



UNIVERSIDAD CENTRAL DE CHILE
FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA DE INGENIERÍA

MÉTODO DE ESTIMACIÓN DEL ÍNDICE DE AUTORREGULACIÓN CEREBRAL UTILIZANDO SVM

MEMORIA PARA OPTAR AL TÍTULO PROFESIONAL DE INGENIERO
CIVIL EN COMPUTACIÓN E INFORMÁTICA.

PROFESOR GUÍA:

CLAUDIO HENRÍQUEZ BERROETA

PROFESORES INFORMANTES:

FRANCISCO CRUZ NARANJO

HERNAN OLMÍ REYES

IVÁN FERNANDO SMITH JIMÉNEZ

SANTIAGO – CHILE

2021



UNIVERSIDAD CENTRAL DE CHILE
FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA DE INGENIERÍA

MÉTODO DE ESTIMACIÓN DEL ÍNDICE DE AUTORREGULACIÓN CEREBRAL UTILIZANDO SVM

MEMORIA PREPARADA BAJO LA
SUPERVISIÓN DE LA COMISION
INTEGRADA POR LOS PROFESORES:

CLAUDIO HENRÍQUEZ BERROETA
FRANCISCO CRUZ NARANJO
HERNAN OLMÍ REYES

QUIENES RECOMIENDAN QUE SEA
ACEPTADA PARA COMPLETAR LAS
EXIGENCIAS DEL TÍTULO DE
INGENIERO CIVIL EN
COMPUTACIÓN E INFORMÁTICA.

SANTIAGO – CHILE

2021

En agradecimiento a todas las personas quienes me acompañaron en este largo camino, mis padres, mi abuela, mi novia y hermano, siendo los pilares fundamentales para lograr este proyecto de título.

IVÁN FERNANDO SMITH JIMÉNEZ

LISTA DE CONTENIDO

LISTA DE CONTENIDO	4
ABSTRACT	9
RESUMEN.....	10
GLOSARIO Y RELACIONES	11
INTRODUCCIÓN.....	12
1. PLANTEAMIENTO DEL PROBLEMA.....	14
1.1 OBJETIVOS.....	18
1.1.2 OBJETIVO GENERAL	18
1.1.3 OBJETIVOS ESPECÍFICOS	18
1.2 HIPÓTESIS	19
1.3 SOLUCIÓN PROPUESTA	19
1.4 ALCANCES Y LIMITACIONES	20
1.4.2 ALCANCES	20
1.4.3 LIMITACIONES	21
2 MARCO TEÓRICO.....	22
2.1 SISTEMA DE AUTORREGULACIÓN CEREBRAL	24
2.2 ÍNDICE DE AUTORREGULACIÓN CEREBRAL.....	26
2.3 ENFERMEDADES COMPROMETIDAS.....	26

2.4	ALGORITMOS DE APRENDIZAJE AUTOMATICO.....	27
a.	APRENDIZAJE SUPERVISADO.....	28
b.	APRENDIZAJE NO SUPERVISADO	28
c.	APRENDIZAJE POR REFUERZO.....	29
2.5	SUPPORT VECTOR MACHINE.....	29
2.6	MODELO DE AASLID-TIECKS	32
2.7	RESPUESTA FINITA AL IMPULSO (FIR)	35
2.8	METODOLOGÍA.....	36
2.9	ENFOQUE CUANTITATIVO.....	36
3.	ESTADO DEL ARTE.....	39
4.	DESARROLLO Y PROCEDIMIENTO.....	43
4.1	DEFINICIÓN DE VARIABLES	43
4.2	DESCRIPCIÓN DEL MODELO ACTUAL	46
4.3	PROPUESTA DE MODELO	48
4.4	DESCRIPCIÓN DEL FLUJOGRAMA DE CÁLCULO.....	49
4.5	ANÁLISIS DE DATOS	50
a.	DISTRIBUCIÓN DE DATOS.....	51
b.	DATOS DE ENTRENAMIENTO.....	51
c.	DATOS DE PRUEBA	52
4.6	CONFIGURACIÓN DE HIPERPARAMETROS SVM.....	52
4.7	PROCESO DE APRENDIZAJE DE SVM.....	54

5. ANÁLISIS Y RESULTADOS	56
5.1 CALCULO DE ARI SEGÚN MODELO PROPUESTO	56
a. CONFIGURACIONES DE MODELO PROPUESTO	57
b. SINTESIS DE RESULTADOS	65
5.2 EVALUACIÓN DE RESULTADOS FINALES	67
5.3 COMPARACIÓN DE MODELOS EXISTENTES	72
6. CONCLUSIÓN Y RECOMENDACIONES.....	73
6.1 CONCLUSIÓN.....	73
6.2 RECOMENDACIONES	76
BIBLIOGRAFÍA	78
ANEXO I.....	81
ANEXO II	85
ANEXO III.....	87
ANEXO IV.....	88
ANEXO V	90

LISTA DE TABLAS

Tabla 1: Asociación entre parámetros para obtención de ARI	34
Tabla 2: <i>Mejores configuraciones para modelo propuesto</i>	57
Tabla 3: <i>Matriz de confusión en base a configuración 1</i>	58
Tabla 4: <i>Matriz de confusión en base a configuración 2</i>	59
Tabla 5: <i>Matriz de confusión en base a configuración 3</i>	60
Tabla 6: <i>Matriz de confusión en base a configuración 4</i>	61
Tabla 7: <i>Matriz de confusión en base a configuración 5</i>	62
Tabla 8: <i>Matriz de confusión en base a configuración 6</i>	63
Tabla 9: <i>Matriz de confusión en base a configuración 7</i>	64

LISTA DE FIGURAS

Figura 1: Hiperplano de separación optima en dos dimensiones – SVM (Fernández-Sainz, 2011).	30
Figura 2: Aplicación de una función kernel modificando el espacio de dimensiones representado por la línea continua en verde, separando el conjunto de datos rojos y azules (Baudat, G., Anouar F, 2000).	31
Figura 3: Los 10 niveles de ARI basado en la respuesta PA Y VFSC por Aaslid-Tiecks (Simpson, 2017).	35
Figura 4: Proceso del enfoque cuantitativo. (Sampieri, Fernández Collado, C., & Baptista, 2010).	37
Figura 5: Curvas ROC para mediciones de Autorregulación (Simpson Henríquez, Panera,).	47
Figura 6: Modelo de estimación del Indicé de autorregulación Cerebral, propuesto por alumno.	48
Figura 7: Estructura de procesos para la obtención del ARI.	49
Figura 8: Distribución de datos utilizado para el análisis.	51
Figura 9: Curva de ROC perteneciente a configuración 1.	58
Figura 10: Curva de ROC perteneciente a configuración 2.	59
Figura 11: Curva de ROC perteneciente a configuración 3.	60
Figura 12: Curva de ROC perteneciente a configuración 4.	61
Figura 13: Curva de ROC perteneciente a configuración 5.	62
Figura 14: Curva de ROC perteneciente a configuración 6.	63
Figura 15: Curva de ROC perteneciente a configuración 7.	64

ABSTRACT

The brain engages a series of essential functions for the functioning of various processes within the body, having an impact on both relevant systems and organs. In particular, the Brain Self-Regulation System allows checking the functioning of blood self-regulation within the brain, establishing certain states in which the human body reacts positively or negatively, depending on how the blood is self-regulating. For this reason, there is the brain self-regulation index, in charge of checking how self-regulation is behaving within the brain at a given time, intending to infer whether the subject in question is in perfect condition or you may be prone to neurodegenerative disease. It should be noted that this index varies between 0 and 9, depending on the state of the person. In case the subject is self-regulating in a good way, the highest value is automatically assigned.

The present research aims to create a predictive model to calculate the Brain Self-Regulation Index using Support Vector Machines as the main algorithm. Through a study in which 27 patients were analyzed, and 7 coefficients were obtained for each of the subjects that allowed establishing a mathematical relationship to determine the ARI in a less invasive way, compared to the existing procedures. In this way, the model capable of predicting each ARI in states of Normocapnia and Hypercapnia was created, referring to the induction of carbon dioxide in the blood of people.

Using an evaluation metric called ROC curve, where the specificity and sensitivity of the model are measured, it is established that the results were positive compared to the model proposed by the researcher Simpson, obtaining a 94% chance that the results were consistent with those expected, concerning the 92% raised by the model of the 7 coefficients. Therefore, a better result is generated, compared to the previous model.

RESUMEN

El cerebro compromete una serie de funciones esenciales para el funcionamiento de diversos procesos dentro del organismo, teniendo un impacto directo tanto para sistemas como para órganos relevantes. En particular, el sistema de autorregulación cerebral permite comprobar el funcionamiento de la autorregulación sanguínea dentro del cerebro, estableciendo ciertos estados en los que el cuerpo humano reacciona de manera positiva o negativa, dependiendo de la forma en la cual se esté autorregulando la sangre. Por esta razón, es que existe el índice de autorregulación cerebral, encargado de comprobar la forma en la cual se está comportando la autorregulación dentro del cerebro en un momento determinado, esto con la intención de inferir si el sujeto en cuestión está en perfectas condiciones o puede estar propenso a sufrir alguna enfermedad neurodegenerativa. Cabe destacar que este índice varía entre 0 y 9, dependiendo del estado de la persona. En caso de que el sujeto este autorregulando de buena manera, se le asigna automáticamente el mayor valor.

La presente investigación tiene como objetivo desarrollar un modelo predictivo para calcular el índice de Autorregulación Cerebral utilizando Support Vector Machine como algoritmo principal. Mediante un estudio donde se analizaron 27 pacientes, se obtuvieron 7 coeficientes para cada uno de los sujetos que permitieron establecer una relación matemática para determinar el ARI de forma menos invasiva, en comparación a los procedimientos existentes. De esta forma, se creó el modelo capaz de predecir cada índice de autorregulación cerebral en estados de Normocapnia e Hipercapnia, haciendo referencia a la inducción de dióxido de carbono en la sangre de las personas.

Utilizando una métrica de evaluación denominada curva de ROC, en donde se mide la especificidad y sensibilidad del modelo, se establece que los resultados fueron positivos en comparación al modelo planteado por el investigador Simpson, obteniendo un 94% de posibilidades que los resultados fueran acordes a los esperados, en relación con el 92% planteado por el modelo de los 7 coeficientes. Por lo tanto, se genera un mejor resultado, en comparación al modelo anterior.

GLOSARIO Y RELACIONES

<i>SAC</i>	<i>Sistema de Autorregulación Cerebral.</i>
<i>ARI</i>	<i>Índice de Autorregulación, de su sigla en inglés Autoregulation Index.</i>
<i>CARNet</i>	<i>Cerebral Autoregulation Research Network.</i>
<i>SÍSTOLE</i>	<i>Ciclo donde el musculo cardiaco se contrae.</i>
<i>DIÁSTOLE</i>	<i>Ciclo donde el musculo cardiaco se relaja.</i>
<i>SVM</i>	<i>Support Vector Machines.</i>
<i>FSC</i>	<i>Flujo Sanguíneo Cerebral.</i>
<i>PAM</i>	<i>Presión Arterial Media.</i>
<i>PPC</i>	<i>Presión de Perfusión Cerebral.</i>
<i>PIC</i>	<i>Presión Intracraneal.</i>
<i>VFSC</i>	<i>Velocidad del Flujo Sanguíneo Cerebral.</i>
<i>PSA</i>	<i>Presión Sanguínea Arterial.</i>
<i>PCC</i>	<i>Presión Critica de Cierre.</i>
<i>RAP</i>	<i>Producto Resistencia-área.</i>
<i>FIR</i>	<i>Filtro de respuesta finita al impulso.</i>

INTRODUCCIÓN

Hoy en día la ciencia es capaz de otorgar facilidades que permiten visualizar diversas enfermedades en los seres humanos, la capacidad de descubrimiento en diferentes áreas y el incremento exponencial que poseen las máquinas en relación con el cómputo de información, hacen posible la utilización de herramientas para la detección de patologías o simplemente conocer el estado de salud de una persona.

En muchas ocasiones estas herramientas ya mencionadas, poseen una gran exactitud en los resultados para realizar análisis más complejos dentro del cuerpo humano como son las biopsias, exámenes relacionados con diferentes sistemas como respiratorio, digestivo, endocrino, nervioso, entre otros. Por otro lado, y en relación con lo anterior, es que en algunos casos se incurre a procedimientos más invasivos lo que con lleva a una desestabilización dentro del cuerpo humano, ya que se manipula de manera directa con órganos que pueden verse comprometidos mediante elementos usados en aquellos procedimientos o la interacción de sustancias o radiaciones que pueden alterar la composición y llevar a ocasionar enfermedades o deterioros, tanto a corto como a largo plazo.

El sistema de autorregulación provee la capacidad de mantener un equilibrio constante en el flujo de sangre, lo que permite diversos funcionamientos físicos y cognitivos dentro del organismo, es por esto que la importancia de analizar el índice que contribuye al conocimiento del estado actual en el que se encuentra tal flujo dentro un ser humano, toma tanta relevancia hoy en día, ya que a pesar de todos los descubrimientos científicos hasta la fecha, aun no se logra encontrar una manera menos invasora en relación a los procedimientos ya existentes.

La obtención mediante un modelo de aprendizaje automático basado en una tasa de error inferior a los modelos existentes hoy en día será el objetivo a cumplir, en base al análisis del estado del arte, comparaciones y contrastes con variables utilizadas en diversos modelos pertenecientes al mundo de la neurología e interpretaciones mediante las iteraciones que se realicen día a día. El uso de algoritmos de aprendizaje automático, tendrán una implicancia en los efectos de esta investigación, en este caso será utilizado Máquinas de Soporte Vectorial (*del inglés Support Vector Machine, SVM*) correspondiente a un conjunto de algoritmos para resolver problemas tanto de clasificación como de regresión.

El proyecto cumple con la labor de contribuir a una problemática existente en la ciencia, como es la obtención de métodos no invasivos para poder establecer un índice de autorregulación cerebral, que pueda indicar el estado actual en el cual se encuentra el flujo sanguíneo dentro del sistema cerebral y como estos resultados podrían influir en la toma de decisiones en base a los cuidados que un agente deba tener.

Mediante bases de datos en donde se exponen trabajos dentro del área de autorregulación cerebral, se obtendrá información para ser expuestas y comparadas de diversos científicos, los cuales juegan un papel determinante en el presente trabajo investigativo cumpliendo el rol de participantes claves, para poder realizar todos los objetivos que se plantean.

Se espera conseguir resultados eficaces, con tasa de errores mínimas en comparación a estudios anteriores, para poder representar y evaluar de una mejor manera este índice de autorregulación cerebral encargado de establecer los parámetros para dar cuenta de un buen o mal rendimiento acerca de la autorregulación sanguínea que posee el cerebro y no continuar utilizando equipamientos que inducen a un deterioro en la salud de las personas.

1. PLANTEAMIENTO DEL PROBLEMA

Como se conoce, la complejidad que posee el cerebro aún no está establecida mediante parámetros finitos, ya que siempre se descubren nuevos elementos que cambian los paradigmas que enfrentan los especialistas del área neurológica. Si bien la tecnología ha contribuido a un avance en cuanto a las mediciones de diversas variables dentro del comportamiento cerebral, la determinación sobre si estos componentes al proveer información presentan ciertas incongruencias, aún están en velo. Por esta razón es que los científicos hacen hincapié en el estudio del sistema cerebral en conjunto con sus subsistemas, para comprender de forma efectiva los diversos padecimientos que pueda contraer una persona.

El sistema de autorregulación cerebral (SAC) permite el correcto funcionamiento del flujo sanguíneo al presentarse diversos cambios de presión en la circulación arterial de la sangre. Su función principal consiste en la modificación de los diámetros de las paredes de los vasos sanguíneos, para establecer un constante flujo de sangre durante esta respectiva circulación. Este proceso posee una alta reactividad vascular, lo que con lleva a una emisión de energía espontánea y de gran nivel, esto produce que la importancia y sincronización que posea este sistema al momento de entregar una constante y prolongada circulación de la sangre al interior del cerebro sea tan relevante para todas las funciones que posee el sistema nervioso. *(Ortiz-Prado, Banderas León, Unigarro y Santillán, 2018)*

El estado del conjunto de funciones que conforman al SAC convergen hacia ciertos síntomas en el caso que la autorregulación no se realice de buena manera, lo que conllevaría a diversas enfermedades que se han demostrado como Fibromialgia en donde se presentan dolores crónicos en los tejidos blandos del cuerpo humano y en los músculos, Accidentes cerebrovasculares, Alzheimer, entre otros. Por esta

razón es la importancia de encontrar un punto exacto que permita analizar de una forma simple esta regulación.

Para conocer el comportamiento de este sistema, es que se utiliza el índice de autorregulación cerebral (ARI), valor que permite establecer si se está comportando de forma positiva o negativa el SAC. Su representación se expresa mediante un rango de números en donde mientras vaya aumentando tal variable, se considera como un buen comportamiento. Existen otras variables que definen este índice como la velocidad del flujo sanguíneo, que representa la velocidad a la que atraviesa una cantidad determinada de sangre en un periodo específico, y la presión arterial, indicada como la fuerza que ejerce la sangre contra las paredes de los vasos sanguíneos. La variación de estas dos variables representa un cambio continuo en el denominado índice de autorregulación. (*Chacón, M. Jara, JL and Panerai, 2014*)

Siguiendo con lo anterior, existen variables externas e internas que involucran una relación directamente proporcional al funcionamiento de este flujo, los cuales han sido aplicados a diversos modelos para comprender la forma en que actúa este sistema y así abarcar ciertos parámetros para tomar resguardos en relación con los resultados obtenidos de cada paciente. Como se mencionó al principio, el nivel de conocimiento que se posee actualmente en el siglo XXI es amplio en comparación a la información que se manejaba hace unos 30 años¹, sin embargo, la complejidad que posee el cerebro no permite validar completamente estos sistemas implantados en el área de salud.

¹Revista Tendencias Científicas, Artículo “Complejidad de nuestro cerebro tiene su lado oscuro”, 2019.

Cabe destacar que el estado actual dentro del campo neurocientífico encargado de analizar y ejecutar nuevos descubrimientos acerca de un índice de autorregulación cerebral se ve afectado por la poca participación de agentes que impiden realizar avances comprometedores en cuanto a la exactitud de resultados. Cerebral Autoregulation Research Network (CARNet) es una organización encargada de representar y apoyar a investigadores en el área de la autorregulación cerebral, con el objetivo de poder compartir conocimientos, mediante papers o conferencias para ir presentando nueva información relevante. Esto habla de la importancia de mantenerse activo en cuanto a los nuevos descubrimientos que se vayan presentado sobre el tema expuesto.

Por otra parte, el énfasis a este descubrimiento de nuevos métodos está asociado a una problemática clínica directa, ya que la única forma de demostrar el estado real en el cual se encuentra este sistema, se basa en equipamientos clínicos, que muestran ciertos parámetros como se nombró anteriormente, presión arterial, velocidad del flujo sanguíneo, niveles de sístole y diástole, entre otros, con gran eficacia pero a la misma vez son métodos muy invasivos para el estado de salud de diversas personas, por lo que es difícil ir monitoreando constantemente a alguien, ya que podría afectar tanto a corto como a largo plazo al individuo.

Haciendo referencia a puntos anteriores, efectivamente existen métodos creados por científicos, basados en cálculos matemáticos a través de herramientas de análisis, que consiguen establecer tal índice de una manera menos invasiva que los equipamientos clínicos, y ayudan a contribuir a la ciencia de la neurología. El problema se destaca, en la poca efectividad o gran tasa de error que puedan poseer esos modelos, ya que utilizan datos de entrenamiento que pueden ir variando con el tiempo y a la larga representar generalizaciones que no siempre son las más adecuadas, dependiendo del contexto y del agente que será analizado. (*Chacón, Jara JL. N. Varas, N. and Panerai, RB, 2014*)

Es por aquello que el enfoque que tendrá el presente trabajo de investigación será conseguir mediante un estudio deductivo, un método más exacto para la obtención del índice de autorregulación cerebral y así poder determinar el estado actual de la persona, como también ir verificando si existen cambios en el sistema nervioso, en cuanto al flujo de sangre que entra y sale del cerebro, con el objetivo de poder prevenir y ayudar a las personas que lo necesiten. Por esta razón es que, mediante un modelo, se tratará de representar este ARI, basado en modelos anteriores y reduciendo la tasa de error que existe actualmente, a través de un estudio justificado en donde se presentaran ciertas aristas que comprometen a una correcta evaluación de este índice tan relevante.

1.1 OBJETIVOS

Para llevar a cabo el proyecto planteado se definen los siguientes objetivos.

1.1.2 OBJETIVO GENERAL

Crear un modelo de estimación mediante algoritmo de máquina de vectores de soporte (SVM) para la obtención del índice de autorregulación cerebral.

1.1.3 OBJETIVOS ESPECÍFICOS

1. Reconocer la función que tiene el sistema de autorregulación cerebral y el índice de autorregulación.
2. Identificar variables que afectan directamente al sistema de autorregulación cerebral.
3. Comparar resultados de modelos de aprendizaje automático existentes en el área del índice autorregulatorio expuestos por investigadores.
4. Definir un modelo de aprendizaje automático para la obtención del índice de autorregulación cerebral.
5. Interpretar los modelos de aprendizaje para conseguir valores certeros en cuanto a la obtención del índice de autorregulación cerebral.

1.2 HIPÓTESIS

“A través de un algoritmo de aprendizaje automático correspondiente a Support Vector Machine, se logra conseguir un modelo para la obtención del índice de autorregulación cerebral con una tasa de fallo menor a los modelos ya existentes.”

1.3 SOLUCIÓN PROPUESTA

Tras la problemática señalada, se puede apreciar una necesidad absoluta de conseguir métodos más eficaces para la obtención del índice de autorregulación cerebral, con el fin de contribuir a la ciencia neurológica para el uso de métodos menos invasivos y más exactos.

En relación con lo anterior, mediante un conjunto de herramientas tecnológicas y centrándose en un algoritmo de aprendizaje automático denominado Support Vector Machines, se realizará un entrenamiento que consiste en el descubrimiento de valores en base a variables de entradas y de salidas. Estos datos corresponden principalmente a la Presión Sanguínea Arterial y la Velocidad del Flujo Sanguíneo Cerebral que representaran la entrada y la salida respectivamente, con el fin de realizar un procesamiento de datos que permita tanto entrenar como validar un modelo que sea capaz de discriminar estas clases con relación a las señales reales de cada una de estas variables.

El modelo enfocado en SVM tendrá la particularidad de seguir el modelo propuesto por Aaslid-Tiecks en los años 80, para analizar de forma completa la resolución y la obtención del índice autorregulatorio que permite esclarecer y mejorar los falsos positivos que fueron demostrados en investigaciones anteriores. Este modelo fenomenológico mediante el uso de regresión permite establecer nuevos resultados que podrán ser comparados y retroalimentados por estudios anteriores, con el fin de obtener un ARI aceptable en base a los datos de salida correspondiente como ya se mencionó con anterioridad a la Velocidad del Flujo Sanguíneo Cerebral.

1.4 ALCANCES Y LIMITACIONES

1.4.2 ALCANCES

La investigación responderá al descubrimiento de un nuevo modelo para la obtención del índice de autorregulación cerebral utilizando un algoritmo de aprendizaje supervisado, que otorgará resultados diferentes a los modelos ya planteados, mediante la utilización de Support Vector Machine el cual resuelve problemas sin ningún inconveniente de clasificación y regresión dentro de un conjunto de datos. A diferencia de los métodos anteriores, se utilizará un modelo de clasificación utilizando herramientas de aleatorización de hiperparametros. Esto para contribuir de manera positiva a la disminución de la tasa de errores que existe actualmente en los diferentes modelos aplicados en base a otros algoritmos de aprendizaje.

Si bien, el trabajo finaliza en la creación de un modelo predictivo, la posible continuación de este proyecto implicará mayores avances en cuanto a los elementos que se utilicen en el presente trabajo, es por esto que se delimita el estudio hasta conseguir interpretaciones claras y concisas, sobre los resultados que se obtengan, cumpliendo tanto con los objetivos específicos como generales.

1.4.3 LIMITACIONES

En consecuencia, al realizar siempre una investigación que implique el descubrimiento de nuevos conocimientos, las limitantes en base al entorno que rodea a este trabajo constituirán una forma de visualizar de qué manera se van a plantear los objetivos y que información se tiene al respecto del tema principal.

En este caso, los datos juegan un rol importante, adquiriendo la labor de ser procesados para un correcto manejo de información y para la obtención de conocimientos en cuanto a la utilización de métodos más exactos dentro de esta ciencia y menos dañinos. Dicho lo anterior, la variabilidad de los datos en conjunto con las características que posean las personas, serán determinantes a la hora de representar los resultados ya que si bien se puede realizar una correcta transformación y limpieza del conjunto de datos, aspectos como la calidad de vida de la persona, estado de salud actual o simplemente el hecho de poseer una capacidad física superior a la otra, implica que la generalización se verá afectada, por ende el modelo sufrirá efectos adversos, y por consiguiente se debe tener total cuidado al interpretar cada una de estas variables.

2 MARCO TEÓRICO

Basado en el conocimiento necesario para cumplir con los objetivos planteados, es de suma relevancia hacer énfasis en cada uno de los tópicos a analizar con el fin de poder esclarecer cada aspecto que será de ayuda para notar el comportamiento intrínseco del modelo a desarrollar por parte del alumno.

Los puntos más importantes tendrán que ver con la descripción del sistema de autorregulación cerebral indicando función y características, índices de autorregulación cerebral, enfermedades comprometidas para notar el impacto del estudio presente, algoritmos de aprendizaje automáticos haciendo hincapiés en Máquinas de Soporte Vectorial, modelos de aprendizaje automático de referencias y finalmente metodología a utilizar.

El sistema de autorregulación cerebral compromete una serie de funciones referidas al constante flujo de sangre en movimiento, permitiendo que este proceso se vaya adaptando mediante el cambio continuo de los vasos sanguíneos en relación a los cambios del diámetro que estos posean lo que hace posible el ingreso de sangre en diferentes cantidades para continuar con un correcto funcionamiento que involucra muchas veces funciones tanto cognitivas como físicas, las cuales fueron mencionadas anteriormente.

El índice de autorregulación cerebral corresponde al elemento principal de esta investigación, ya que permite establecer el comportamiento actual de la Autorregulación Sanguínea Cerebral, esclareciendo aristas o patologías que posiblemente el agente pueda tener en momentos específicos.

Las enfermedades dan cuenta del impacto que puede tener el descubrimiento de un nuevo método científico que facilite el conocimiento de un comportamiento anómalo dentro del mismo organismo, con el fin de poder prevenir ciertas patologías o enfermedades que la persona pueda poseer, independiente de factores externos o internos que existan.

Los algoritmos de aprendizaje automático serán los responsables del cumplimiento de los objetivos, teniendo en cuenta que la capacidad de cómputo y procesamiento que se les aplique a los datos a analizar, tendrán que poseer ciertos patrones para identificar un comportamiento y realizar un análisis posterior que ayude a establecer parámetros para la creación de un modelo más exacto en la obtención de un índice de autorregulación cerebral.

La metodología por utilizar establecerá las aristas y pasos a seguir para el cumplimiento de una correcta investigación, con el fin de poder trabajar de una manera más ordenada y sistemática que facilite la obtención del objetivo a cumplir.

2.1 SISTEMA DE AUTORREGULACIÓN CEREBRAL

Una tarea imprescindible para el organismo es la labor que cumple el sistema autorregulatorio del cerebro en donde su comportamiento define el flujo correcto de sangre que entra y sale de este órgano para recibir nutrientes representados principalmente por oxígeno y glucosa, lo que con lleva a un funcionamiento constante que permite el buen estado de las funciones que cumple el cerebro en todo el cuerpo humano.

El sistema de autorregulación cerebral (SAC) es un conjunto de procesos constantes que permiten la perfusión en un rango amplio de valores. La perfusión se puede definir como la acción que ejerce el sistema circulatorio en cuanto al traspaso capilar de sangre hacia los tejidos. Cabe destacar que la variación de esta perfusión define los umbrales máximos o mínimos de la autorregulación, lo que con lleva a un estado correcto o fallido en algún momento del proceso de traspaso de la sangre al cerebro. *(Ortiz-Prado, Banderas, Unigarro & Santillán, 2018)*

Este proceso de autorregulación o también conocido como vasorreactividad cerebral, representa un mecanismo de protección para el individuo. Los vasos sanguíneos modifican su diámetro mediante este sistema para mantener un flujo constante durante el traspaso de sangre hacia los diversos tejidos ,en este caso particular el cerebro, en base a esto permite que el capilar no se inunde ni se seque, lo que con llevaría en el primer caso a una rotura de la barrera hematoencefálica con edema cerebral vasogénica y sangrado pudiendo causar la creación de algún tumor o neoplasma, y en el segundo caso a la isquemia cerebral o mejor conocido como un accidente cerebrovascular que está representado por la falta de sangre a una zona del cerebro.

Existen variables involucradas en el mecanismo de la autorregulación del flujo sanguíneo cerebral (FSC), las cuales fundamentan el comportamiento de una vasorreactividad cerebral acorde a lo necesario para sostener un nivel de vida continuo,

estos factores pueden ser externos como internos del organismo, permitiendo describir la variabilidad en el nivel de autorregulación en cada una de las personas. Estos factores fisiológicos que alteran el flujo sanguíneo directamente son: Temperatura, presión Arterial Media (PAM), presión de perfusión Cerebral (PPC), presión Intracraneal (PIC), presión parcial de oxígeno arterial, presión parcial de dióxido de carbono, vasos dilatadores como el óxido nítrico (NO), vasoconstrictores como la adrenalina, la viscosidad sanguínea y la actividad simpática y parasimpática del sistema nervioso en general. (*Aaslid, Lindegaard, Sorteberg, Nornes, 1989*)

Tras la identificación de los actores involucrados en el proceso de autorregulación cerebral, se considera importante el hecho de describir la relación directa de la velocidad del flujo sanguíneo cerebral junto con la obtención del estado de este sistema autorregulatorio, por lo que se define la siguiente ecuación₁:

$$VFSC = \frac{PSA - PCC}{RAP} \quad (1)$$

En donde:

- PSA corresponde a la Presión Sanguínea Arterial.
- PCC corresponde a la Presión Crítica de Cierre.
- RAP corresponde Producto Resistencia-área.

En base a ecuación₁ se describe la relación entre la velocidad del flujo y la presión sanguínea arterial, lo que permite obtener un valor de velocidad final que involucra directamente a la obtención del indicador ARI a estudiar.

Por consiguiente, en relación con estas variables mencionadas, el equilibrio adecuado entre la perfusión cerebral, presión arterial media y la presión intracraneal aseguran un comportamiento eficaz de este sistema para así contribuir en el transporte de oxígeno hacia los tejidos del cerebro, teniendo en consideración que la importancia de un conveniente mecanismo está asociado al gasto cardíaco que necesita el cerebro equivalente entre un 15% y un 25% para su correcto funcionamiento.

2.2 ÍNDICE DE AUTORREGULACIÓN CEREBRAL

La representación del comportamiento del sistema autorregulatorio cerebral permite explicar cuantitativa y cualitativamente el real estado de la regulación sanguínea hacia el cerebro, lo que implica un conocimiento previo para la toma de decisiones futuras en algún agente.

El índice de autorregulación cerebral (ARI) corresponde a un valor discreto que posee una numeración entre intervalos de 0 a 9, en donde un valor $ARI=0$ representa una ausencia de autorregulación y un $ARI=9$ representa una respuesta rápida y por ende al mismo tiempo un buen comportamiento de la autorregulación sanguínea.

2.3 ENFERMEDADES COMPROMETIDAS

El correcto funcionamiento en la autorregulación sanguínea del cerebro compromete no solo aspectos fisiológicos dentro del organismo, el cerebro como ya es sabido, es un órgano complejo de analizar, a pesar de las diversas tecnologías que existen hoy en día, las variantes acerca del comportamiento cerebral aun no son descritas en su totalidad y están muy lejos de hacerlo. Es por esto que su complejidad hace que el estudio de enfermedades dentro de la neurociencia sea tan poco explorado.

Las enfermedades comprometidas dentro de este ámbito se denominan neurodegenerativas, y corresponden a enfermedades de causas desconocidas, inicio insidioso y que tienen en común el hecho de ser progresivas. Estas enfermedades producen alteraciones en el nivel de conciencia o en los mecanismos de la deglución que con frecuencia hacen necesario un soporte nutricional especializado.

Las enfermedades neurodegenerativas más comunes comprometidas por el funcionamiento de la autorregulación cerebral corresponden a esclerosis lateral amiotrófica, Alzheimer, Parkinson, y otras alteraciones de la hemodinámica cerebral, incluidos los accidentes cerebrovasculares cerebrales, hemorragia subaracnoidea, entre otras.

Cabe recalcar que, si bien el sistema de autorregulación permite fundamentar la aparición de estas enfermedades, elementos internos del organismo como la previa alimentación de una persona o estado de salud físico, son variables para considerar dentro de este tipo de anomalías, a pesar de que en un gran porcentaje todos estos factores mencionados dan cabida a una mala regulación del flujo sanguíneo cerebral.

2.4 ALGORITMOS DE APRENDIZAJE AUTOMATICO

El conocimiento junto con el aprendizaje, son dos aspectos claves dentro de la vida cotidiana. En base a la experiencia los seres vivos van aprendiendo a tomar decisiones o efectuar acciones para tener una vida activa y duradera. Gracias al nivel computacional se han efectuado grandes avances en el área de Inteligencia Artificial, en donde las capacidades cognitivas han podido ser representadas por máquinas que mediante una gran cantidad de datos pueden aprender en diferentes contextos.

Aprendizaje automático es un área dentro de la computación utilizada primordialmente por la Inteligencia artificial para encontrar patrones y predecir ciertas acciones en base a un conjunto de datos adquiridos, con el fin de poder cumplir con un objetivo en particular. Cabe destacar que este concepto no solo está dentro del área computacional, sino que se relaciona tanto con la filosofía, psicología y biología en algunos aspectos. Es de suma importancia recalcar que el aprendizaje por sí solo permite incrementar el conocimiento o habilidades de un agente para cumplir con una tarea en específica.

Los algoritmos de aprendizajes automáticos, basados principalmente en una lógica matemática permiten analizar millones y millones de datos, con el fin de poder encontrar como se mencionó anteriormente patrones que permitan identificar una relación directa entre las variables involucradas y las decisiones u objetivos que se quieran abordar, dependiendo del problema o el contexto en cuales están inmersos. La información utilizada para poder realizar un aprendizaje correcto debe cumplir con ciertas características que permitan a la máquina realizar una generalización en cuanto a un comportamiento, con el fin de aplicar el modelo adquirido y que este sea adaptable a los mayores casos posibles.

Para argumentar de mejor manera el objetivo que poseen estos algoritmos de aprendizaje automático, es de suma importancia conocer los 3 tipos de aprendizaje que se realizan actualmente.

a. APRENDIZAJE SUPERVISADO

El aprendizaje se realiza gracias a ejemplos proporcionados como entradas los cuales son necesarios para cumplir con las metas impuestas. Tiene una relación directa con la relación de profesor-alumno. En este tipo de método se dan ejemplos y se especifica de que concepto son.

b. APRENDIZAJE NO SUPERVISADO

El aprendizaje se realiza mediante el descubrimiento de nuevos conocimientos y regularidades de los datos, sin enfocarse en un objetivo en particular. A la máquina no se le entrega al resultado al cual se quiere llegar.

c. APRENDIZAJE POR REFUERZO

Este aprendizaje corresponde a un intermedio entre los dos anteriores. El sistema busca la solución a un problema en particular, siguiendo una lógica conductista en la cual mediante un tercero que juega el papel de entrenador, además de un entorno creado por este mismo, se busca conseguir un objetivo en base a prueba y error, estimulando al agente con ganancias en caso de no equivocarse, o pérdidas (castigos) en el caso de equivocarse.

2.5 SUPPORT VECTOR MACHINE

Máquina de vectores de soporte o Support Vector Machine es un método matemático, desarrollado en los años 90 dentro del área de la computación para aplicarlo a diversos modelos con el fin de poder contribuir a problemas de clasificación múltiple y regresión tanto en áreas matemáticas como fundamentalmente en problemas de cómputo directo.

Su objetivo es la optimización o maximización dentro del descubrimiento de patrones, con el fin de poder esclarecer y distinguir características propias de cada uno de los elementos involucrados en el modelo, a cada objeto con su característica se le denomina clase. El algoritmo es capaz en base a una serie de elementos matemáticos de poder identificar en base a entrenamiento a que clase pertenece cada dato entregado por un agente, para así luego poder ir identificando de manera automáticamente las entradas correspondientes.

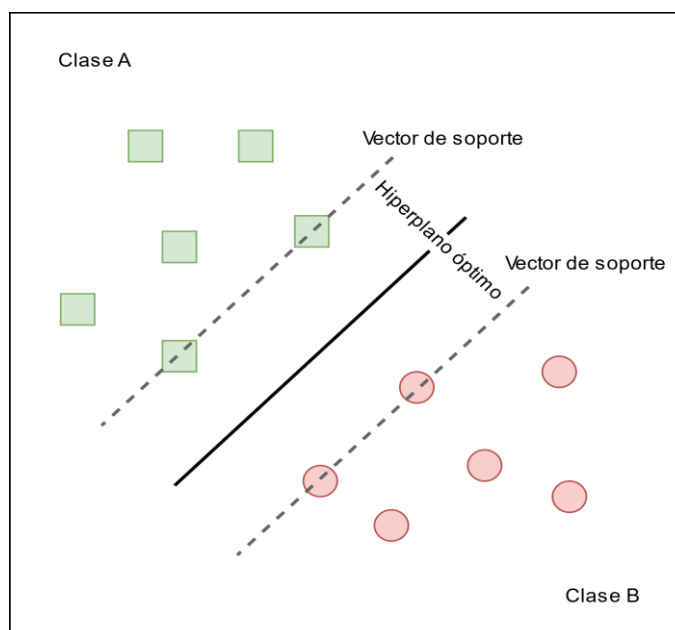


Figura 1: Hiperplano de separación óptima en dos dimensiones – SVM (Fernández-Sainz, 2011).

En el caso de un problema donde existen dos clases diferentes, un SVM se entrena de manera que la función maximice la habilidad de generalización. Esto se representa en la figura 1, donde se trata de encontrar una forma de separar las figuras cuadradas y redondas con el fin de identificar sus características propias y finalmente encasillarlas.

Como se puede apreciar, el hiperplano óptimo engloba todos los elementos fundamentales dentro del SVM, lo que permite determinar la separación maximizada representada por el margen que a su vez siempre busca la mayor división, para apreciar y clasificar de mejor manera en base a datos entregados. Los puntos más próximos a la recta, son los que permiten localizar de buena manera los vectores de soporte que fundamentan la localización del hiperplano de separación. Cabe destacar que todo dato de entrenamiento debe satisfacer la condición de tener una distancia al hiperplano, mayor que la distancia de los vectores de soporte. (Banf, 2019)

Haciendo nuevamente referencia a la figura 1, la forma más sencilla de realizar una separación es mediante una línea recta, un plano recto o un hiperplano N-dimensional, como se denota en la figura. Sin embargo, problemas computacionales referentes a tipos de aprendizajes actuales, exigen un trabajo de complejidad mayor como los casos de mayor cantidad de clases, curvas no lineales de separación, clasificación en más de dos categorías o casos donde los conjuntos de datos no pueden ser separados completamente. Para esto se utiliza la función kernel, la cual pertenece a una amplia gama de funciones matemáticas que permiten convertir un problema de clasificación no lineal a un problema de clasificación en un espacio dimensional mayor.

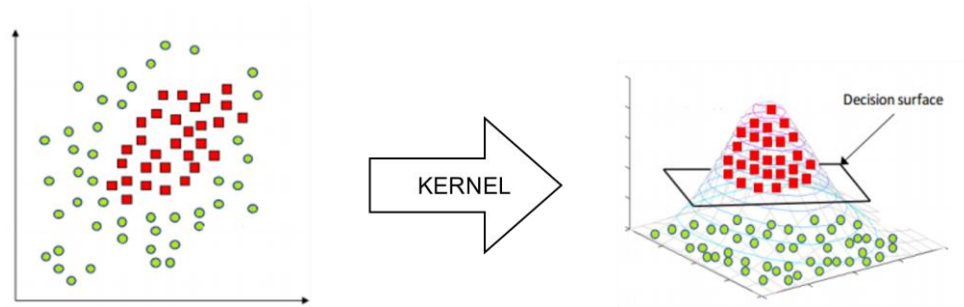


Figura 2: Aplicación de una función kernel modificando el espacio de dimensiones representado por la línea continua en verde, separando el conjunto de datos rojos y azules (Baudat, G., Anouar F, 2000).

Como se representa en la figura 2, la dimensionalidad puede ser adaptada a cada problema en particular, independiente de la complejidad que tenga la clasificación de elementos o en su justa medida la separación de estos mismos.

2.6 MODELO DE AASLID-TIECKS

Diversos estudios dentro del área neurológica han podido demostrar ciertos métodos para la medición de variables en el cuerpo humano, como es el caso de la presión, frecuencia cardíaca, entre otros. Es de suma importancia conocer la raíz de todo origen para comprender el valor real de cada descubrimiento científico.

El modelo de Aaslid-Tiecks permite fundamentar la obtención del índice de autorregulación Cerebral, con la maniobra de oclusión de la circulación en los muslos que permite estudiar el comportamiento dinámico de la autorregulación del flujo sanguíneo cerebral. En base a esto mediante variables como la velocidad del flujo sanguíneo cerebral en respuesta a los cambios que se efectuaban en la presión sanguínea arterial, se logró realizar una correlación entre estos factores que permitieron demostrar y efectuar un modelo matemático capaz de representar este ARI.

Mediante tres parámetros de una ecuación diferencial de segundo orden, se permite calcular un índice discreto con intervalos de valores entre el 0 y 9 lo cual representa claramente una ausencia de autorregulación o un comportamiento constante y gradual de este mismo. (*Aaslid, Lindegaard, Sorteberg, Nornes, 1989*)

El modelo plantea que en base a cambios de presión arterial representado por $\mathbf{P}(t)$ se logra obtener la Velocidad del Flujo sanguíneo Cerebral (VFSC) correspondiente a \mathbf{V}' , utilizando cuatro ecuaciones de estado.

$$\mathbf{dP}(t) = \frac{P(t)}{1-CCP_r} \quad (1)$$

$$\mathbf{X}_1(t) = X_1(t-1) + \frac{dP(t-1)-X_2(t-1)}{f \times T} \quad (2)$$

$$\mathbf{X}_2(t) = X_2(t-1) + \frac{X_1(t-1)-2 \times D \times X_2(t-1)}{f \times T} \quad (3)$$

$$\mathbf{V}'(t) = 1 + dP(t) - K \times X_2(t) \quad (4)$$

Donde $\mathbf{dP}(t)$ se encarga de normalizar la presión vía línea base, \mathbf{CCP}_r es un valor introducido por Tiecks, el cual hace referencia a la presión de cierre critica, \mathbf{f} corresponde a la frecuencia de muestreo, \mathbf{K} representa un parámetro de ganancia en la ecuación, \mathbf{T} es la constante de tiempo y \mathbf{D} es el factor de amortiguamiento. Por otra parte, $X_1(t)$ y $X_2(t)$ son las variables de estado de un sistema diferencial.

Este modelo combina los parámetros ***K***, ***D***, ***T***, los cuales son asociados un valor ARI específico con el fin de poder representar el estado del individuo analizado. Esto se puede apreciar en la siguiente Tabla 1.

K	D	T	ARI
0.00	1.70	2.00	0
0.20	1.60	2.00	1
0.40	1.50	2.00	2
0.60	1.15	2.00	3
0.80	0.90	2.00	4
0.90	0.75	1.90	5
0.94	0.65	1.60	6
0.96	0.55	1.20	7
0.97	0.52	0.87	8
0.98	0.50	0.65	9

Tabla 1: Asociación entre parámetros para obtención de ARI.

El modelo produce diez curvas que representan cada ARI en base a la velocidad $V'(t)$ para cada parte de medida $P(t)$ real de un sujeto, mediante las comparaciones con la velocidad real del sujeto, se miden por error cuadrático mínimo o correlación Máxima, entre la velocidad real y la velocidad estimada por el modelo. Finalmente, la asignación del valor de ARI se consigue mediante el ajuste entre la velocidad real y las 10 curvas estimadas por el modelo mediante un error o correlación directa representados por figura3

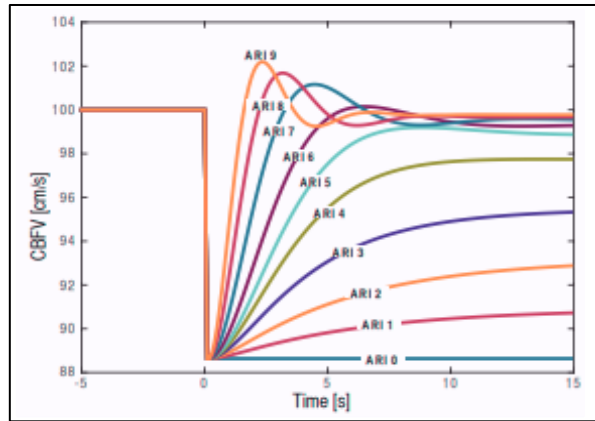


Figura 3: Los 10 niveles de ARI basado en la respuesta PA Y VFSC por Aaslid-Tiecks (Simpson, 2017).

2.7 RESPUESTA FINITA AL IMPULSO (FIR)

En sistemas computacionales la conversión de señales analógicas como entradas para convertirse en señales digitales permite describir ciertos comportamientos a partir de los objetivos que se quieran cumplir. Es posible hoy en día representar todo tipo de señal para ser analizada y estudiada a profundidad por los computadores. Esto es gracias a lo que se conoce como un filtro digital lo cual permite realizar la conversión señalada anteriormente.

FIR o respuesta finita al impulso corresponden a un tipo de filtro digital cuya salida tendrá una cantidad finita de términos en base a entradas representadas por señales de impulso. Este es un sistema no recursivo y de fase lineal lo que permite siempre funcionar de forma estable en todo momento. Es importante agregar que un filtro FIR tiene una implementación sencilla y es fácil de comprender, además de poseer memoria finita.

2.8 METODOLOGÍA

La metodología de investigación será utilizada para cumplir con los objetivos planteados, empleando dos enfoques fundamentales, el enfoque cualitativo y enfoque cuantitativo, recalando que este último será utilizado en la presente investigación.

Trabajar de una forma planificada y constante dependerá de cómo se establezcan los estándares de trabajo, por aquello que la metodología de investigación conlleva una serie de características que la hacen relevante para la ejecución de un plan de trabajo que involucre elementos tangibles y superficiales. El campo del conocimiento y la abstracción de información fundamentan un buen desarrollo investigativo. El objetivo de investigar es conseguir conocimiento en base a la búsqueda de respuestas de problemas empíricos y contribuir a la realidad de las ciencias para producir diversos cambios o mejoras dentro de este entorno. *(Sampieri, Fernández Collado, C., & Baptista, 2010)*

Esta metodología posee un paradigma basado en una fundamentación epistemológica, instrumentos o técnicas que contribuyen al objeto de estudio y diseños metodológicos.

2.9 ENFOQUE CUANTITATIVO

Esta metodología hace referencia a una investigación empírica-deductiva que indaga mediante datos reales, generalmente variables estadísticas que pueden responder a preguntas específicas. Los datos que se recolecten deben ser medibles, observables y documentados, además de poseer un origen estandarizado. El estudio debe ser lo más objetivo posible, obviando tendencias de personas o del propio investigador, para así generalizar resultados y encontrar patrones relevantes.

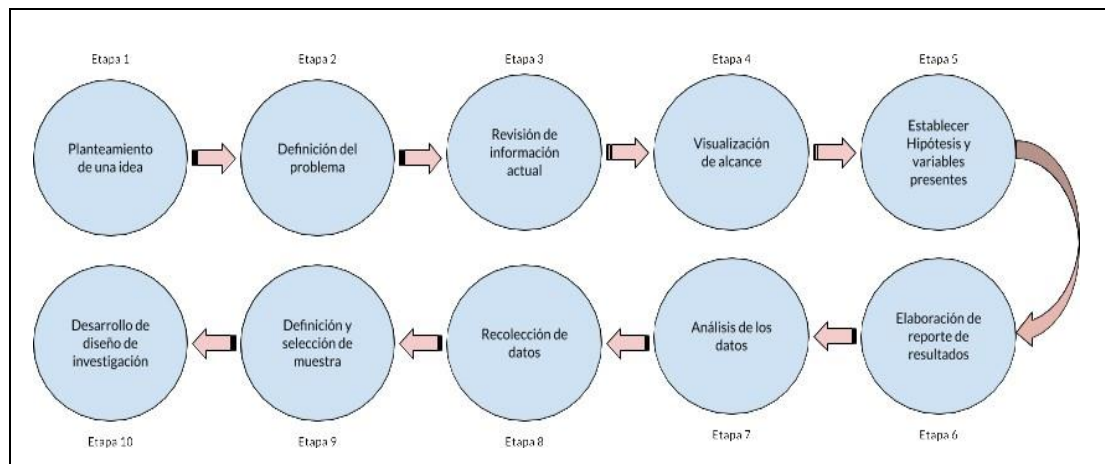


Figura 4: Proceso del enfoque cuantitativo. (Sampieri, Fernández Collado, C., & Baptista, 2010).

Los pasos a seguir son:

1. Plantear una idea a investigar, además de conocer fuentes que inspiren aquella investigación.
2. Planteamiento del problema. Se establecen los objetivos de la investigación, se desarrollan preguntas y se evalúan las deficiencias en el conocimiento del problema.
3. Revisión de estado del arte y desarrollo de Marco teórico. Se obtiene información actual de la investigación a realizar y se construye el marco teórico.
4. Visualización del alcance de estudio. Se define la investigación y el alcance.

5. Elaboración de Hipótesis y definición de variables. Si se considera conveniente se realiza una formulación de hipótesis y decisión de variables de la hipótesis.
6. Desarrollo del diseño de investigación. Consiste en la definición de escoger el diseño que sea más apropiado para esta investigación, esto de acuerdo con el planteamiento del problema e hipótesis propuesta.
7. Definición y selección de la muestra. Emplear de manera minuciosa el procedimiento de selección mediante una muestra para la investigación, esto conlleva hacer énfasis en el tamaño de la muestra requerida.
8. Recolección de los datos. Definición de la forma en que se recolectaran los datos de acuerdo con las etapas previas de la investigación, aplicando métodos o instrumentos para la codificación y obtención de estos.
9. Análisis de los datos. Se decide la forma de analizar los datos para realizar una interpretación mediante pruebas estadísticas las hipótesis planteadas, luego preparar los resultados para ser presentados.
10. Elaboración del reporte de resultados. Se escoge el tipo de reporte a presentar, para elaborar una presentación de la información deducida.

3. ESTADO DEL ARTE

La importancia de analizar estudios y fuentes de información actuales sobre investigaciones específicas, permiten otorgar un amplio rango de conocimientos para ser aplicados en el desarrollo de un trabajo investigativo, además de suplir la poca experiencia que una persona pueda tener sobre un tema en particular. Señalado esto, los análisis demostrados hasta la fecha permiten realizar una introspección en relación con el área neurocientífica y dentro de aplicaciones clínicas.

En consideración con lo anterior, los estudios dentro del área del índice de autorregulación cerebral indican que en base a los descubrimientos propuestos por Aaslid-Tiecks en donde utilizando variables como el cambio de la presión arterial y la relación directa que posee con la velocidad del flujo sanguíneo cerebral, permitió colocar los cimientos de nuevos descubrimientos. En el año 1989 la tecnología no había avanzado en relación a la actualidad, lo que con lleva a que esta investigación no esté exenta de fallas, en consideración con los resultados de estos investigadores, más adelante Panerai, James y Potter, en su trabajo denominado "*Variability of time-domain indices of dynamic cerebral autoregulation*" propuesto en el año 2003, demostraron que la variación que poseía el índice de autorregulación cerebral era mucha, en comparación a los agentes involucrados en el experimento, los cuales eran personas sin algún tipo de enfermedad neurodegenerativa y poseían buena salud, lo que también con llevaría un sesgo por parte del modelo propuesto dando como resultado falsos positivos. Si bien este nuevo trabajo permitía realizar una crítica exhaustiva acerca de la poca exactitud, también es importante señalar que la investigación estuvo enfocada en los cambios de presión arterial induciendo tal variabilidad mediante CO₂, lo cual era suministrado por profesionales del área a personas que estaban calificadas y además con sus respectivos consentimientos.

Por consiguiente, en este trabajo se presenta un índice sustituto de autorregulación cerebral derivado del modelado de la media móvil autorregresiva que consiste en la relación dinámica entre la Presión arterial y la Velocidad del Flujo Sanguíneo Cerebral, lo que contribuyó a que el concepto de variabilidad dentro del campo de regulación cerebral se viese afectado.

También es relevante hacer referencia al trabajo de Chacón, Núñez, Henríquez y Panerai denominado *"Unconstrained parameter estimation for assessment of dynamic cerebral autoregulation"* en donde proporcionan una visión aún más amplia en relación a las variables a modificar dentro del modelo para la obtención del ARI, viéndose afectados parámetros como la constante de tiempo (T), factor de amortiguamiento (D) y parámetro de ganancias (K), los cuales establecen una relación directa con el modelo planteado por Aaslid, y trabajando de igual manera con la inducción a cambios en la presión sanguínea en base a la fluctuación de CO₂ en la sangre. Anteriormente estos autores en una investigación previa ya habían mencionado la importancia que poseían los cambios inducidos de presión para la obtención directa de un Índice de Autorregulación Cerebral, en donde trabajaron con las variables más importantes correspondiente a Presión Sanguínea como entrada y la Velocidad del Flujo Sanguíneo como salida.

La fluctuación espontánea de presión arterial permite también conocer un poco más acerca de cómo es posible conseguir un mejor resultado de ARI, haciendo énfasis en la variabilidad de las pruebas y de los factores que contribuyen a esto, en el trabajo de Liu, Simpson y Allen denominado *"High spontaneous fluctuation in arterial blood pressure improves"* permiten describir el uso de 7 coeficientes para la obtención de este índice autorregulatorio, señalando que los criterios que permitieron conseguir un mejor resultado fueron gracias al hecho de poseer una mayor correlación con la presión de CO₂ inducida en las personas, en comparación a la importancia que se le daba en investigaciones anteriores a la variabilidad de los parámetros que afectan tanto a Velocidad del Flujo Sanguíneo Cerebral Y Presión arterial.

Para contrastar lo anterior, es importante estudiar las diferentes propuestas de casos que combinan el uso de SVM con una aplicación clínica para la obtención de este índice, para esto se encuentra la investigación establecida por Chacón, Jara, Miranda, Katsogridakis, Panerai denominada *"Non-linear models for the detection of impaired Cerebral Blood Flow Autoregulation"* y publicado en el año 2018, trabajo más reciente que propone el uso de un modelo no lineal en base a la variable de entrada representada por la presión arterial sanguínea y la variable de salida representada por la velocidad del flujo sanguíneo cerebral, este experimento se basó en dividir las fluctuaciones espontaneas de cada una de estas variables en registros para entrenamiento que almacenaba los primeros dos minutos y medio de cada una de las señales y el segmento de validación con los últimos dos minutos y medio. Cabe destacar que la eficiencia del modelo fue determinada gracias a correlaciones hechas empíricamente en base a la señal real de la velocidad del flujo sanguíneo y la estimación propuesta por el modelo.

Este estudio tan revolucionario determino la real importancia de la variabilidad en el estudio de sujetos, y además la gran capacidad que poseen los modelos no lineales para el aprendizaje en base a búsqueda de patrones, lo que permite establecer un gran dinamismo en base al uso de Support Vector Machine y además demostrar una gran capacidad de discriminación lo que permitió conseguir valores más certeros, en relación con otros experimentos anteriores. Sin embargo, es importante mencionar que el uso de este modelo no lineal da cuenta del positivismo que implica trabajar con algoritmos de aprendizaje automáticos, sin embargo, estas características "no-lineales" se refieren al ajuste apropiado del modelo a los datos y no necesariamente a la característica propia de la autorregulación cerebral.

A pesar de haber mencionado estudios propuestos por científicos experimentados en el área, el estado actual referente a los trabajos de proyecto de títulos en base a la determinación de un modelo capaz de mejorar la obtención del Índice de autorregulación cerebral está basado en tesis de alumnos anteriores en donde la utilización de

mecanismos de aprendizaje automáticos han sido relevantes, como el uso de caja gris combinándolo con redes neuronales en donde se consigue una leve mejoraría en los resultados, pero por sobre todo utilizar un nuevo modo de aplicar el aprendizaje a estos datos de muestras.

Estos análisis indican un patrón de estudios generalmente aplicados a pacientes con buena salud en donde las enfermedades neurodegenerativas no están presentes, sería importante denotar que la ampliación de investigación en cuanto al rango de pacientes en base a su edad y con algún tipo de daño neurológico, contribuirá a la mejora de discriminación por parte del estudio en base a los datos adquiridos.

Finalmente es importante acotar que la mejora que se busca obtener está basada en la disminución del porcentaje de error que entregan estos algoritmos a la hora del aprendizaje, variables como la Presión Arterial Sanguínea, Velocidad del Flujo Sanguíneo Cerebral o la presencia de factores externos como el aumento o disminución de presión en el aire no están en tela de juicio, pero si el modo de estudiar estas variables será de suma relevancia para el cumplimiento de los objetivos finales.

4. DESARROLLO Y PROCEDIMIENTO

En concordancia con el presente trabajo investigativo, hay que recalcar la importancia de la ejecución de un buen plan de desarrollo que permita establecer las bases para el cumplimiento de objetivos planteados con anterioridad. Por esta razón, que se expondrán y efectuarán un conjunto de actividades necesarias para conseguir resultados acordes, que permitan demostrar la efectividad del modelo a plantear por parte del alumno en cuestión.

Los procedimientos y acciones estarán divididas y enfocadas en el contexto del caso de estudio, en particular la investigación realizada por M, Simpson. Henríquez, Claudio, basándose en una problemática de clasificación, las cuales establecen las directrices que permitirán conseguir un modelo adecuado que represente a cabalidad el comportamiento del sistema de Autorregulación Cerebral en base a la obtención del índice autorregulatorio, esclareciendo las variables y los análisis correspondientes para una ejecución investigativa satisfactoria.

4.1 DEFINICIÓN DE VARIABLES

La importancia de factores, tanto externos como internos que afectan a un sistema en particular, permiten conocer el real comportamiento y la relación directa que poseen, para así establecer patrones certeros, que aporten a la obtención de resultados, esto es aplicable a cualquier tipo de modelo establecido.

El cerebro conforma un sistema indispensable para el ser humano, el cual posee funciones fisiológicas determinantes que permiten la subsistencia de este mismo, es por esto que identificar cada variable que afecta a su autorregulación sanguínea, juega un

papel determinante para reconocer los posibles estímulos que provocan cambios de manera positiva o negativa dentro del organismo.

En cuanto al SAC o sistema Autorregulatorio, se puede señalar que el efecto producido por determinadas variables cuantificadas, permiten establecer un modelo no exacto que acerque a la neurociencia a procedimientos más certeros y menos invasivos, relacionándose directamente con el objetivo planteado en el presente caso de estudio.

El índice de autorregulación cerebral representa a cabalidad este sistema en cada organismo, por lo que la volatilidad de las variables involucradas y el hecho de establecer de manera concisa cada parámetro que afecte a este índice, representaran un papel primordial para el desarrollo del modelo a trabajar.

Mediante los estudios ya mencionados por diversos científicos y académicos, los modelos propuestos establecen variables medibles y comparables, con el objetivo de aportar al análisis exhaustivo para la obtención del cálculo del ARI, que permita demostrar el comportamiento de forma precisa de este sistema autorregulatorio.

Para realizar un modelo es importante contextualizar las variables a utilizar que afectaran de manera directa tanto a los resultados del cálculo de ARI, como al desarrollo experimental del presente trabajo.

- **VFSC (V):** Correspondiente a la velocidad del flujo sanguíneo circulatorio por unidad de tiempo.
- **PSA (p):** Correspondiente a la presión sanguínea arterial.

- **COEFICIENTES (h):** Valores representativos en base a la relación VFSC Y PSA, propuestos por modelo mencionado con anterioridad.
(M, Simpson. Henríquez, Claudio)
- **HIPERCAPNIA:** Cantidad excesiva de dióxido de carbono en la sangre.
- **NOMOCAPNIA:** Cantidad normal de dióxido de carbono en la sangre.

Los parámetros mencionados, establecen un panorama específico sobre el uso de factores a analizar, para la conformación del modelo requerido, con el fin de no caer en análisis inciertos o equívocos, que no representan a cabalidad los resultados finales.

4.2 DESCRIPCIÓN DEL MODELO ACTUAL

La representación en base a un modelo existente permite establecer una pauta para un trabajo más disciplinado y orientado a objetivos reales. El modelo de Simpson denominado " *A Parametric Approach to Measuring Cerebral Blood Flow Autoregulation from Spontaneous Variations in Blood Pressure* ", plantea una serie de ideas y acondicionamientos para la obtención de un ARI respectivo, teniendo en cuenta diversas variaciones dentro del organismo, de cada sujeto estudiado.

En relación con encontrar una buena estimación para el cálculo de ARI, se estableció una relación entre las variables de la Velocidad del Flujo Sanguíneo Cerebral y la Presión Arterial Sanguínea.

En consecuencia, se desarrolló una ecuación representativa que ayude a entender de mejor forma su comportamiento:

$$V(i) = h(0)p(i) + h(5)p(i-1) + \dots + h(6)p(i-6) + e(i) \quad (5)$$

- **V (i)** representa la velocidad del flujo sanguíneo cerebral
- **p** es la presión sanguínea arterial
- **h** corresponde a los coeficientes resultantes a la respuesta del impulso de un filtro FIR, el cual tuvo como entrada y salida las variables mencionadas anteriormente.

En cuanto a los coeficientes establecidos por Simpson, se puede mencionar que permiten correlacionar los parámetros de velocidad y presión, creando una relación directa que facilite la obtención de un ARI en base a los valores de presión arterial sanguínea, obteniendo la velocidad del flujo sanguíneo directamente, utilizando el modelo de A-T.

Los resultados del modelo actual contribuyen a una comparación posterior, es por esta razón que se presenta la siguiente ilustración.

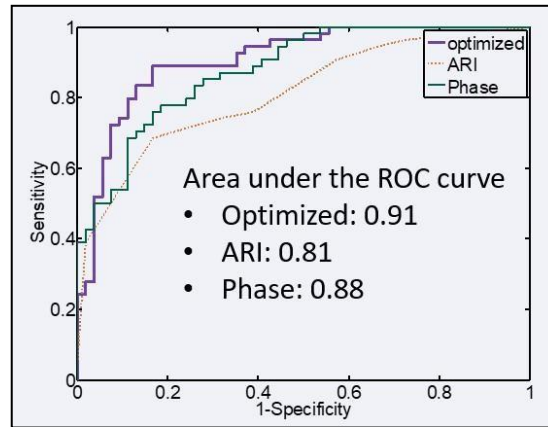


Figura 5: Curvas ROC para mediciones de Autorregulación (Simpson, 2014).

En cuanto a los resultados del presente modelo, es importante señalar que la curva de ROC implica valores positivos alcanzando un área bajo la curva de 0.91 optimizado, además de alcanzar un ARI sobre un 0.81, lo que justifica su buena implementación.

En base a lo anterior se puede apreciar que los resultados pueden estar expuestos a sesgos estadísticos por lo que se necesita una mayor precisión a la hora de analizar cada resultado, además de utilizar validaciones cruzadas para bajar el porcentaje de error resultante.

La utilización de Máquinas de vectores de soportes, permitirán identificar estos coeficientes, encontrando un patrón representativo que permita determinar en base a estos "h", el estado actual de la persona en cuestión en base al resultado del índice autorregulatorio.

4.3 PROPUESTA DE MODELO

La construcción del modelo es esencial para el cumplimiento de los objetivos, en base a la investigación realizada en el presente informe, se puede demostrar una serie de parámetros cuantitativos que involucran directamente al sistema de autorregulación cerebral, lo que permite esclarecer la posibilidad existente de desarrollar un modelo capaz de demostrar y mejorar los resultados mencionados con anterioridad.

El modelo por construir se basará en la utilización de Máquinas de Vectores de Soporte o SVM, utilizando las capacidades de clasificación de clases que permitirán mediante la ecuación de los coeficientes, determinar el ARI respecto para cada uno de los involucrados dentro del experimento.

SVM o Máquinas de vectores de soportes, mediante su capacidad computacional de encontrar patrones e identificar las relaciones en conjunto de cada muestra de datos a analizar, permite corroborar su gran efectividad en el área de la neurociencia, utilizando configuraciones establecidas por hiperparametros, además de trabajar con sus dimensionalidades dentro de un hiperplano establecido.

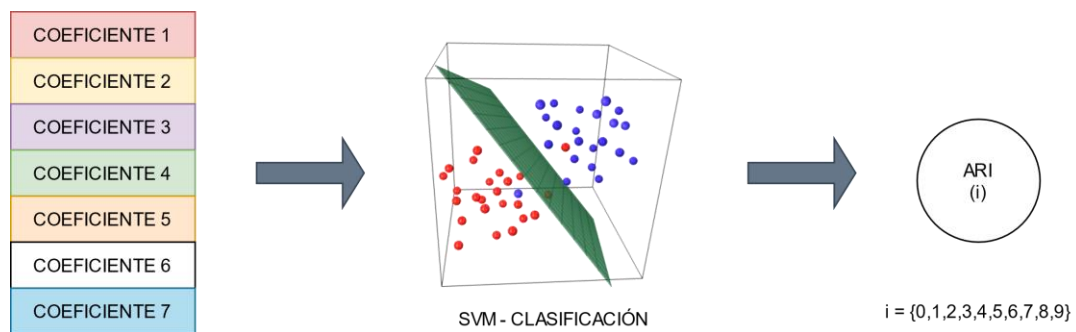


Figura 6: Modelo de estimación del Índice de autorregulación Cerebral, propuesto por alumno.

4.4 DESCRIPCIÓN DEL FLUJOGRAMA DE CÁLCULO

En base al conjunto de datos obtenidos, los parámetros de entrada corresponden a los coeficientes de la ecuación, en donde cada valor establece una relación en conjunto con el individuo en cuestión y su valor de ARI. Estos coeficientes son procesados por el SVM, para dimensionalizar sus valores y calcular cada clase correspondiente, prediciendo cada valor.

Cabe destacar que intrínsecamente la PSA y la VFSC, participan en este proceso de manera fenomenológico, creando un modelo de cálculo en donde la variable de presión corresponde a la entrada en conjunto con los determinados "**h**", siendo computarizados por las máquinas de vectores de soporte, estableciendo una generalidad y aleatoriedad en cuanto al análisis de estos datos.

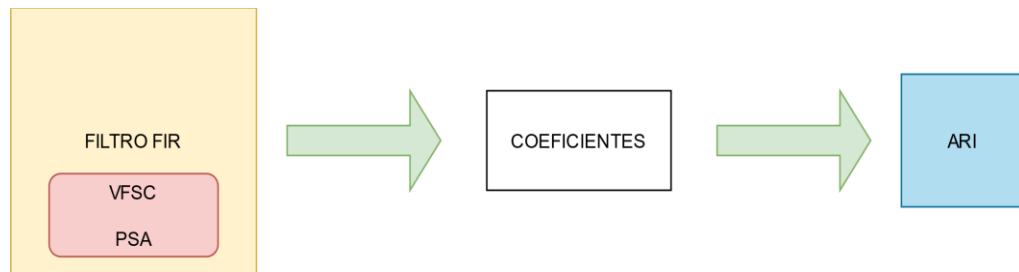


Figura 7: Estructura de procesos para la obtención del ARI.

Los coeficientes se dividen en 7 valores (H1...H7), dando como resultados los índices de autorregulación cerebral o ARI correspondientes, pudiendo tener valores que fluctúan entre el 0 y el 9.

4.5 ANÁLISIS DE DATOS

Para realizar un análisis certero es primordial tener en consideración una muestra representativa en base a elementos que conformen un comportamiento claro dependiendo del problema a tratar.

En este caso particular, el conjunto de datos fue obtenido en la Universidad de Leicester, en Inglaterra. Estas mediciones fueron aprobadas por el Comité de ética del Royal Infirmary Hospital de Leicester, los cuales apoyaron el estudio previo en todo momento para el desarrollo de cada experimento realizado.

La muestra contempla datos de 27 voluntarios adultos de edades que fluctúan entre los 18 a 55 años. Estas personas fueron expuestas a cambios de dióxido de carbono en la sangre (Correspondiente a un 5% de incremento), estableciendo estados de Hipercapnia y Normocapnia, en donde cada registro fue almacenado y procesado en relación a la velocidad del flujo sanguíneo cerebral obtenido mediante un Doppler transcraneal y la presión arterial sanguínea, las cuales fueron expuestas a un filtro FIR para la obtención de una base de datos formada por los coeficientes correspondientes a la *ecuación*₅

a. DISTRIBUCIÓN DE DATOS

La estructura de análisis está conformada mediante siete valores de $H(n)$ para cada individuo, realizando 100 muestras para cada ARI correspondiente, en donde cada paciente posee 1000 registros en consideración a los posibles índices a calcular desde el 0 al 9. Es importante señalar que la base de datos a analizar consta de 27.000 registros, en donde cada valor queda registrado y normalizado en base a las necesidades del modelo a desarrollar.

$H(1)$	$H(2)$	$H(3)$	$H(4)$	$H(5)$	$H(6)$	$H(7)$
27.000 REGISTROS DE 27 PACIENTES						

Figura 8: Distribución de datos utilizado para el análisis.

b. DATOS DE ENTRENAMIENTO

Se utilizará el conjunto de los 27.000 registros, en donde en base al estudio, se calculará un porcentaje óptimo de división para los datos, que permitan realizar un entrenamiento meticuloso, siendo capaz de representar todas las variables y generalizar el comportamiento de los coeficientes para el cálculo del ARI en específico.

Utilizando ley de Pareto o también conocida como regla de 80/20, se aproximarán las divisiones del conjunto de data set para conformar un estudio capaz de evaluar todas las posibilidades existentes dentro del modelo, y utilizar las capacidades del SVM de clasificación para obtener resultados más exactos a los ya propuestos.

c. DATOS DE PRUEBA

Corresponde a un conjunto de datos de 108 registros, en donde se tomaron 4 muestras para cada paciente en estados de Normocapnia e Hiperapnia respectivamente. Estos estados fueron medidos equitativamente, para poder representar un ARI en estado de cambios de dióxido de carbono en la sangre, para así atribuir deformaciones al modelo propuesto, y que este sea capaz de identificar si una persona posee un estado positivo o negativo en base al índice de autorregulación cerebral.

4.6 CONFIGURACIÓN DE HIPERPARAMETROS SVM

Para una correcta implementación del modelo, es importante denotar la importancia que posee una configuración adecuada de las Máquinas de Vectores de Soporte, en donde los parámetros que influyen tanto en entrenamiento como la validación corresponden a factores relevantes que se deben tener en consideración en todo momento.

Como se mencionó con anterioridad, SVM es un algoritmo de aprendizaje que posee funcionalidades en base a clasificación y regresión, por lo que la variabilidad de los datos y los hiperparámetros a escoger para su configuración, afectan de manera directa a la distribución dentro del hiperplano para cada una de las clases a predecir.

A continuación, se mencionarán los hiperparámetros a utilizar, y que permitirán explorar diversos resultados para la obtención de resultados exactos:

- **Kernel:** Basada en funciones matemáticas, permite transformar la dimensionalidad de un hiperplano en base a los datos de entrada y de salida requeridos, para su correcta clasificación.

Existen Kernel de tipo:

- Lineal.
 - No lineal.
 - Polinomial.
 - RBF (Función de base radial).
-
- **C:** Parámetro encargado de controlar el error dentro de la clasificación. A mayor valor de C se le da una mayor connotación a los valores correctos y a menos valor de C se le quita relevancia al error, maximizando el margen correspondiente.
 - **Gamma:** Permite establecer el límite de la curvatura, en cuanto a la decisión que se quiere obtener antes del modelo de entrenamiento dándole una mayor o menos complejidad a las separaciones de observaciones.

Es importante señalar que la variación de pruebas en base a la elección de una configuración jugara un rol importante dentro de la elección del modelo y de los resultados a obtener, por lo que era importante contextualizar de alguna forma, cada uno de los factores involucrados.

4.7 PROCESO DE APRENDIZAJE DE SVM

SVM posee un comportamiento dimensional, en donde el hiperplano representado mediante una ecuación va adaptándose a los datos procesados para su correcta clasificación en base a funciones que implican una modificación en la dimensión a trabajar valga la redundancia. Este aprendizaje supervisado, mediante el cómputo de las variables, permitirá establecer un porcentaje de aprendizaje que generalice el comportamiento de los coeficientes para la obtención del ARI en cada uno de los casos a estudiar.

Es importante señalar que este proceso consta de tres fases fundamentales, las cuales se dividen en entrenamiento, validación y la comprobación con el conjunto aleatorio.

1. En la fase de entrenamiento se utilizó un porcentaje aproximado al 80% de la muestra total, para dar una representatividad y así obtener un comportamiento general del conjunto, de tal manera que SVM Clasificación, sea capaz de aprender los patrones requeridos para dividir las diversas clases (ARIs), e identificar mediante los coeficientes, los valores señalados.
2. En la fase de validación se utilizó el resto de la muestra total dependiendo directamente de la variación del porcentaje de training, para demostrar en situs la efectividad del modelo propuesto, predicción en base a valores aleatorios y no dejándose influenciar por valores posiblemente memorizados que podrían inducir a una falta de aprendizaje y error por parte de la clasificación, dando falsos positivos correspondientemente.

3. Finalmente se utilizó un set aleatorio, en donde el conjunto de datos corresponde a una toma muestra independiente al data set utilizado tanto para la clasificación como para la validación, lo que ayuda a comprobar su perfecto funcionamiento y representatividad en cuanto a la identificación de un estado de autorregulación bueno o malo, dependiendo del índice autorregulatorio y el estado en Normocapnia e Hiperapnia.

5. ANÁLISIS Y RESULTADOS

Tras el desarrollo de las actividades propuestas, es indispensable realizar un análisis y una representación de los resultados dados, para así comprobar la positividad en cuanto a los objetivos esperados en comparación con el modelo realizado anteriormente. Tras la existencia de un contraste, el modelo creado en el presente trabajo permitirá establecer nuevos parámetros y valores que permitan entender de mejor manera el comportamiento del uso de nuevas herramientas o tecnologías establecidas, con el fin de aportar al ámbito científico en particular.

5.1 CALCULO DE ARI SEGÚN MODELO PROPUESTO

Para comenzar con la visualización de los resultados, es importante señalar que cada uno de los métodos utilizados para el análisis de los datos, correspondiente tanto al proceso de entrenamiento como validación, serán expuestos en base los valores obtenidos mediante un estudio meticulado, en donde se trabajó tanto con los hiperparámetros del SVM, además del conjunto de datos formado por los 7 coeficientes obtenidos de la ecuación matemática propuesta por el investigador Simpson en su trabajo investigativo(Simpson, 2014) y sus respectivos valores de cada ARI correspondiente.

Por medio del uso de las librerías de análisis de datos pertenecientes al lenguaje de programación Python, es posible establecer las mejores 7 configuraciones para denotar la capacidad de clasificación del modelo, utilizando una grilla de búsqueda de hiperparámetros mediante la herramienta de *GridSearchCV*, la cual se encarga de encontrar el valor de los mejores hiperparámetros mediante iteraciones compuestas por elementos primordiales, como la cantidad de Split que corresponde a cada una de las iteraciones junto con aleatorizaciones y estimadores para el cálculo de los mejores valores de aciertos, para un resultado más certero

a. CONFIGURACIONES DE MODELO PROPUESTO

Como se mencionó anteriormente, las configuraciones se propusieron en base a un estudio de hiperparametros en donde cada una de las variables fueron expuestas e interpretadas mediante un análisis realizado por el uso de herramientas estadísticas, en este caso en particular la librería *GridSearchCV* permite aportar a la elección de estos valores, con la intención de obtener resultados acordes a los objetivos planteados.

En consecuencia, se presentan los mejores resultados establecidos, para la etapa de entrenamiento y validación del modelo en base a los 7 coeficientes otorgados y el valor de ARI para cada uno

CONFIGURACIÓN	KERNEL	PARAMETRO C	PARAMETRO GAMMA
CONFIGURACIÓN 1	<i>POLY</i>	<i>1000</i>	<i>'AUTO'</i>
CONFIGURACIÓN 2	<i>RBF</i>	<i>600</i>	<i>'AUTO'</i>
CONFIGURACIÓN 3	<i>RBF</i>	<i>10</i>	<i>1</i>
CONFIGURACIÓN 4	<i>RBF</i>	<i>500</i>	<i>0.5</i>
CONFIGURACIÓN 5	<i>RBF</i>	<i>2000</i>	<i>2</i>
CONFIGURACIÓN 6	<i>RBF</i>	<i>8000</i>	<i>'AUTO'</i>
CONFIGURACIÓN 7	<i>RBF</i>	<i>3000</i>	<i>1</i>

Tabla 2: Mejores configuraciones para modelo propuesto.

A continuación, se exponen las predicciones realizadas en base a las curvas de ROC para cada ARI obtenido, en donde cada clase hace referencia a los valores de ARI desde el 0 al 9, junto con su matriz de confusión para denotar la capacidad de clasificación que posee el modelo.

CONFIGURACIÓN 1

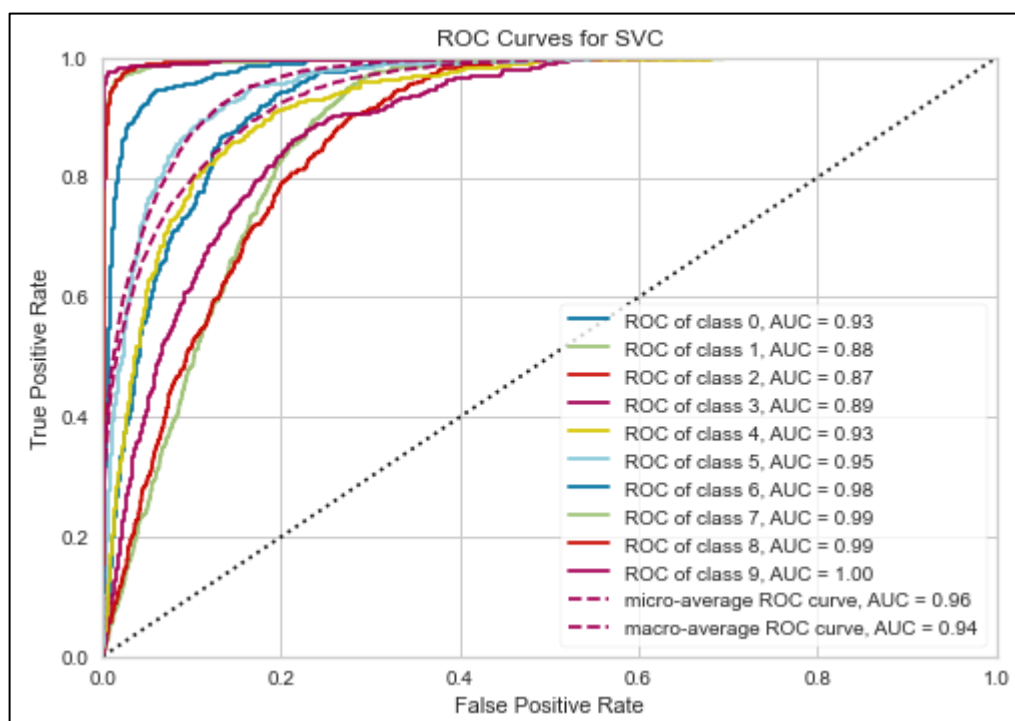


Figura 9: Curva de ROC perteneciente a configuración 1.

OBSERVACIÓN	PREDICCIÓN									
	0	1	2	3	4	5	6	7	8	9
0	539	165	88	31	5	0	1	0	0	0
1	322	242	160	77	11	0	0	0	0	0
2	154	156	259	177	35	3	2	0	0	0
3	34	49	145	410	129	20	3	1	0	0
4	9	7	17	129	480	133	22	0	0	0
5	2	2	4	33	145	540	87	4	2	1
6	2	0	0	11	12	75	681	39	4	1
7	0	0	0	1	1	0	18	740	27	0
8	0	0	1	1	1	1	1	24	760	31
9	0	0	0	0	0	0	1	0	25	807

Tabla 3: Matriz de confusión en base a configuración 1.

Se estima un área bajo la curva de 0.94 presentando un comportamiento razonable, lo cual establece un buen resultado para el proceso de entrenamiento y validación. A pesar de eso, es importante señalar que el valor de aciertos es de un 67% haciendo referencia a la matriz de confusión dada.

CONFIGURACIÓN 2

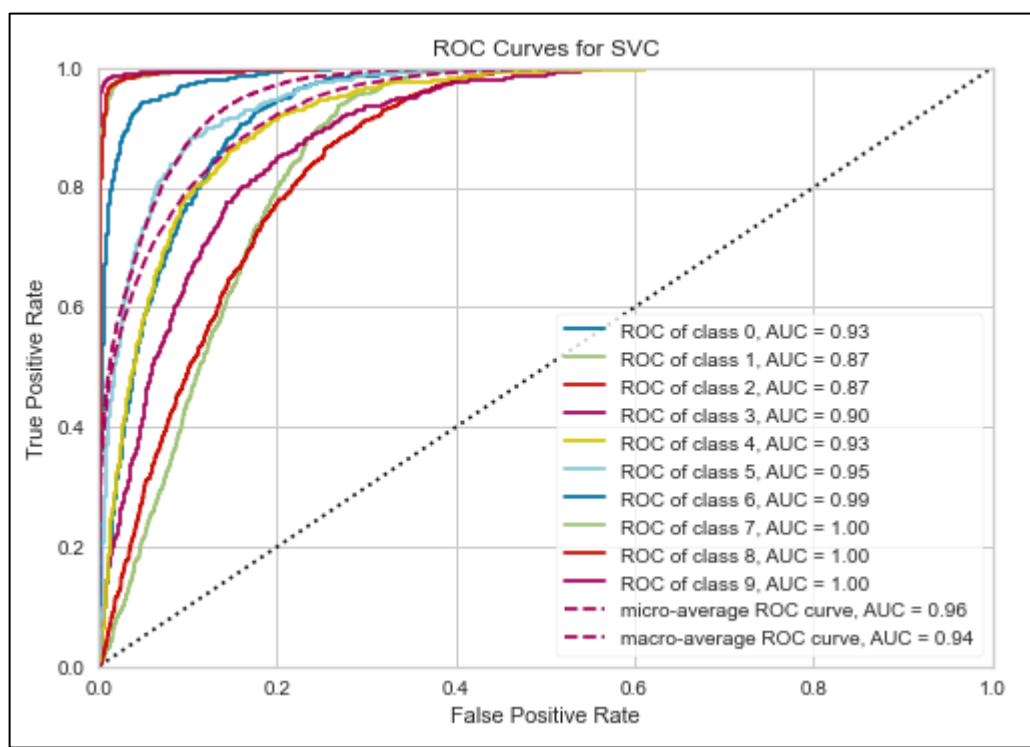


Figura 10: Curva de ROC perteneciente a configuración 2.

OBSERVACIÓN	PREDICCIÓN									
	0	1	2	3	4	5	6	7	8	9
0	425	161	82	25	8	0	0	0	0	0
1	228	200	154	76	16	1	0	0	0	0
2	102	154	223	159	41	5	1	0	0	0
3	20	42	119	334	133	19	3	0	0	0
4	4	7	15	107	390	162	12	0	0	0
5	0	1	0	19	110	457	72	4	1	0
6	0	0	0	3	8	73	547	27	2	0
7	0	0	0	2	1	0	22	630	23	0
8	0	0	0	0	0	0	2	19	573	18
9	0	0	0	1	0	0	0	1	27	679

Tabla 4: Matriz de confusión en base a configuración 2.

Se estima un área bajo la curva de 0.94 presentando un comportamiento razonable, lo cual establece un buen resultado para el proceso de entrenamiento y validación. A pesar de eso, es importante señalar que el valor de aciertos es de un 63.2% haciendo referencia a la matriz de confusión dada.

CONFIGURACIÓN 3

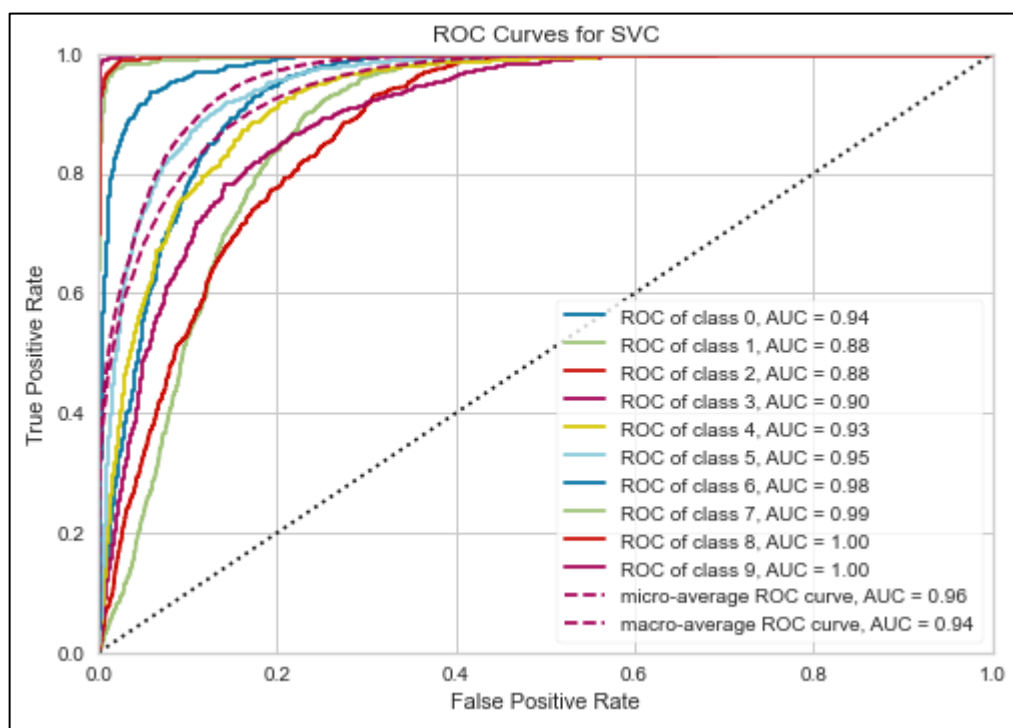


Figura 11: Curva de ROC perteneciente a configuración 3.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	428	149	69	28	3	0	0	1	0	0
	1	268	196	115	62	12	1	4	0	0	0
	2	108	169	217	148	33	4	2	0	0	0
	3	30	41	97	358	118	18	5	1	1	1
	4	2	4	10	119	396	138	29	2	0	0
	5	0	1	2	25	110	476	76	3	1	1
	6	0	0	0	5	10	69	578	22	2	0
	7	0	0	0	1	0	1	20	617	28	2
	8	0	0	1	0	0	0	0	16	628	16
	9	0	0	0	0	0	0	0	0	10	642

Tabla 5: Matriz de confusión en base a configuración 3.

Se estima un área bajo la curva de 0.94 presentando un comportamiento razonable, lo cual establece un buen resultado para el proceso de entrenamiento y validación. A pesar de eso, es importante señalar que el valor de aciertos es de un 60% haciendo referencia a la matriz de confusión dada.

CONFIGURACIÓN 4

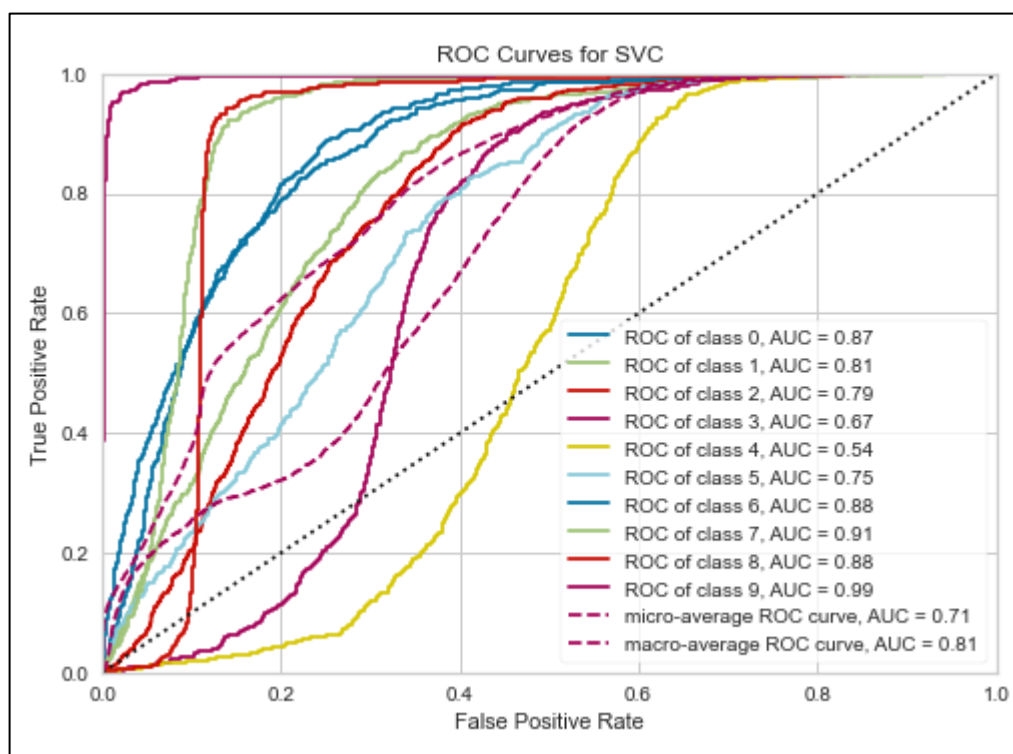


Figura 12: Curva de ROC perteneciente a configuración 4.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	0	0	0	141	264	66	60	4	0	0
	1	0	0	0	72	211	161	101	2	0	0
	2	0	0	0	45	127	194	140	5	0	0
	3	0	0	0	18	24	212	305	10	0	0
	4	0	0	0	0	4	133	375	7	0	0
	5	0	0	0	0	2	65	484	17	0	0
	6	0	0	0	0	1	13	454	66	0	0
	7	0	0	0	0	0	2	39	491	0	0
	8	0	0	0	0	2	0	19	525	5	0
	9	0	0	0	0	9	0	21	236	267	1

Tabla 6: Matriz de confusión en base a configuración 4.

Se estima un área bajo la curva de 0.81 presentando un comportamiento negativo, lo cual establece un mal resultado para el proceso de entrenamiento y validación. Es importante señalar que el valor de aciertos es de un 30% haciendo referencia a la matriz de confusión dada.

CONFIGURACIÓN 5

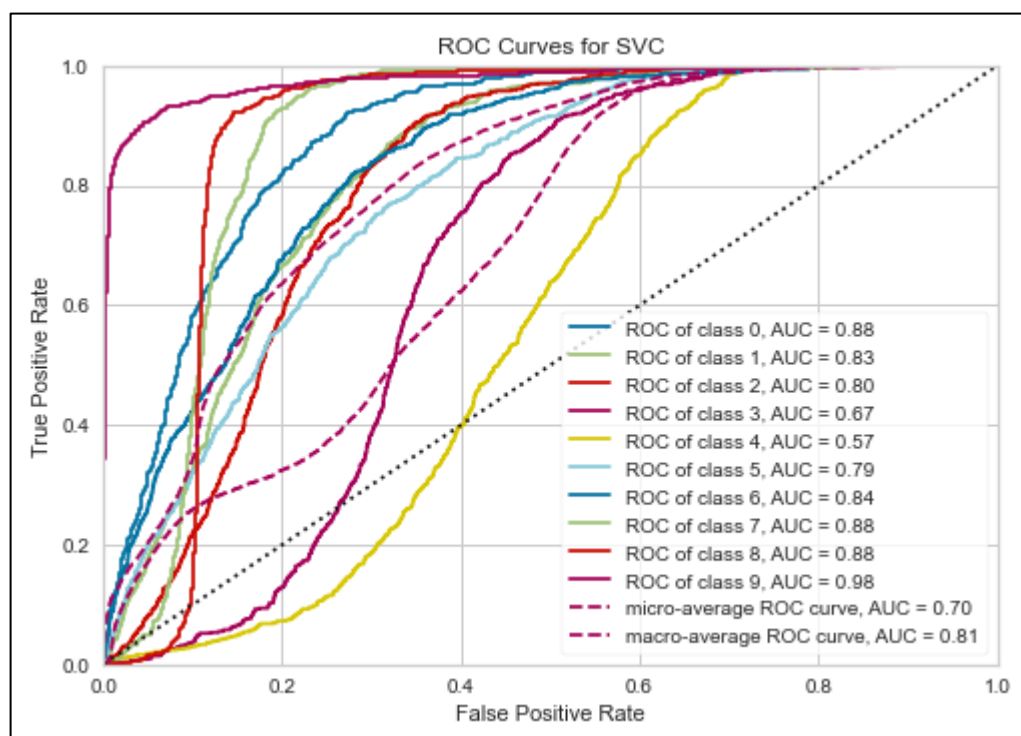


Figura 13: Curva de ROC perteneciente a configuración 5.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	0	2	0	156	372	82	60	7	0	0
	1	0	3	0	97	305	174	94	8	0	0
	2	0	1	0	67	232	249	82	9	0	0
	3	0	0	0	25	69	413	162	22	0	0
	4	0	0	0	4	13	345	286	12	0	0
	5	0	0	0	0	4	214	464	31	0	0
	6	0	0	0	0	6	195	362	115	0	0
	7	0	0	0	0	5	14	34	595	0	0
	8	0	0	0	0	5	0	16	684	4	0
	9	0	0	0	0	20	0	2	430	199	0

Tabla 7: Matriz de confusión en base a configuración 5.

Se estima un área bajo la curva de 0.81 presentando un comportamiento negativo, lo cual establece un mal resultado para el proceso de entrenamiento y validación. Es importante señalar que el valor de aciertos es de un 36% haciendo referencia a la matriz de confusión dada.

CONFIGURACIÓN 6

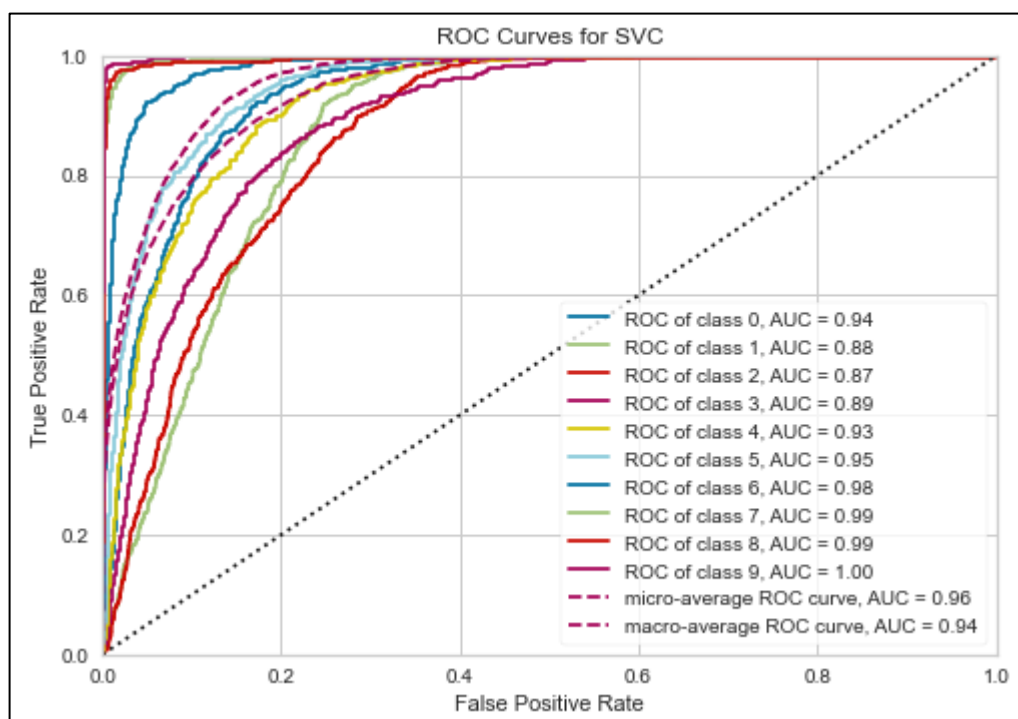


Figura 14: Curva de ROC perteneciente a configuración 6.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	333	84	83	15	4	1	0	0	0	0
	1	201	121	147	61	11	2	2	0	0	1
	2	94	76	237	113	30	3	0	1	0	0
	3	21	15	119	266	110	29	3	1	0	0
	4	1	8	15	68	306	123	15	1	0	0
	5	0	1	0	20	93	360	70	1	0	0
	6	0	1	0	3	7	63	417	15	0	1
	7	0	0	0	1	0	0	27	483	21	1
	8	0	0	0	0	0	0	1	15	522	16
	9	0	0	0	0	0	0	1	0	14	525

Tabla 8: Matriz de confusión en base a configuración 6.

Se estima un área bajo la curva de 0.94 presentando un comportamiento positivo, lo cual establece un buen resultado para el proceso de entrenamiento y validación. Es importante señalar que el valor de aciertos es de un 78% haciendo referencia a la matriz de confusión dada, estableciendo la mejor configuración representada.

CONFIGURACIÓN 7

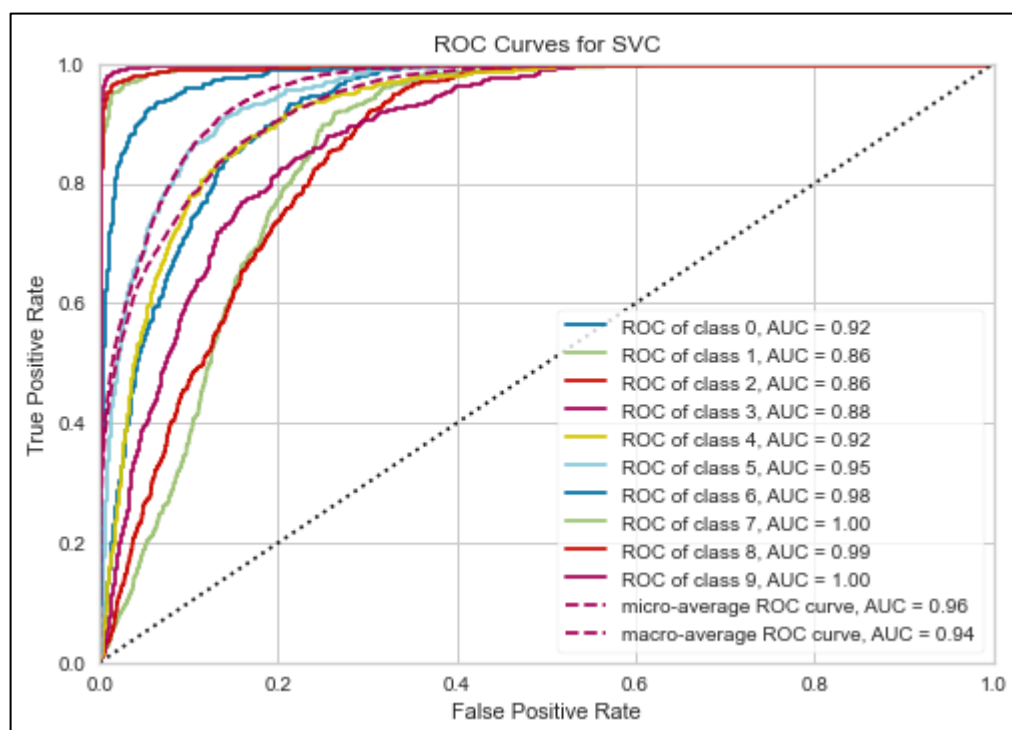


Figura 15: Curva de ROC perteneciente a configuración 7.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	291	159	77	23	2	0	0	0	0	0
	1	155	164	144	54	11	1	0	0	0	0
	2	95	124	207	121	16	6	0	0	0	0
	3	26	45	106	237	92	15	1	0	1	0
	4	7	7	11	101	306	121	13	0	0	0
	5	2	3	0	21	93	355	53	3	0	0
	6	0	3	2	0	11	55	426	27	0	0
	7	0	0	1	1	0	3	20	483	17	0
	8	0	0	0	0	0	2	2	12	522	16
	9	0	0	0	1	0	0	0	2	18	497

Tabla 9: Matriz de confusión en base a configuración 7.

Se estima un área bajo la curva de 0.94 presentando un comportamiento positivo, lo cual establece un buen resultado para el proceso de entrenamiento y validación. Es importante señalar que el valor de aciertos es de un 72% haciendo referencia a la matriz de confusión dada.

b. SINTESIS DE RESULTADOS

Un análisis exhaustivo determina el impacto que cualquier modelo, tanto predictivo como descriptivo poseerá a la hora de darle un uso específico, dentro del mundo real.

En el presente trabajo, cada una de las configuraciones fueron estudiadas a cabalidad con el fin de exponer las mejores opciones dentro de un conjunto de interacciones establecidas y procesadas en un tiempo determinado.

En base, a los resultados de entrenamiento y validación se puede señalar que la curva de ROC juega un papel determinante en este trabajo clínico, por lo que el valor de aciertos que posea el modelo tiene una relevancia intermedia al basarse en el impacto que poseerá este trabajo en cada una de las personas, pudiendo establecer de tal manera la especificidad y la sensibilidad del modelo valga la redundancia. Es por esta razón, que encontrar un punto intermedio para escoger el mejor modelo y predecir los variables en el conjunto de datos oficiales de validación entregados por la Universidad de Leicester, se hace tan relevante en cuanto a las variantes en el estado de Normocapnia a Hiperapnia para cada uno de los pacientes.

Las configuraciones 1 y 2 presentan un comportamiento constante durante el tiempo, en donde cada ARI clasificado, expresa resultados positivos variando las curvas de ROC para cada clase en un porcentaje mínimo. Sin embargo, el valor de aciertos que se logra obtener es muy bajo en comparación con modelos establecidos, por lo que no se denota una mejora considerable para escoger alguno de estos modelos como un elemento fundamental. También hay que recalcar, que los hiperparametros expresan un comportamiento similar, al poseer un parámetro C bajo en donde la distancia entre el hiperplano y los vectores de soportes no logran maximizar de buena manera el resultado dado, no implica una mayor diferencia la elección de la función de Kernel, ya que en ambos casos se utilizaron ecuaciones diferentes para realizar la clasificación correspondiente, comprendiendo lo mismo para el valor de Gamma donde la

flexibilidad de la función matemática no tendrá efecto, dando como resultados valores parecidos.

Al igual que el análisis anterior, la configuración 3 se comporta de la misma manera, en base a un valor C extremadamente bajo, no se logra realizar la clasificación de manera satisfactoria. A pesar de eso el impacto del modelo es positivo en cuanto a las curvas de ROC obtenidas. Sin embargo, las siguientes configuraciones 4 y 5, representan los peores modelos al obtener un área bajo la curva no satisfactoria con un valor de 0.81 en cada uno de los casos, dando como valor de aciertos un 30% y un 36% respectivamente lo que permite establecer una notoria diferencia con los otros modelos planteados.

La configuración 6 presenta los mejores resultados obtenidos con un área bajo la curva de 0.94 y con un valor de acierto de un 78%. Haciendo hincapié en esta relación de encontrar un valor intermedio entre los dos métodos de medición, se puede señalar que este modelo representa la mejor opción hasta el momento con un valor de C elevado (8000) y con una función gaussiana acorde para realizar las parametrizaciones de cada una de las clases correspondientes. Finalmente, la configuración 7, al igual que las primeras configuraciones, demuestra un comportamiento positivo, sin embargo, no logra alcanzar a superar al caso anterior con un valor de área bajo la curva de un 94% y un valor de aciertos de 72%.

Con relación a los análisis establecidos, hay que denotar una relación directa entre el análisis de cada ARI, en donde las variaciones según las curvas en todo lo expuesto, fueron positivas y ayudan a establecer un correcto funcionamiento del modelo como tal, lo que con lleva a pensar que el mejoramiento del modelo dependerá exclusivamente por una forma más certera de obtener los datos respectivos de cada paciente.

5.2 EVALUACIÓN DE RESULTADOS FINALES

Tras el proceso de entrenamiento y validación con datos generales, en donde cada uno de los coeficientes formo parte del modelo, para dar lugar al resultado final correspondiente al índice de autorregulación cerebral, es importante realizar una evaluación final en donde se compruebe que el modelo propuesto es un método confiable para determinar el ARI de una persona en diversas situaciones o contextos.

En relación a lo anterior, se procede a utilizar una base de datos para la validación final, en donde cada uno de los 27 pacientes fueron expuestos a diversos cambios en su organismo para la medición del estado autorregulatorio cerebral, pasando de estados de Normocapnia a Hipercapnia, en base a la inducción de CO_2 , con la intención de obtener resultados más certeros, agregando ruido al conjunto de datos, y permitiendo verificar si los modelos se adaptan a estos cambios repentinos, siendo capaces de identificar el paso de un estado a otro y además corroborar la situación del organismo en cuestión, basándose en el ARI correspondiente.

A continuación, se visualizan los índices autorregulatorios obtenidos mediante el estudio clínico realizado, haciendo énfasis en los cambios de Normocapnia e Hipercapnia que cada paciente posee, con la intención de visualizar la diferencia de un estado a otro.

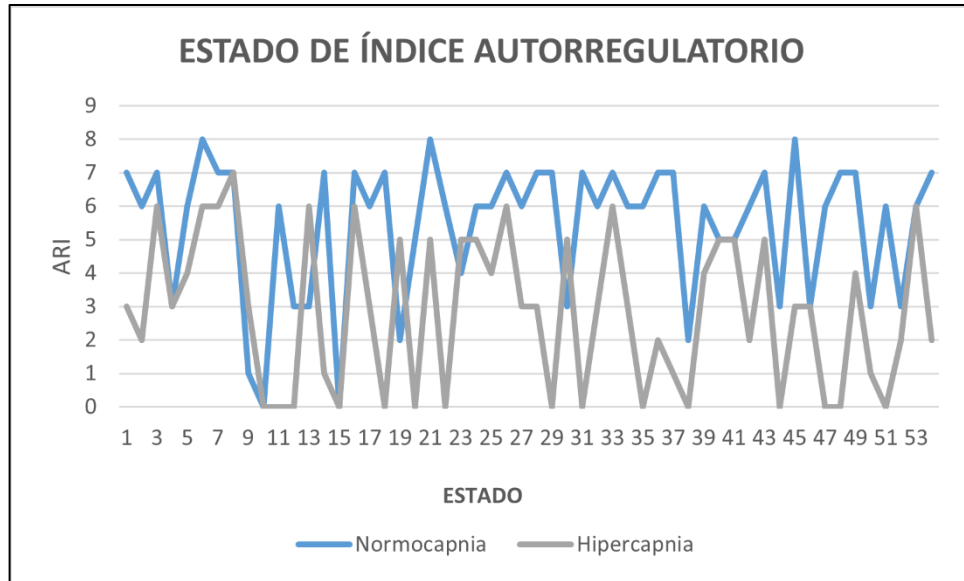


Figura 17: Grafico de datos para validación en estado de Normocapnia e Hiper-capnia.

Es posible visualizar un claro comportamiento de un estado a otro, lo que con lleva a deducir que las posibles variaciones que sufre cada paciente son directamente relacionadas con el cambio de oxígeno, el cual fue inducido por los profesionales a cargo del estudio realizado.

En consecuencia, para verificar el modelo propuesto y obtener resultados acordes a los esperados es que se realiza la elección de los 3 modelos con resultados positivos en base a elementos de análisis, los cuales corresponden al mejor modelo obtenido en cuanto a correlación, error cuadrático medio y curva de ROC respectivamente. Cada uno de estos, será aplicado al conjunto de datos final, para su análisis y validación correspondiente.

MEJOR MODELO EN CUANTO A CORRELACIÓN

En cuanto a los modelos aplicados a los datos de validación, es importante señalar que la relación directa que posee cada cambio de estado de Normocapnia a Hiperapnia se ve reflejado en los resultados establecidos. La correlación de las diversas pruebas fluctuó entre 0.55 a 0.79, dando como resultado final un modelo capaz de poseer una correlación alta y positiva cercano a 0.80, lo que compromete que cada uno de los estados son capaces de clasificarse de una manera correcta.

A continuación, se visualiza lo explicado.

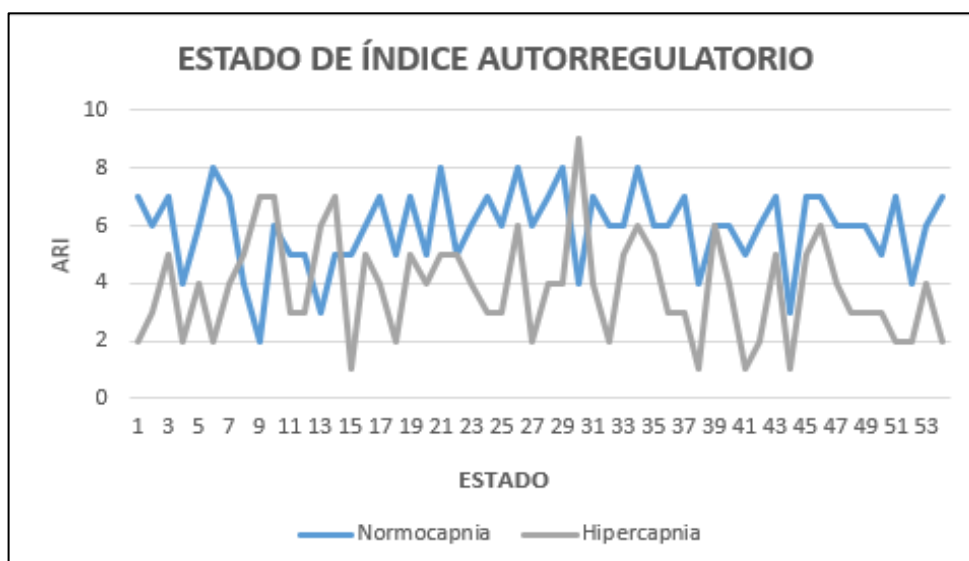


Figura 18: Grafico de predicción en base a correlación en estado de Normocapnia e Hiperapnia.

MEJOR MODELO EN CUANTO A ERROR CUADRÁTICO MEDIO

En base al análisis del error cuadrático medio, se pudo demostrar que el mejor modelo en base a este método fue capaz de obtener un error cuadrático medio acorde a lo esperado, correspondiente a un 0.80, lo que establece un resultado correcto, aun que de igual manera es importante señalar que es posible seguir mejorando en base a este método estadístico.

A continuación, se visualiza lo explicado.

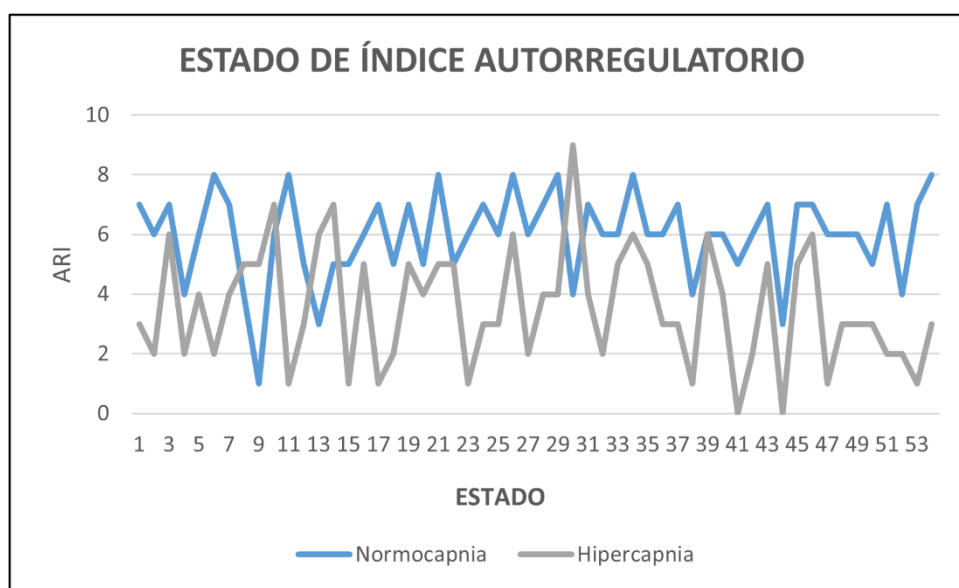


Figura 19: Gráfico de predicción en base a error cuadrático medio en estado de Normocapnia e Hiperapnia.

MEJOR MODELO EN CUANTO A CURVA DE ROC

Siendo el análisis por comparar según modelos propuestos, esta medición permite establecer un análisis exhaustivo para cada ARI analizado, con la intención de obtener un modelo capaz de predecir valores en cuanto a una alta sensibilidad y especificidad, por esta razón que el modelo más certero, dio un área bajo la curva de 0.94, expresando casi un 94% de aciertos correspondientemente.

A continuación, se visualiza lo explicado.

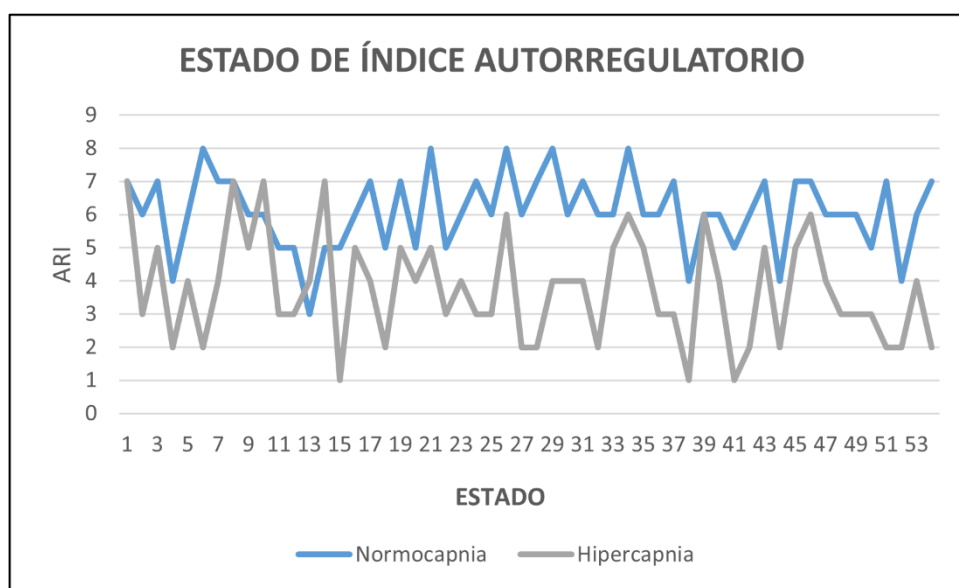


Figura 20: Grafico de predicción en base a curva ROC en estado de Normocapnia e Hiperapnia.

5.3 COMPARACIÓN DE MODELOS EXISTENTES

Es importante tener noción de las diferencias que pueden existir dentro de una investigación, y más cuando pequeños resultados hacen la diferencia. En estudios clínicos, la sensibilidad y especificidad que posea el modelo, permiten acotar la brecha de error por parte de resultados dados. Establecer un cierto estado en un ser humano, puede conllevar tanto a aspectos positivos como negativos, por lo que hacer hincapié en las pequeñas diferencias dentro de un estudio clínico, puede relacionarse directamente con el hecho de salvar una vida o peor aún, perderla.

Dicho lo anterior, la comparación del presente modelo se basa en el estudio investigativo realizado por Simpson (*Simpson, 2014*), en donde sus resultados fueron explicitados de forma meticulosa. El estudio del área bajo la curva (ROC), permitió establecer una asociación en base a los 7 coeficientes determinados por el investigador, lo que contribuyó a un posible método menos invasivo para obtener el índice de autorregulación cerebral para cada ser humano.

El estudio realizado, permitió demostrar valores de ROC para una cantidad limitada de pacientes, correspondiente a un área bajo la curva de 0.91 en donde se logra apreciar la gran efectividad del modelo, en donde distingue cada uno de los estados de los pacientes en base a la inducción de oxígeno, para complementar el modelo hecho y así dar una capacidad de adaptación al estudio, logrando una efectividad mayor.

Siguiendo con lo anterior, fue posible obtener valores altos, tras un exhaustivo análisis, pudiendo obtener un valor de 0.94 de área bajo la curva, mejor que el modelo propuesto por Simpson, lo que con lleva a pensar que el uso de las máquinas de vectores de soporte, aportan en el área clínico.

Sin embargo, es importante destacar que este modelo, está expuesto a sesgos, por lo que hay que tener en consideración esta variable, para determinar si su implementación sería acorde o no a lo esperado.

6. CONCLUSIÓN Y RECOMENDACIONES

6.1 CONCLUSIÓN

Para finalizar, es importante hacer énfasis en la hipótesis y en cada uno de los objetivos planteados, con el fin de corroborar la efectividad y la correcta aplicación de metodologías y técnicas implementadas por el alumno en cuestión.

En consecuencia, se logra utilizar el algoritmo de Support Vector Machine (SVM) para la creación de un modelo de estimación, para la obtención del índice de autorregulación cerebral, obteniendo resultados acordes a los esperados, en base a una serie de procesos de análisis de datos, utilizando un conjunto de variables que impactaron directamente en el desarrollo de este modelo predictivo, validando la hipótesis planteada con anterioridad.

Entender en que consiste cada uno de los conceptos que abarcaban este trabajo, permite reconocer de manera exhaustiva el real funcionamiento que posee el sistema de autorregulación cerebral junto con el índice de autorregulación, en cuanto a la capacidad y al funcionamiento que poseen dentro del organismo, estableciendo las directrices en base a la autorregulación sanguínea que el cerebro posee en momentos determinados, permitiendo conocer su estado real para así poder prevenir enfermedades neurodegenerativas.

Para la creación del modelo, fue importante entender que diversos factores juegan un papel primordial para su desarrollo, los coeficientes establecidos por el investigador Simpson compromete el gran porcentaje de variables que conducen a obtener resultados de confianza en base al estudio del ARI, en diversos estados dentro del organismo de cada paciente estudiado. Por lo que, detectar los cambios que se producen en base a los estados de Normocapnia hacia Hipercapnia facilitaron entender de mejor manera, como afecta a la autorregulación teniendo en consideración la inducción de elementos inesperados dentro de la sangre de cada una de las personas.

En cuanto a la creación del modelo, fue necesario plantear ciertas comparaciones, permitiendo tener una visión más clara, para la justificación de los resultados, con el objetivo principal de conseguir un modelo que fuera capaz de igualar o incluso superar, los resultados expuestos por los diversos investigadores del estudio ya mencionado anteriormente. Por esta razón, es que se logró cumplir con este objetivo, en base a una serie de pruebas dentro del modelo, lo que permitió esclarecer a ciencia cierta, que el presente modelo posee una capacidad considerable con relación a los modelos planteados.

Por lo tanto, se logró definir un modelo único de aprendizaje para la obtención del índice de autorregulación cerebral, en base al conjunto de datos analizados, haciendo referencia a los resultados expuestos y dejando en claro, que a pesar de ser un desarrollo propenso a diversos sesgos de parte del estudiante y de los propios datos, se pudo comprobar su eficacia en relación con los valores dados por el modelo final.

La interpretación del modelo de aprendizaje desarrollado, expuso resultados positivos, logrando superar el área bajo la curva establecida por Simpson en su investigación (*Simpson, 2014*), dando como resultado una curva de ROC de 0.94 en comparación al valor de 0.91 expuesto por el investigador, haciendo énfasis en los diversos estudios en donde se comprobó mediante el uso de correlación y error cuadrático medio, que sus valores correspondían a 0.79 y 0.80 respectivamente, justificando el valor de aciertos correspondiente a un 94%, en base a la especificidad y sensibilidad del modelo planteado, por lo que resulta relevante tener en consideración el uso del algoritmo Support Vector Machine para su implementación en el campo clínico. Es importante, tener en consideración la configuración de cada uno de los hiperparámetros, lo que facilitó la obtención de un modelo más certero, en cuanto al uso de herramientas y librerías pertenecientes al lenguaje de programación establecido, para su correcto funcionamiento.

Lo anterior puede justificarse, haciendo referencia a la flexibilidad que permite otorgar SVM en este contexto, pudiendo establecer una clasificación más certera en base a los vectores de soporte en donde la función de Kernel y el hiperparametro C, ofrecían un comportamiento que favorecía al valor de precisión dado. Según la configuración establecida, C hace referencia a un valor alto, por lo que la maximización en cada una de las clasificaciones de los distintos índices autorregulatorios permitió dar una mayor connotación a los valores correctos, controlando de mejor forma los errores, al igual que el uso del Kernel de base radial gaussiana que permite trabajar con datos desconocidos como era el caso del presente desarrollo.

Finalmente, se señala que cada uno de los objetivos planteados fueron cumplidos a cabalidad, interpretando cada uno de los resultados de manera exhaustiva y facilitando su implementación para aplicaciones futuras o en la continuación de algún trabajo investigativo, demostrando la efectividad de Support Vector Machine para la implementación dentro de un modelo predictivo para un caso clínico respectivamente.

6.2 RECOMENDACIONES

Existen diversos factores que muchas veces no son considerados en desarrollos o investigaciones científicas, por lo que denotar cada variable que pueda ser implementada a futuro, permite abarcar nuevas aristas que beneficiaran al campo en el cual se desempeña cada trabajo o investigador correspondiente.

Uno de los objetivos a futuro debiese apuntar a la utilización de diversas herramientas de análisis de datos que permitiesen comprobar la efectividad de las variables utilizadas, en este caso los coeficientes que fueron variables indispensables para conseguir el índice de autorregulación cerebral, puede ser abarcado de una manera más minuciosa estableciendo tareas de preprocesamiento, que permitieran conseguir mejores resultados haciendo hincapié en la importancia que poseen los procesos previos a realizar un entrenamiento o validación de un modelo.

También, el uso de diversos algoritmos dentro de Support Vector Machine, como el de regresión establecen una idea clara a futuro de realizar una comparación, para conocer si efectivamente la clasificación se comporta de mejor manera que un modelo de regresión, al tener que percibir ciertos datos numéricos que favorecen a la comprensión de la problemática planteada. Siguiendo con lo anterior, continuar con el uso de diversas configuraciones para establecer nuevos hiperparametros y abarcar ciertos elementos, no considerados en el presente trabajo.

Tomando en cuenta los coeficientes utilizados, también se puede señalar que es posible considerar nuevas variables para el presente estudio, como la edad de los pacientes o el tipo de enfermedades neurodegenerativas que puedan poseer las personas evaluadas, al igual que considerar incluir a pacientes con problemas asociados a deficiencias que permitan generalizar aún más el modelo.

Finalmente, en relación con los datos trabajados, es importante señalar que estos poseían ciertos errores en cuanto a la toma de muestra, por lo que afectaban tanto al entrenamiento de los datos como a la predicción de los resultados. Es por esta razón, que es importante tener en consideración una toma de muestra más certera que involucre valores verídicos y no posean errores inesperados, que puedan afectar al modelo generando sesgos en los resultados finales.

BIBLIOGRAFÍA

1. *Chacón, M. Jara JL, Panerai RB (2014). A New Model-Free Index of Dynamic Cerebral Blood Flow Autoregulation.*
2. *Czosnyka, Smielewski, Lavinio, Pickard, Panerai, Ronney (2008). An Assessment of Dynamic Autoregulation from Spontaneous Fluctuations of Cerebral Blood Flow Velocity: A Comparison of Two Models, Index of Autoregulation and Mean Flow Index.*
3. *Aaslid R, Lindegaard KF, Sorteberg W, Nornes H. (1989). Cerebral autoregulation dynamics in humans. Stroke.*
4. *Chacón, M. Jara JL. N. Varas, N. and Panerai, RB (2015). Analysis of the influence of systemic and intracranial pressure in patients with severe head injury using linear and non-linear models.*
5. *D. Nikolić, A. A. Birch, R. B. Panerai and D. M. Simpson. (2015). Assessing cerebral blood flow control from variability in blood pressure and arterial CO2 levels.*
6. *Sebastian Schlag, Matthias Schmitt, Christian Schulz. (2020). Faster Support Vector Machines.*
7. *Michael Banf. (2019). "Learning Theory and Support Vector Machines."*
8. *Penny Eames, Penelope J. Eames, Melanie J. Blake, Ronney B. Panerai, John F. Potter (2003). Cerebral autoregulation indices are unimpaired by hypertension in middle aged and older people.*

9. *Kontos, H. A. (1989). Validity of cerebral arterial blood flow calculations from velocity measurements.*

10. *Panerai, R.B. (2008). Cerebral Autoregulation: From Models to Clinical Applications.*

11. *R. B. Panerai, A. W. R. Kelsall, J. M. Rennie and D. H. Evans (1996) Analysis of cerebral blood flow autoregulation in neonates in IEEE Transactions on Biomedical Engineering.*

12. *Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2010). Metodología de la investigación.*

13. *Esteban Ortiz-Prado, Alfredo Banderas León, Luis Unigarro y Pablo Santillán. (2018). Oxigenación y Flujo Sanguíneo Cerebral, Revisión Comprensiva de la Literatura.*

14. *Fernández-Sainz, A., & Llaugel, F. (2011). ¿Bancos con Problemas? Un Sistema de Alerta Temprana para la Prevención de Crisis Bancarias.*

15. *Chacon, Max & Contreras, R. & Acuña, Gonzalo & Panerai, Ronney. (2008). Getting a direct index of self-regulation for cerebral blood flow changes induced pressure.*

16. *Chacón M, Nuñez N, Henríquez C, Panerai RB. (2008). Unconstrained parameter estimation for assessment of dynamic cerebral autoregulation.*

17. *Liu, J., Simpson, D. M., & Allen, R. (2005). High spontaneous fluctuation in arterial blood pressure improves the assessment of cerebral autoregulation.*

18. Simpson, D., Henriques Berroeta, C., Katsogridakis, E., & Panerai, R. (2014). *Quantifying autoregulation from estimated model parameters: an optimization approach.*
19. Mahdi, A., Nikolic, D., Birch, A. A., Olufsen, M. S., Panerai, R. B., Simpson, D. M., & Payne, S. J. (2017). *Increased blood pressure variability upon standing up improves reproducibility of cerebral autoregulation indices.*
20. Kumar, R., Indianan, A. (2011). *Receiver operating characteristic (ROC) curve for medical researchers.*
21. Rakotomamonjy, A. (2004). *Optimizing Area Under Roc Curve with SVMs. ROCAI.*
22. Yubo Li, Lei Wang, Liang Ju, Haoyue Deng, Zhenzhu Zhang, Zhiguo Hou, Jiabin Xie, Yuming Wang, Yanjun Zhang. (2016). *A Systematic Strategy for Screening and Application of Specific Biomarkers in Hepatotoxicity Using Metabolomics Combined with ROC Curves and SVMs.*
23. Galindo Huerta, Adrián (2018). *Algoritmos de clasificación para data sets desequilibrados: análisis y comparativa.*
24. A. Mathur and G. M. Foody. (2008). *Multiclass and Binary SVM Classification: Implications for Training and Classification Users.*
25. G. F. Smits and E. M. Jordaán. (2002). *Improved SVM regression using mixtures of kernels.*

ANEXO I

A continuación, se muestran las diferentes curvas de ROC analizadas.

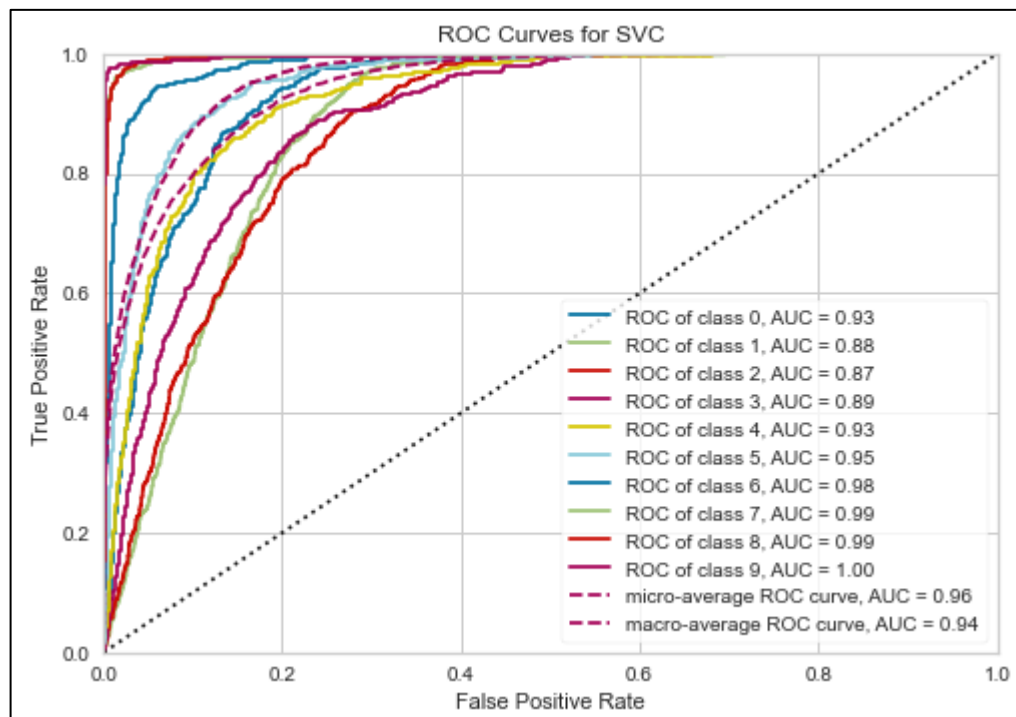


Figura 9: Curva de ROC perteneciente a configuración 1.

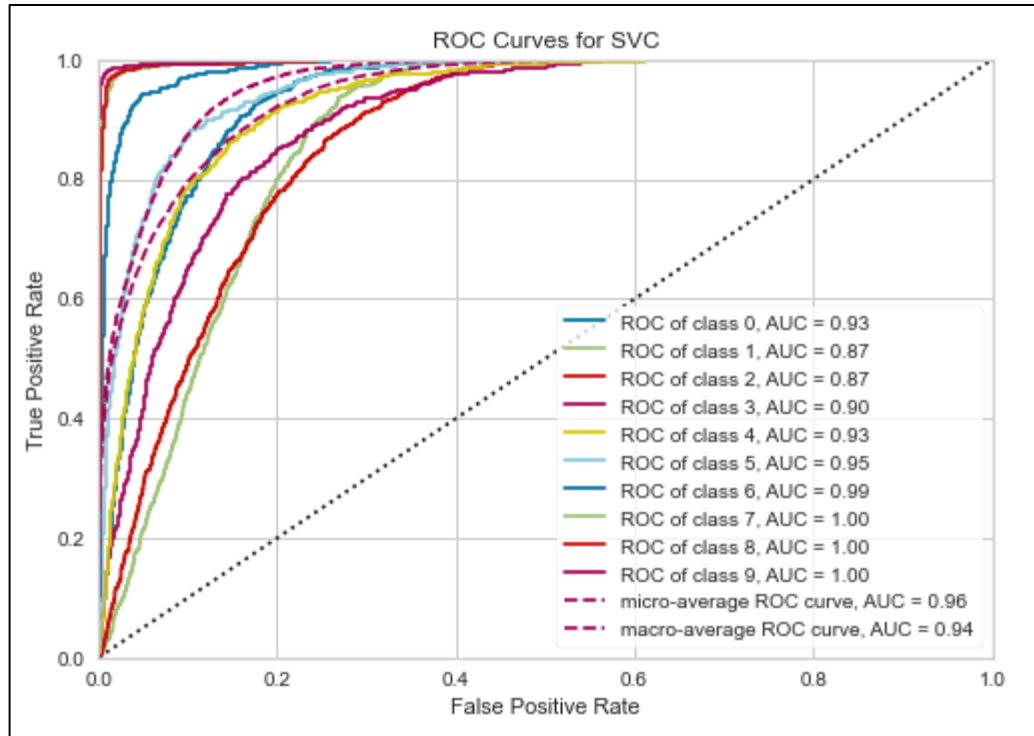


Figura 10: Curva de ROC perteneciente a configuración 2.

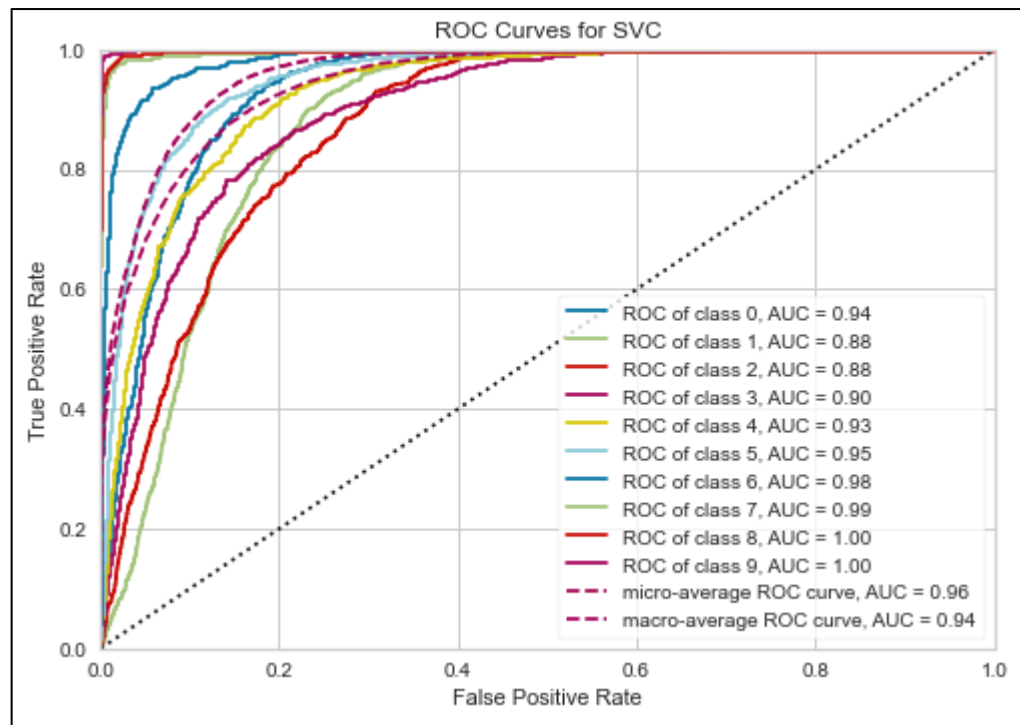


Figura 11: Curva de ROC perteneciente a configuración 3.

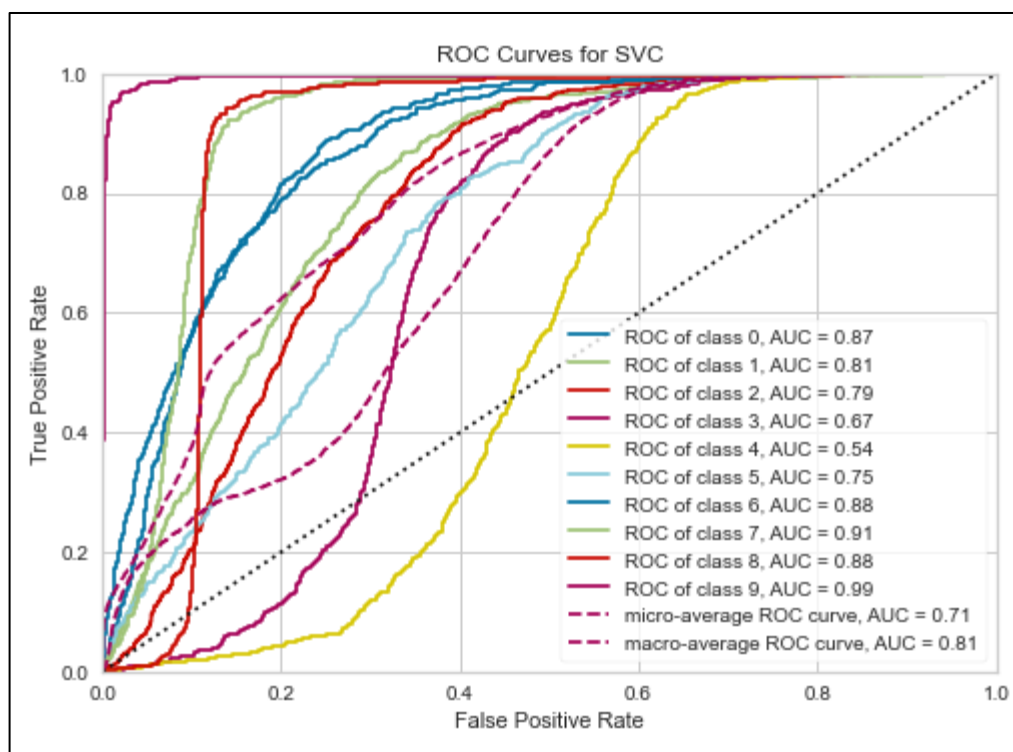


Figura 12: Curva de ROC perteneciente a configuración 4.

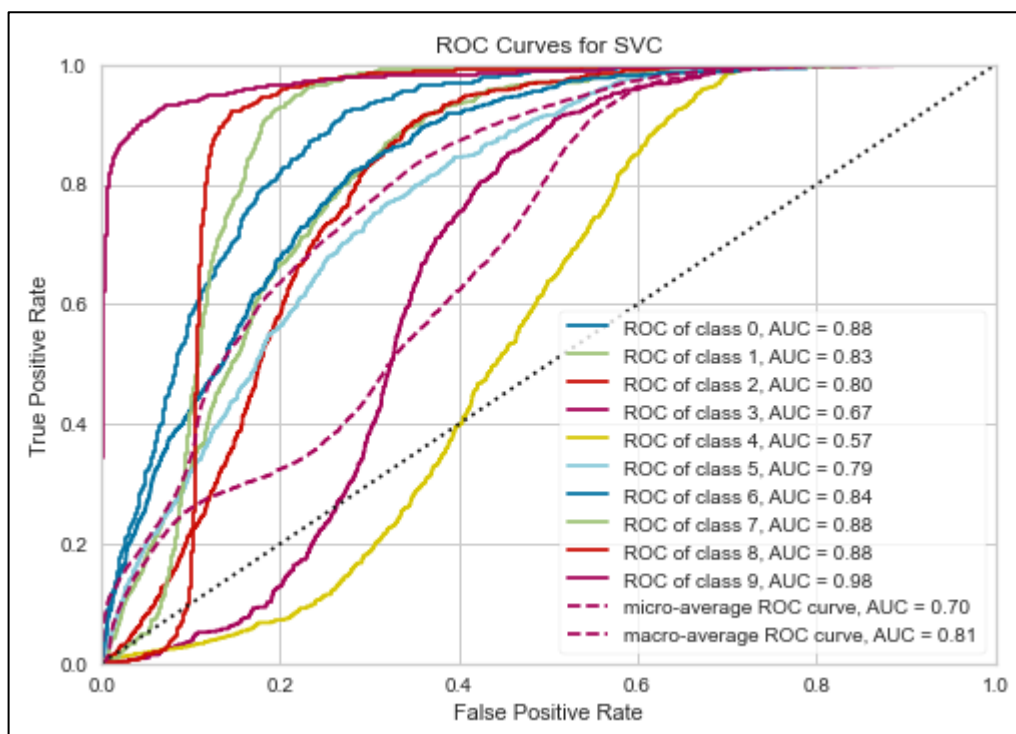


Figura 13: Curva de ROC perteneciente a configuración 5.

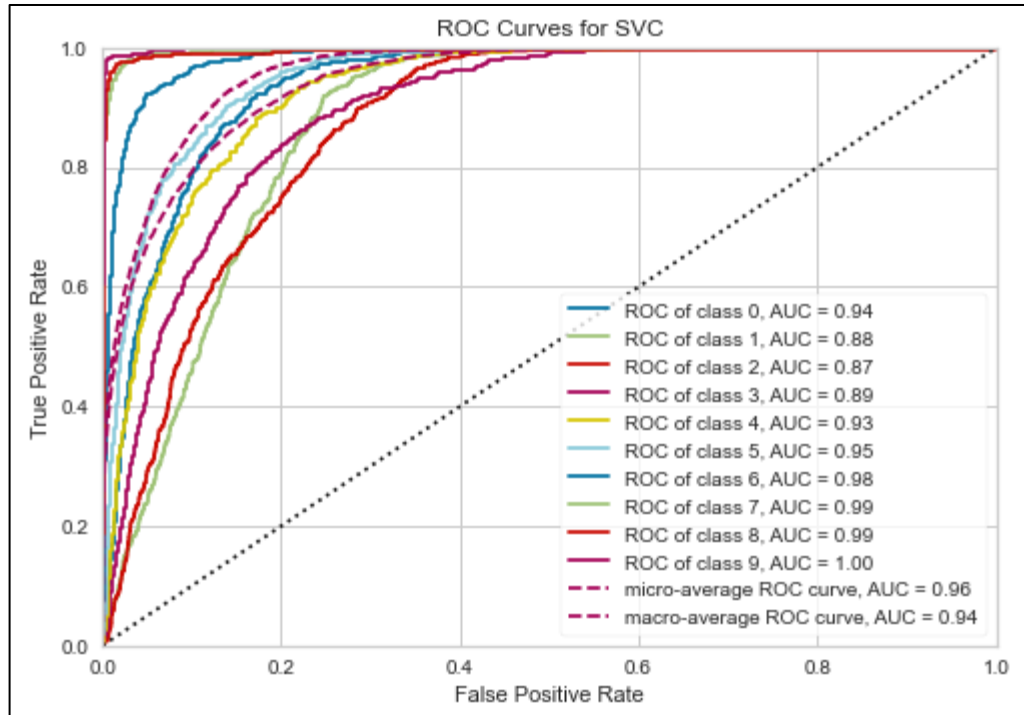


Figura 14: Curva de ROC perteneciente a configuración 6.

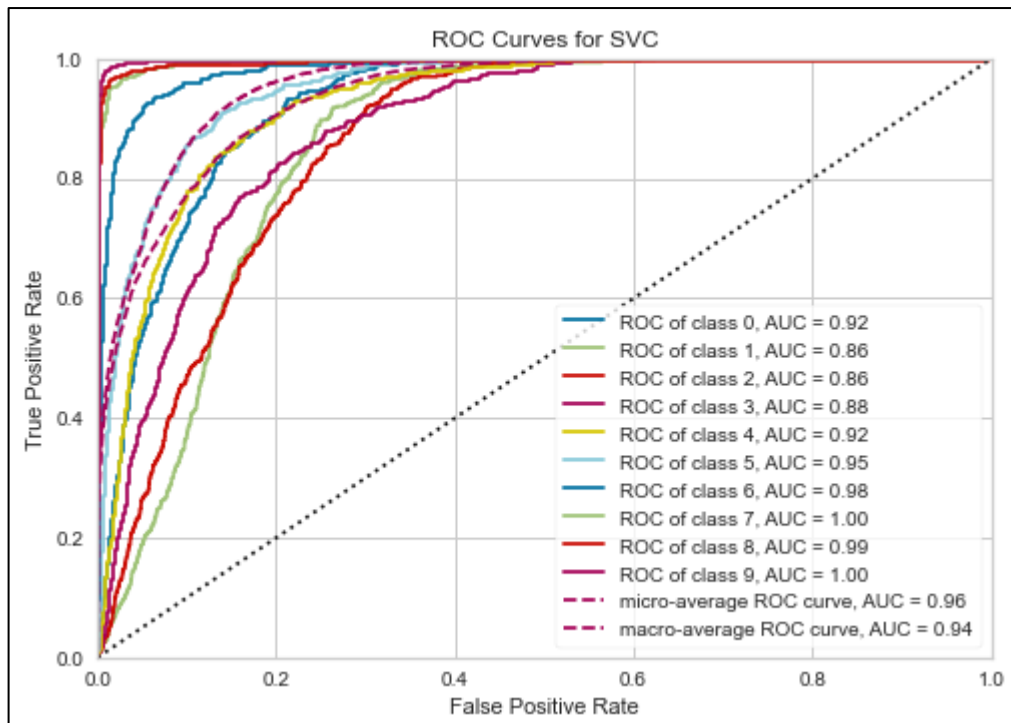


Figura 15: Curva de ROC perteneciente a configuración 7.

ANEXO II

A continuación, se muestran las matrices de confusión para cada análisis.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	539	165	88	31	5	0	1	0	0	0
	1	322	242	160	77	11	0	0	0	0	0
	2	154	156	259	177	35	3	2	0	0	0
	3	34	49	145	410	129	20	3	1	0	0
	4	9	7	17	129	480	133	22	0	0	0
	5	2	2	4	33	145	540	87	4	2	1
	6	2	0	0	11	12	75	681	39	4	1
	7	0	0	0	1	1	0	18	740	27	0
	8	0	0	1	1	1	1	1	24	760	31
	9	0	0	0	0	0	0	1	0	25	807

Tabla 3: Matriz de confusión en base a configuración 1.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	425	161	82	25	8	0	0	0	0	0
	1	228	200	154	76	16	1	0	0	0	0
	2	102	154	223	159	41	5	1	0	0	0
	3	20	42	119	334	133	19	3	0	0	0
	4	4	7	15	107	390	162	12	0	0	0
	5	0	1	0	19	110	457	72	4	1	0
	6	0	0	0	3	8	73	547	27	2	0
	7	0	0	0	2	1	0	22	630	23	0
	8	0	0	0	0	0	0	2	19	573	18
	9	0	0	0	1	0	0	0	1	27	679

Tabla 4: Matriz de confusión en base a configuración 2.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	428	149	69	28	3	0	0	1	0	0
	1	268	196	115	62	12	1	4	0	0	0
	2	108	169	217	148	33	4	2	0	0	0
	3	30	41	97	358	118	18	5	1	1	1
	4	2	4	10	119	396	138	29	2	0	0
	5	0	1	2	25	110	476	76	3	1	1
	6	0	0	0	5	10	69	578	22	2	0
	7	0	0	0	1	0	1	20	617	28	2
	8	0	0	1	0	0	0	0	16	628	16
	9	0	0	0	0	0	0	0	0	10	642

Tabla 5: Matriz de confusión en base a configuración 3.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	0	0	0	141	264	66	60	4	0	0
	1	0	0	0	72	211	161	101	2	0	0
	2	0	0	0	45	127	194	140	5	0	0
	3	0	0	0	18	24	212	305	10	0	0
	4	0	0	0	0	4	133	375	7	0	0
	5	0	0	0	0	2	65	484	17	0	0
	6	0	0	0	0	1	13	454	66	0	0
	7	0	0	0	0	0	2	39	491	0	0
	8	0	0	0	0	2	0	19	525	5	0
	9	0	0	0	0	9	0	21	236	267	1

Tabla 6: Matriz de confusión en base a configuración 4.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	0	2	0	156	372	82	60	7	0	0
	1	0	3	0	97	305	174	94	8	0	0
	2	0	1	0	67	232	249	82	9	0	0
	3	0	0	0	25	69	413	162	22	0	0
	4	0	0	0	4	13	345	286	12	0	0
	5	0	0	0	0	4	214	464	31	0	0
	6	0	0	0	0	6	195	362	115	0	0
	7	0	0	0	0	5	14	34	595	0	0
	8	0	0	0	0	5	0	16	684	4	0
	9	0	0	0	0	20	0	2	430	199	0

Tabla 7: Matriz de confusión en base a configuración 5.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	333	84	83	15	4	1	0	0	0	0
	1	201	121	147	61	11	2	2	0	0	1
	2	94	76	237	113	30	3	0	1	0	0
	3	21	15	119	266	110	29	3	1	0	0
	4	1	8	15	68	306	123	15	1	0	0
	5	0	1	0	20	93	360	70	1	0	0
	6	0	1	0	3	7	63	417	15	0	1
	7	0	0	0	1	0	0	27	483	21	1
	8	0	0	0	0	0	0	1	15	522	16
	9	0	0	0	0	0	0	1	0	14	525

Tabla 8: Matriz de confusión en base a configuración 6.

OBSERVACIÓN	PREDICCIÓN										
		0	1	2	3	4	5	6	7	8	9
	0	291	159	77	23	2	0	0	0	0	0
	1	155	164	144	54	11	1	0	0	0	0
	2	95	124	207	121	16	6	0	0	0	0
	3	26	45	106	237	92	15	1	0	1	0
	4	7	7	11	101	306	121	13	0	0	0
	5	2	3	0	21	93	355	53	3	0	0
	6	0	3	2	0	11	55	426	27	0	0
	7	0	0	1	1	0	3	20	483	17	0
	8	0	0	0	0	0	2	2	12	522	16
	9	0	0	0	1	0	0	0	2	18	497

Tabla 9: Matriz de confusión en base a configuración 7.

ANEXO III

A continuación, se muestran las configuraciones utilizadas para creación de modelo.

CONFIGURACIÓN	KERNEL	PARAMETRO C	PARAMETRO GAMMA
CONFIGURACIÓN 1	<i>POLY</i>	<i>1000</i>	<i>'AUTO'</i>
CONFIGURACIÓN 2	<i>RBF</i>	<i>600</i>	<i>'AUTO'</i>
CONFIGURACIÓN 3	<i>RBF</i>	<i>10</i>	<i>1</i>
CONFIGURACIÓN 4	<i>RBF</i>	<i>500</i>	<i>0.5</i>
CONFIGURACIÓN 5	<i>RBF</i>	<i>2000</i>	<i>2</i>
CONFIGURACIÓN 6	<i>RBF</i>	<i>8000</i>	<i>'AUTO'</i>
CONFIGURACIÓN 7	<i>RBF</i>	<i>3000</i>	<i>1</i>

Tabla 3: Mejores configuraciones para modelo propuesto.

ANEXO IV

A continuación, se muestran los gráficos de resultados para los estados de Normocapnia hacia Hipercapnia.

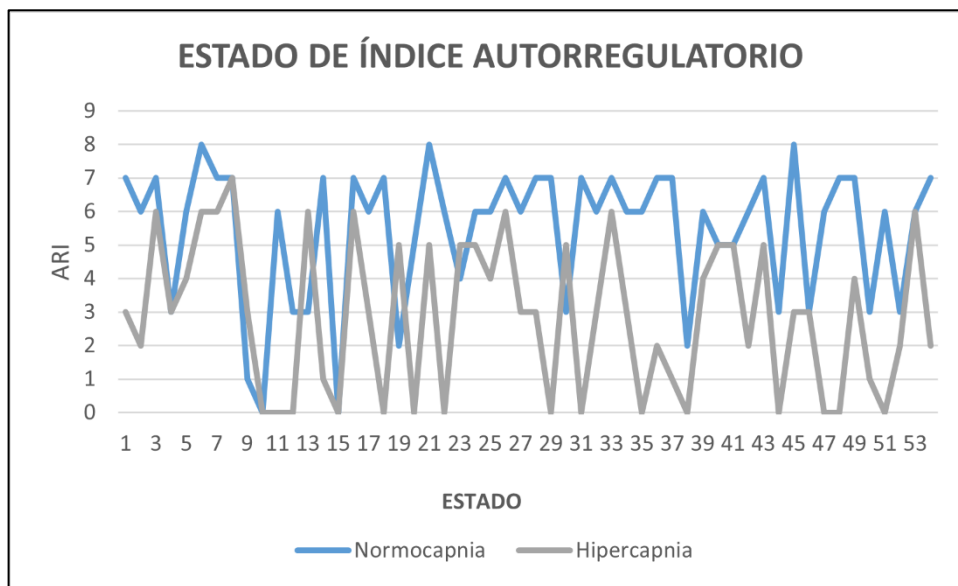


Figura 17: Gráfico de datos para validación en estado de Normocapnia e Hipercapnia

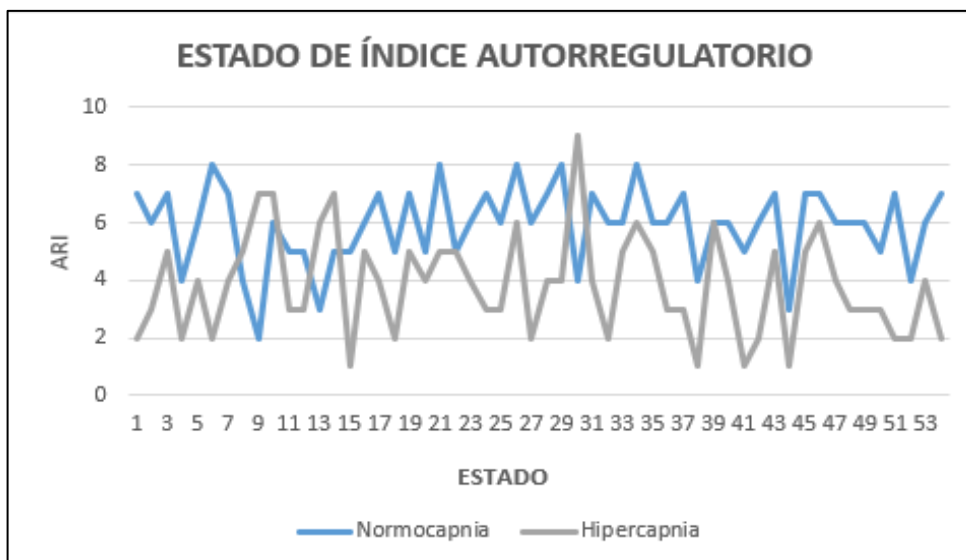


Figura 18: Gráfico de predicción en base a correlación en estado de Normocapnia e Hipercapnia.

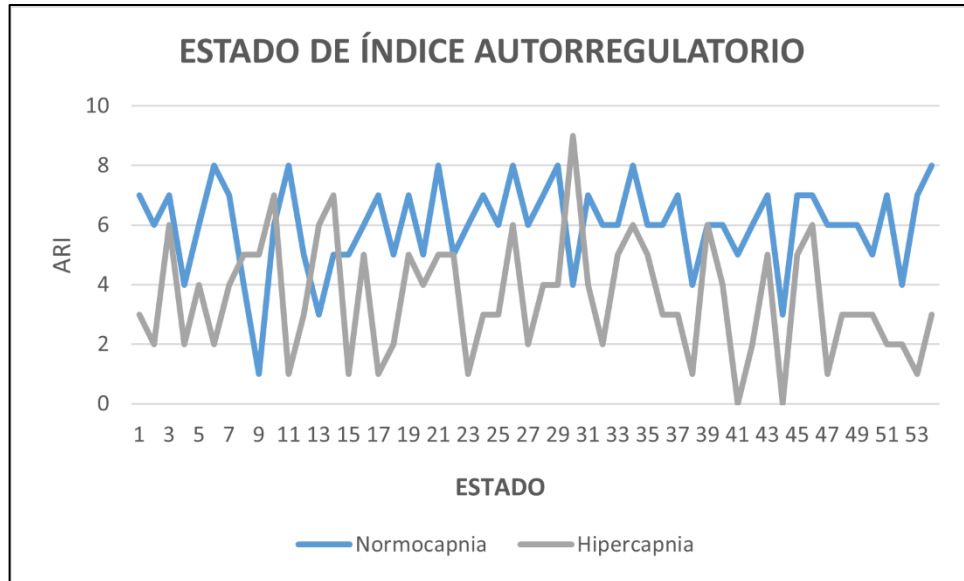


Figura 19: Grafico de predicción en base a error cuadrático medio en estado de Normocapnia e Hiperapnia.

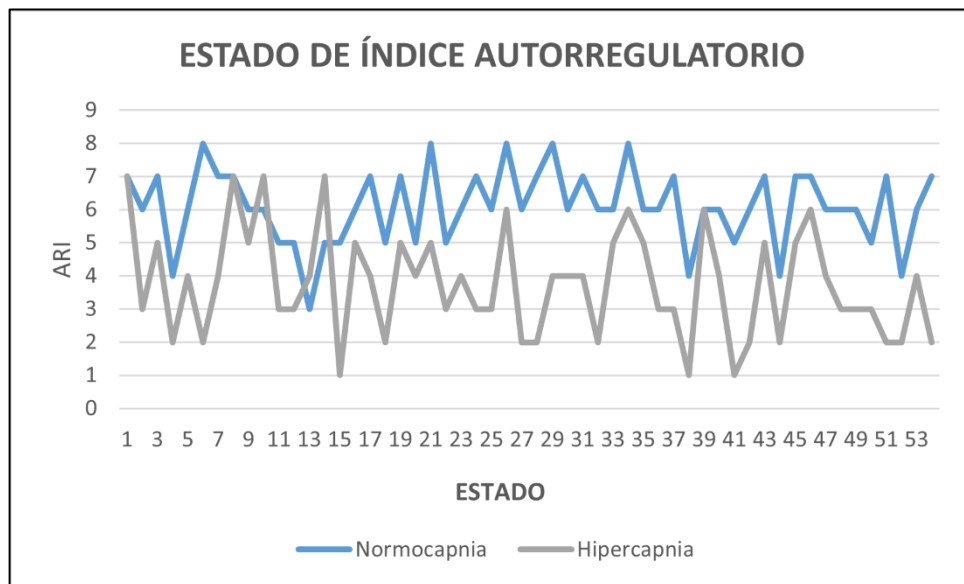


Figura 20: Grafico de predicción en base a curva ROC en estado de Normocapnia e Hiperapnia.

ANEXO V

Código para la creación y validación del modelo.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.metrics import mean_squared_error, r2_score
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score, make_scorer
from sklearn.model_selection import GridSearchCV
from sklearn.metrics import roc_auc_score
from sklearn.metrics import confusion_matrix

# Cargando Datos Base de datos de entrenamiento.
datos = pd.read_csv("bdtesis.csv")
# Renombre de Columnas
datos.columns = ["c1", "c2", "c3", "c4", "c5", "c6", "c7", "ARI"]

# Cargando Datos Base de datos de validación.
datvalid = pd.read_csv("bdvalidacion.csv")
xvalid = datvalid.iloc[:, 0:7].values
yvalid = datvalid.iloc[:, 7].values

xval = pd.DataFrame(xvalid)
yval = pd.Series(yvalid)
# Definiendo datos de entrenamiento y test
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(
    datos.drop('ARI', axis = 'columns'),
    datos['ARI'],
    train_size = 0.80,
    shuffle = True
)

# Normalización de datos

from sklearn.preprocessing import StandardScaler
from sklearn import preprocessing
# Estandarización de datos (Z-score normalization)
st_x = StandardScaler()
st_y = StandardScaler()

x_trainprueba = st_x.fit_transform(X_train)
y_trainprueba = st_y.fit_transform(y_train.values.reshape(-1,1))

# Estableciendo modelo y entrenamiento.

modelo = GridSearchCV(SVC(gamma='auto', probability=(True)), {
    'C': [1000, 600, 10, 500, 2000, 3000, 8000],
    'kernel': ['rbf', 'poly']}, cv=10, return_train_score=False)

modelo.fit(X_train, y_train)
```

```

#Predicción con BD Entrenamiento.
prediccion = modelo.predict(X_test)
df_predicciones = pd.DataFrame({'ARI': y_test, 'Predicción': prediccion})
accuracy1 = accuracy_score(y_test, prediccion)
print(f"Valor accuracy BD entrenamiento: {accuracy1}")

#Error cuadrático medio BD Entrenamiento.
rmse = mean_squared_error(
    y_true = y_test,
    y_pred = prediccion,
    squared = False
)

print(f"Valor Error cuadrático Medio BD Entrenamiento: {rmse}")

#Coeficiente de determinación BD Entrenamiento.
r2 = r2_score(
    y_true = y_test,
    y_pred = prediccion)

print(f"Valor coeficiente de determinación BD Entrenamiento: {r2}")

#Correlación BD Entrenamiento.
out = pd.DataFrame()
out['y'] = y_test
out['yp'] = prediccion
corr=out.corr()
#corr['y'][1]
print(f"Correlación BD Entrenamiento: %.4f %corr['y'][1])

y_pred_proba = modelo.predict_proba(X_test)
score = roc_auc_score(y_test,y_pred_proba,multi_class="ovr")

print(f"Valor area bajo la curva / ROC: {score} \n")

print("----")
print("---- \n")

```

```

#Predicción con BDValidación.
prediccionval = modelo.predict(xvalid)
df_prediccionval = pd.DataFrame({'ARI': yvalid, 'Predicción': prediccionval})
accuracy2 = accuracy_score(yvalid, prediccionval)
print(f"Valor accuracy BD validación: {accuracy2}")

#Error cuadrático medio BDValidación.
rmse = mean_squared_error(
    y_true = yvalid,
    y_pred = prediccionval,
    squared = False
)

print(f"Valor Error cuadrático Medio BDValidación: {rmse}")

#Coeficiente de determinación BDValidación.
r2 = r2_score(
    y_true = yvalid,
    y_pred = prediccionval)

print(f"Valor coeficiente de determinación BDValidación: {r2}")

#Correlación BD Entrenamiento.
out = pd.DataFrame()
out['y'] = yvalid
out['yp'] = prediccionval
corr=out.corr()
#corr['y'][1]
print('Correlación BDValidación: %.4f'%corr['y'][1])

#Matriz de confusion
confusion_matrix(y_test, prediccion)

#ROC
from yellowbrick.classifier import ROCAUC
visualizer = ROCAUC(modelo, classes=[0, 1, 2, 3, 4, 5, 6, 7, 8, 9])

visualizer.fit(X_train, y_train) # Fit the training data to the visualizer
visualizer.score(X_test, y_test) # Evaluate the model on the test data
visualizer.show()
y_pred_proba = modelo.predict_proba(X_test)
from sklearn.metrics import roc_auc_score
score = roc_auc_score(y_test,y_pred_proba,multi_class="ovr")

```

