
Figura 17 Matriz de correlación entre las 8 variables involucradas para entrenamiento del modelo Bayesiano	45
Figura 18. Gráficas de densidad para las 8 variables. A) Sexo; B) Edad; C) IMC; D) Frecuencia cardiaca; E) Presión arterial (Sistólica y diastólica); F) Saturación de oxígeno.....	47
Figura 19 Padecimiento de la diabetes. A) Distribución de las muestras; B) Clasificación otorgada de acuerdo a las muestras con diabetes	49
Figura 20 Distribución de las variables correspondientes a los antecedentes quirúrgicos y el tipo de ASA asignadas	49
Figura 21 Padecimiento de tiroides A) Distribución de las muestras; B) Clasificación otorgada de acuerdo a las muestras con Tiroides	50
Figura 22 Padecimiento de la hipertensión. A) Distribución de las muestras; B) Clasificación otorgada de acuerdo a las muestras con hipertensión.....	50
Figura 23 Padecimiento de ASMA. A) Distribución de las muestras; B) Distribución según el uso de inhalador; C) Clasificación otorgada de acuerdo a las muestras con ASMA	51
Figura 24 Metodología para el preprocesamiento de los datos correspondientes a la anamnesis, para el modelo de árbol de clasificación	52
Figura 25 Resultado del modelo de árbol de clasificación creado e implementado a la anamnesis	56
Figura 26 Curvas ROC para el modelo con ambas normalizaciones. A) Normalización estándar; B) Normalización de límites definidos.....	58

1. INTRODUCCIÓN.

La visita médica que realiza un paciente antes de cualquier tipo de cirugía suele ser su primer encuentro con el respectivo anestesiólogo, y es vital para que dicho procedimiento se realice de una manera segura y satisfactoria, ya sea por temas médicos u organizativos. Por lo general las consultas preanestésicas son realizadas externamente, pero hay que tener en cuenta que todos los pacientes deben ser valorados independientemente del tipo de cirugía que se le realizará, antes de llevar a cabo la anestesia [1], ya que esta evaluación permite obtener y conocer información sobre el estado físico y riesgo del paciente, para establecer un plan anestésico de acuerdo con su análisis [2].

En el área de la evaluación preoperatoria se ha demostrado que el juicio clínico es importante, pero por sí solo no es suficiente para predecir eventos adversos postoperatorios. Por lo tanto, se han desarrollado una variedad de herramientas estándares de predicción del riesgo para la clasificación de los pacientes antes de la cirugía [3], como la clasificación ASA (American Society of Anesthesiologists), además se han implementado diferentes protocolos de dicha evaluación para reducir los errores y que el progreso del paciente en el perioperatorio y postoperatorio estén lo más acorde posible con lo planteado en la consulta externa realizada con el anestesiólogo.

Actualmente se encuentran algunas dificultades producto de dicha consulta preoperatoria, en relación con los pacientes, estas dificultades pueden ir desde tiempo laboral perdido y necesidad de permisos, costos de desplazamiento desde lugares rurales lejanos para la consulta, duración corta de la consulta, lo que no permite resolver las dudas del paciente., así como otros efectos cuando se trata de pacientes con limitaciones físicas, madres de familia, etc. Y lo más importante, es que existen diferencias de criterios en la clasificación ASA entre el anestesiólogo de consulta y el anestesiólogo de cirugía (hasta 42% según datos reportados por el HUSVF)[4], algunos estudios que se han llevado a cabo comprueban que la ausencia de la valoración del estado clínico y físico de los pacientes, o una ineficiente evaluación influye en un 11.6% de los accidentes durante la realización de las cirugías y es la causa de un 40% de los fallecimientos [2]. Además, el tiempo limitado de la consulta no permite aplicar los índices de riesgo apropiados para cada paciente, finalmente debido a como se realiza la consulta presencial, no se realiza construcción de bases de datos, es más difícil unificar la información y establecer protocolos de manejo para la evaluación.

Con el fin de contribuir a una evaluación integral prequirúrgica se propuso realizar un modelo predictivo de la clasificación ASA a partir de los datos obtenidos en una consulta preoperatoria. En este proyecto se llevó a cabo el entrenamiento de dos modelos de aprendizaje automático supervisado (Naive Bayes y Árboles de decisión) en el entorno de desarrollo de Python, usando una base de datos obtenida de las consultas realizadas a

diferentes pacientes, para evaluar si alguno de los modelos podría ser capaz de clasificar y estimar de manera automática el riesgo que puede generar la anestesia para los diferentes estados del paciente (ASA).

El desarrollo de este proyecto es un aporte a la comunidad tanto del sector salud, como a los pacientes en general, ya que podría contribuir como una herramienta para ahorrar el desplazamiento de las personas al poder realizar la evaluación de manera digital, y lo más importante ayudar a crear una clasificación única del ASA entre la consulta y la cirugía para solucionar la diferencia de criterios, permitiendo obtener un índice apropiado de clasificación, además contribuiría a reducir el tiempo y el costo de la consulta preoperatoria. Es por esta razón que se pretende dar respuesta a la pregunta ¿Es posible clasificar de manera correcta y con un gran porcentaje de confiabilidad el riesgo preoperatorio de un paciente (ASA) a través de un modelo de aprendizaje automático e inteligencia artificial, a partir de una consulta preanestésica a distancia?

En la Sección 3 del presente informe se presentan los conceptos necesarios para entender lo realizado en el proyecto. En la sección 4 se encuentra la metodología desarrollada para cumplir los objetivos propuestos, en la sección 5 se presentan los resultados obtenidos y en la sección 6 se analizan los resultados, finalmente en la sección 7 se presentan las conclusiones respecto a los objetivos.

1. OBJETIVOS.

1.1 Objetivo General

Desarrollar un modelo predictivo con inteligencia artificial para identificar la clasificación ASA de pacientes en una consulta preanestésica, mediante un algoritmo bayesiano

1.2 Objetivos Específicos

- Identificar y establecer la información necesaria que sea clasificable para la adquisición de los datos con los cuales se entrenará el algoritmo de clasificación.
- Realizar el algoritmo para el modelo de clasificación bayesiano, el cuál determinará los tipos de ASA I y II, teniendo en cuenta tanto la anamnesis como la exploración física del paciente.
- Calcular el porcentaje de acierto y desacierto en la clasificación, a través de la validación del algoritmo creado.

3. MARCO TEÓRICO.

3.1 Evaluación Preanestésica.

La valoración preanestésica se define como todo el procedimiento de estudio que se lleva a cabo antes de la administración de una anestesia, específicamente en procedimientos quirúrgicos, es realizada por los médicos anesestesiólogos y aplicada a personas de todas las edades que vayan a ser sometidos a cualquier tipo de intervención quirúrgica y que conlleve una anestesia general o regional [5]. La evaluación preanestésica se ha convertido en un protocolo importante en el cual se evalúa el estado físico y riesgo del paciente para implementar un plan anestésico como resultado de su análisis [2], esta evaluación se efectúa en la mayor parte de los casos a través de una consulta médica especializada, siendo un requisito obligatorio para la realización de una intervención quirúrgica, en periodos de tiempo variables, ya que depende de algunos factores como las características demográficas del paciente, las condiciones clínicas en las que se encuentre, el tipo y la invasividad del procedimiento al que será cometido [6].

Consideraciones en la evaluación preanestésica: En el momento en el que se hace una consulta preanestésica a cualquier paciente independiente de su patología, el médico anesestesiólogo debe tener presente las siguientes variables, para que la valoración sea efectiva[6].

- **Recopilación de la información médica sobre el paciente:** realizar las interconsultas que se requieran y obtener los exámenes de laboratorio necesarios para determinar el riesgo perioperatorio, a partir del conocimiento sobre el estado físico en el que se encuentra el paciente
- **Optimización de condiciones médicas del paciente:** Implementar un plan anestésico y postoperatorio adecuado para el paciente.
- **Educación al paciente:** Disminuir la ansiedad que pueda manifestar y brindar un apoyo para que tome decisiones diferentes alternativas de técnicas anestésicas y manejo del dolor postoperatorio.

3.1.1 Factores de Evaluación.

Durante la consulta preanestésica se recomienda emplear una serie de preguntas al paciente, con el propósito de enfocarse en los temas relevantes que den mayor información sobre el estado físico del paciente y que sean determinantes para su respectiva clasificación en el riesgo de la cirugía que se llevará a cabo, dichos temas se presentan a una continuación:

3.1.1.1 Anamnesis: Es el proceso en el cual el médico anesestesiólogo obtiene información a través del paciente por medio de un diálogo, en el que se entera de los antecedentes médicos y quirúrgicos, además del historial de anestias efectuadas[2]. La aplicación de escalas de

riesgo prequirúrgico permite asignar valoración a la gravedad de la afección[7]. En este cuestionario se debe incluir como mínimo:

- **Procedimiento programado:** Se tiene en cuenta el motivo y el tipo de procedimiento quirúrgico
- **Historia médica:** Es quizá uno de los pasos determinantes para la clasificación del paciente, porque el anestesiólogo indaga sobre antecedentes y complicaciones quirúrgicas, antecedentes y complicaciones anestésicas, alergias e intolerancias a medicamentos y otras sustancias, uso de medicamentos (prescritos, de venta libre, herbales, nutricionales, etc.), antecedentes patológicos (estado cardiovascular, estado pulmonar, tiroides, diabetes, etc.), clase funcional (estado físico), posibilidad de embarazo (mujeres en edad fértil), historia personal y familiar de complicaciones anestésicas, tabaquismo, consumo de alcohol y otras sustancias [7].

La anamnesis es la fase más beneficiosa para detectar las enfermedades, a partir de los datos recolectados en la consulta preanestésica puede sospecharse del 50% al 75% de los diagnósticos [2]. La experiencia que manifiesta el paciente respecto a sus antecedentes anestésicos, así como la producción de complicaciones, consiente ajustar toda la información y ayuda a la elección de la técnica que será empleada.

3.1.1.2 Exploración Física: Es el conjunto de mediciones que realiza el médico anestesiólogo al paciente durante la consulta preanestésica para obtener información sobre su estado de salud, a través de los valores de las variables fisiológicas, la exploración física sólo contribuye entre el 10-25% de las afecciones crónicas [2], en este procedimiento se incluyen las variables correspondientes al peso, estatura, índice de masa corporal (IMC), signos vitales como la presión arterial, la saturación de oxígeno frecuencia cardiaca y respiratoria, además características del paciente como el sexo y la edad [7].

Las pruebas complementarias que pueden realizarse para la consulta con el anestesiólogo se recomiendan si permiten cuantificar el riesgo o cambia en algo el concepto o conducta médica. Se solicitan en función de la anamnesis y la exploración física, o bien, según los protocolos adecuados para cada institución [5]. Estas pruebas pueden ser: ecocardiograma de reposo, pruebas de esfuerzo no invasivas, coronariografía, exploraciones respiratorias [2], etc.

3.1.2 Escalas de clasificación de riesgo preoperatorio.

Existen dos estándares principales en la clasificación del riesgo multifactorial, las cuales son las más usadas en las consultas preanestésicas por los anestesiólogos, estos índices son la American Society of Anesthesiology (ASA) y la clasificación funcional New York Heart Association (NYHA), sin embargo dichas escalas no ofrecen un resultado que dé cuenta después de la cirugía, consecuentemente su técnica predictiva durante la realización de la cirugía es limitada e inconsistente [8], pero son el medio más adecuado para basarse en una

clasificación cercana a la realidad, antes de realizar cualquier procedimiento quirúrgico que necesite la presencia de un médico anesthesiologo.

3.1.2.1 Clasificación ASA: Es un sistema que recopila y tabula datos estadísticos en el área de anestesia, que pueden ser aplicables en cualquier circunstancia para estandarizar lo que es considerado “Riesgo quirúrgico”, el ASA establece una clasificación que da cuenta del estado preoperatorio del paciente según el padecimiento de ciertas enfermedades, otorgando una muy buena correlación entre la clasificación y el índice de mortalidad respecto al acto anestésico [8], esta valoración preanestésica se basa en una evaluación y un registro subjetivo, dependiendo del criterio del médico anesthesiologo antes de realizar la intervención, en el cual se determina del estado general del paciente, con la clasificación de la gravedad de la cirugía en seis categorías (Tabla 1. Clasificación ASA según estado del paciente.), para proporcionar al equipo médico una terminología común [9].

Tabla 1. Clasificación ASA según estado del paciente.

Tipo de ASA	Definición	Variables
ASA I	Sin alteraciones orgánicas, fisiológicas, bioquímicas o psiquiátricas (Paciente sano)	Ninguna enfermedad, sin limitaciones de actividad Física
ASA II	Enfermedad sistémica leve o moderada, producida por el proceso que se interviene o por otra patología	Hipertensión controlada, ASMA, Anemia, Diabetes controlada, Obesidad leve y tabaquismo.
ASA III	Alteración sistémica grave o enfermedad de cualquier etiología aunque no sea posible definir un grado de discapacidad	Hipertensión mal controlada, ángor, obesidad severa, patología respiratoria sintomática (EPOC)
ASA IV	Alteraciones sistémicas graves que amenazan contantemente la vida del paciente, no siempre corregible con la intervención	Ángor inestable, enfermedad respiratoria incapacitante, fallo hepatorenal
ASA V	Paciente moribundo, con pocas posibilidades de supervivencia, sometido a la intervención como único recurso para salvarle la vida	
ASA VI	Paciente donante de órganos para trasplante, en estado de muerte cerebral	

Respecto a la relación que existe entre la mortalidad que se predice y la que es observada en una cualquier institución, esta puede variar a través de la escala de riesgo que se use y entre la subjetividad o criterio de los médicos cirujanos. Esto puede ser tenido en cuenta al valorar

el riesgo preoperatorio. Los cirujanos de manera individual pueden tener eficacia o inconsistencia cuando son relacionados con estratificación de riesgo preoperatorio [8].

3.2 Base de Datos

Una base de datos se refiere a una colección organizada de información estructurada, almacenada típicamente en un sistema computacional. Una base de datos es controlada por un sistema de gestor de base de datos (DBMS). Los datos, el DBMS y las aplicaciones asociadas a estos se conocen como el sistema de base de datos, frecuentemente mencionado solo como base de datos.

Lo más común es encontrar los datos organizados en filas y columnas en una serie de tablas, que hace eficiente el preprocesamiento y consulta de los datos. Los datos pueden ser fácilmente consultados, modificados, gestionados, actualizados, controlados y organizados. La mayoría de las bases de datos usa un lenguaje estructurado de consulta para escribir y consultar datos [10].

3.2.1 Archivos .CSV

Consiste en un formato de texto estándar para el intercambio y organización de los datos, esta organización se da por filas, en la cual cada una de ellas que se encuentra en estos archivos corresponde a los diferentes valores que pueden tomar las distintas variables que se manejan en este tipo de base de datos, donde dichos valores están separados por coma [11]. Se pueden manipular dos tipos de variables principalmente: numéricos en los que el separador decimal es representado por el punto y variables tipo texto o categóricas, las cuales se diferencian porque se recogen con comillas (dobles o simples). Estos formatos se utilizan principalmente para importar o exportar diferentes bases de datos para ser trabajadas en determinadas aplicaciones [12].

3.3 Aprendizaje Automático.

La inteligencia artificial (IA) es definida como la habilidad de un sistema para interpretar correctamente datos externos, de aprender de estos datos y usar estos aprendizajes para lograr metas específicas a través de adaptación flexible [13].

Los algoritmos de aprendizaje de máquina “aprenden” directamente de los datos y mejoran su rendimiento a medida que el conjunto de estos datos es cada vez mayor y son representativos del problema que se quiere imitar [14].

Hay dos técnicas en el aprendizaje automático: El aprendizaje supervisado en el cual se entrena un modelo con los datos de entrada y de salida conocidos para obtener salidas futuras, y el aprendizaje no supervisado, que encuentra patrones ocultos o estructuras específicas en los datos de entrada [15].

Además de los diferentes tipos de aprendizaje automático también hay diversas técnicas que se ilustran en la Figura 1. Tipos y técnicas de aprendizaje de máquina.

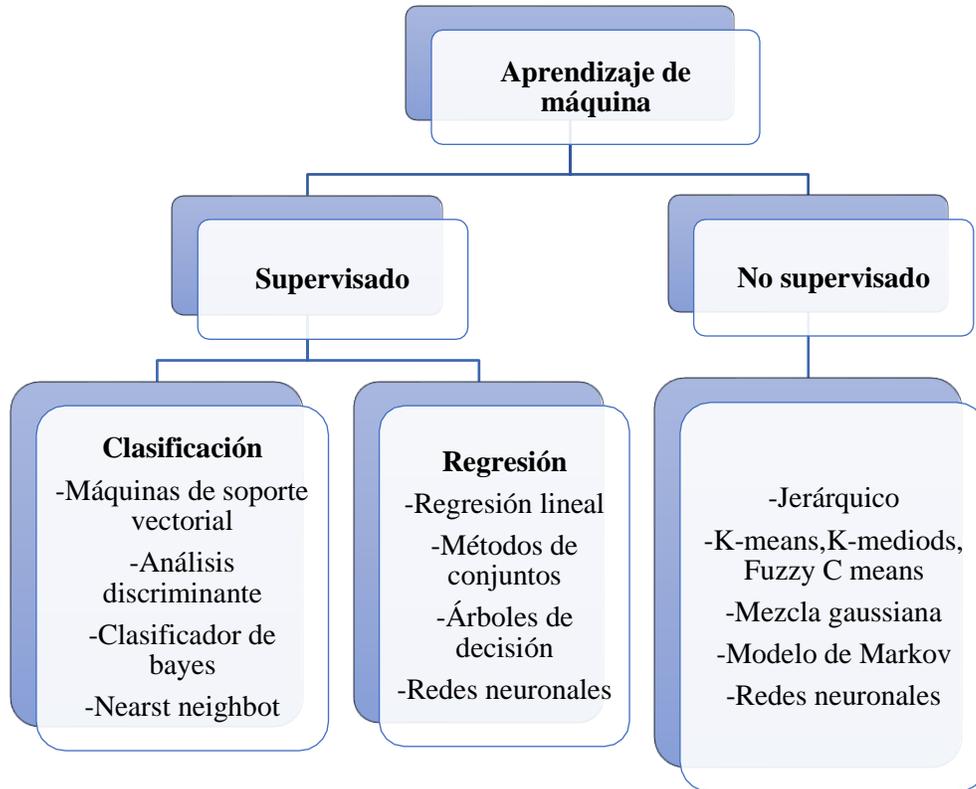


Figura 1. Tipos y técnicas de aprendizaje de máquina.

3.3.1 Preprocesamiento de los datos.

Antes de tratar de clasificar un conjunto de datos por medio de métodos de inteligencia artificial es importante realizar un análisis estadístico básico de los datos, esto ayuda a caracterizarlos y encontrar patrones que ayudan a decidir qué modelo y parámetros usar de acuerdo con la naturaleza de los datos. Por lo general este análisis incluye tabular la cantidad, graficar su distribución y calcular la media y desviación estándar.

Después del análisis estadístico se debe realizar una preparación de los datos minimizando así errores cuando se entrene el modelo escogido, generalmente se siguen los siguientes pasos antes de entrenar un modelo:

- **Normalización.**

Normalizar los datos es escalarlos a una misma distancia, es decir que el valor de todas las variables se encuentre en un mismo rango, comúnmente entre 0 y 1, esto se hace con el propósito de que todas las variables tengan el mismo peso en el modelo, si por ejemplo una variable se midiera en cientos y otra en miles, solo por su valor la de los miles pesaría más al momento de entrenar el modelo [16].

Las dos técnicas de normalización más utilizadas son la normalización mínimo-máximo (Ec.1) donde x_{min} es el mínimo valor de una variable en un conjunto de datos, $x_{m\acute{a}x}$ es el máximo valor de esta variable y x' es el valor normalizado.

$$x' = \frac{x - x_{min}}{x_{m\acute{a}x} - x_{min}} \quad (\text{Ec. 1})$$

La segunda técnica de normalización es la denominado Z-Score (Ec. 2), cuando se usa esta técnica los valores no estarán entre 0 y 1 si no que se obtendrán valores negativos, pero se seguirá trabajando con una misma escala para todas las variables trabajadas.

$$x' = \frac{x - \mu}{\sigma} \quad (\text{Ec. 2})$$

Donde μ es la media y σ corresponde a la desviación estándar.

- **Agregar datos faltantes**

Por diversas razones, muchas bases de datos poseen valores faltantes, lo que conlleva a descartar una entrada completa a la que le falte un solo valor, esto significa perder datos valiosos que puedan contribuir a que el modelo se desempeñe de una manera mucho mejor. Con el fin de darle solución a esta problemática se han creado diversas estrategias para inferir valores a partir de valores ya existentes en la base de datos [17].

Una de las técnicas más sencillas para obtener nuevos datos es generar valores aleatorios a partir del cálculo de las métricas estadísticas correspondientes a la media y la desviación estándar, igualmente se puede implementar otra estrategia que usa el valor más repetido para reemplazar los valores de los datos faltantes [18].

Se han desarrollado diferentes técnicas mucho más avanzadas como los *K nearest neighbours* [19] o los MICE (*Multivariate imputation by Chained Equation*) [20].

- **Balanceo de clases**

En algunos problemas de clasificación (aprendizaje supervisado) hay una diferencia importante entre la cantidad de clases, lo que significa que hay una clase que tiene atribuido muchas más muestras que la otra, si estas cantidades se ingresaran en estas proporciones iniciales a un modelo de clasificación, este tendría una alta probabilidad de aprender a clasificar las nuevas muestras ingresadas con la clase que tiene la mayor cantidad de datos, generando un problema conocido como el desbalanceo de los clases [21].

Cuando se presenta este tipo de problema en la base de datos trabajada, es posible darle solución a través de la implementación de técnicas principales de remuestreo. En primer lugar, se encuentra sobre muestrear, la cual crea una base de datos replicando ciertas muestras originales y creando nuevas con base en las ya existentes; la segunda técnica es el

submuestreo, la cual consiste en eliminar de la base de datos original, algunos datos que correspondan a la clase mayoritaria [22]. Finalmente cabe destacar que existen otras técnicas más robustas como “Synthetic Minority Oversampling Technique” (SMOTE) [23], la cual realiza una interpolación de algunos de los datos que se acerquen a la clase de menor cantidad.

- **Transformación de datos categóricos**

Los tipos de datos más comunes usados en los algoritmos de machine learning son los categóricos, estos datos tienen un número finito de categorías o etiquetas. Por ejemplo, los estados de EE. UU. o una lista de los tipos de animales que se encuentran en un conjunto de imágenes, estos hacen referencia a alguna característica que puede tener la variable incluida en la base de datos, es por esto que en el preprocesamiento es importante asignarle un valor numérico a cada una de las etiquetas, para que puedan ser útiles al generar algún modelo de aprendizaje automático [24].

Existen varias técnicas para manipular y trabajar con datos categóricos de una base de datos, la más simple consiste en asignar un valor entero para que represente cada una de las categorías, y están asociados a los pesos que se le asignan a cada una de estas, de acuerdo a la cantidad que exista en los datos, significa que la etiqueta con mayor frecuencia en la base de datos se le otorgará el valor entero más grande; la otra es la codificación frecuente que lo que hace es que toma un conjunto finito de valores y los asigna a enteros cuya representación binaria tiene un valor 1 único en posiciones únicas en la cadena [25].

3.3.2 Aprendizaje automático supervisado.

Es un algoritmo que basa su aprendizaje en un conjunto de muestras de las que conoce los parámetros de entrada y sus valores de salida válidos y ciertos, a partir de los cuales pretende encontrar el patrón correcto para adecuarse a cualquier nueva entrada y ser capaz de predecir su valor de salida (problemas de regresión lineal) o incluirla en una categoría conocida (problemas de clasificación) [26]. En sí, este tipo de solución busca diseñar un modelo que realice predicciones de los valores respuesta a partir de un conjunto de datos.

3.3.3 Clasificador bayesiano.

Es uno de los métodos más usados de aprendizaje automático, se basa en el teorema de probabilidad de Naïve Bayes, el cual tiene como principio la independencia condicional de las variables predictoras conocida cierta clase, este clasificador aparece por primera vez en la literatura del aprendizaje automático a finales de los años ochenta (Cestnik y col (1987)), enfocado en comprobar su capacidad de predicción comparándolo con otros métodos de clasificación [27]

Los clasificadores bayesianos realizan un análisis tanto cualitativo como cuantitativo, ya que evalúa los atributos y valores que son usados para resolver el problema de interés, y además da su respectiva importancia en términos numéricos, cuando se habla de la forma cualitativa

se refiere al hecho de que puede dar una representación de cómo es la relación existente entre los datos, ya sea causal o indicando el porcentaje de correlación que hay entre las diferentes variables involucradas. El método cuantitativo es el que le da peso al método de clasificaciones bayesianas, ya que da un porcentaje probabilístico de la importancia de las variables o atributos en el problema, y por ende otorga una probabilidad explícita de la hipótesis que se formula para dar solución al problema, es por esto que a partir del uso de este método de clasificación o aprendizaje automático se puede conseguir una excelente solución por medio de todas las distribuciones que pueden tener los datos observados y analizados [2].

Un clasificador de patrones se basa en que teniendo cualquier conjunto de datos (que se dividen en entrenamiento y evaluación) que son representados por pares <atributo o variable-valor>, la solución a dicho problema sería encontrar una función que es denominada como su hipótesis y la cual clasifique los ejemplos, el objetivo de usar el teorema de Bayes en algún problema de aprendizaje automático, es poder obtener una estimación de la probabilidad a posteriori de alguna hipótesis en el conjunto de datos usados para su entrenamiento, para escoger de esta manera la hipótesis más acertada o probable. Para esto se ha propuesto usar el algoritmo de Naïve Bayes [28].

Esto en términos matemáticos se puede explicar de la siguiente forma, dando a conocer que el enfoque bayesiano en una clasificación supervisada está basada en asignar a un conjunto de atributos o características, X_1, X_2, \dots, X_n una de las m clases posibles, C_1, C_2, \dots, C_m , con el objetivo que la probabilidad calculada para cada clase existente, a través de los atributos sea la máxima posible (Ec. 3) [29].

$$Arg_c[Max P(C | X_1, X_2, \dots, X_n)] \text{ (Ec. 3)}$$

Si los atributos del problema se representan como $X = \{X_1, X_2, \dots, X_n\}$, la Ec 3. Se puede reducir a: $Arg_c[Max P(C | X)]$. La base matemática de los clasificadores bayesianos, usan la regla de Bayes (Ec. 4) para hallar la probabilidad posterior de cada una de las clases trabajadas dependiendo de los atributos [29].

$$P(C | X_1, X_2, \dots, X_n) = \frac{P(C) * P(X_1, X_2, \dots, X_n | C)}{P(X_1, X_2, \dots, X_n)} \text{ (Ec. 4)}$$

Que se puede reducir a (Ec.5) :

$$P(C | X) = \frac{P(C) * P(X | C)}{P(X)} \text{ (Ec.5)}$$

De modo que el problema de clasificación representado en la Ec 5 se puede expresar como (Ec.6):

$$Argc[Max_{C|X} \frac{P(C)*P(X|C)}{P(X)}] \text{ (Ec.6)}$$

El denominador de la Ec 6, es independiente de las diferentes clases, por lo que es posible considerarlo como una constante dentro del modelo matemático, para enfocarse en maximizar la probabilidad de cada clase (Ec 7.).

$$Argc[Max[P(C|X) = \alpha P(C) * P(X|C)]] \text{ (Ec 7.)}$$

Con base en la ecuación 7 y para darle solución a un problema de clasificación con el enfoque bayesiano, se requiere como se mencionó anteriormente conocer la probabilidad a priori de cada una de las clases, $P(C)$, y la probabilidad de las características conociendo la clase $P(X|C)$, denominada como verosimilitud, para hallar la probabilidad posterior $P(C|A)$. Para que este clasificador aprenda del conjunto de datos implementados, es necesario calcular estas probabilidades, a partir de dichos datos, conocidos como los parámetros del clasificador [29].

Al aplicar directamente lo que representa la ecuación 7 se obtiene un sistema complejo para implementarlo en un sistema computacional, debido a que el término que se refiere a la probabilidad de verosimilitud, hace que se incremente de forma exponencial su tamaño dependiendo del número de atributos que se encuentren en el problema, exigiendo una alta capacidad de memoria para el almacenamiento en el computador, y de igual manera crece enormemente el número de operaciones para calcular su probabilidad. Una de las alternativas para esto es considerar algunas relaciones de dependencia a través de un algoritmo llamado clasificador bayesiano simple.

Clasificador Bayesiano Simple: Este método de clasificación se basa en suponer que todos los atributos o características son independiente de la clase, es decir que cada uno de los atributos es independiente del resto dada la clase (Ec. 8).

$$P(X_i|X_j, C) = P(X_i|C) \text{ donde } \forall j \neq i \text{ (Ec 8).}$$

Es por esto que la Ecuación 4 se puede representar como (Ec. 9):

$$P(C|X_1, X_2, \dots, X_n) = \frac{P(C)*P(X_1|C)*P(X_2|C)*\dots*P(X_n|C)}{P(X)} \text{ (Ec. 9)}$$

En la cual $P(X)$ se considera como una constante de normalización.

Este método de clasificación simple reduce lo complejo del clasificador bayesiano tanto en espacio como en tiempo de cálculo. Referente a la memoria, es necesario conocer la probabilidad de las m clases existentes y las n probabilidades de cada atributo conociendo la clase, es decir que el espacio que se requiere incrementa linealmente con la cantidad de atributos, e igualmente el cálculo de la probabilidad a posteriori se vuelve eficiente, porque

solo son necesarias n multiplicaciones para hallar la probabilidad posterior de las clases dadas las características (complejidad lineal) [29].

En la figura 2 se puede observar la representación de un clasificador bayesiano simple (CBS), usando el principio de modelos gráficos probabilísticos, en la cual la independencia condicional entre los atributos es representada a través de un grafo, en donde se evidencia que los atributos dependen solo de la clase y son totalmente independientes entre ellos.

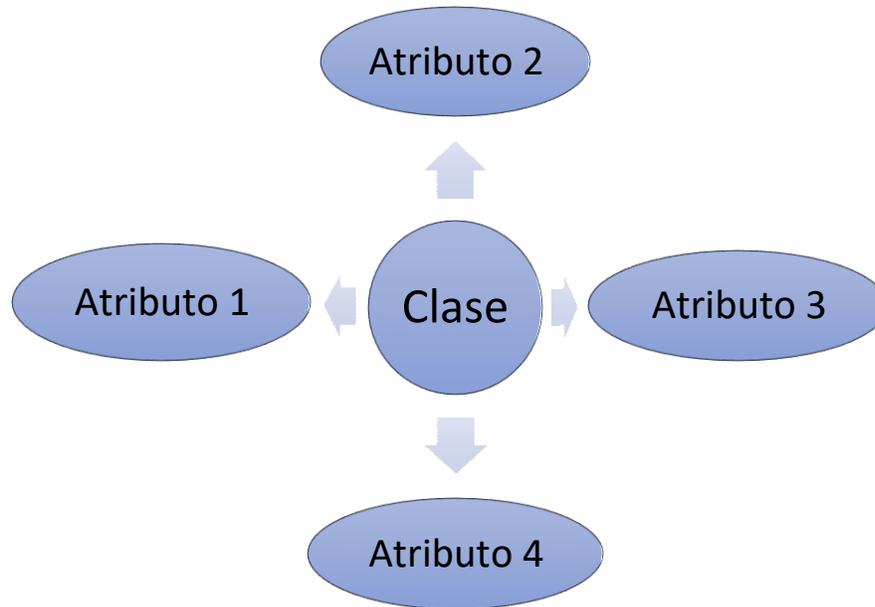


Figura 2 Representación gráfica de un CBS, con la variable clase, C, y 4 atributos.

Para que un CBS aprenda es necesario la probabilidad de las clases y la probabilidad condicional de cada variable conociendo la clase.

Los clasificadores bayesianos simples, generan una estructura muy eficiente para la construcción de métodos de clasificación, el cual da excelentes resultados en diversas aplicaciones.

En resumen los clasificadores bayesianos modelan problemas mediante un conjunto de datos calculando la relación de dependencia entre ellas, realizando una inferencia bayesiana basada en estimar la probabilidad posterior de todas las variables no conocidas del problema, en base a las conocidas, eligiendo la máxima probabilidad encontrada para poder aplicarla en una de las tantas aplicaciones, como es el caso de predicción, clasificación o diagnóstico [3].

3.3.4 Árboles de Decisión.

El enfoque de árboles tanto de clasificación como de regresión (CART) fue desarrollado por Breiman a principios de los años ochenta (1984), es una alternativa para analizar de manera

tradicional la clasificación y la predicción (regresión), una de las ventajas que se pueden destacar de este método de clasificación es su robustez, la invarianza que se presenta en la estructura de los árboles que se construyen a las transformaciones monótonas de sus variables independientes, y algo que hace que sea elegido es su fácil interpretación [31].

Específicamente un árbol de decisión es de clasificación cuando la variable dependiente o la variable que define la clase es de tipo cualitativa o discreta [32], su funcionamiento se base en una partición recursiva con el dominio de las variables predictoras y se puede conocer como es el problema trabajado a través de una estructura de árbol [33]. En esencia se trata de encontrar la variable independiente o la variable que representa una de las etiquetas del conjunto de datos, la cual haga una separación de todos los datos en grupos que correspondan a las categorías en la que está dividida la variables objetivo, esto se hace para cada una de los grupos resultantes después de obtener una regla o condición de separación, hasta que ya no es posible obtener una mejor, esto es representado a través de un esquema de múltiples divisiones o dicotomías, que al seguir cada una de estas ramas del árbol se obtenga al final la mejor predicción o clasificación para la clase de pertenencia de los casos o muestras que cumplan con las condiciones o propiedades que se eligieron en las distintas subdivisiones del árbol [32].

Una de las características más relevantes para tener en cuenta es que si una variable ha sido elegida para separar el conjunto de datos, ya no es usada por el algoritmo en los grupos que se ha creado, siempre procura buscar diferentes variables para encontrar una mejor división de los datos [32].

En términos matemáticos los árboles de clasificación se pueden explicar de la siguiente forma, si C es una variable de respuesta y A es un conjunto de variables predictoras X_1, X_2, \dots, X_A , en la cual el conjunto A son etiquetas fijas y la variables Y es aleatoria el problema estadístico a resolver se basa en encontrar y establecer una relación entre C y A , de tal forma que pueda ser posible predecir el valor de C a través de los valores que tiene A , hallando una función que estudie la probabilidad condicional que existe de la variable aleatoria C (Ec. 10) [31].

$$A[C = c|X_1, X_2, \dots, X_A] \text{ (Ec. 10)}$$

Elementos del árbol.

Cualquier árbol de decisión se divide básicamente en tres niveles, el primero corresponde al **nodo raíz** el cual es la primera división que hace el modelo y es la cima del árbol, el segundo nivel corresponde a uno o varios **nodos internos** y está asociado a uno de los atributos y tiene 2 o más ramas que salen de él, cada una de estas representan los diferentes valores que puede tomar el atributo correspondiente, finalmente se encuentra el tercer y último nivel que pertenece a los **nodos terminales**, el cual da la clasificación adecuada y retorna la decisión

del árbol con respecto a los datos de entrada, entre los elementos que conforman el árbol es importante dar ciertas definiciones para un correcto análisis del árbol [31].

- **Homogeneidad**

Es asociada a los nodos terminales y el ideal es que las variables resultantes de estos nodos tengan la más alta homogeneidad que es medida mediante la noción de impureza [31].

- **Impureza del nodo**

Se define diciendo que C es una variable dicotómica que solo toma los valores de 0 y 1, un nodo tiene una impureza mayor cuando su impureza es máxima con $P\{C=\text{correcto}\}=1/2$. La función de impureza para los nodos adquiere una forma cóncava y se define como (Ec. 11.) [31].

11) Donde τ se refiere a la impureza y φ tiene propiedades como:

- $$i(\tau) = \varphi(P\{C = \text{correcto}\}) \text{ (Ec. 11.)}$$
- (i) $\varphi \geq 0$ y (no negativa)
 - (ii) Para cualquier $p \in (0,1)$, $\varphi(p) = \varphi(1 - p)$ (simétrica) y $\varphi(0) = \varphi(1) < \varphi(p)$ mínima para el éxito o el fracaso absoluto

Las elecciones más comunes para la función de impureza en la construcción de un árbol de clasificación son:

- $\varphi(p)=\min(p,1-p)$: Mínimo error o error de Bayes
 - $\varphi(p)=-p \cdot \log(p)-1-p \cdot \log(1-p)$: Entropía
 - $\varphi(p)= p(1-p)$: Índice de Gini
- Donde se define $0\log(0)=0$.

- **División de un nodo:**

El nodo raíz se divide en dos nodos homogéneos, estos nodos son seleccionados hallando el valor entre el rango de las variables predictoras que más cerca este al límite de pureza para cada uno de los nodos internos, el objetivo es que si el nodo N se divide en dos, N_L y N_R , la pureza de estos dos nodos deben ser mayor que la pureza del nodo N , o en efecto tener una impureza menor, que suele medirse con la mínima probabilidad, el índice de Gini o la entropía [31].

La estructura general de un árbol de decisión se muestra mediante la Figura 3, en la cual se puede observar que las variables de predicción se representan a través de círculos, y las hojas del árbol las cuales significa la terminación completa en la división del árbol donde los datos pertenecen a una sola clase, son representados con un rectángulo.

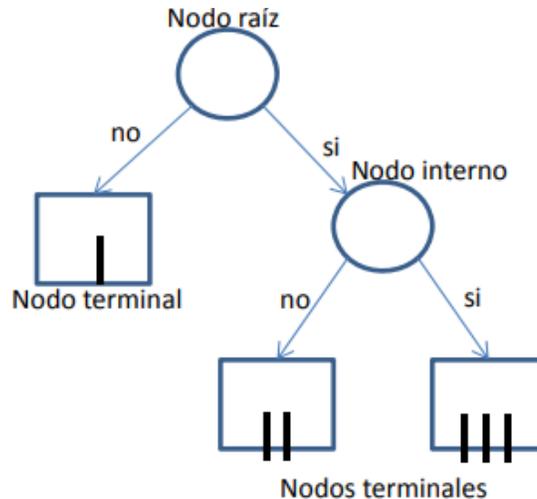


Figura 3. Estructura general de un árbol de decisión.

Algoritmo ID3.

Significa Iterative Dichotomizer 3. Este algoritmo fue creado por Ross Quinlan. En el cual se construye un árbol de decisión a partir de un conjunto fijo de datos y el árbol final se usa para clasificar muestras futuras.

El principal funcionamiento se basa en crear el árbol de decisiones realizando una búsqueda exhaustiva de arriba a abajo a través de los conjuntos de datos dados para evaluar cada atributo en cada nodo del árbol [34].

Existen diferentes métodos de selección para la división y construcción del árbol entre las que se encuentra:

- **Entropía.**

En ciencias de la información, la entropía mide la incertidumbre sobre una fuente de ejemplos o muestras. Otorga el grado de desorden en el que se encuentran los datos de entrenamiento [34].

Si se tiene una muestra completamente homogénea (es decir, en la que todos se clasifican igual) se le relaciona una mínima incertidumbre, significa que no se presenta ninguna duda sobre cuál es la clasificación de cualquiera de los elementos involucrados, es decir que se sabrá cual es el resultado si se elige cualquiera al azar, para este caso se asigna una entropía=0., si por el contrario un conjunto de datos está distribuido, es decir, que se tiene la misma cantidad de muestras para cada posible clasificación, su incertidumbre es máxima en el sentido que es difícil conocer con anterioridad a cuál de las clases posibles pertenece cualquiera de los ejemplos si alguno es elegido al azar, así pues en este cada la entropía del sistema sería igual a 1 [35].

La definición matemática para la entropía proporcionada por Shannon que mide el grado de incertidumbre se muestra en la Ec. 12.

$$E(S) = \sum_{i=1}^C -P_i \log_2(P_i) \text{ (Ec. 12)}$$

Donde S es el conjunto de datos del sistema analizado, C corresponde al número de diferentes clases que se tiene en el problema y P_i es la proporción de datos que existen en la clasificación i de la muestra.

En el caso de una clasificación binaria (Datos divididos entre positivos y negativos) la Ecuación 12 quedaría de la siguiente manera (Ec. 13).

$$E(S) = -P \log_2(P) - N \log_2(N) \text{ (Ec. 13)}$$

Donde P y N, pertenecen a la proporción de los datos positivos y negativos respectivamente, En la Figura 4 se puede observar la relación entre la entropía y una de las dos clases.

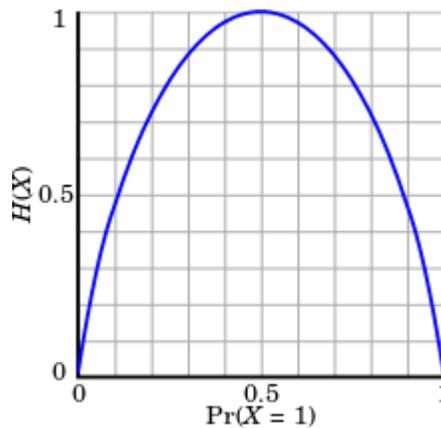


Figura 4. Relación de la entropía con las clases

- **Ganancia de información.**

Es una medición que da cuenta de la reducción esperada en entropía. Elige cuál de los atributos se asigna como un nodo de decisión o nodo interno, su método de selección se basa en el atributo que se le asocia la mayor reducción de entropía [34]. La ganancia de información de un atributo A con respecto a un conjunto de datos S se define como (Ec. 14).

$$Gain(S, A) = Entropia(S) - \sum_{v \in \text{valores}(A)} \frac{|S_v|}{|S|} * Entropia(S_v) \text{ (Ec. 14)}$$

Donde:

S= cada valor de v de todos los posibles valores del atributo A

S_v =Subconjunto de S en el cual el atributo A tiene el valor de v.

$|S_v|$ =Cantidad de elementos en S_v

$|S|$ = Cantidad de elementos en S

Restricciones tamaño del árbol (pre-pruning)

Es uno de los procesos que se llevan a cabo para evitar el sobreajuste en el momento del proceso de la modelación de cualquier árbol de decisión, estos parámetros es importante definirlos para obtener un buen rendimiento del modelo y son independientes del entorno de desarrollo donde se realice [36].

- **Mínimo de observaciones para dividir un nodo:** Es el mínimo número de muestras que requiere un nodo para que se considere realizar una división o ramificación, si se eligen valores muy altos puede resultar en un pobre ajuste del modelo y permite que aprenda relaciones muy específicas [36].
- **Mínimo número de observaciones para un nodo terminal:** Es el mínimo número de muestras que se tienen en cuenta para resultar en un nodo hoja o terminal, es decir que a cualquier profundidad se considera un punto de división si hay al menos mínimo de muestras para entrenamiento tanto al lado derecho como al izquierdo [37].
- **Máxima profundidad del árbol (vertical):** Es la profundidad máxima que tiene el árbol, indicando la cantidad de expansiones que se genera en los nodos hasta que los nodos terminales sean puros, si no se indica ningún número para su profundidad este se dividirá hasta que los nodos terminales contengan menos del mínimo número de muestras para realizar su división [37].

3.3.5 Evaluación del modelo.

Luego de la construcción de algún modelo y elegir los parámetros, es de gran importancia evaluar su desempeño, para escoger el más adecuado respecto a cómo está clasificando los datos. Aunque existen diversas técnicas para evaluar los algoritmos, se ha hecho una selección de las métricas más usadas para la evaluación del rendimiento de los modelos de clasificación

Exactitud y precisión en la clasificación.

La **exactitud** es el número total de predicciones correctas que realiza el modelo creado entre todas las muestras del modelo (Ec. 15), Esta métrica tiene un buen rendimiento si el número de muestras por cada clase es igual.

$$Exactitud = \frac{\text{Número de predicciones correctas}}{\text{Número total de predicciones}} \quad (\text{Ec. 15})$$

Un ejemplo es considerar que el 98% de las muestras pertenecen a la clase 1 y el 2% a la clase 2, después el modelo creado obtendrá el 98% de exactitud prediciendo solamente las muestras de la clase 1, pero si este porcentaje cambia respecto a la relación de ambas clases esta métrica de exactitud cambiará [38].

La **precisión** Es el número de muestras identificadas correctamente como positivas de un total de datos identificados como positivos en los datos de desempeño, es decir que solo da cuenta de las muestras positivas (Ec. 16).

$$Presición = \frac{Número\ de\ predicciones\ positivas\ correctas}{Número\ total\ de\ muestras\ positivas} \quad (Ec. 16)$$

Si se tiene en cuenta solo estas dos métricas no se obtiene información suficiente que dé cuenta sobre el rendimiento del modelo [39].

Matriz de confusión.

Es una de las métricas más relevantes en el aprendizaje automático ya que describe el rendimiento de cualquier modelo implementado sobre todos los datos de prueba, en los cuales son desconocidas las clases a las cuales pertenecen, adquiere el nombre de matriz de confusión ya que hace referencia a la facilidad de detección en donde el sistema está confundiendo las clases evaluadas [40]. En la Figura 5 se muestra la forme general de la matriz de confusión.

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	Verdadero Positivos	Falsos Negativos
	Negativos	Falsos Positivos	Verdaderos Negativos

Figura 5. Matriz de confusión.

Al momento de realizar una predicción en esta matriz de confusión se pueden observar cuatro resultados diferentes, **verdaderos positivos (VP)** son los casos que son predichos de manera acertada por el modelo, es decir que son positivos y efectivamente se clasificaron como positivos; **falsos positivos (FP)** son los casos que el modelo clasificó como positivos pero realmente son negativos; **falsos negativos (FN)** es el número de muestras que el modelo le otorgó la etiqueta de negativo y realmente son positivos y **verdaderos negativos (VN)** son los casos que el modelo predice como negativo y efectivamente lo son.

Con estos cuatro resultados se puede calcular diferentes métricas, de las cuales hay dos fundamentales para evaluar el rendimiento del modelo, las cuales son la especificidad y la sensibilidad, aunque también se considera la métrica correspondiente al puntaje F.

La **Sensibilidad** (Ec. 17) representa el número de predicciones positivas y que realmente lo son respecto al número total de predicciones reales catalogadas como positivas, es decir que esta métrica da cuenta de que tan bueno es el modelo para predecir los casos positivos [38].

$$\text{Sensibilidad} = \frac{VP}{VP+FN} \text{ (Ec. 17)}$$

La **especificidad** (Ec. 18) corresponde a la proporción del número total de casos negativos que son tomados como positivos por el modelo, significa que tanto es el error del modelo prediciendo casos negativos [38].

$$\text{Especificidad} = \frac{VN}{VN+FP} \text{ (Ec. 18)}$$

El **recall** es la métrica que da información sobre el rendimiento del modelo con respecto al número de muestras catalogados como falsos negativos, es decir cuantas predicciones fallaron, esto también se conoce como el recuerdo del clasificador, significa que si se desea minimizar los falsos negativos el porcentaje de recuerdo del modelo debe estar cerca del 100% [41].

El **puntaje F1** (Ec. 19) es la medida armónica entre la precisión del modelo y su memoria (recall), y la relación se basa que a mayor porcentaje de F1 el modelo es mucho mejor [42].

$$F1 = 2 * \frac{\text{Precisión} * \text{recall}}{\text{Precisión} + \text{recall}} \text{ (Ec. 19)}$$

Generalmente, de acuerdo con el objetivo del modelo desarrollado se le da una mayor relevancia a alguna de estas métricas, un ejemplo de esto es que en el área de clasificación de riesgo es muy importante que el modelo clasifique correctamente los casos de la clase correspondiente al riesgo más alto [41].

Curva ROC

La curva ROC por sus siglas en ingles Receiver Operating Characteristic, es una representación gráfica de la sensibilidad, contra el inverso de la especificidad (1-especificidad) para un clasificador binario y de acuerdo a cierto umbral, esto se mide con el área bajo la curva (AUC) que dice que tan bien funciona el modelo al diferenciar y predecir entre las dos clases [43]

De la curva ROC se puede realizar tres tipos de interpretaciones o análisis , el primero es si el AUC es de 1 (Figura 6) quiere decir que el modelo distingue perfectamente todas las muestras o características de cada una de las clases(clase 1 y clase 2), asignándoles la clase correcta, esto denota que el modelo está sobreentrenado y solo es útil únicamente para los datos que se usaron para su entrenamiento, provocando que su rendimiento no sea el mejor para nuevos datos, un valor entre 0,8 y 0,9 para una curva ROC es en general un excelente

resultado, si el AUC es de 0,5 (Figura 7), el modelo es incapaz de diferenciar entre ambas

clases y finalmente si el $AUC=0$ (Figura 8), representa que el modelo confunde entre las dos clases, a la clase 1 le otorga la clase 2 y viceversa

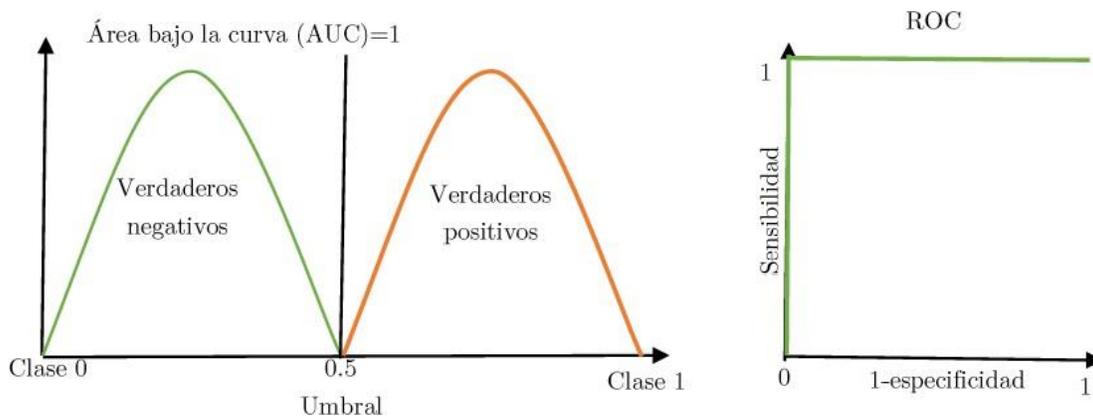


Figura 6. Curva ROC, con $AUC=1$

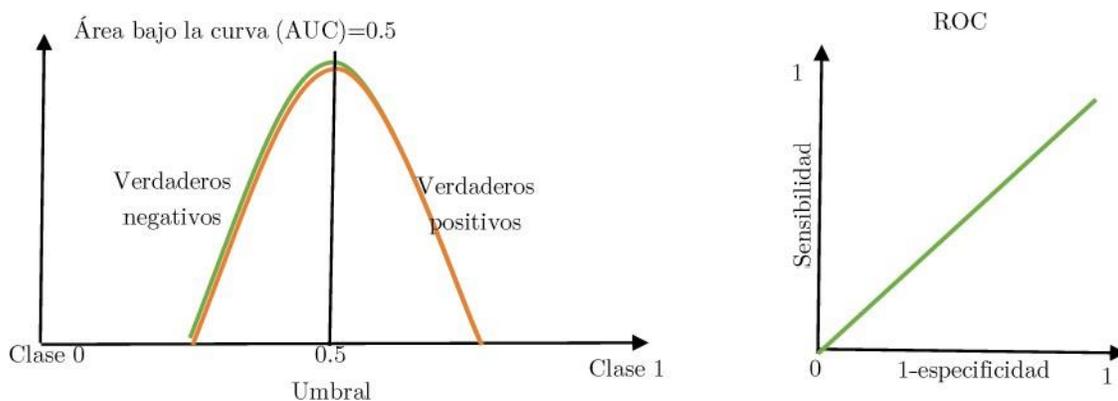


Figura 7. Curva ROC, con $AUC=0.5$

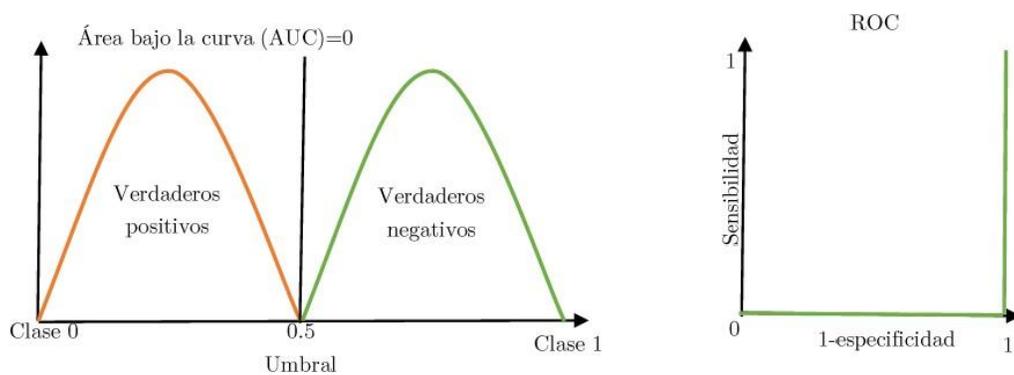


Figura 8. Curva ROC, con $AUC=0$

3.3.6 Scikit-Learn

Es un modelo de Python que integra un amplio rango de algoritmos de machine learning de última generación para problemas de aprendizaje automático supervisado y no supervisado de mediana escala. El enfoque de este paquete es hacer de fácil acceso todas las técnicas que involucran machine learning a personas que nos son muy especializadas, usando un lenguaje de alto nivel, con el enfoque de que sea de fácil uso, alto rendimiento, documentación y consistencia en la API [44].

Scikit-Learn cuenta con diferentes algoritmos de clasificación, regresión y análisis de máquinas de soporte vectorial, bosques aleatorios, K-means y otros.

Para usar los clasificadores bayesianos y los árboles de decisión que facilita Scikit-learn se usan las funciones `GaussianNB`, `train_test_split`, `DecisionTreeClassifier`, estas reciben parámetros como el criterio de división del algoritmo, el mínimo demuestras que toma el algoritmo para tomar una decisión y la profundidad del árbol.

Scikit-learn también provee las herramientas para evaluar el desempeño de los modelos, con funciones como `accuracy_score`, `precision_score`, `classification_report`, entre otras. Para conocer todas las funcionalidades que posee esta librería, diríjase a la documentación oficial disponible en: <https://scikit-learn.org/stable/>.

4. METODOLOGÍA

A continuación, se presenta el diagrama (Figura 9) de la metodología general propuesta para el desarrollo del proyecto. Se plantean dos metodologías diferentes, la primera es para la construcción de la base de datos con la información recolectada, la segunda para el desarrollo del modelo, la cual fue dividida en las construcción de dos tipos de algoritmo, uno correspondiente al algoritmo de Naive Bayes, y otro perteneciente a los árboles de decisión (clasificación), estas a su vez contienen subactividades que se desarrollaron con el propósito de cumplir los objetivos definidos, una actividad principal inicia al realizar la anterior; excepto la revisión bibliográfica, que se realizó de manera paralela a todas las actividades.

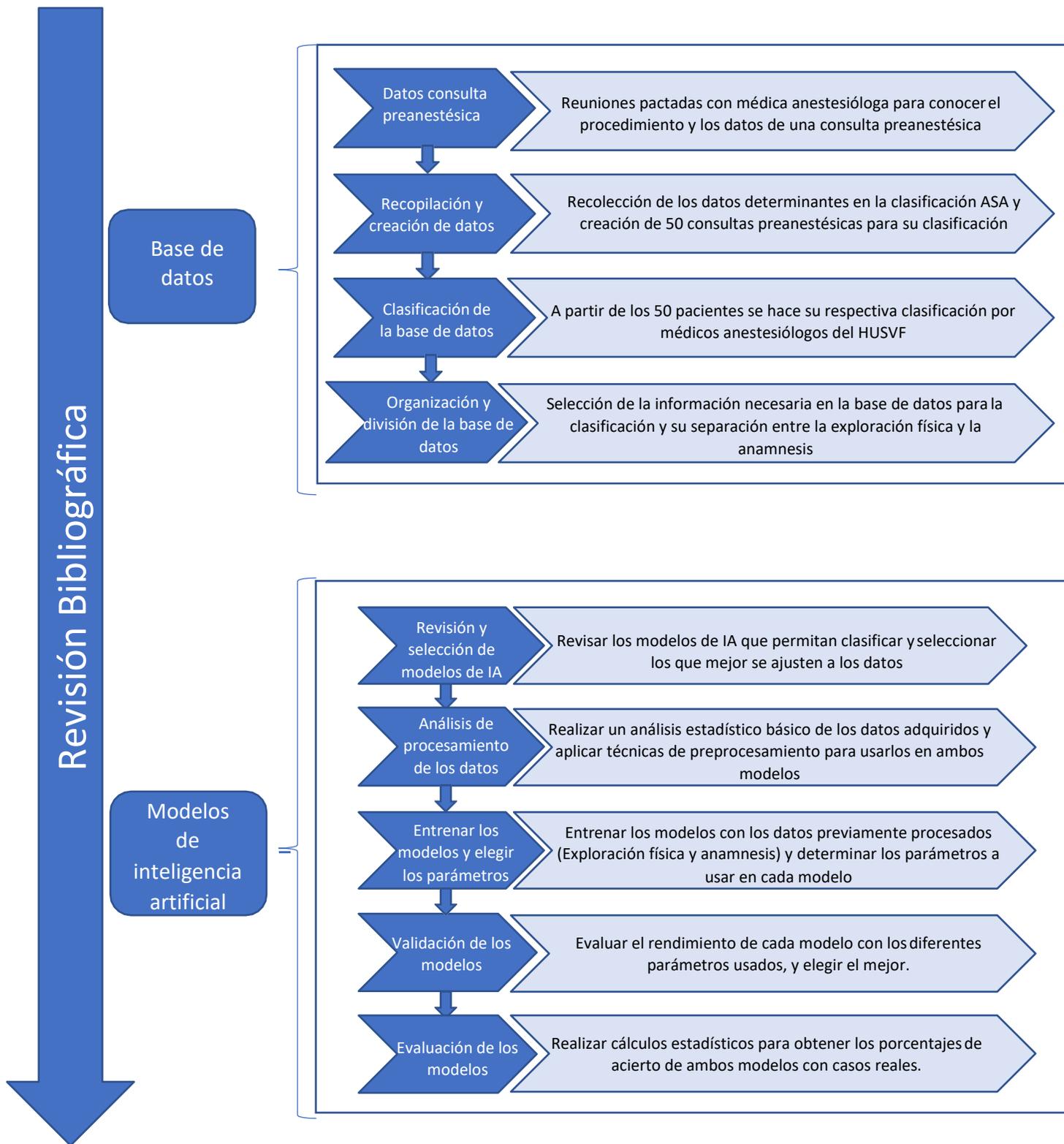


Figura 9. Metodología implementada en el proyecto.

4.1 Construcción Base de Datos.

4.1.1 Datos consulta preanestésica.

Se realizaron diferentes mesas de diálogo con la Anestesióloga Olga Giraldo del HUSVF (Hospital Universitario San Vicente Fundación) en las que se trató el tema de cómo se lleva a cabo una consulta preanestésica detalladamente, de las cuales se identificaron los datos más relevantes para la clasificación ASA de un paciente que será intervenido quirúrgicamente, a partir de esto se realizó el establecimiento de los parámetros y variables necesarias para obtener una información objetiva y clasificable dentro de los posibles valores que pudiera tomar la variables a tener en cuenta.

4.1.2 Recopilación y creación de datos.

Se unieron todos los datos de interés que se determinaron en las reuniones para conocer el protocolo de una consulta preoperatoria realizada a un paciente, los cuales son consideradas como determinantes para otorgar una clasificación ASA, según el personal médico; esto se realizó a partir de un formato en excel, para la recolección de los datos, dichos datos se establecieron a partir de 50 casos diferentes de pacientes ficticios inicialmente, para construir toda la base de datos usada para el entrenamiento del aprendizaje automático supervisado que se llevó a cabo, y se procuró tener diferenciabilidad y variabilidad en cada uno de los valores de las variables para obtener una base de datos balanceada para su posterior clasificación.

4.1.3 Clasificación de la base de datos.

Con el objetivo de tener una clasificación verdadera y real de los 50 pacientes creados, se solicitó a dos médicos anestesiólogo realizar la evaluación preanestésica y clasificar a los pacientes en una categoría ASA correspondiente, con base en la información diseñada, con lo cual cada uno de ellos evaluaron todos los pacientes de forma separada, dando su criterio médico, y logrando así obtener dos opiniones para realizar un balance de coincidencia entre la subjetividad médica.

4.1.4 Organización y división de la base de datos.

Una vez obtenido la clasificación de todos los pacientes, se separó la base de datos en dos archivos .csv, porque se decidió crear dos modelos de aprendizaje supervisado, debido a los tipos de atributos que surgieron de la consulta preanestésica creada, es por esto que se eligió trabajar por separado las variables correspondientes a la exploración física, las cuales son numéricas, y las variables de la anamnesis, las cuales son de tipo categóricas y pueden tomar diferentes valores.

4.2 Modelos de Inteligencia Artificial

4.2.1 Revisión de modelos de inteligencia artificial.

Se realizó una revisión de modelos de Inteligencia artificial que sirvieran para la clasificación de los datos, de acuerdo con las características de cada modelo y las características de los datos recopilados en la base de datos se eligieron dos que se consideraran apropiados para la clasificación de las dos clases de ASA, los atributos correspondientes a la fase de exploración física fueron trabajados con el algoritmo de Naive Bayes, y los pertenecientes a la etapa de la anamnesis con los algoritmos de árboles de decisión o clasificación.

4.2.2 Preprocesamiento de datos.

Para caracterizar los datos recolectados a partir de las reuniones y el protocolo médico para las consultas preanestésicas, se hizo un recuento de datos, se balancearon las clases, se graficó la distribución de los datos para cada variables, y sus gráficos de densidad para observar la diferenciación entre clases, haciendo un análisis estadístico básico, se encontró la correlación entre las variables, se normalizaron los datos y se convirtieron de categóricos a numéricos, con base en esta información se eligieron las variables que podrían ser más determinantes en la discriminación de las clases otorgadas por cada modelo trabajado.

4.2.3 Entrenamiento de los modelos.

Luego de la elección de los modelos de clasificación que pudieran funcionar bien con los datos que se tienen, cambiando para el modelo de Bayes los datos de entrenamiento que se normalizaron con dos técnicas diferentes, se entrenó con cada conjunto de datos y se obtuvo así los mejores resultados posibles en las clasificación, e igualmente para los árboles de decisión se eligieron ciertos parámetros de entrada para su respectivo entrenamiento y adquirir una división y clasificación de los datos lo mejor posible.

4.2.4 Validación de los modelos.

Para validar ambos modelos se encontró para cada uno los valores de exactitud, sensibilidad y especificidad a partir del cálculo de la matriz de confusión, además para el algoritmo de Bayes se graficaron las curvas ROC para cada conjunto de datos usado para su entrenamiento, y de esta manera se evaluó que tan efectivos eran los modelos en clasificar el riesgo preoperatorio entre ASA I Y II.

4.2.5 Evaluación de los modelos.

Luego de su respectiva validación se evaluaron ambos modelos con datos reales obtenidos de las consultas preanestésicas que se realizan en el HUSVF, Y se encontraron los porcentajes de acierto y desacierto de cada uno de los modelos, acompañados de la sensibilidad y la especificidad a partir del cálculo y la construcción de la matriz de confusión para los datos reales, como método de evaluación también fueron usados para su clasificación los datos trabajados para su entrenamiento.

5. RESULTADOS

5.1 Base de Datos

A continuación, se presentan los resultados obtenidos en el mismo orden en que se desarrolló en la metodología.

5.1.1 Datos consulta preanestésica.

Los datos otorgados por la médica y apoyados en la bibliografía encontrada respecto al tema de interés, y los cuales se tuvieron en cuenta a la hora de establecer los parámetros y variables para poder crear la base de datos y usarla para la construcción de los modelos se muestra en la Tabla 2, donde se muestran los datos elegidos.

Tabla 2. Datos recolectados y elegidos para la realización de la base de datos referente a una consulta preanestésica

MACRO TEMA	DESCRIPCIÓN
0. Cirugía a realizar	Se refiere al procedimiento quirúrgico que se le realizará al paciente y por el cual asiste a la consulta preanestésica.
1. Exploración física	Involucra las variables fisiológicas que se le toman en el momento de la consulta, las cuales son: Edad, Sexo, Estatura, Peso, Frecuencia cardíaca, Presión arterial, Saturación de Oxígeno e Índice de masa corporal.
2. Antecedentes quirúrgicos	Cuestionario sobre la historia clínica del paciente, referente a si ha sido intervenido quirúrgicamente, que tipo de cirugía y anestesia a la cual ha sido sometido.
3. Antecedentes patológicos	Preguntas sobre padecimientos de alguna enfermedad en particular, como tiroides, Diabetes, Hipertensión y enfermedad pulmonar, aquí se tiene en cuenta el tiempo de evolución de la enfermedad, el control y las complicaciones de esta.
4. Antecedentes tóxicos	Información sobre el consumo de alucinógenos, alcohol o tabaquismo y con qué frecuencia se realiza.
5. Antecedentes alérgicos	Tipos de alergias que puede presentar el paciente, ya sea a algún medicamento o al látex
6. Evaluación estado físico	Serie de preguntas que se le realizan al paciente sobre su capacidad para respirar bajo ciertas actividades cotidianas, como caminar y subir escaleras, dando información sobre su fatiga y el funcionamiento del sistema cardiovascular y respiratorio.
7. Medicamentos	Tipos de medicamentos que está ingiriendo el paciente o si recientemente suspendió alguno.
8. Embarazo	Conocimiento sobre la etapa gestante en caso de que el paciente sea una mujer y este en los años de fertilidad y tengo vida sexual activa, incluyendo el tipo de planificación familiar y riesgo de embarazo.
9. Antecedentes familiares	Información sobre complicaciones anestésicas o quirúrgicas que se hayan presentado en el círculo familiar del paciente.

Cada uno de los datos elegidos se transformaron de tal manera que al ser usados en el modelo fueran clasificables, es decir, que cada una de las variables que son importantes para la clasificación ASA se convirtieron en preguntas cerradas o con múltiples respuestas determinadas para que las entradas del algoritmo fueron objetivas y dirigidas a las categorías de interés, y se hiciera más fácil el entrenamiento y la evaluación del modelo de inteligencia artificial.

5.1.2 Recopilación y creación de datos.

De los datos y parámetros que se muestran en la Tabla 2 se establecieron las variables para la consulta preanestésica que se llevó a cabo, de la cual resultaron 54 variables independientes, que se distribuyeron inicialmente como se muestra en la Tabla 3, todas estas variables incluyen los diferentes valores que puede tomar cada pregunta realizada. En la Figura 10 se muestra la composición de la base de datos con todas las variables y sus respectivos valores.

Tabla 3. Cantidad de variables independientes obtenidas inicialmente para la consulta preanestésica y su respectiva distribución de acuerdo a los parámetros establecidos.

PARÁMETROS	CANTIDAD DE VARIABLES
Exploración física	8 variables
Antecedentes quirúrgicos	46 variables independientes distribuidas entre los 9 macro temas
Antecedentes patológicos	
Antecedentes tóxicos	
Antecedentes alérgicos	
Procedimiento quirúrgico a realizar	
Evaluación estado físico	
Medicamentos	
Embarazo	
Antecedentes familiares	

5.1.3 Clasificación de la base de datos.

En la Figura 11 se muestra la clasificación otorgada por cada uno de los anesthesiólogos a todos los pacientes creados, y en la Figura 12 se observa el promedio de dichas clasificaciones y el cual fue el valor usado por la variable ASA para la realización de los dos modelos de aprendizaje automático implementados.

5.1.4 Organización y división de la base de datos.

Se dividió la base de datos en dos, según el tipo de variables o atributos al que perteneciera, es decir que con el macro tema correspondiente a la exploración física que todas sus variables

son de tipo numéricas se creó una base de datos, la cual se compuso de 8 atributos para cada uno de los 50 casos de la base de datos global, e igualmente para la etapa de anamnesis se construyó otra base de datos con las 46 variables restantes independientes que son categóricas o booleanas.

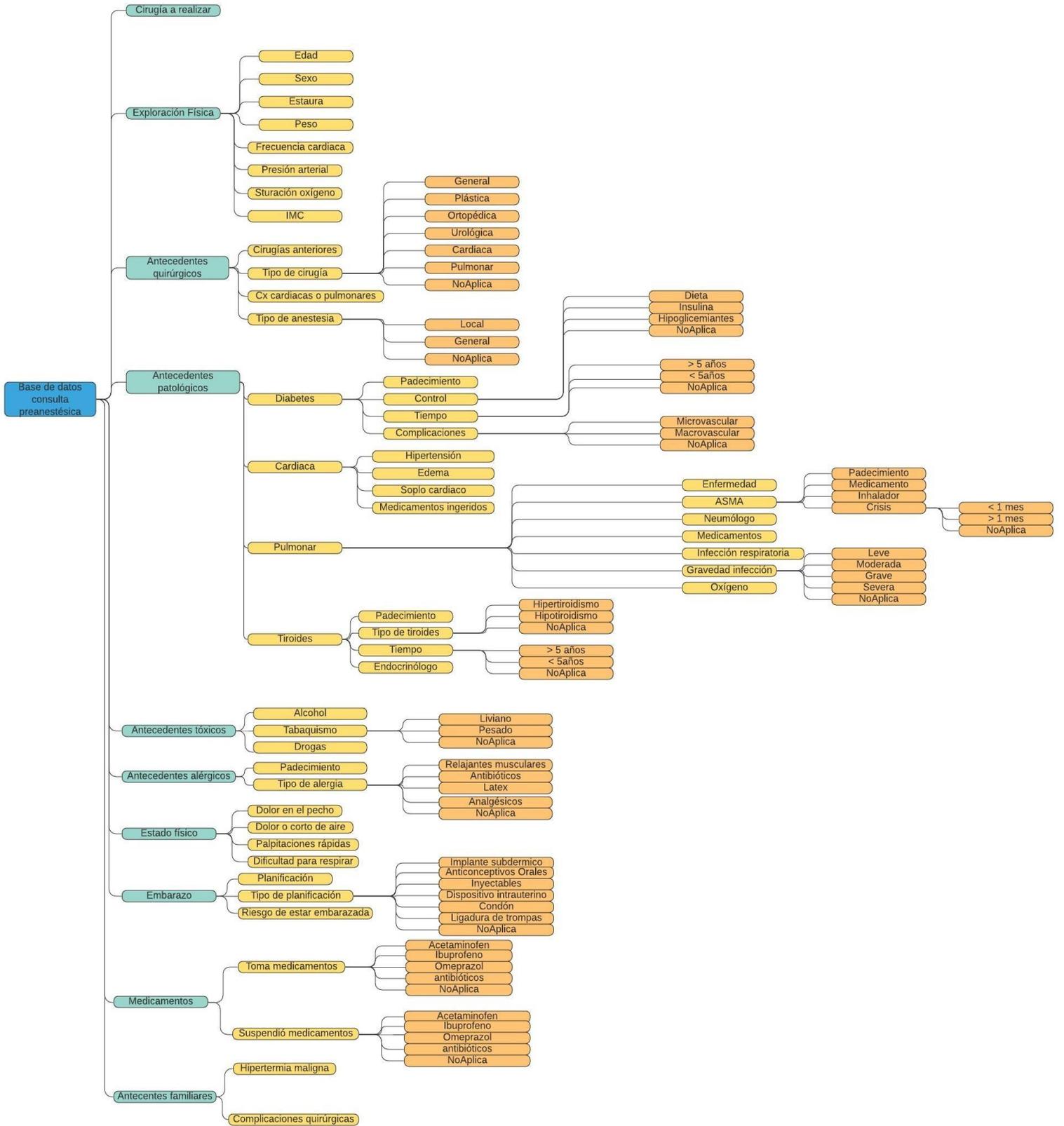


Figura 10. Composición de la base de datos

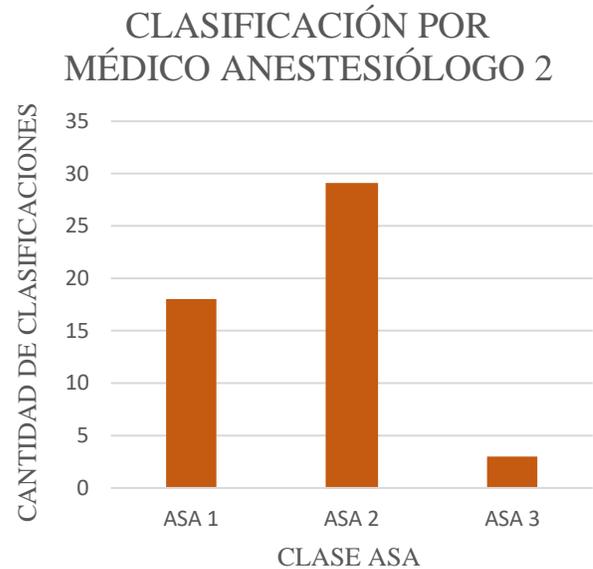
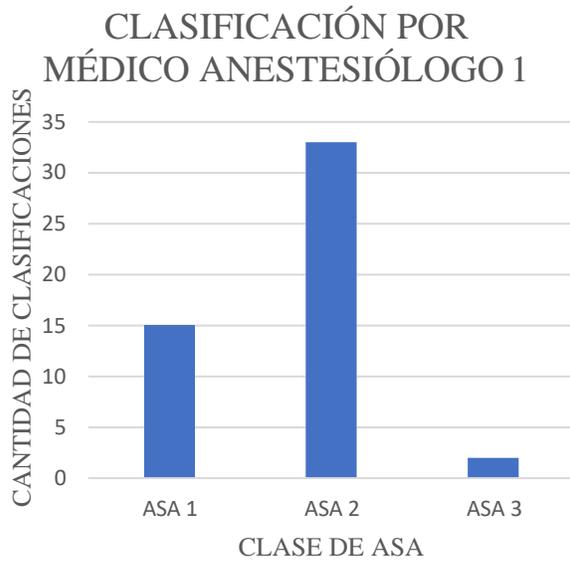


Figura 11. Número de clasificaciones ASA otorgadas por cada uno de los médicos anestesiólogos..

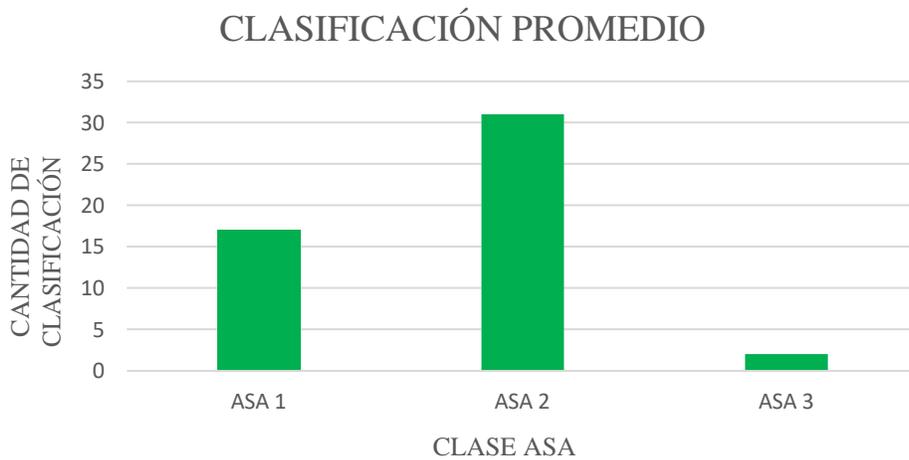


Figura 12. Número de clasificaciones promediadas y usadas para el entrenamiento de los dos modelos de inteligencia artificial.

5.2 Modelos de Inteligencia Artificial

5.2.1 Preprocesamiento de datos y entrenamiento de modelos.

Descripción de los datos

Los datos usados para el análisis son los construidos o creados a partir de la información recolectada, los datos consisten en una historia que es conformada por respuestas a preguntas cerradas o con diferentes opciones de respuesta que se le hacen al paciente al momento de su consulta y dan cuenta de su estado de salud y como se encuentran sus signos vitales en ese momento, este formulario tiene un total de 54 variables independientes, en estos muestras se consideró solo personas entre los 18 y los 60 años de edad, para los modelos se usaron 50

datos de consultas preanestésicas inicialmente, pero como se muestra en la Tabla 4 la distribución de los datos iniciales estaba un poco desproporcionada, en la cual había un mayor número de pacientes clasificados como ASA II, y a dos pacientes se les otorgó un nivel de ASA que no se estaba trabajando en el proyecto, por esta razón se vio la necesidad de balancear la base de datos con el mismo número de pacientes para cada clase, en este caso se optó por obtener 30 pacientes ASA II Y ASA I, para esto se crearon 13 pacientes adicionales, con características similares a los 17 clasificados con ASA I anteriormente, y se les asignó el mismo valor de ASA para completar el número requerido, finalmente se procedió a eliminar el paciente adicional clasificado como ASA II y los dos restantes como ASA III, y se realizó la construcción de las bases de datos para usarlas en los dos modelos creados.

Tabla 4. Distribución de la clasificación de los 50 pacientes creados inicialmente.

Tipo de clasificación ASA	Cantidad de sujetos
ASA I	17
ASA II	31
ASA III	2

La distribución porcentual de las clases se observa en la Figura 13.

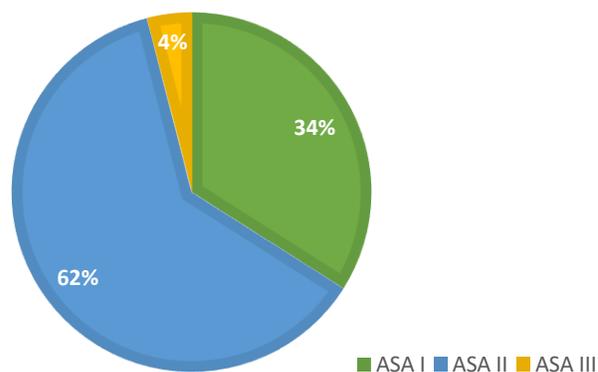
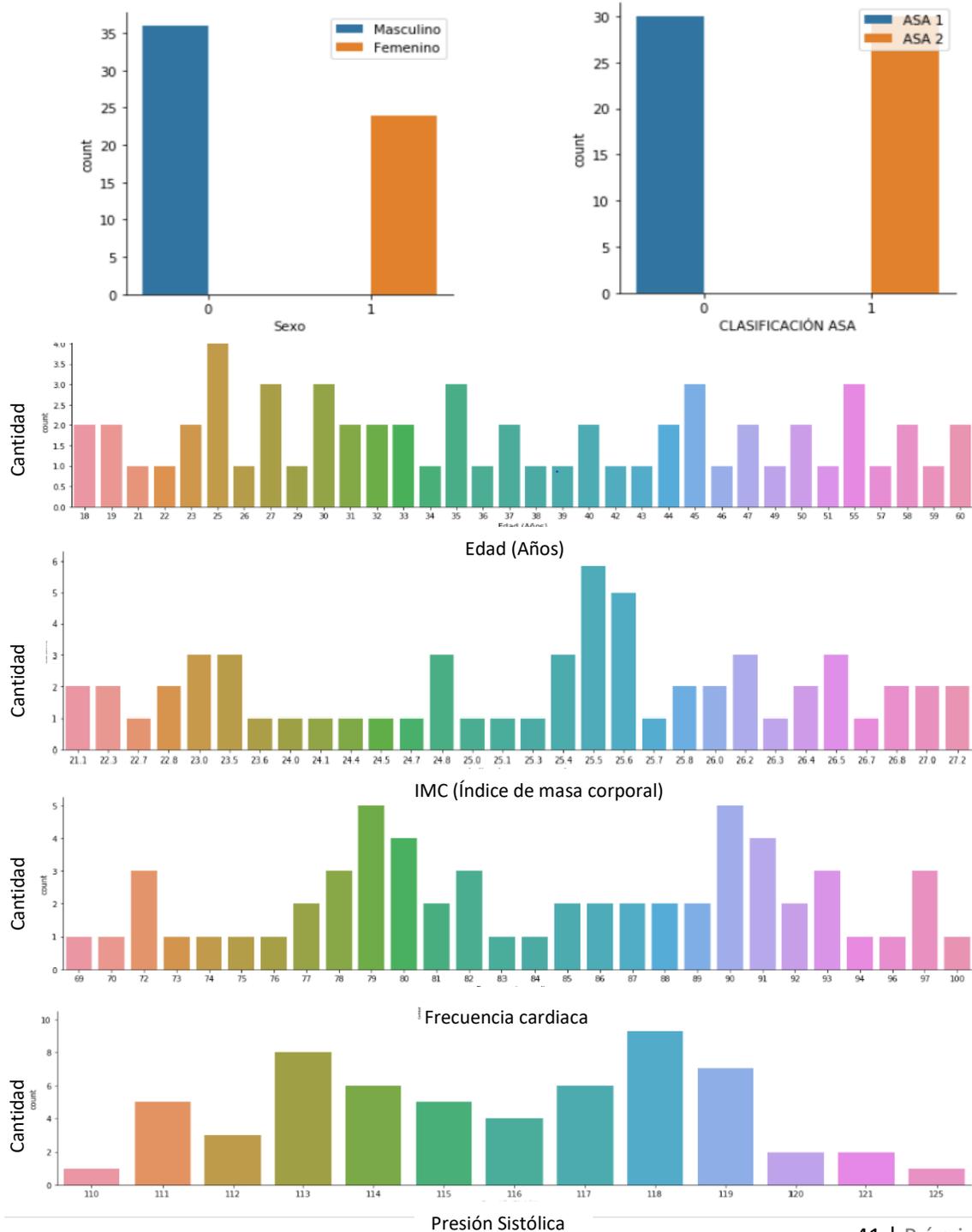


Figura 13. Porcentajes de las clases de ASA

A continuación, se muestran algunas gráficas y tablas donde se observa cómo es la distribución de los datos para las variables de cada modelo, correspondiente a la etapa del preprocesamiento de la base de datos en cualquier algoritmo de machine learning.

5.2.1.1 Preprocesamiento de datos para el algoritmo de Naive Bayes (Exploración Física)

En la Figura 14 se observa la distribución de cada una de las variables usadas en la exploración Física, para analizar si hay una cantidad significativa para cada uno de los valores posibles, entre las 60 muestras.



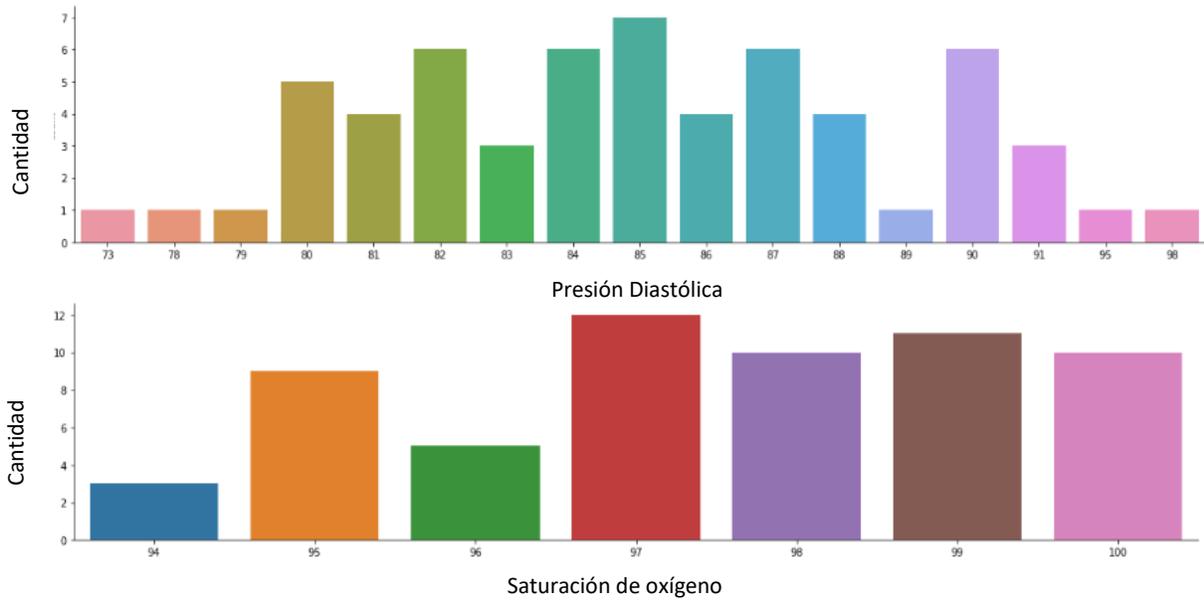
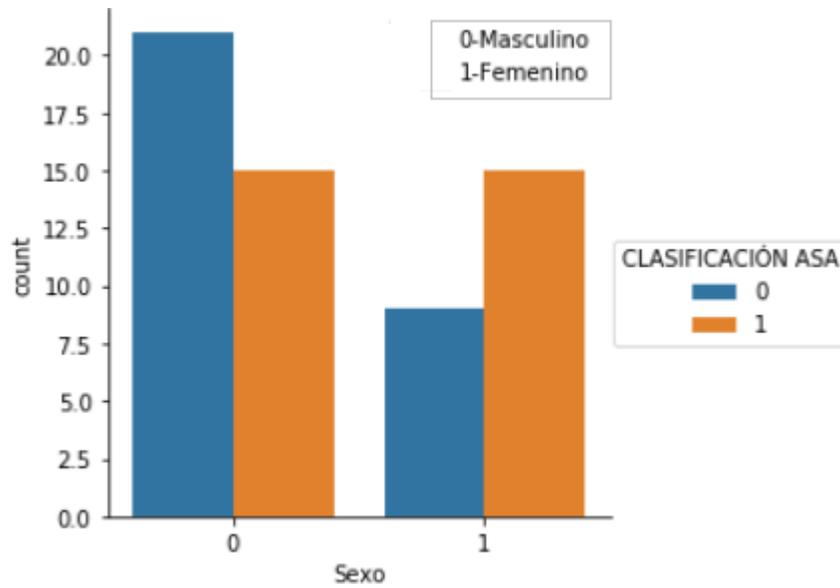


Figura 14. Distribución de cada variables perteneciente a la exploración física de la consulta preanestésica, A)Sexo (0-H,1-M); B) ASA;C) Edad; D) IMC (índice de masa corporal); E) Frecuencia cardiaca; F) Presión arterial, G) Saturación de oxígeno

En la Figura 15 se evidencia la cantidad de datos correspondiente a cada clase, dependiendo del atributo de la exploración física, donde la etiqueta de color azul hace alusión a los casos clasificados como ASA I, y la etiqueta de color naranja representa a los clasificados con ASA II, el 0 es el sexo masculino y el 1 el sexo femenino.



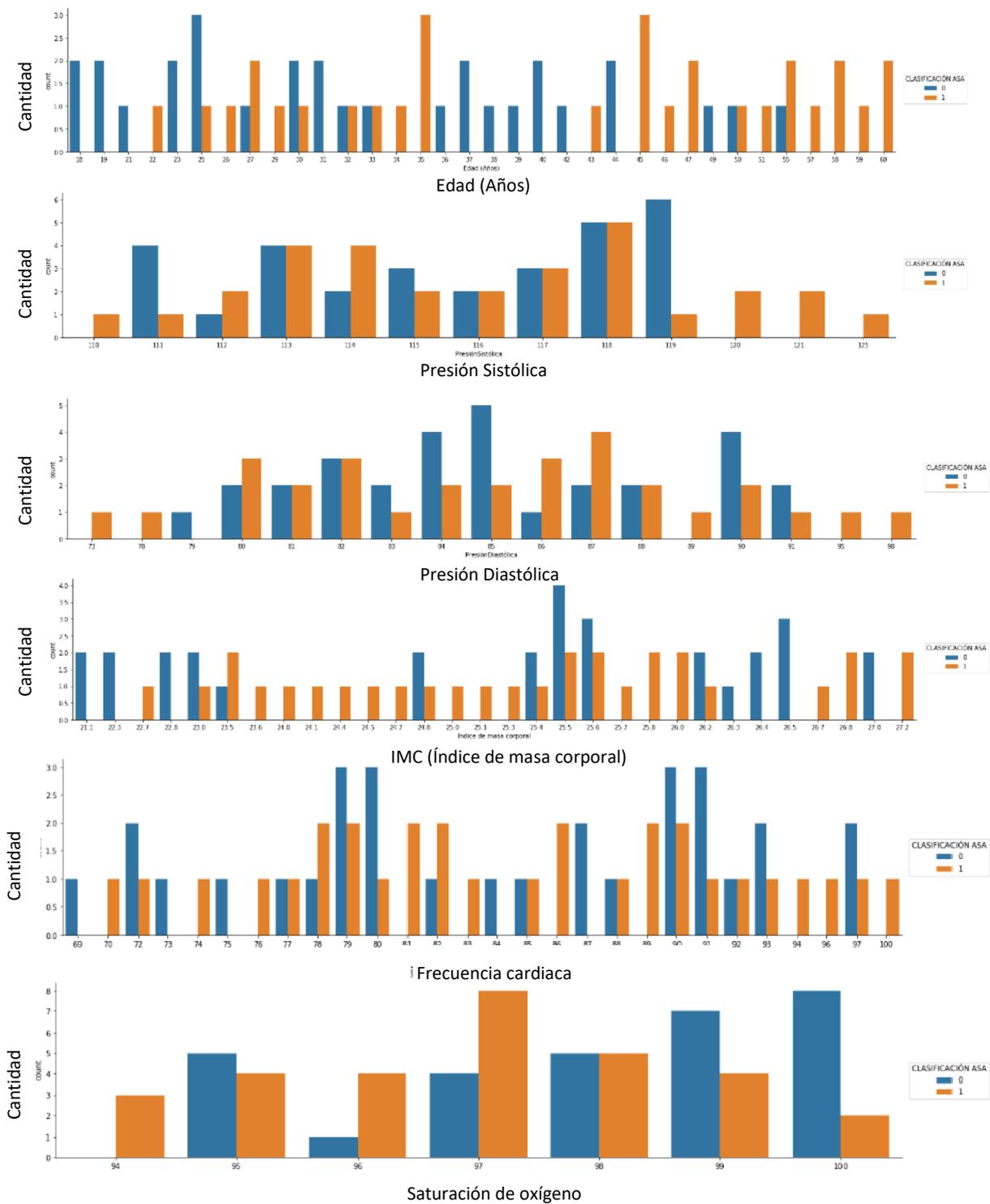


Figura 15. distribución de las variables de la exploración física según el tipo de ASA, A) Sexo; B) Edad; C) Presión arterial; D) IMC (índice de masa corporal); E) Frecuencia cardiaca; F) Saturación de oxígeno.

En la Figura 16 se muestra la metodología utilizada, la cual contiene cada proceso realizado a los datos usados en el modelo correspondiente a la exploración física y el posterior entrenamiento del algoritmo de Bayes.

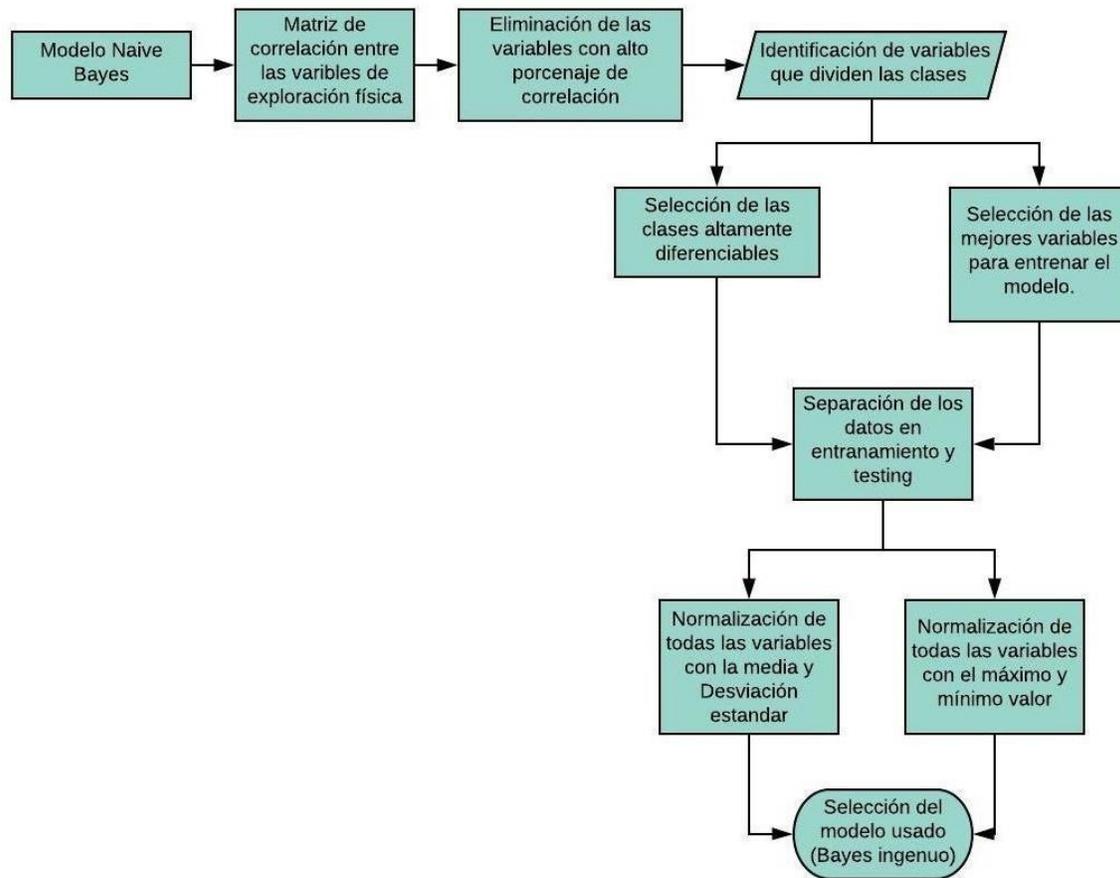


Figura 16. Metodología del preprocesamiento de los datos de la exploración física para el modelo de Naive Bayes

- **Cálculo de la matriz de correlación de las variables:** Se procedió a realizar la matriz de correlación para conocer como es el porcentaje de correlación entre la información otorgada por cada una de las variables a la hora de clasificar los pacientes, y evitar que sea redundante la información para entrenar el modelo de Bayes.

En la Figura 17 se presenta la matriz de correlación entre las variables creadas para la exploración física.

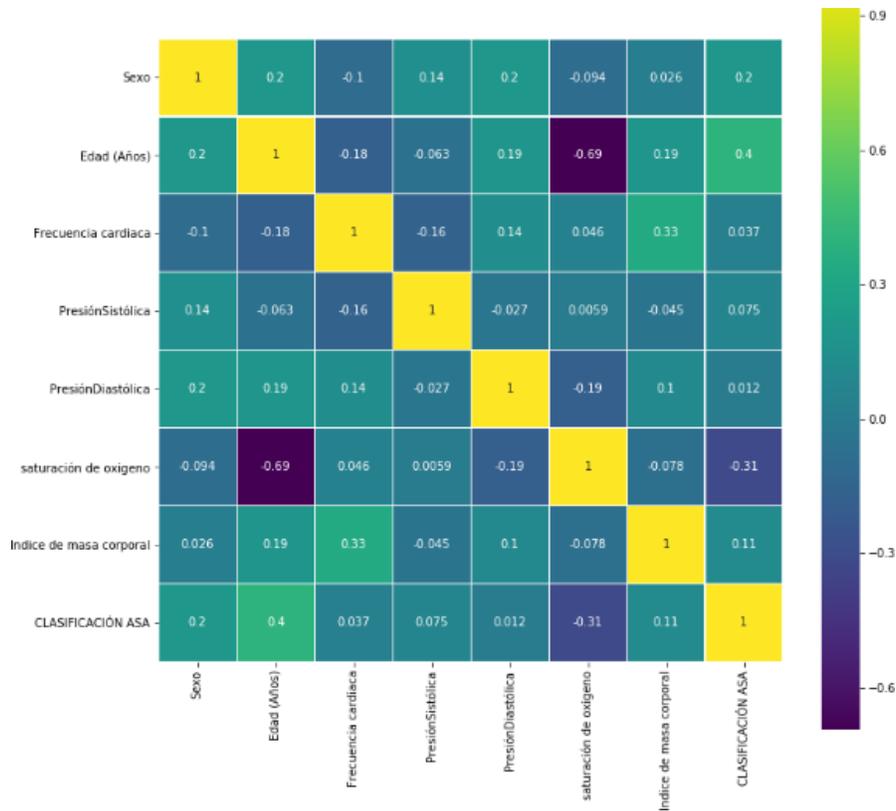
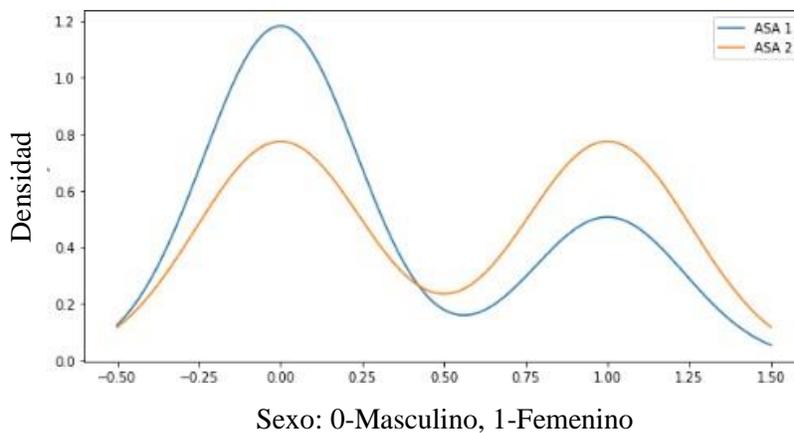
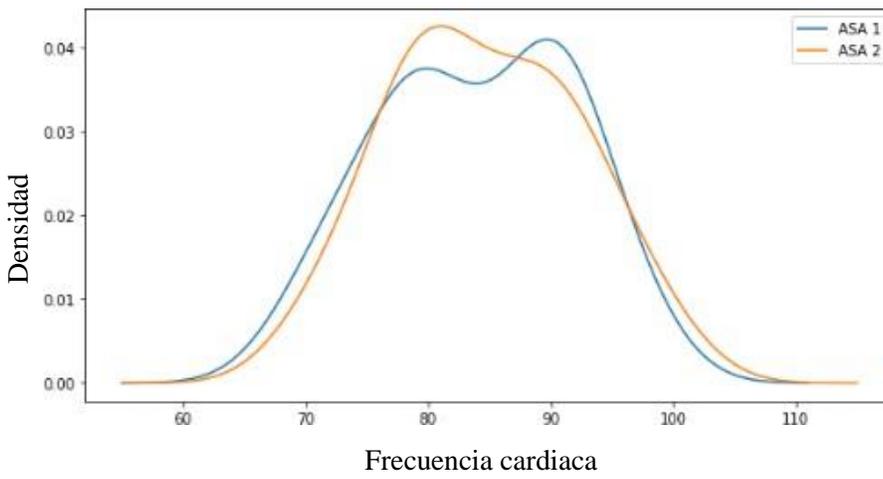
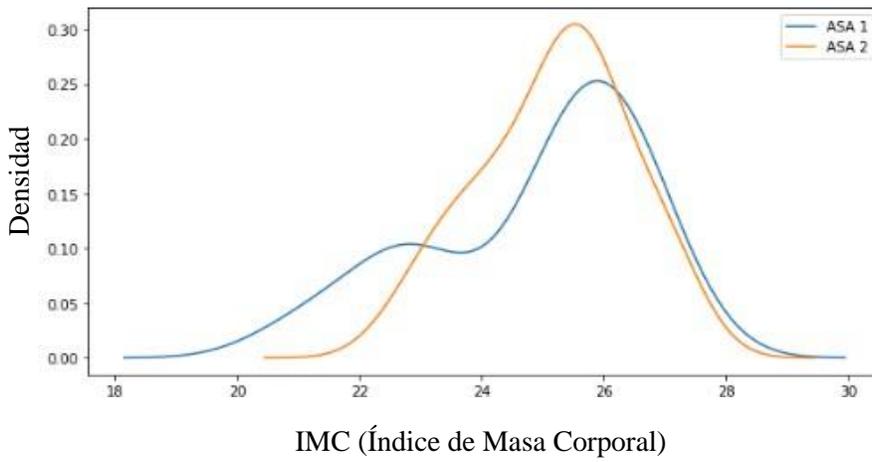
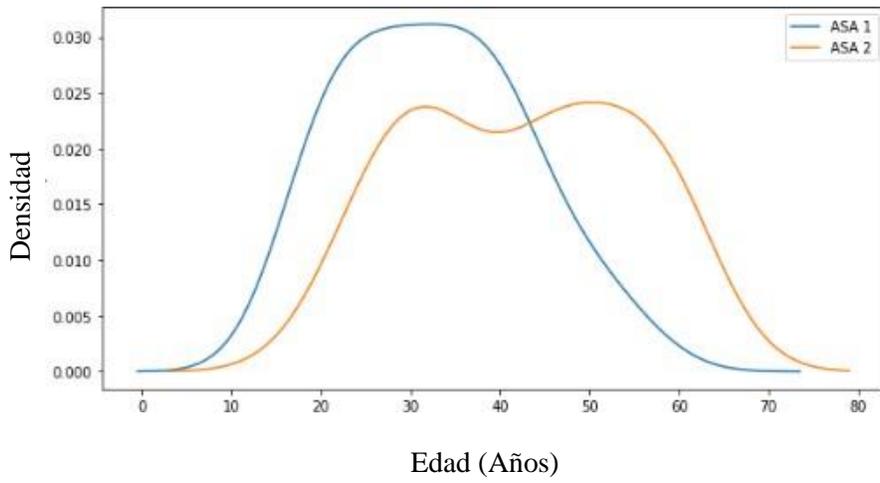


Figura 17 Matriz de correlación entre las 8 variables involucradas para entrenamiento del modelo Bayesiano

- Identificación de variables que dividen las clases:** Se realizaron las gráficas de densidad de cada una de las variables de la exploración física de acuerdo con cada una de las clases de ASA, para conocer su distribución de probabilidad en la clasificación, y conocer si la variable representaba una diferencia notable entre dichas clases a la hora de usarla para entrenar el modelo.

En la Figura 18 se observan cada una de las gráficas que representan la distribución de las variables en estudio, de acuerdo con el tipo de ASA.





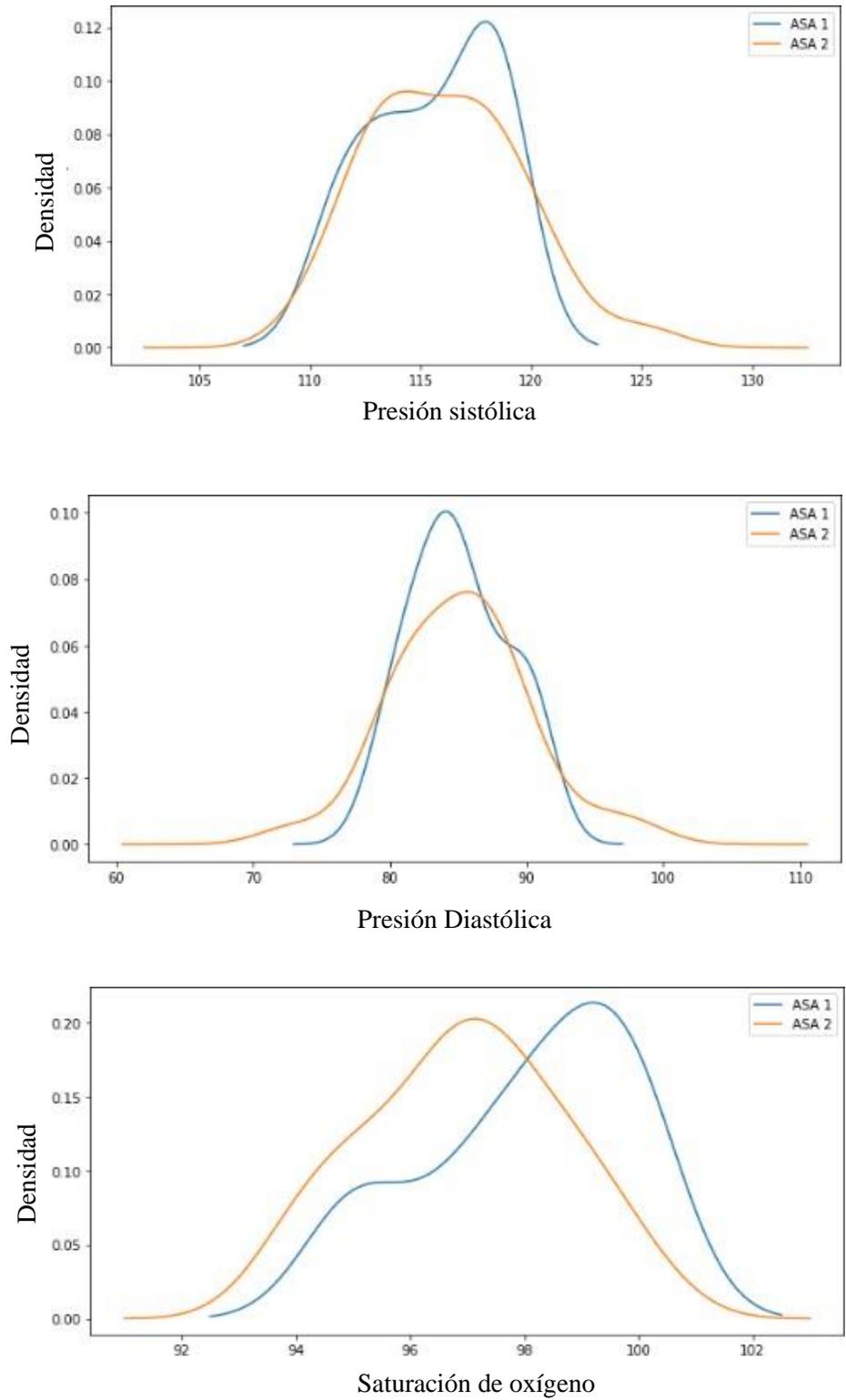


Figura 18. Gráficas de densidad para las 8 variables. A) Sexo; B) Edad; C) IMC; D) Frecuencia cardiaca; E) Presión arterial (Sistólica y diastólica); F) Saturación de oxígeno.

- **Selección del modelo y su entrenamiento:** Para la construcción de este modelo se eligió el algoritmo ingenuo de Bayes, el cual trabaja bajo el supuesto de que existe independencia entre los predictores usados para el entrenamiento, luego de esta elección se procedió a separar los datos predictores de las etiquetas de clasificación correspondiente a cada una de los 60 pacientes, después se realizaron dos tipo de normalización a los datos, para saber con cuál de ellas se obtenía un mejor rendimiento del modelo a lo hora de entrenarlo, dicha normalización en ambos casos fue aplicada a las 8 variables trabajadas, y al momento de entrenarlo solo fueron seleccionadas Sexo, Edad y Saturación de oxígeno.

Los dos tipo de normalización usados fueron: Escalado de variables (Feature Scaling o MinMax Scaler), el cual normaliza cada entrada entre unos límites definidos, y Escalado estándar (Standard Scaler) que a cada dato se le resta la media de la variable y se le divide por la desviación típica, con esta normalización se obtuvo valores entre -1 y 1 por lo que se procedió a calcular el valor absoluto de cada dato para obtener valores entre 0 y 1.

5.2.1.2 Preprocesamiento de datos para el algoritmo árboles de clasificación (Anamnesis).

En la Figura 19 se presenta la distribución de las cirugías antecesoras con la anestesia aplicada a los pacientes con su respectiva clasificación ASA, para esto se debe explicar, que como se hizo para la exploración física la etiqueta 0 corresponde al tipo ASA I y la etiqueta 1 al tipo ASA II, y finalmente se muestra el tipo de anestesia aplicada para antecedentes quirúrgicos según el tipo de cirugía realizada.

En las Figuras 20, 21, 22 y 23 se observan las distribuciones de los pacientes que padecen alguna enfermedad, sea Diabetes, Tiroides, Hipertensión o Asma respectivamente y como esto influye en la clasificación otorgada, las etiquetas en las enfermedades representan solo dos tipos de respuesta, ya que dan cuenta de un dato de entrada booleana, es decir la etiqueta 0 representa los casos negativos y la entrada 1 los positivos, esto se realizó con todos los parámetro de entrada de la base de datos perteneciente a la anamnesis que representaban datos de salida booleanos.

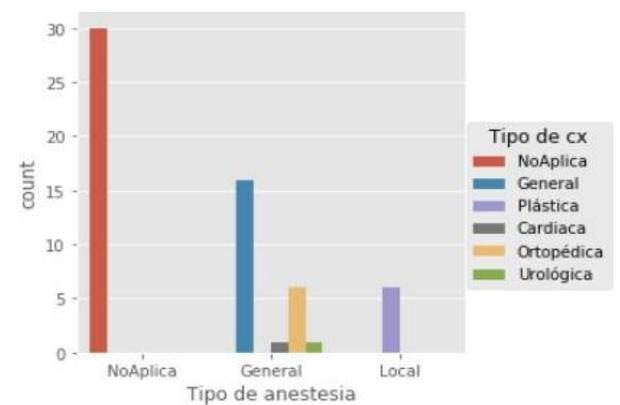
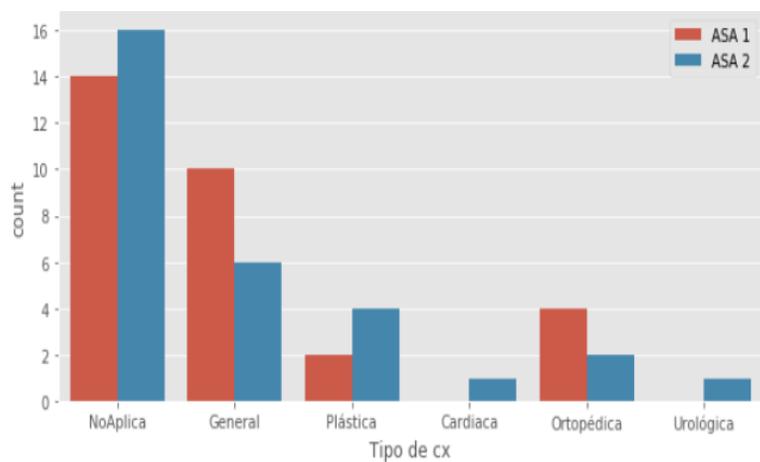
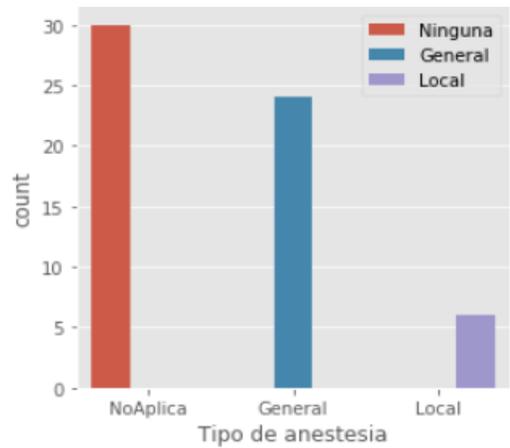
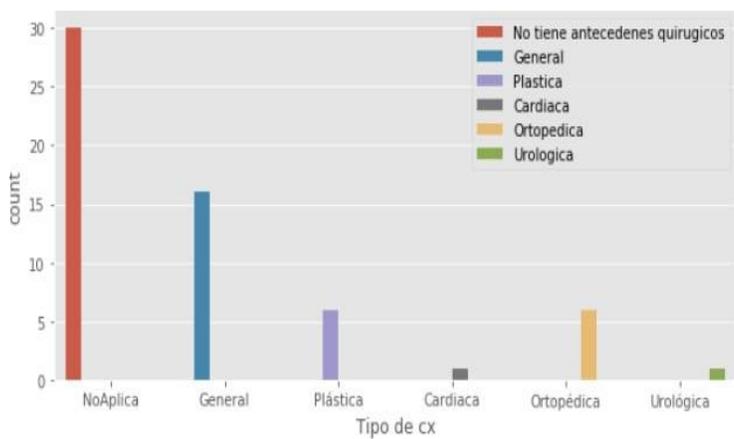


Figura 20 Distribución de las variables correspondientes a los antecedentes quirúrgicos y el tipo de ASA asignadas

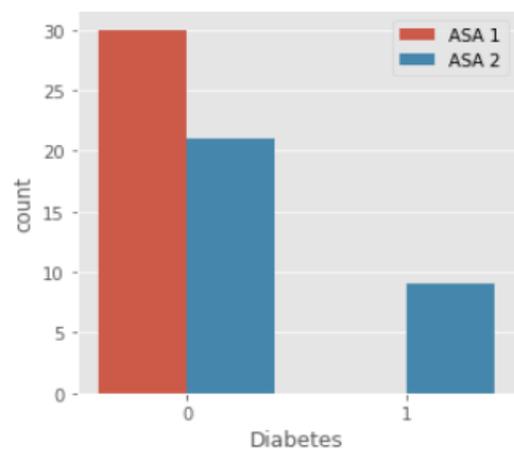
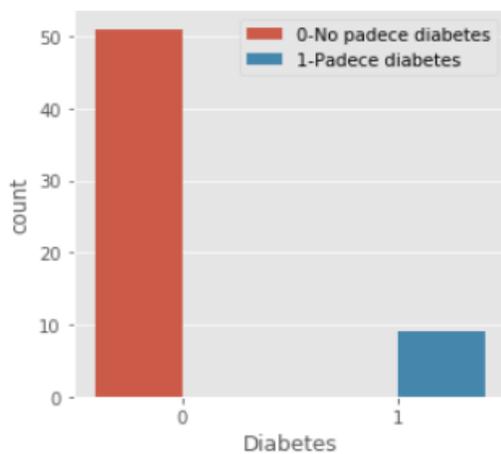


Figura 19 Padecimiento de la diabetes. A) Distribución de las muestras; B) Clasificación otorgada de acuerdo a las muestras con diabetes

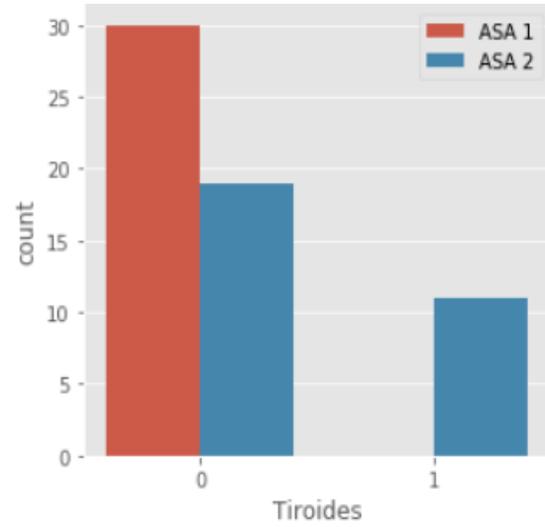
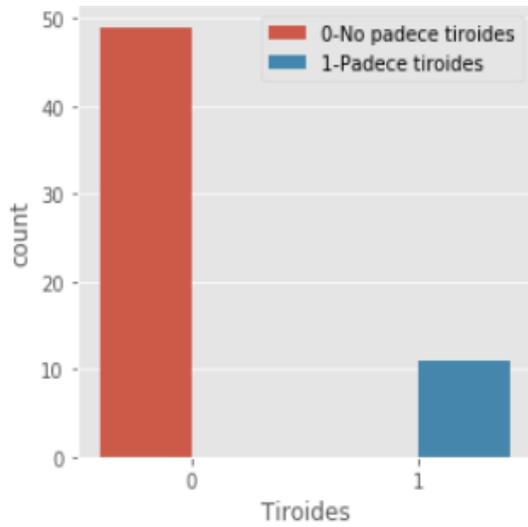


Figura 22 Padecimiento de tiroides A) Distribución de las muestras; B) Clasificación otorgada de acuerdo a las muestras con Tiroides.

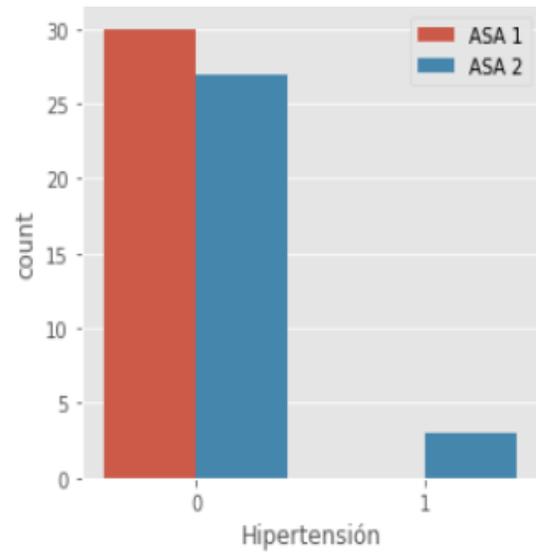
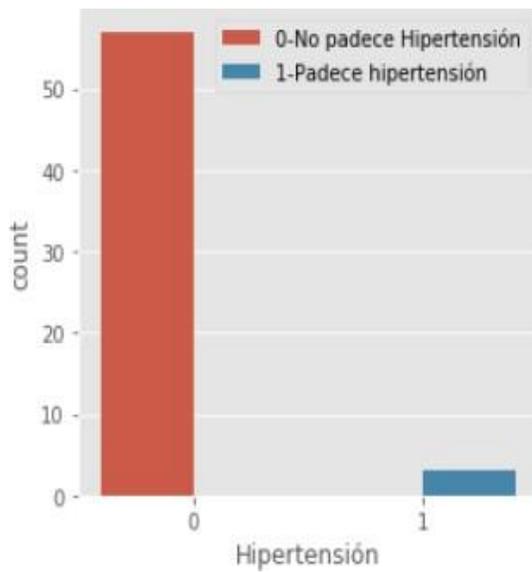


Figura 21 Padecimiento de la hipertensión. A) Distribución de las muestras; B) Clasificación otorgada de acuerdo a las muestras con hipertensión

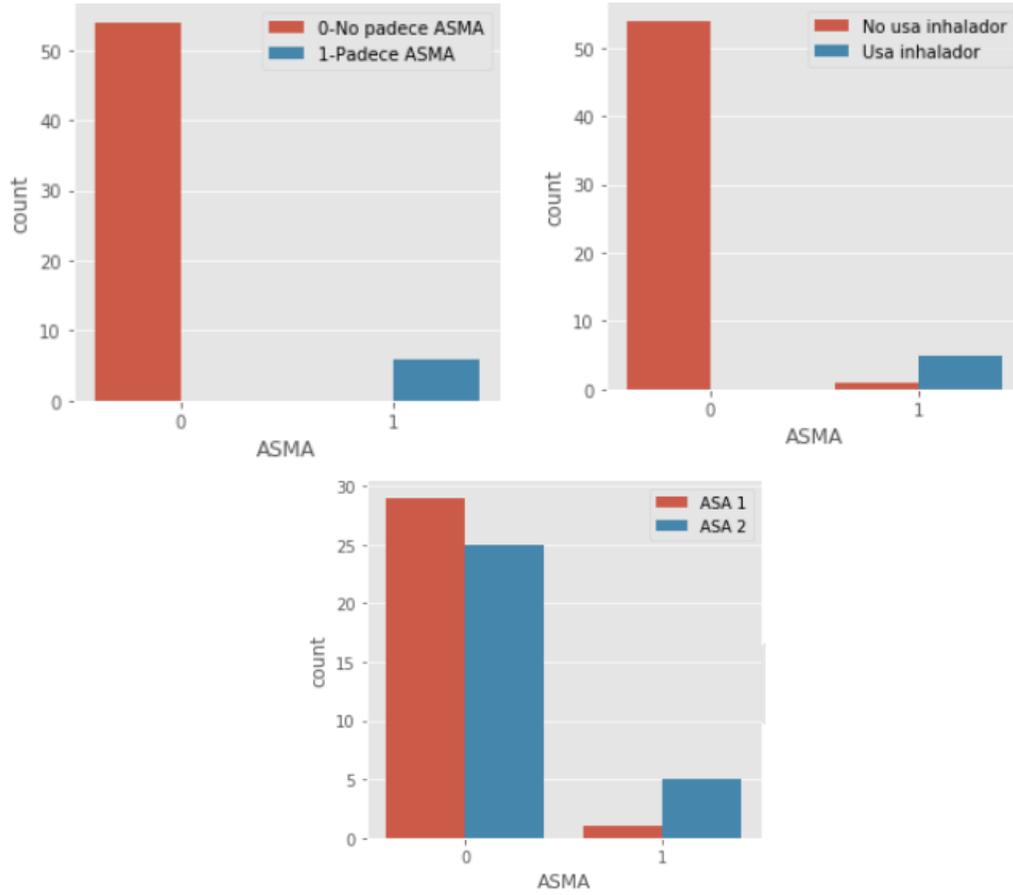


Figura 23 Padecimiento de ASMA. A) Distribución de las muestras; B) Distribución según el uso de inhalador; C) Clasificación otorgada de acuerdo a las muestras con ASMA

En la Figura 24 se muestra la metodología utilizada con cada proceso realizado a los datos para el entrenamiento del algoritmo de árbol de decisión, correspondiente a la anamnesis.

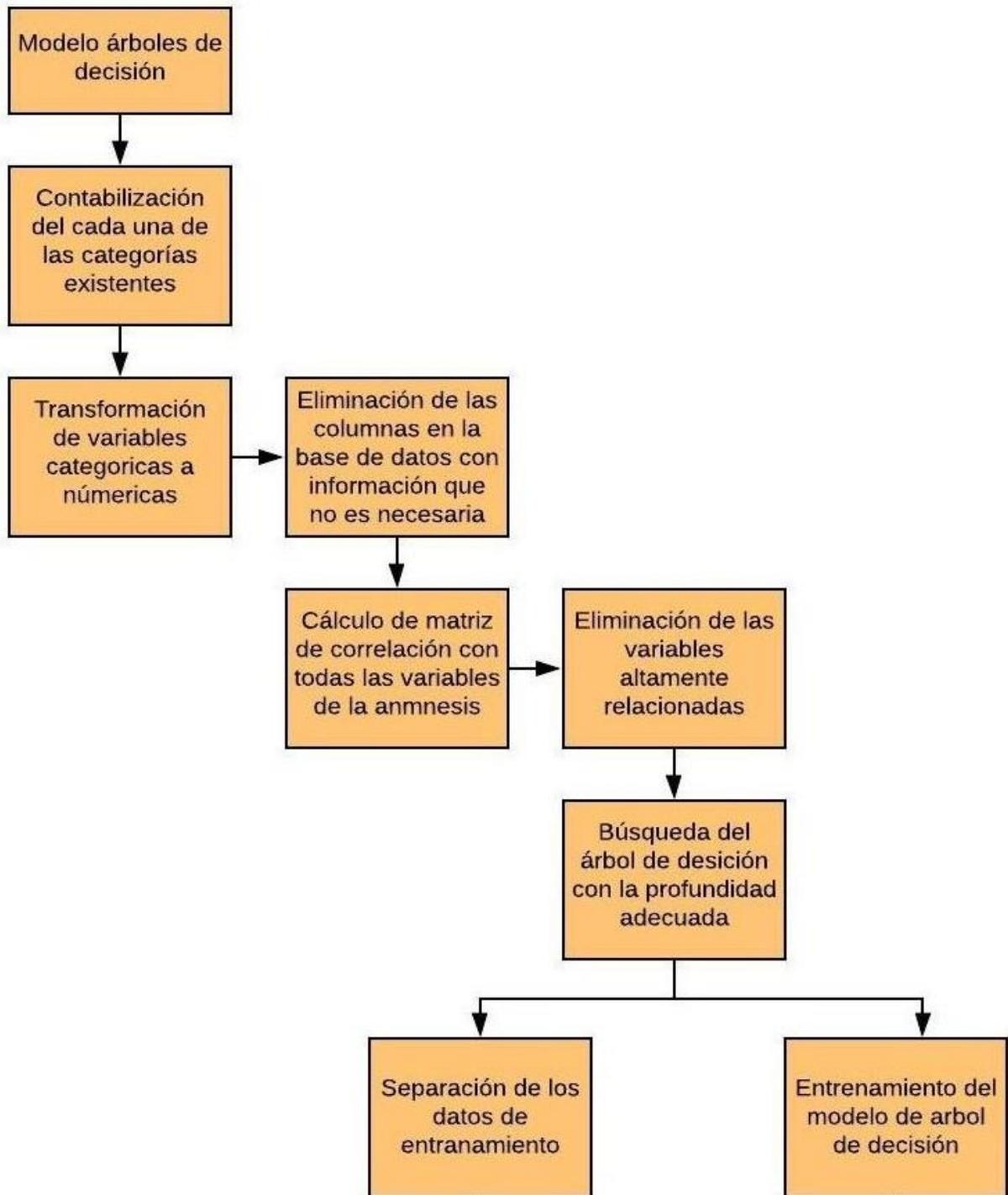


Figura 24 Metodología para el preprocesamiento de los datos correspondientes a la anamnesis, para el modelo de árbol de clasificación

- **Transformación de variables categóricas a numéricas:** Se agruparon las categorías de las variables asignando un valor a cada una, teniendo presente su respectiva importancia de acuerdo a la cantidad existente en ella, es decir que la categoría con mayor número de datos, se le asignó el valor con la más alta denominación dentro del número total de opciones que puede tomar la variable, esto representa los pesos que se le dan para que el algoritmo o modelo cuando prediga la clasificación de un nuevo paciente tenga presente las categorías más representativas en la base de datos con las cuales fue entrenado el modelo.

Se realizó el conteo total por cada categoría presente en la base de datos, perteneciente a las variables que dentro de la anamnesis pueden tomar diferentes valores de respuesta, en la Tabla 5 se muestra dicha contabilización con el valor otorgado para la creación del modelo a modo de ejemplificación, solo para tres de las 14 variables categorizables.

Tabla 5. Recuento de la cantidad de pacientes existentes por cada clase para 3 de las 14 variables categóricas

Variable Base de Datos	Categorías	Cantidad	Valor Asignado
Cirugía	General	24	3
	Plástica	15	2
	Ortopédica	14	1
	Urológica	7	0
Tipo de Cirugía	No Aplica	30	5
	General	16	4
	Plástica	6	3
	Ortopédica	6	2
	Urológica	1	1
	Cardíaca	1	0
Tipo de anestesia	No Aplica	30	2
	General	24	1
	Local	6	0
Control de diabetes	No Aplica	51	3
	Dieta	6	2
	Insulina	2	1
	Hipoglicemiantes	1	0

- **Eliminación de la información innecesaria:** Debido a que se obtuvo un nuevo conjunto de datos, el cual fue conformado por los variables definitivas para crear el árbol, se eliminaron todas las columnas que no eran necesarias, como las que tenían respuestas categóricas ya que sus opciones de respuesta se pasaron a numérico, además se tomó la decisión de eliminar las variables que en todos los pacientes tenían la misma respuesta, la cual era un No, pero que representa un 0 en la base de datos, quedando así la base de datos con un total de 33 atributos o datos de entrada para la anamnesis.

- **Matriz de correlación:** Se calculó la correlación entre los 35 atributos para verificar y analizar que otro dato de entrada era pertinente eliminar, debido a que dieran información altamente similar, efectivamente se encontró que algunas de las variables tenían un alto porcentaje de relación por lo que se usó un umbral del 80% para eliminar aquellas que tuvieran este o un valor mayor de similitud, con el fin de finalizar la depuración de los datos y obtener un conjunto de datos lo más libre de ruido posible, y empezar con el entrenamiento del algoritmo de árboles de decisión para la clasificación.

En la Tabla 6 se muestra todas las 33 variables que quedaron en la base de datos usada para el entrenamiento del árbol de clasificación, luego de la eliminación de información redundante.

Tabla 6. Recopilación de todas las variables eliminados de la base de datos inicial para la anamnesis

Variables de entrada finales para la anamnesis	
Cirugías anteriores	Suspendió medicamento
Cirugías cardiacas o pulmonares	Planificación
Diabetes	Hipertermia maligna
Hipertensión	Complicaciones Familiares
ASMA	Procedimiento a realizar
Toma medicamentos	Tipo de cirugía antecesoras
Neumólogo	Tipo de anestesia antecesora
Infección respiratoria	Control diabetes
Tiroides	Tipo de infección respiratoria
Alcohol	Tipo de tiroides
Tabaquismo	Tipo de tabaquismo
Drogas	Tipo de alergia
Alergias	Tipo de medicamento consumido
Dolor pecho actividad física	Tipo de medicamento suspendido
dolor o corto de aire al subir 3 o 4 pisos	Método de planificación
Palpitaciones rápidas	Última crisis de ASMA
Algún medicamento	

En la tabla 7 se ilustran las variables que presentaron un porcentaje de correlación mayor o igual al 80%, y cuales se decidieron eliminar.

Tabla 7. Variables altamente correlacionadas de la anamnesis con su porcentaje y las variables seleccionadas para la eliminación

Variables relacionadas	Porcentaje de correlación	Variable eliminada
ASMA- Uso de inhalador	90%	Uso de inhalador
Control de diabetes-Tiempo de padecimiento de diabetes	89%	Tiempo de padecimiento de diabetes
Tipo de tiroides-Tiempo de padecimiento de tiroides	94%	Tiempo de padecimiento de tiroides
Hipertensión- Medicamentos para la presión o corazón	99%	Medicamentos para la presión o corazón

- **Búsqueda de la profundidad del árbol de decisión y ajuste de parámetros:** Se encontró con cuántos niveles de profundidad se debía entrenar el modelo, para ello, se hizo uso de la función de K-Fold que crea varios subgrupos con el conjunto de datos de entrada para validar y valorar los árboles con diversos niveles de profundidad. Entre ellos, se escogió el mejor resultado.

El entrenamiento se llevó a cabo con la separación del conjunto de datos, entre los datos que corresponden a las características de entrada (variables independientes) y las variable de clasificación o distinción de clases (variable dependiente), además se escogió el valor de los parámetros de entrada para la creación del modelo.

Se obtuvo que luego de 7 niveles de división, un porcentaje de precisión del 97%. La configuración de los parámetros para construir el árbol de decisión fue: Criterion=entropy se utiliza para entradas categóricas, Min_samples_split= 3.

Se realizó el diagrama del árbol mostrando cual fue su ruta de decisión y los valores de sus parámetros en cada nodo (Figura 25).

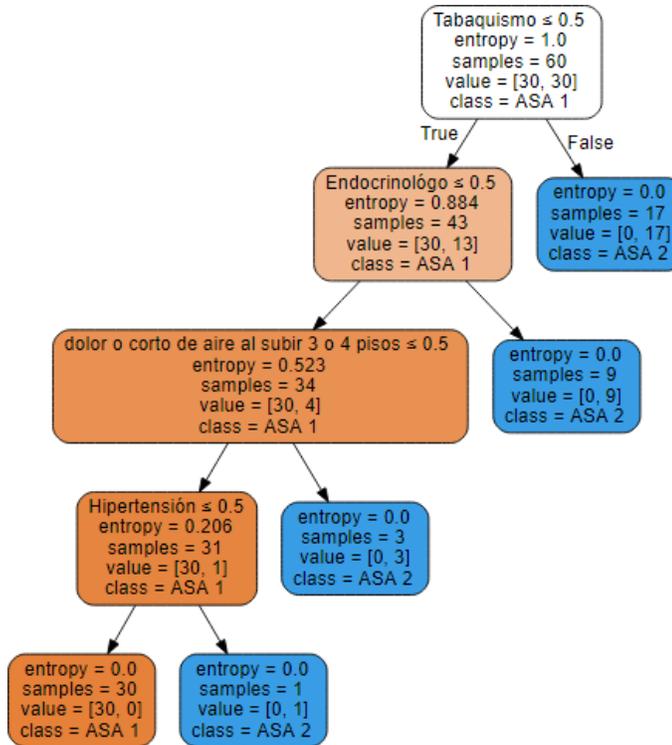


Figura 25 Resultado del modelo de árbol de clasificación creado e implementado a la anamnesis

5.2.2 Validación de los modelos.

Modelo de Naive Bayes.

Antes de entrenar el modelo se separaron los datos de entrenamiento y evaluación con una relación de 80/20 respectivamente, y así poder evaluar su viabilidad para la clasificación ASA en una consulta preanestésica con los datos de la exploración física del paciente, a través de un algoritmo de aprendizaje supervisado de Bayes.

Una vez entrenado el modelo ingenuo de Bayes con las tres características seleccionadas y los datos normalizados con dos técnicas diferentes se realizó el cálculo de las métricas para su evaluación de rendimiento, y conocer con cual método de normalización hubo mayor exactitud en la clasificación de los datos de evaluación. En las Tablas 8 y 9 se muestran las matrices de confusión resultantes de la predicción del modelo con cada uno de los datasets normalización.

Tabla 8. Matriz de confusión con el método de normalización estándar

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	3	3
	Negativos	4	2

Tabla 9. Matriz de confusión con el método de normalización de límites definidos.

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	4	2
	Negativos	1	5

En la Tabla 10 se muestran las métricas obtenidas para el modelo con la normalización de límites definidos, y en la Tabla 11 se muestran los resultados para la normalización estándar.

Tabla 10. Métricas calculadas al modelo bayesiano con el método de normalización estándar

Métricas	Valor
Precisión en el set de Entrenamiento	67%
Exactitud del modelo de Bayes en la exploración Física	42%
Promedio ponderado del modelo de Bayes en la exploración Física	37%
Sensibilidad del modelo	50%
Especificidad del modelo	33%
Total de Muestras en Test	12
Fallos	7

Tabla 11. Métricas calculadas al modelo bayesiano con el método de normalización de límites definidos.

Métricas	Valor
Precisión en el set de Entrenamiento	65%
Exactitud del modelo de Bayes en la exploración Física	75%
Promedio ponderado del modelo de Bayes en la exploración Física	77%
Sensibilidad del modelo	66%
Especificidad del modelo	83%
Total de Muestras en Test	12
Fallos	3

En la Figura 26 se muestran las curvas ROC con su área bajo la curva(AUC), como otra métrica para la evaluación del modelo con cada método de normalización.

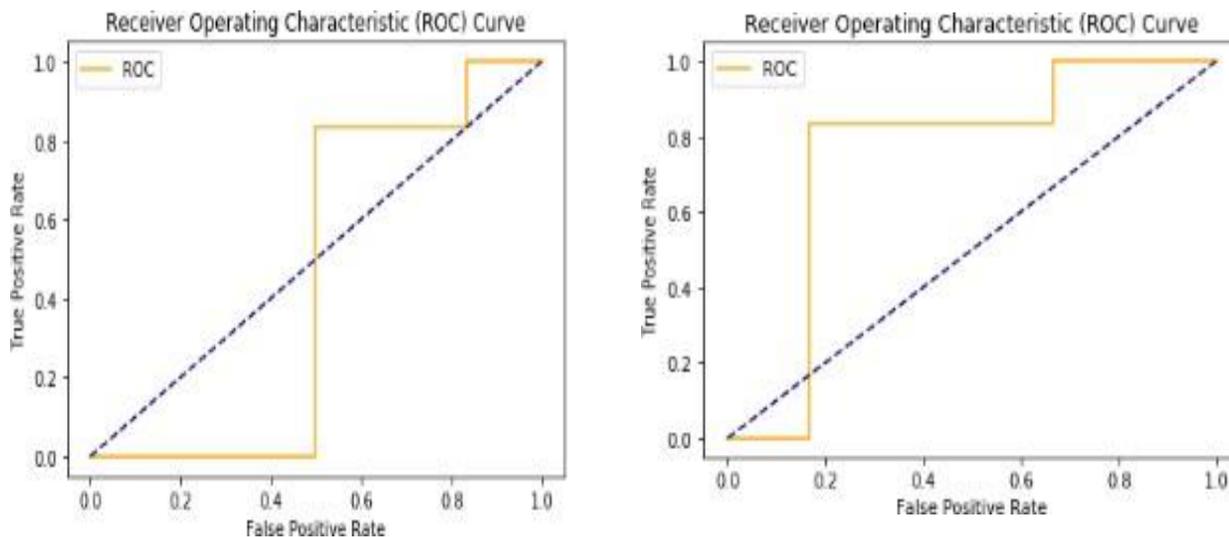


Figura 26 Curvas ROC para el modelo con ambas normalizaciones. A) Normalización estándar; B) Normalización de límites definidos

Modelo árbol de clasificación.

Antes de entrenar el modelo se separaron los datos de entrenamiento y evaluación con una relación de 75/25 respectivamente, y así poder evaluar su viabilidad para la clasificación ASA en una consulta preanestésica con los datos de la anamnesis del paciente, a través de un algoritmo de aprendizaje supervisado de árbol de clasificación.

Una vez entrenado el modelo de árbol de clasificación con los 33 atributos seleccionadas se realizó el cálculo de las métricas para su evaluación de rendimiento. En la Tabla 12 se presenta la matriz de confusión obtenida con los datos de testing que fueron sacados de la base de datos usada para el entrenamiento, y en la Tabla 13 se observa los valores tomados por el modelo para cada una de las métricas calculadas.

Tabla 12. Matriz de confusión para el modelo de árbol de decisión

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	8	0
	Negativos	0	7

Tabla 13. Métricas calculadas para el modelo del árbol de clasificación.

Métricas	Valor
Precisión en el set de Entrenamiento	100%
Exactitud del modelo de árbol de clasificación en la anamnesis	100%
Promedio ponderado del modelo árbol de clasificación en la anamnesis	100%
Sensibilidad del modelo	100%
Especificidad del modelo	100%
Total de Muestras en Test	15
Fallos	0

5.2.3 Evaluación de los Modelos.

Se evaluaron ambos modelos con dos tipos de conjuntos de datos, para conocer que tan bien o mal se da la clasificación en cada uno de los métodos implementados de machine learning, tanto Naive Bayes como árboles de decisión-clasificación, el primer conjunto de datos fue el que se usó para su respectivo entrenamiento y donde se encuentran las 60 muestras iniciales que se crearon y se clasificaron, y el segundo data set fueron 10 casos de consultas preanestésicas reales realizadas en el HUSVF, los cuales fueron recolectados de su sistema, obteniendo todos los valores de los atributos considerados en cada modelo, y su respectiva clasificación ASA, otorgada por el anesthesiólogo a cargo de la consulta.

Clasificador Naive Bayes. Para este modelo se evaluaron los datos con dos técnicas, la primera consistió en la clasificación de los atributos de entrada sin realizar algún tipo de preprocesamiento, y la segunda se basó en normalizar todos los datos de entrada correspondiente a los pacientes, antes de realizar la clasificación.

En la Tabla 14 se muestra la matriz de confusión resultante para la clasificación final que el modelo dio a las 10 muestras reales de consultas preanestésicas, para los datos sin normalizar y los normalizados respectivamente, y en el Tabla 15 se observan los respectivos porcentajes de acierto, desacierto, sensibilidad y especificidad calculados, para cada set de datos con los que fue evaluado el clasificador.

Tabla 14. Matriz de confusión de la clasificación bayesiana para las muestras reales. A) Datos sin normalizar; B) Datos normalizados.

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	4	2
	Negativos	3	1

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	6	0
	Negativos	4	0

Tabla 15. Reporte de los porcentajes de las métricas para la evaluación del modelo de Bayes en las 10 muestras reales

Métrica	Valor (%)	
	Sin normalizar	Normalizado
Acierto	50%	60%
Desacuerdo	50%	40%
Sensibilidad	67%	100%
Especificidad	25%	0%

La matriz de confusión hallada para los resultados obtenidos de los 60 casos de entrenamientos sin normalizar y normalizados se ilustran en la Tabla 16, y en la Tabla 17 finalmente se presentan los porcentajes correspondientes al acierto, desacuerdo, sensibilidad y especificidad del modelo.

Tabla 16. Matriz de confusión del clasificador Bayesiano, para las 60 muestras de entrenamiento. A) Datos sin normalizar; B) Datos normalizados

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	27	3
	Negativos	24	6

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	16	14
	Negativos	10	20

Tabla 17. Reporte de los porcentajes de las métricas para la evaluación del modelo de Bayes en las 60 muestras de entrenamiento

Métrica	Valor (%)	
	Sin normalizar	Normalizado
Acierto	55%	60%
Desacierto	45%	40%
Sensibilidad	90%	53%
Especificidad	20%	67%

Clasificador árbol decisión. En la Tabla 18 se presenta la matriz de confusión resultante correspondiente a los 10 casos reales y la Tabla 19 pertenece a las 60 muestras usadas para su respectivo entrenamiento y que también fueron usados para su evaluación. En las Tablas 20 y 21 se observan los valores pertenecientes al acierto, desacierto, sensibilidad y especificidad que resultaron de la clasificación otorgada por el modelo, para los 10 casos reales y los 60 creados respectivamente.

Tabla 18. Matriz de confusión resultante del clasificador de árbol para los 10 datos reales

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	4	2
	Negativos	0	4

Tabla 19. Matriz de confusión resultante del clasificador de árbol para las 60 muestras del set de entrenamiento

		Predicción	
		Positivos	Negativos
Verdaderos	Positivos	30	0
	Negativos	0	30

Tabla 20. Reporte de las métricas obtenidas del clasificador de árbol para las 10 muestras reales

Métricas	Valor(%)
Acierto	80%
Desacierto	20%
Sensibilidad	67%
Especificidad	100%

Tabla 21. Reporte de las métricas obtenidas del clasificador de árbol para las 60 muestras del set de entrenamiento

Métricas	Valor(%)
Acierto	100%
Desacierto	0%
Sensibilidad	100%
Especificidad	100%

6. ANÁLISIS DE RESULTADOS

6.1 Base de Datos.

6.1.1 Clasificación base de datos

De la Figura 11 se observa que existen unas diferencias entre las clases en término de cantidad de pacientes clasificados, esto es debido a que el criterio médico es subjetivo, y ambos anestesiólogos basaron sus clasificaciones dándole diferente relevancia a algunas variables, aunque la diferencia no sea significativa (2 para ASA I Y 4 para ASA II y 1 para ASA III) es importante saber cuáles factores fueron determinantes para cada uno de los médicos a la hora de dar su clasificación, aunque los dos tuvieron en cuenta la exploración Física, siendo determinante la Edad, la presión arterial y el IMC (Índice de masa corporal), y variables como el embarazo y el uso de oxígeno, el médico anestesiólogo 2 le dio mayor peso a los antecedentes quirúrgicos y patológicos, siendo más específico en el tiempo de padecimiento de alguna enfermedad y el manejo y control que ha tenido el paciente respecto a la enfermedad, es por esta razón que hay un mayor número de pacientes con ASA I y III clasificados por el médico anestesiólogo 2.

Finalmente en la Figura 12 se muestra un promedio de ambos criterios médicos respecto a las clases asignadas a cada paciente, pero como puede observarse hay un desbalance entre las clases I y II, ya que hay 31 pacientes ASA II y 17 ASA I y no sería lo mejor para el modelo a la hora de entrenarlo, porque tendería a irse hacia la clase ASA II cuando clasifique un nuevo paciente, otorgando un porcentaje de sesgo al algoritmo, es por esto que el paso siguiente fue balancear las clases de ASA I Y II, obteniendo el mismo número para ambas; respecto a la clase correspondiente a ASA III se procedió a eliminar esos dos pacientes, ya que son muy pocos y no es uno de los objetivos del presente trabajo predecir dicha clase.

6.2 Modelo de Inteligencia Artificial.

6.2.1 Preprocesamiento de datos y entrenamientos de los modelos

6.2.1.1 Descripción de los datos.

En la Tabla 4 y Figura 13 se observa un desbalance en las clases de ASA bastante pronunciado, la clase de ASA II es un 62% del total de los datos, mientras que la clase perteneciente a ASA I representa un 34%, y el ASA III es apenas un 2% de la cantidad total de los datos, por lo que no pueden ser usados en los modelos, y además este nivel de clasificación no está dentro de los objetivos desarrollados en el proyecto.

6.2.1.2 Modelo Naive Bayes.

En la Figura 14, en general se evidencia que la distribución de los datos creados para los pacientes abarcan todo el rango de valores permitidos por cada signo vital que representan, encontrándose, dentro de los valores normales para tener un buen estado de salud, en la Figura 14 (A) se puede evidenciar que la mayor cantidad de pacientes se encuentran en edades que representa la etapa adulto joven, la cual dentro de los criterios de la clasificación ASA son un indicativo que la persona es saludable y su riesgo en cirugía es menor, entrando

en las categorías de ASA I o II, es por esta razón que dichos valores se justifican con base en el objetivo general del presente trabajo, esta variable es consecuente con los valores que toman las demás, ya que se trató de que los signos vitales diseñados para las 60 muestras no representaran ninguna anomalía o descontrol que llevase a que el riesgo intraoperatorio incrementase y fuese un ASA mayor a II, y poder mantener un balance entre las clases de interés a la hora de la clasificación por parte del equipo médico.

En la Figura 15 se observa la verificación de cómo se distribuyen las dos clases trabajadas, de acuerdo a los valores de los signos vitales que tenga el paciente en su revisión física, cabe destacar algunas de las variables que evidencian algún tipo de patrón para determinar el tipo de ASA, como es el caso del género (Figura 15 A), ya que se observa que en las mujeres prevalece la clase ASA 2 sobre la ASA 1, y en los hombres sucede el caso opuesto; en el caso de la edad (Figura 15 B) se comprueba que cuanto más años tenga la persona la probabilidad de que su clasificación sea ASA 2 es mayor, esto se debe a las complicaciones que puedan presentarse por las patologías que se desarrollan al pasar los años y que es muy probable que representen un mayor riesgo durante la intervención quirúrgica a realizar, otra de estas variables es la presión arterial (Figura 15 C), en la cual se observa que cuando la presión arterial tanto sistólica como diastólica toma valores fuera del rango considerado normal, la persona presenta una patología con la presión arterial, ya sea hipertensión o hipotensión, por lo cual la probabilidad de que sea clasificado como ASA II es mucho mayor y es un indicativo para tomar precauciones al momento de intervenir quirúrgicamente.

Finalmente, la saturación de oxígeno (Figura 15 F) muestra que a mayor porcentaje de oxígeno en sangre mayor probabilidad de ser clasificado como ASA 1, porque da un buen parte de cómo se encuentra su oxigenación representando un menor riesgo.

En la Figura 17 se observa la matriz de correlación, en la cual se evidencian los porcentajes de relación existente entre las variables de la exploración física, y los cuales dan cuenta de la similitud en el tipo de información que da cada una y su respectiva dependencia, según la matriz obtenida se conoce que las 8 variables relacionadas no superan el 20% de dependencia, es decir que todas dan diferente tipo de información que aporta a la construcción del modelo, y esto se dio gracias a que inicialmente en la base de datos en la cual estaban consignadas todas las 10 variables de la exploración física que incluían peso y estatura se obtuvo un porcentaje de relación mayor al 80% de estas variables con la del índice de masa corporal, por lo que se procedió a escoger solo el IMC ya que es una variables que contiene tanto el peso como la estatura y a eliminar posteriormente ambas características de la base de datos.

En la Figura 18 se observan las gráficas de densidad, que representa la distribución de las variables de la exploración física, según el tipo de clasificación ASA, esto es con el objetivo de identificar en cada una de ellas como se diferencian ambas clases de acuerdo el valor que tomen, y en que valores de cada variable se determina a qué tipo de clase pertenece; las variables que tiene una distribución más cercana a una normal tanto pasa ASA I como ASA

II son las correspondientes al Sexo, la Edad y la saturación de oxígeno, y se evidencia que la separación en la distribución de las clases en dichas variables es mayor, obteniendo una diferencia entre los valores determinantes de cada clase al no estar tan superpuestas, lo que significa que dichas variables son las que deberían tener un mayor peso para el modelo al momento de tomar decisiones y clasificar un paciente, debido a que en las demás no se evidencia una separación relevante que lleve a suponer que dan una buena diferenciación entre las clases, al momento de su clasificación.

Finalmente se realizó una selección de las mejores características de la base de datos para construir el modelo, con el uso de uno de los comandos de la librería Scikit-Learn, el cual arroja las características con mejor diferenciación de clases y las cuales aportan mayor información para entrenar el modelo y obtener un mejor rendimiento, de acuerdo a esto se obtuvo que las tres mejores según la librería son exactamente las mismas que se evidencia en las gráficas de densidad (Edad, Sexo y saturación de Oxígeno), por lo que se decidió entrenar el algoritmo solo con estas tres características.

6.2.1.3 Modelo árboles de decisión

La Figura 19 muestra que la mayoría de pacientes a los cuales se le realizó la consulta preanestésica no tienen antecedentes quirúrgicos, esto se relaciona al hecho que no haberle practicado ninguna cirugía anteriormente su clasificación es la que representa el menor riesgo, se comprueba al observar que el número de ASA I para pacientes sin antecedentes quirúrgicos es mayor que los de ASA II, además el tipo de cirugía influye en la clasificación otorgada a cada caso, y es por el riesgo que representan dichas cirugías para el estado de salud de la persona y da cuenta de su historial clínico y las precauciones a tener en cuenta en otra intervención, es por esta razón que el número de cirugía cardíacas y urológicas son tan pocas, por que aumentan la probabilidad de la clasificación fuese mayor a ASA II.

La distribución de la anestesia es un indicativo de cómo ha sido la aceptación del cuerpo de la persona a esta, y alguien que ya haya sido sometido a anestesia general y no haya tenido ningún tipo de complicación en la cirugía es probable que su riesgo intraoperatorio sea menor.

Los datos de las Figuras 20, 21, 22 y 23 se graficaron con el fin de demostrar que el riesgo preoperatorio asignado en la consulta preanestésica aumenta dependiendo de la presencia de patologías, ya que para las cuatro incluidas en este conjunto de datos se evidencia que la mayor parte de los casos que padezcan de alguna de ellas al momento de clasificarlos correspondieron al tipo ASA II según la opinión médica de ambos anesthesiólogos que realizaron la clasificación, esto también depende del tipo de control médico que se le den a estas enfermedades, para este caso todas según los datos creados estaban debidamente controladas, por esta razón no se les asignó un valor mayor de ASA.

Se decidió eliminar las variables que presentaban invariabilidad en su valor de entrada en todo el conjunto de datos, y trabajar con las 33 variables observadas en la Tabla 6, porque aunque representan información importante para la clasificación del riesgo intraoperatorio, en el conjunto de datos, todos los pacientes creados no tenían una respuesta o valor diferentes en dichas variables ya que si hubieran tenido una respuesta diferente, para este caso un si en alguna de estas variables la clasificación ASA del paciente hubiera pasado a ser ASA 3.

De las variables que presentaron un alto porcentaje de correlación (Tabla 7) se decidió eliminar la que menor información relevante daba, o la que según la opinión de los anestesiólogos no eran de alta relevancia o en algunos casos no se consideraban para otorgar una clasificación ASA a un paciente.

Ya que el 97 % fue el porcentaje más alto obtenido en la búsqueda de la profundidad que debía llevar el algoritmo del árbol creado, se decidió trabajar con dicho porcentaje, por lo cual la profundidad para la construcción del árbol de decisión fue de 7.

El parámetro Criterion hace alusión a la medida de selección de atributos, para dividir los datos de la mejor manera posible, y el tipo de división elegido para el árbol de decisión fue la entropía ya que da mejor rendimiento cuando se trabaja con datos categóricos.

El Min_samples_split se halló a través de un proceso de ensayo y error calculando el porcentaje de precisión del modelo, obteniendo con este el mejor rendimiento del modelo.

La precisión alcanzada por el árbol fue muy alta, arrojando que el modelo creado es confiable para la clasificación o por el contrario esta sobreentrenado y no está clasificando de manera adecuada, esto se podrá comprobar con la realización del objetivo tres donde se evalúe cual es el porcentaje de acierto en la clasificación del modelo.

La variables de entrada más significativa para el árbol de decisión que dividió la población en conjuntos para la asignación de las clases fue la condición de Tabaquismo (nodo raíz), la cual da información acerca de si un paciente fuma o no, dándole una probabilidad del 50% a cada una de las dos opciones de respuesta posible (Si o No), de las 33 variables independientes con las cuales fue entrenado el modelo, este escogió para su construcción solo 4 de estas variables, repitiendo el proceso que hizo con el nodo raíz para cada atributo y poder clasificar los datos entre las dos clases, esta ruta tan pequeña puede deberse a que las variables restantes tienen al mismo valor de los atributos seleccionados, a esto también se le asocia el porcentaje de entropía calculado para cada uno de los nodos de decisión, ya que para el atributo de tabaquismo este valor es del 100%, lo que significa que esta variable presenta la mínima ganancia de información en los datos, ya que hay una alta impureza en la información y no permite conocer que tan bien dicha variable separa los ejemplos de entrenamiento de acuerdo con sus clasificación objetivo, terminando su división cuando la entropía sea igual a cero, es decir cuando todos los valores del conjunto de división pertenezcan a una sola clase,

terminando así en el atributo con la mayor ganancia de información que en este caso corresponde a la Hipertensión.

Finalmente se puede concluir que una persona será clasificada como ASA 2 si fuma, o si no ha sido evaluado por un endocrinólogo al padecer de diabetes, si siente dolor en el pecho al subir 3 o 4 pisos y si sufre de hipertensión.

6.2.2 Validación de los modelos

6.2.2.1 Modelo Naive Bayes.

En relación con las Tablas que contienen las matrices de confusión con los resultados de predicción del modelo con cada uno de los dataset normalizados, usando solo los datos del conjunto de test (12), se evidencia que la correspondiente a los datos con normalización estándar (Tabla 8), etiquetó correctamente 5 de las 12 muestras totales elegidas para su evaluación, y el modelo con los datos procesados con la normalización del escalado de variables (Tabla 9) acertó en 7 muestras clasificándolas correctamente.

Existe un paradigma que establece que en un modelo de clasificación es más relevante y de mayor peso la cantidad de falsos positivos (pacientes de ASA 1 y que fueron clasificados como ASA 2), que la cantidad de falsos negativos (pacientes ASA 2 etiquetados como ASA 1), ya que se basa en que es mejor suponer que el riesgo del paciente es mayor en la consulta preanestésica y que luego se compruebe que dicho riesgo es menor, a que suceda lo contrario, porque este último caso implicaría, que el personal médico no esté más prevenido sobre los efectos o complicaciones que puede presentar el paciente en la cirugía, dicho esto y de acuerdo a los resultados obtenidos el que presentó mayor número de falsos positivos fue el modelo entrenado con normalización estándar con 4 casos FP y 3 FN cumpliéndose la premisa de $FP > FN$, pero este modelo arrojó unos porcentajes de desempeño bajos (Tabla 10) con respecto a los obtenidos por el modelo con normalización de escalado de variables es decir en términos porcentuales y de rendimiento lo hace mucho mejor el algoritmo que se resume en la Tabla 11, ya que su capacidad tuvo menor número de fallas en las predicciones y mayor porcentaje en la exactitud y el promedio entre esta y la precisión, en consecuencia este modelo distingue mejor entre las dos clases de ASA trabajadas, ya que su curva de aprendizaje ROC mostrada en la Figura 26, tiene un área bajo la curva superior (75%), lo que significa que este modelo clasifica de manera más adecuada los datos de entrada del dataset, por esta razón se considera que la mejor opción para trabajar los datos y entrenar el modelo es con la normalización del escalado de variables, ya que sus métricas de aprendizaje obtenidas son más altas.

6.2.2.2 Modelo árboles de decisión

En la Tabla 12 Se evidencia que de las 15 muestras tomadas del conjunto de entrenamiento para el testing, todas fueron clasificadas correctamente, obteniendo 8 casos verdaderos positivos y 7 casos verdaderos negativos, es decir que el clasificador predijo el valor de la

variable objetivo (ASA) correctamente, mediante el aprendizaje de reglas de decisión simples inferidas de las características de los datos, esto como se dijo anteriormente puede ser una evidencia de que la ruta escogida por el algoritmo para clasificar entre ambas clases es el óptimo o correcto, y se ha obtenido un excelente resultado, o por el contrario al obtener un 100% en todas las métricas calculadas para la validación de su rendimiento (Tabla 13), lo más factible es que el árbol construido sobre ajustó los datos, ya que la probabilidad de que esto ocurra es mayor cuando se manejan una gran cantidad de características en el dataset, es por esta razón que al no obtener una correcta proporción de muestras con la cantidad de atributos o características totales presente, el árbol al contener pocas muestras en un espacio de alta dimensión se sobreajuste y no esté aprendiendo correctamente para una clasificación confiable, aunque dentro de los parámetros usados para su entrenamiento se usó la máxima profundidad alcanzada para controlar el tamaño del árbol y poder reducir la probabilidad de sobreajuste.

6.2.3 Evaluación de los modelos

6.2.3.1 Clasificación Bayesiano.

En las Tablas 15 y 17 se observa que la clasificación con las variables de entrada tanto para los casos reales como para los creados, los porcentajes de acierto son un poco mayores que los obtenidos con los datos de entrada sin normalizar, con una diferencia del 5%, incrementando la probabilidad de que este clasificador tenga menos casos mal etiquetados cuando los datos se pasan por el proceso de normalización, cabe mencionar que dichos porcentajes son bajos comparados con los normalmente aceptados para catalogar un modelo como de alto rendimiento (mayor al 80%), pero esto es por la poca cantidad de muestras usadas para su evaluación, tanto las reales como los 60 diseñadas; en consecuencia con esto en la Tabla 14 se observa que para los datos normalizados de los 10 reales solo clasificó correctamente 6, y en los datos sin normalización 5, con la diferencia de 1 muestra bien clasificada, esto da cuenta de la diferencia en el porcentaje de acierto entre un método y el otro, y aunque en los datos normalizados no clasificó ningún verdadero positivo, tampoco clasificó falsos negativos, por lo que el porcentaje de sensibilidad es mucho mayor en datos normalizados, otorgándole mayor peso a esta métrica y cumpliendo con la premisa que en un sistema de clasificación es mejor que $FP > FN$ por lo justificado anteriormente

Caso contrario ocurre con las 60 muestras clasificadas, ya que como puede observarse en la Tabla 16 aunque para los datos normalizados se clasificó una mayor cantidad de datos correctamente y que dan el porcentaje de acierto mayor, hubo una mayor cantidad de falsos negativos, los cuales son la razón de que el porcentaje de sensibilidad diera menor, y su especificada mucho mayor, por lo cual para este conjunto de datos el modelo no funciona de la manera más acertada, dando mayor peso a la especificada y clasificando algunos casos que son de un mayor riesgo (ASA II) como ASA I.

6.2.3.2 Clasificador árbol de decisión.

En las Tablas 20 y 21 se puede observar que para este modelo de clasificación se obtuvo un alto porcentaje de acierto, 80% para la clasificación de los 10 casos reales y 100% para los 60 diseñados inicialmente, pero el porcentaje de sensibilidad es menor que el de especificidad, calculados para las muestras reales, debido a que el modelo clasificó 2 casos como falsos negativos y 0 casos como falsos positivos, al analizar estas dos muestras en el clasificador bayesiano, se evidenció que el modelo de Bayes clasificó correctamente estas dos muestras, por lo tanto el error de etiquetación se debe a un desvío en la información de la base de datos correspondiente a la anamnesis, al comparar cada uno de los valores que tomaron los atributos con los demás muestras y las cuales fueron clasificados de manera correcta como ASA 2, se encontró que los verdaderos positivos presentaban alguna de las condiciones que fueron tomadas para construir el árbol y que pertenecen a los nodos de decisión (Tabaquismo, dolor o corto de aire, Hipertensión), mientras que las dos muestras mal clasificadas solo presentaban la condición de estar consumiendo algún tipo de medicamento en el momento en que se les realizó la consulta preanestésica, el cual no es un indicativo para que el árbol resultante lo considere como relevante para la división de clases, aunque según el criterio médico ambos casos pertenecían a ASA II, esta es la razón por la cual estos dos casos se etiquetaron como menos graves o con menor riesgo de una complicación durante la intervención quirúrgica.

7.CONCLUSIONES

Respecto a los modelos de Bayes y árboles de decisión, los rendimientos no son los esperados, a pesar de que se realizó un balanceo en los datos para evitar una tendencia hacia la clase con mayor número de etiquetas resultante inicialmente, la depuración de información redundante y los métodos realizados para obtener una base de datos lo más limpia de ruido posible; esto pudo ser el resultado de la poca cantidad de datos usados tanto para el entrenamiento como para la evaluación de cada uno de los modelos, una mayor cantidad de consultas preanestésicas puede ser una solución para mejorar los porcentajes de rendimiento de ambos modelos, o usar otras técnicas de aprendizaje automático, elaborando un solo modelo que incluya todas las etapas de una consulta, e incluso el uso de aprendizaje no supervisado para encontrar patrones en todas las variables involucradas en una consulta preoperatoria, y detectar de alguna manera de qué forma se relacionan y qué patrón siguen para asignar el tipo de clasificación de riesgo intraoperatorio del paciente.

Aunque es posible clasificar de manera correcta el riesgo preoperatorio de un paciente (ASA) a través de un modelo de aprendizaje automático e inteligencia artificial a partir de una consulta preanestésica, en el proyecto desarrollado no se obtuvo un gran porcentaje de confiabilidad para afirmar o asegurar que es posible obtener una excelente clasificación a través de una consulta preoperatoria a distancia, porque para ello se tendría que construir y trabajar en una base de datos lo suficientemente robusta, incluyendo la mayor cantidad de consultas preanestésicas realizadas a pacientes intervenidos en el HUSVF, para poder evaluar la posibilidad de que efectivamente sea una herramienta de la cual cualquier paciente pueda hacer uso de forma digital, evitando su desplazamiento a las instalaciones del hospital, además de solucionar la diferencia de criterios médicos entre el anestesiólogo de la consulta y el de la intervención quirúrgica, implementando una clasificación global de ASA en todo el hospital.

Aclaración: la finalización de este trabajo de grado se dio durante la pandemia declarada el 11 de marzo del 2020 por la organización mundial de la salud, a causa del coronavirus COVID-19, instando a los estados a tomar acciones urgentes y decididas para el manejo de la enfermedad en su país, para tratar los casos confirmados y mitigar el contagio. Por esta razón a través de la resolución 385 emitida el 12 de Marzo del 2020 por la ministerio de salud y protección social de Colombia se declaró emergencia sanitaria en todo el país hasta el día 30 de Mayo del 2020, adoptando medidas sanitarias para prevenir y controlar la propagación del COVID-19 en todo el territorio nacional; una de las medidas tomadas fue ordenar el aislamiento preventivo obligatorio de todas las personas habitantes del país desde el día 25 de Marzo, hasta el 11 de Mayo, en el marco de la emergencia sanitaria, a través de la resolución 457 del 22 de Marzo del año 2020, para poder preservar la salud y la vida de los ciudadanos.

Por lo anterior no fue posible acceder a los datos reales de evaluaciones preanestésicas llevadas a cabo en el HUSVF, para realizar una evaluación de los algoritmos y construir una base de datos confiable, ya que ello implicaba un desplazamiento hasta sus instalaciones, y por la situación descrita no fue posible el traslado hasta el Hospital para recolectar y construir una base de datos con consultas reales.

BIBLIOGRAFÍA

- [1] B. Solsona. (2013). Evaluación preoperatoria. factor de RISC anestésica, escales de RISC. [En línea]. Available: <http://www.scartd.org/arxiu/preop07.pdf>
- [2] M. de la Rosa, E. Rodríguez-Mota, “Valoración preanestésica. Importancia en el paciente quirúrgico”, *Rev Hosp Jua Mex*, vol. 81, pp. 193-198, 2014.
- [3] Dra. M. Cheresheva, Dra. X. Watson, Dr. M. Hamilton. (2016, Dic 13). Escalas Predictivas de Riesgo Perioperatorio. [En línea]. Available: https://www.wfsahq.org/components/com_virtual_library/media/7afd2bd25c39e1b0259ff4f408873616-343-Escalas-de-Riesgo-Perioperatorio.pdf
- [4] O. Giraldo, “Desarrollo de un modelo predictivo con inteligencia artificial para establecer clasificación ASA y riesgo perioperatorio en pacientes de la consulta preanestésica del Hospital Universitario San Vicente Fundación y la IPS Universitaria”, Universidad de Antioquia, Medellín, Colombia, Tech, 2019.
- [5] Castañeda Pascual, M. “Consulta Preanestésica, ¿Un Trámite Incómodo O Una Necesidad Obligatoria?” [En línea] Zona Hospitalaria. Available: <https://zonahospitalaria.com/consulta-preanestesica-un-tramite-incomodo-o-una-necesidad-obligatoria/> [Último acceso: 26 03 2020].
- [6] Sachile.cl. “Recomendaciones Sobre Evaluación Pre Anestésica.”[En línea] Available: https://www.sachile.cl/upfiles/userfile/files/evaluacion_pre_anestesica.pdf [Último acceso 26 03 2020].
- [7] Rincón Valenzuela, D. and Escobar, B. “Manual de práctica clínica basado en la evidencia: preparación del paciente para el acto quirúrgico y traslado al quirófano”. *Rev Colombiana de Anestesiología*, España: Elsevier España, S.L.U., vol. 43, pp.32-50, 2015.
- [8] Ojeda Gonzales, J. “Elementos predictores en la evaluación del riesgo anestésico por el anesthesiólogo.” *Revista Cubana de Anestesiología y Reanimación*. Habana, Cuba: SciELO. Vol. 12, N° 2, 2013.
- [9] Dra. G.López-Herranz, Dra. O.Torres-Gómez, “ Variabilidad de la clasificación del estado físico de la Sociedad Americana de Anestesiólogos entre los anesthesiólogos del Hospital General de México”, *Rev Mexicana de anestesiología*, vol. 40, pp. 190-194, Septiembre 2017.
- [10] Oracle, “What is a database?”, [En Línea], Available: <https://www.oracle.com/database/what-is-database.html>. [Último acceso: 20 Marzo 2020].

- [11] Estadística-dma.ulpgc.es. “6. Lectura De Datos Desde Archivos .Csv.” [En línea] Available: <http://estadistica-dma.ulpgc.es/FCC/matlab-06-Lectura_de_archivos_csv.html> [Último acceso: 26 03 2020].
- [12] Parra, V. “Qué Es Un Csv, Cómo Se Hace Y Para Qué Sirve.” [En línea] Bitácora de una profesora. Available: <<https://lolap.wordpress.com/2015/01/14/que-es-un-csv-como-se-hace-y-para-que-sirve/>> [Último acceso: 26 03 2020].
- [13] A. Kaplan y M. Haenlein. “Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretation, illustrations and implication of artificial intelligence”, *Business Horizons*, vol 62, N° 1, pp 15-25, 2019.
- [14] I. Goodfellow, Y. Bengio y A. Courville, Deep Learning. Massachussets: MIT, 2016.
- [15] MATHWORKS, “Machine learning: Tres cosas que es necesario saber”. [En línea]. Available: <https://la.mathworks.com/discovery/machine-learning.html> [último acceso: 20 03 2020].
- [16] Code Academy. “Nomalization” [En línea] Available: <https://www.codecademy.com/articles/normalization>. [Último acceso 20 03 2020].
- [17] Scikit-Learn, “Imputation of missing values”. Scikit-learn [En línea]. Available: <https://scikit-learn.org/stable/modules/impute.html#impute>. [Último acceso 20 03 2020]
- [18] W. Badr. “6 Different Ways to Compensate for Missing Values In a Dataset (Data Imputation with examples)”. Towards Data Science, 5 Enero 2019. [En línea]. Available: <https://towardsdatascience.com/6-different-ways-to-compensate-for-missing-values-data-imputation-with-example-6022d9ca0779>. [Último acceso: 20 03 2020].
- [19] R. Malarvizhi, “K-nearest Neighbor in Missing Data Imputation”, *International Journal of Engineering Research and Development*, vol. 5, N°. 1, pp. 5-7, 2012.
- [20] S. Van Buuren, Flexible multivariate imputation by MICE. TNO. 1998.
- [21] X. Wu, X. Zhu, G. Wu y W. Ding. “Data mining with big data.” *IEEE Trans Knotel Data Eng*, vol. 26, N°. 1, pp. 97-107, 2014.
- [22] S. Garcia, S. Ramirez-Gallego y J. Luengo, “Big data preprocessing: methods and prospects.” *Big Data Analytics*, 2016.
- [23] N. Chawla, K. Bowyer, L. Hall y W. Kegelmeyer, “SMOTE:synthetic minority oversampling technique.” *J Artif Intell Res*, vol.16, N°. 1, pp. 31-57, 2002.
- [24] Docs.microsoft.com. “Preparación De Los Datos Para La Compilación De Un Modelo - ML.NET.” [En línea] Available: <<https://docs.microsoft.com/es-es/dotnet/machine-learning/how-to-guides/prepare-data-ml-net>> [Último acceso: 26 03 2020].

- [25] Briega, R. “Análisis De Datos Categóricos Con Python.” [En línea] Relopezbriega.github.io. Available: <<https://relopezbriega.github.io/blog/2016/02/29/analisis-de-datos-categoricos-con-python/>> [Último acceso: 26 03 2020]. Febrero 2016.
- [26] J. Zambrano. “¿Aprendizaje supervisado o no supervisado? Conoce sus diferencias dentro del machine learning y la automatización inteligente” [En línea]. Available: <https://medium.com/@juanzambrano/aprendizaje-supervisado-o-no-supervisado-39ccf1fd6e7b>. [Último acceso: 26 03 2020]. Marzo 2018.
- [27] MIng Enrique Fernández, “ANÁLISIS DE CLASIFICADORES BAYESIANOS.”, Lanoratorio de sistemas inteligentes, [En Línea]. Available: www.fi.uba.ar/laboratorios/lsi. [Último acceso: 23 03 2020].
- [28] C. Malagón Luque, “Clasificadores bayesianos. El algoritmo Naïve Bayes.” [En línea]. Available: https://www.nebrija.es/~cmalagon/inco/Apuntes/bayesian_learning.pdf. [Último acceso: 23 03 2020].
- [29] Sucar, L. “Clasificadores Bayesianos: De Datos A Conceptos.” [En línea] Ai.org.mx. Available: <http://www.ai.org.mx/ai/images/sitio/2014/05/ingresos/less/trabaj_final_dr._sucar.pdf> [Último acceso: 23 03 2020].
- [30] O. J. Prieto, I. Raúl, and C. Díaz, “Aprendizaje Bayesiano.” [En línea]. Available: <https://www2.infor.uva.es/~isaac/doctorado/AprendizajeBayesiano.pdf>. [Último acceso: 23 03 2020].
- [31] Uv.es. <https://Www.Uv.Es/Mlejarza/Actuariales/Tam/Arbolesdecision.Pdf>. [En línea] Available: <<https://www.uv.es/mlejarza/actuariales/tam/arbolesdecision.pdf>> [Último acceso: 24 03 2020].
- [32] Bosco Mendoza, J. “Arboles De Decisión Con R — Clasificación”. [En línea] Medium. Available: <<https://medium.com/@jboscomendoza/arboles-de-decisi%C3%B3n-con-r-clasificaci%C3%B3n-c6c583b16125>> [Último acceso: 24 03 2020].
- [33] Larranaga, P., Inza, I. and Moujahid, A., 2020. “Tema 10. Arboles De Clasificación.” [En línea] Sc.ehu.es. Available: <<http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t10arboles.pdf>>. Último acceso: 24 03 2020].
- [34] Tecnologias-informacion.com. “Aprendizaje Con Árboles De Decisión: Algoritmos Y Ejemplos.” [En línea] Available: <<https://www.tecnologias-informacion.com/arboles.html>> [Último acceso: 24 03 2020].

- [35] Caparrini, F. and Work, W. “Aprendizaje Inductivo: Árboles De Decisión - Fernando Sancho Caparrini”. [En línea] Cs.us.es. Available: <<http://www.cs.us.es/~fsancho/?e=104>> [Último acceso: 24 03 2020].
- [36] johanna.orellana@ucuenca.edu.ec, J. “Árboles De Decision Y Random Forest.” [En línea] Bookdown.org. Available: <<https://bookdown.org/content/2031/arboles-de-decision-parte-i.html>> [Último acceso: 24 03 2020].
- [37] Scikit-learn.org. Sklearn.Tree.DecisionTreeClassifier — Scikit-Learn 0.22.2 Documentation. [En línea]. Available: <<https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html>> [Último acceso: 23 03 2020].
- [38] A. Mishra, “Metrics to Evaluate your Machine Learning Algorithm.”, Towards Data Science, 24 Febrero 2018. [En línea]. Available: <https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234>. [Último acceso: 23 03 2020].
- [39] Heras, J. “Precision, Recall, F1, Accuracy En Clasificación - Iartificial.Net.” [En línea] IArtificial.net. Available: <<https://iartificial.net/precision-recall-f1-accuracy-en-clasificacion/>> [Último acceso: 23 03 2020].
- [40] Interactive Chaos. “Matriz de confusión”, [En línea]. Available: <https://www.interactivechaos.com/manual/tutorial-de-machine-learning/matriz-de-confusion>. [Último acceso: 23 03 2020].
- [41] sitiobigdata.com. “Machine Learning: Selección Métricas De Clasificación - Sitiobigdata.Com.” [En línea]. Available: <<http://sitiobigdata.com/2019/01/19/machine-learning-metrica-clasificacion-parte-3/>> [Último acceso: 23 03 2020].
- [42] Ligdi González. “Métricas De Evaluación Clasificación Con Scikit Learn - Ligdi González.” [En línea]. Available: <<https://ligdigonzalez.com/metricas-de-evaluacion-clasificacion-con-scikit-learn-machine-learning>>. [Último acceso: 23 03 2020]
- [43] S. Narkhede, “Understanding AUC-ROC Curve”, Towards Data Science, 26 Junio 2018. [En línea]. Available: <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>. [Último acceso: 23 03 2020]
- [44] F. Pedregosa, G. Varoquaux, A. Gramfort y V. Michel, “Scikit-learn: Machine Learning in Python”. JMLR, vol.12, pp. 2825-2830, 2011.