

**"ANÁLISIS DE DATOS EN PACIENTES CON ACV-ISQUÉMICO, USANDO TÉCNICAS CLÁSICAS DE MACHINE LEARNING"**

***Autor:*** *Abraham Marianjel Sepúlveda*

***Profesora Guía:*** *Carola Figueroa Flores*

# Introducción

# nindent=0.6em]{E}{n} nuestros tiempos los computadores o similares son indispensables para nuestro día a día, ayudando a mejorar la calidad de los trabajos y llevándolos a una mejor producción, siendo un gran impacto en la sociedad hoy más que nunca. Las múltiples ramas de las ciencias de la computación han ayudado al desarrollo progresista de otras áreas de trabajo e investigación, como lo es la medicina, economía, estadística o social, todo lo anterior mencionado se debe a la gran cantidad de bancos de memoria que puede procesarse y su rapidez, esto hace que las posibilidades de combinación sean altísimas entre las áreas.\\

# \par Durante las últimas décadas se ha incrementado enormemente nuestra capacidad para recoger información en cualquier actividad, por ejemplo, en los negocios, recopilando información sobre procesos de producción, ventas, servicios, campañas de marketing, entre varios. La gran cantidad de información ha crecido el interés de las personas por utilizarlos para extraer la información que represente una ventaja competitiva para las organizaciones.\\

# \par El campo interdisciplinario de la ciencia de la computación involucra métodos científicos, procesos y sistemas para extraer conocimiento o mejor entendimiento de datos en sus diferentes formas, ya sean información útil o no útil, idealmente estructurado. Dentro de las ciencias de la computación, existen especialidades que solucionan un problema específico, de las cuales nosotros nos enfocaremos en la Inteligencia Artificial, que posee una variedad de metodologías para resolver los problemas de forma eficaz y eficiente. \\

# \par Actualmente la Inteligencia Artificial se ha disparado en ascenso, ya que nos ayuda a resolver problemas con mucha certeza (aplicando bien el método) y es aplicable para un gran número de disciplinas de áreas de trabajo. El mercado se ha disparado con herramientas que utilizan aplicaciones con Inteligencia Artificial, en opciones comerciales, salud, social, etc.\\

# \par Procesando los datos con Inteligencia Artificial llegamos a una de sus principales ramas y una de las más conocidas que es el Machine Learning, tomando un rol importantísimo en tareas con aprendizaje, análisis de datos, minería de datos, reconocimiento de patrones, entre varias. Todo lo anterior mencionado del Machine Learning, lo vuelve una de las ramas más atractivas para los sectores que trabajan grandes cantidades de datos.\\

# \par Este trabajo se abordará una problemática referente al área de la salud, en una de las principales causas de muerte a nivel mundial y específicamente dentro de nuestra región es la segunda causa de muerte. La finalidad es analizar algunos factores que se producen al evaluar un Accidente Cerebro Vascular con Machine Learning.\\

# Descripción del Problema

Las enfermedades de Accidente Cerebro Vascular llamados ACV, hoy en día es una de las principales causas de muerte, asimismo las personas que logran sobrevivir quedan con secuelas o discapacidades en la mayoría de los casos, presentándose con más frecuencia sorpresivamente en los adultos jóvenes, pero también aumentando en los adultos mayores \cite{Ortiz-Galeano2020}. \\

\par En Chile, los Accidentes Cerebro Vascular llamados ACV son unos de las principales causas de muerte, solo el año 2021 hubo 29.542 egresos hospitalarios por ACV siendo la segunda causa de mortalidad a nivel país, después de las enfermedades isquémicas del corazón y no considerando la pandemia por el SARS-CoV-2. El registro de defunciones por ACV llego a los 7.501 casos ese mismo año, lo que equivale a 1 muerte cada 1 hora y 12 minuto \cite{Minsal2022}.\\

\par La salud es vital para estimar nuestro bienestar y disminuir las posibles enfermedades en el menor plazo posible, vinculado a esto, el ACV posee una alta tasa de morbimortalidad, las muertes por esta enfermedad son muy rápidas en poco tiempo, una vez detectada por el ACV en sí o por las secuelas \cite{Gaudiano2019}. El impacto negativo de las secuelas en los pacientes es un alto costo sanitario, físico y social, encontrando en lo sanitario lo intrahospitalario como extra hospitalario, en lo físico la perdida de la movilidad de una parte de su cuerpo (discapacidad) y en lo social, quizás la pérdida de un sentido (hablar o escuchar) o disminución de la calidad de vida de la persona. Ante lo anteriormente expuesto, la aparición de la enfermedad y su posible evolución se debería a los factores de riesgos para esta enfermedad \cite{Cabrera2020}.\\

\par En Ñuble, se cuenta con una red hospitalaria que lleva un registro de las enfermedades y pacientes que ingresan a los hospitales, de este modo se determinó que una de las principales causas de muerte en la región es consecuencia de los ACV, además una persona muerta al día y represente la primera causa en la zona y el país, hablando en términos de probabilidades. Esta enfermedad se trata de una urgencia, donde acceder a un tratamiento oportunamente, puede establecer la diferencia en el pronóstico de salud. Las acciones deben estar destinadas a preservar la integridad del tejido cerebral \cite{ServicioSaludNuble}.\\

\par Para obtener resultados más óptimos, los médicos deben realizar una serie de exámenes para ayudar al pronóstico del paciente, ayudando a conservar la mayor parte de salud en la persona. Para una predicción médica con más apoyo, la Inteligencia Artificial y el Machine Learning ayuda a obtener resultados en el menor tiempo posible, lo que gracias a exámenes que se deben llevar a cabo. Es por esta razón que el rol de la inteligencia artificial jugará un papel importante a futuro en la salud, es así como esta disciplina está explorándose y ejecutándose en los países más desarrollados. Al mismo tiempo, los errores de diagnósticos en los países en vías de desarrollo, las herramientas ligadas a la inteligencia artificial pueden cumplir con el rol de optimizar el diagnóstico de enfermedades \cite{Curioso2020}.\\

# Descripción del Proyecto

En la presente sección tiene como propósito dar a conocer la hipótesis y los objetivos del proyecto.\\

# Hipótesis

Es posible predecir con un alto nivel de precisión, en un sistema de clasificación mediante comparación con técnicas clásicas de Machine Learning, las secuelas que se originan en un accidente cerebro vascular isquémico.\\

# Objetivo General

Detectar la mejor técnica de Machine Learning para identificar los factores de mal pronóstico en adultos mayores con diagnóstico de accidente cerebrovascular isquémico a través del análisis comparativo de su precisión.\\

## Objetivos Específicos: Distintos

1. Estudiar las distintas técnicas de Machine Learning que existen en la literatura y la bibliografía relacionada con modelos basado en Machine Learning aplicado a la salud.
2. Analizar comparativamente tres técnicas clásicas de Machine Learning que permita abordar el estudio de accidentes ACV.
3. Desarrollar sistema de clasificación, que evaluando exámenes, estadísticas e historial médico pueda ser capaz de predecir las secuelas en pacientes con ACV.
4. Implementar sistema de clasificación que con una base de datos clínica pueda ser capaz de predecir secuelas en pacientes con ACV.

# Marco Teórico

## \lettrine[lines=4, slope=0.2em, findent=0.2em, nindent=0.6em]{E}{l} presente capitulo tiene como finalidad promover al lector para dar conocimiento de algunos conceptos sobre la Inteligencia Artificial, específicamente Machine Learning como aplicación al problema con variadas técnicas que podrían llevar a posibles soluciones, ya que algunas técnicas son necesarias conocerlas para saber cuáles se escogerán al momento de resolver el problema. Además, estará presente algunos conceptos de salud que es necesario conocer, todo esto para saber que es, por que, como y que se realiza con los ACV Isquémicos.\\

## \par La organización de este capítulo está elaborada de la siguiente forma, primeramente, se dará a conocer lo que se refiere a Machine Learning con algunas de sus técnicas clásicas, más adelante algo de Deep Learning, que sin lugar a dudas es muy eficaz con algunos problemas del mundo contemporáneo. Finalmente, el capítulo se enfocará en lo que son algunos términos de salud que nos ayudará a resolver y entender de mejor forma la raíz del problema.\\

## Inteligencia Artificial

Las matemáticas nos han ayudado a interpretar y entender nuestro medio ambiente y poder ser capaz de predecir algunos sucesos en áreas como la cosmología o naturaleza o problemas incomprendidos por los humanos \cite{Grazia2022}. Hoy en día contamos con el área de las ciencias de la computación que fue creada hace solo un par de años atrás, siendo una de las ciencias más nuevas y que tiene gran valoración hoy en día. En las ciencias de la computación podemos encontrar varias áreas aplicadas, de las cuales la Inteligencia Artificial es una de las más cotizadas, porque involucra en su relación la matemática con filosofía, biología, lingüística u otra, pudiendo resolver los problemas con distintos tipos de algoritmos \cite{Grazia2022}.\\

\par La Inteligencia Artificial llegó para resolver tareas que el ser humano realiza como tareas cotidianas básicas y complejas. Los algoritmos de IA, se basan en el aprendizaje automático, siendo que cada vez las máquinas pueden aprender por ellas mismas algunas cosas, es así que los humanos dejaremos de perder nuestro tiempo programando reglas para lidiar con muchas combinaciones de datos y situaciones que aparecen en nuestro mundo. Pues es así, que aprender los algoritmos, es fijarse como aprender desde pequeños, como el aprendizaje por refuerzos, que involucra una serie de técnicas de aprendizaje automático que las máquinas usan en sistemas artificiales \cite{Carola}.\\

\par En la IA, para que los modelos funcionen existen modelos neuronales y estos son los que determinan como se conectan los datos predictores con los objetivos a través de capas ocultas, y estas a su vez contienen unidades no observables \cite{IBM}. Dentro de la Inteligencia Artificial, convive un sub conjunto llamada Machine Learning y su vez, dentro de Machine Learning vive un subconjunto llamado Deep Learning. Estos sub conjuntos interactúan con lo que es el Data Science y el Big Data, haciendo de las aplicaciones una variada gama de aplicaciones \cite{moreno2021diseno}.\\

\begin{figure}[h]

\centering

\includegraphics[scale=0.51]{img/Marco Teorico/Expertiz.png}

\caption{Diagrama de Venn con los subcojuntos de la IA}

\end{figure}

\doublespacing

## Machine Learning

\spacing{1.5}

Los problemas por computadora siempre son solucionado con un algoritmo que le indica a la maquina los pasos a seguir para dar solución a dicho problema, en un ejemplo clásico podemos usar un algoritmo de búsqueda, donde el dato a buscar seria la entrada y su salida seria la respuesta a que si se encontró o no y en cuanto tiempo, es bien conocido que para resolver el problema de búsqueda existen variados algoritmos, algunos más eficientes que otros, aplicables a un contexto u otro de mejor forma, solucionando el mismo problema.\\

\par Asimismo, existen tareas cotidianas como lo son la publicidad en el celular, como muestra cosas que me gustan o que he estado pensando en comprar, esto es gracias a algoritmos de Machine Learning que captan la información de navegación de nuestro celular y nos muestra lo que nosotros estamos deseando o es de nuestro interés.\\

\par En el año 1959, Arthur Samuel, definió el concepto de Machine Learning como:

\begin{quote}

\emph{"Un campo de estudio que le entrega a los computadores la habilidad de aprender sin haber sido explícitamente programado para eso"}

\end{quote}

Tom M. Mitchell define el Machine Learning en uno de sus libros como:

\begin{quote}

\emph{“El estudio de algoritmos de computación que mejoran automáticamente su rendimiento gracias a la experiencia. Se dice que un programa informático aprende sobre un conjunto de tareas, gracias a la experiencia y usando una medida de rendimiento, si su desempeño en estas tareas mejora con la experiencia”}

\end{quote}

\par Es decir, estos algoritmos aprenden y mejoran solos gracias a la experiencias pasadas a diferencia de modelos que un experto puede asignar reglas y modela gracias a sus conocimientos.\\

\par El Machine Learning (ML) es el subconjunto de la Inteligencia Artificial, utilizándose para que las máquinas aprendan de grandes volúmenes de datos. Una vez adquirido el conocimiento, este puede ser empleado a otro conjunto de datos, recibiendo una atención cada vez mayor de un modelo que es capaz de interpretar los datos con mejor capacidad cada vez \cite{murdoch2019interpretable}.\\

\doublespacing

### Machine Learning: Tipo de Algoritmos

\spacing{1.5}

Para Machine Learning existen algunos enfoques para trabajar, los cuales son los más usados el aprendizaje supervisado y el aprendizaje no supervisado.\\

\doublespacing

#### Aprendizaje Supervisado

\spacing{1.5}

Este enfoque es un método de análisis de datos que necesita de algoritmos que aprendan a través de entrenamiento, en el cual el programa es alimentado con datos etiquetados tanto con los atributos como la variable objetivo de cada iteración.\\

\par En el aprendizaje supervisado del algoritmo existen dos tareas comunes que son la clasificación y regresión. En las tareas de clasificación las variables objetivo, son valores etiquetados de manera nominal, mientras que en las tareas de regresión la variable objetivo tiene un dominio dentro de los números reales y por lo tanto pueden tomar infinitos valores \cite{murdoch2019interpretable}. \\

\doublespacing

#### Aprendizaje No Supervisado

\spacing{1.5}

El aprendizaje no supervisado tiene datos sin etiquetar que el algoritmo tiene que entender por si mismo, el propósito es descubrir patrones ocultos en ellos. Si nosotros le pedimos a nuestro programa \textit{"predecir Y para nuestros datos A"} nosotros deberíamos \textit{“pedirle que nos provea de la información de nuestros datos A"} \cite{murdoch2019interpretable}.\\

\par Las tareas más comunes dentro del aprendizaje no supervisado son el \emph{clustering} y \emph{dimensión reduction}, los cuales no son objeto de nuestro estudio por lo que sólo son mencionados en esta parte del documento.\\

\doublespacing

### Técnicas de Clasificación

\spacing{1.5}

Para poder que nuestro modelo de Machine Learning funcione, debemos mediante la entrada ya la salida de datos, reconocer atributos del elemento a clasificar, utilizando el conocimiento adquirido durante el entrenamiento del algoritmo para asignar un valor a la variable objetivo de dicho elemento.

\par La clasificación de datos es un proceso que posee dos etapas:

\par La primera es la etapa de aprendizaje, la cual los datos ingresan al modelo son entrenados con una serie de datos que sirven de base del algoritmo, logrando que se entrene con un set de entrenamiento que no son mas que tuplas de datos etiquetados tanto en sus atributos como en su variable objetivo.

\par La segunda etapa, el algoritmo ya ha sido entrenado con los datos de entrenamiento y esta listo para su clasificación de una variable cualquera, siendo capaz de tomar algunos atributos de la variable, analizarlos y tomar una decisión basándose en los datos que fueron entregados anteriormente.

En la clasificación veremos algunas de las técnicas clásicas de Machine Learning, pertenecientes al enfoque de los algoritmos de aprendizaje supervisado, definiendo brevemente en que consiste, como se lleva a cabo las técnicas a mencionar.

\doublespacing

### Clasificación

\spacing{1.5}

Para poder que nuestro modelo de Machine Learning funcione, debemos mediante la entrada ya la salida de datos, reconocer atributos del elemento a clasificar, utilizando el conocimiento adquirido durante el entrenamiento del algoritmo para asignar un valor a la variable objetivo de dicho elemento.\\

\par La clasificación de datos es un proceso que posee dos etapas:\\

\par La primera es la etapa de aprendizaje, la cual los datos ingresan al modelo son entrenados con una serie de datos que sirven de base del algoritmo, logrando que se entrene con un set de entrenamiento que no son mas que tuplas de datos etiquetados tanto en sus atributos como en su variable objetivo.\\

\par La segunda etapa, el algoritmo ya ha sido entrenado con los datos de entrenamiento y esta listo para su clasificación de una variable cualquera, siendo capaz de tomar algunos atributos de la variable, analizarlos y tomar una decisión basándose en los datos que fueron entregados anteriormente.\\

\par En la clasificación veremos algunas de las técnicas clásicas de Machine Learning, pertenecientes al enfoque de los algoritmos de aprendizaje supervisado, definiendo brevemente en que consiste, como se lleva a cabo las técnicas a mencionar \cite{murdoch2019interpretable}.\\

\doublespacing

Decision Tree (Árbol de Decisión)

\spacing{1.5}

Este algoritmo que se utiliza como herramienta de apoyo gráfico o modelo de decisiones con sus posibles consecuencias, también a veces son representados los costos y su posible utilidad (CART, Classification and Regression Trees). Este método se crea particionando la entrada recursivamente en distintas ramas, siendo que la idea es crear un camino desde la raíz hasta las hojas, donde cada nodo podría ser una condición, así si se cumple la condición se sigue por el camino de decisión o por la otra rama si no se llegará a cumplir la condición \cite{Teknomo2015}.\\

\par La creación del modelo se puede representar con la siguiente ecuación:\\

\begin{equation}

f(x)=E[y|x]=\sum\_{m=1}^{M}w\_{m}I(x \in R\_{m})=\sum\_{m=1}^{M}w\_{m}\phi(x;v\_{m})

\end{equation}

\par $R\_{m}$ representa la región de $m$.

\par $w\_{m}$ es respuesta media a esa región.

\par $v\_{m}$ codifica la elección de la variable por la que dividir y el valor límite de la división.\\

\par La forma gráfica del modelo se puede representar de la siguiente forma en condiciones:\\

\begin{figure}[h]

\centering

\includegraphics[scale=1]{img/Marco Teorico/arbol de desicion.png}

\caption{Forma Gráfica de un Árbol de Desiciones}

\end{figure}

\par Las condiciones se dan como alternativas de caminos a seguir.\\

\doublespacing

\subsubsection{Usos para los Decision Tree}

\spacing{1.5}

Al ser de fácil implementación, los Decision Tree son usados en diversas áreas, siendo las instituciones financieras las más comunes, ayudando a clasificar clientes, estableciendo sus riesgos o posibilidades financieras.\\

\par En el área de la salud son empleados para diagnósticos de infecciones a la sangre o predicción de ataque al corazón en pacientes de alto riesgo.\\

\par En el área de entretenimiento son ampliamente usados como métodos de seguimiento de movimiento y a su vez como método de reconocimiento facial.

\doublespacing

\subsubsection{Ventajas y Desventajas}

\spacing{1.5}

Principalmente las ventajas de esta técnica de Machine Learning son que los resultados son de fácil interpretamiento por las personas, inclusive trabaja bien con ruidos en los datos, además incluyendo un bajo costo computacional.\\

\par La mayor desventaja es que la técnica es propensa a caer en el sobreajuste, siendo a veces posible manipulada por quien lo implemente.\\

## Naïve Bayes

La clasificación de Naïve Bayes es un algoritmo probabilístico basados en la probabilidad condicional y teorema de Bayes con atributos fuertes y de independencia entre características \cite{Shafe}. Este modelo es de fácil construcción y funciona bien en comparación a los métodos más sofisticados de clasificación, lo que lo hace particularmente útil en los diferentes campos investigativos \cite{Vembandasamy2015}.\\

\par La fórmula condicional se expresa:\\

\begin{Large}

\begin{equation}

P(A|B)=\frac{P(A \cap B)}{P(B)}= \frac{(\frac{\#casos favorables A \cap B}{\#casos posibles})}{(\frac{\#casos favorables B}{\#casos posibles})}

\end{equation}

\end{Large}\\

\par Desarrollando la ecuación nos queda:\\

\begin{Large}

\begin{equation}

P(A|B)=\frac{P(B|A)\*P(A)}{P(B)}

\end{equation}

\end{Large}\\

\par Donde cada evento es tomado como:\\

\begin{large}

\begin{equation}

P(A|B)= P(B\_{1}|A) \times P(B\_{2}|A) \times … \times P(B\_{n}|A) \times P(A)

\end{equation}

\end{large}\\

\par $P(A|B)$ es la probabilidad posterior de la clase (objetivo) dado el predictor (atributo).

\par $P(A)$ es la probabilidad previa de clase.

\par $P(B|A)$ es la probabilidad, que es la probabilidad de la clase dada del predictor.

\par $P(x)$ es la probabilidad previa del predictor.\\

\par Las Redes de Bayes se generan de reglas de decisión donde participan activamente las probabilidades que ocurren referente a eventos, siendo que en base a esas probabilidades y resultados obtenidos, se toman decisiones sobre cual arco de red moverse, al final del proceso, el resultado será dado por el valor del nodo final de la red.\\

\doublespacing

\subsubsection{Ventajas y Desventajas}

\spacing{1.5}

La gran sencilles para enfrentar problemas lo hace ser una de las técnicas clásicas de Machine Learning más fácil de interpretar al momento de los resultados, además debemos agregar al valor que son fáciles de construir y toma evidencia de muchos atributos para realizar la predicción final.\\

\doublespacing

\subsection{Logistic Regression (Regresión Logística)}

\spacing{1.5}

El algoritmo Logistic Regression modela la relación entre distintas variables utilizando una medida de error que se intentará minimizar en un proceso iterativo para poder realizar predicciones acertadas, llevando a cabo una clasificación binaria con una distribución Bernoulli en vez de Gaussian y después realiza una combinación lineal de las variables en un rango de 0 a 1 \cite{ Stoltzfus2011}.\\

\par La ecuación lineal se debe ajustar:\\

\begin{Large}

\begin{equation}

y(X)=W^{T}X + \epsilon = \sum\_{j=1}^{D}w\_{j}x\_{j} + \epsilon

\end{equation}

\end{Large}\\

\par $W^{T}X$ representa el producto escalar de entrada X.

\par $W$ e $Y$ son vectores de pesos $\in$ $\lbrace 0,1 \rbrace$.\\

\par Ahora mostrando la distribución de Bernoulli:\\

\begin{Large}

\begin{equation}

p(y|x,w) = Ber(y|u(x))

\end{equation}

\end{Large}

\begin{center}

El resultado del intervalo quedaría $0 \leq u(x) \leq 1$.\\

\end{center}

\par El resultado de la ecuación nos ayudará a predecir valores con la mejor respuesta a partir del menor error posible, teniendo un valor continuo entre 0 y 1. Si existe un valor mayor o igual a 0.5, la clase será 1, en cambio si es menor será 0. Todo esto ocurre porque el algoritmo de Logistic Regression predice un valor en vez de una clase en función de las variables utilizadas.\\

\doublespacing

\subsubsection{Ventajas y Desventajas}

\spacing{1.5}

La Regresión Logística al igual que las técnicas anteriores es fácil de implementar, interpretar y muy eficiente al momento de entrenar, incluyendo que no hace suposiciones sobre distribuciones de clases en el espacio de características y es muy rápido para clasificar registros desconocidos.\\

\par Las principales desventajas de esta técnica es si el número de observaciones es menor que el número de características, no se debe utilizar la regresión logística; de lo contrario, puede provocar un sobreajuste y la difícil obtención de relaciones complejas mediante regresión logística. Los algoritmos más potentes y compactos, como las redes neuronales, pueden superar fácilmente a este algoritmo.\\

\doublespacing

### Support Vector Machine (Máquina de Vectores de Soportes)

\spacing{1.5}

\par Si bien es cierto que los algoritmos para dar un resultado de aprendizaje automático deben ser capaces de aprender de su pasado y así poder predecir su futuro, de manera que las Support Vector Machine (SVM), están para ayudarnos a la regresión a base de un conjunto de muestras, donde existen dos valores generales, uno lo compone una matriz con variables explicativas de la muestra y el otro valor general está con el valor que se espera de la muestra. En este método, una función elige la predicción del valor esperado del caso mediante una entrada de datos \cite{Dantas2021}.\\

\doublespacing

### Artificial Neural Networks (Redes Neuronales Artificiales)

\spacing{1.5}

Aunque no será parte de las técnicas para analizar los datos, importante conocer esta técnica. La Artificial Neural Networks (ANN), se caracteriza por su aprendizaje supervisado y no supervisado, eso quiere decir que es un modelo que trabaja con ML en su forma base, aunque de igual forma se ocupa para DL, posee una arquitectura de procesadores múltiples interconectados \cite{salas2004redes}. Mediante la técnica señalada para el aprendizaje de una ANN, podemos guiarnos de dos formas de organizar las capas ocultas, la cuales son los perceptores multicapa (PMC) \cite{Borracci2021}, encargado de las relaciones más complejas vía mayor coste de tiempo de entrenamiento, y la función de base radial (RBF) con una potencia de predicción reducida. De este modo, los métodos de aprendizaje que se emplean por lo regular son arquitecturas de redes neuronales tradicionales, donde solo se contienen dos o tres capas ocultas, imitando la operación que realiza el cerebro (Azath et al., 2020), en cambio, el Deep Learning aprenden sobre la marcha y su arquitectura puede llegar a tener 150 capas ocultas. \\

\doublespacing

### Deep Belief Network (Red de creencias profundas)

\spacing{1.5}

Aunque no será parte de las técnicas para análizar los datos, importante conocer esta técnica. Esta red conlleva a un método de Machine Learning que nos ayuda con el rol probabilístico, que se ejecuta en los múltiples niveles de capas y variables ocultas, conectándose entre las capas visibles y ocultas, pero no en las capas visibles – visible u oculta – oculta \cite{PuertaBarrera2015}.\\

\doublespacing

### Feedforward Artificial Neural Network (Red Neuronal de Retroalimentación)

\spacing{1.5}

Aunque no será parte de las técnicas para análizar los datos, importante conocer esta técnica. Las Feedforward Artificial Neural Network (FANN) es la sucesora de ANN, trabajándose en diversos campos y aplicándose más a lo que es el Deep Learning. La diferencia de la FANN en contraste con su predecesora es su arquitectura con un conjunto de neuronas organizadas por capas y su algoritmo de entrenamiento que se ajusta a la muestra, ejecutando el proceso de aprendizaje que la hace crecer más \cite{salas2004redes}.\\

\doublespacing

### Recurrent Neural Networks (Redes Neuronales Recurrentes)

\spacing{1.5}

Aunque no será parte de las técnicas para análizar los datos, importante conocer esta técnica. Las Recurrent Neural Networks (RNN), son reconocidas por obtener información de datos secuenciales como lo son el procesamiento del lenguaje natural, videos y subtitulación de imágenes. Las RNN su modelo asigna parámetros únicos para representar a cada dato en una secuencia \cite{arana2021redes}, para poder tener el control y no sufrir interrupciones sobre la secuencia. Su arquitectura es multicapa que comparte pasos entre los datos espaciados secuencialmente, para poder unir la información.\\

\par Su arquitectura se va incrementando con la conexión de nodos adyacentes a través de la adición de ciclos dentro de la red.\\

\doublespacing

## Deep Learning (Aprendizaje Profundo)

\spacing{1.5}

Por falta de tiempo y tipo especial de captura de información, no será posible analizar esta técnica, pero es importante conocerla y profundizar un poco en ella. La tecnología está al servicio del hombre e impulsa muchos aspectos de la sociedad moderna. Las técnicas de Machine Learning, están limitadas en el procesamiento de los datos naturales en forma cruda y para dar solución a la problemática se creó el aprendizaje profundo. La comprensión de la Inteligencia Artificial y como puede llegar a igual a los comportamientos humanos, inclusive en el aprendizaje, a veces superándonos, fue un gran acontecimiento que se debe a la gran contribución de Alan Turing \cite{Carola}. El Deep Learning (DL) como subconjunto del Machine Learning entra en acción cuando los datos tienen demasiadas características, son enormes, se requiere de un nivel de precisión altísima y el ML no puede ofrecer completamente los resultados deseados.\\

\doublespacing

### Entrenamiento de Redes Neuronales del Deep Learning

\spacing{1.5}

La forma para poder crear y entrenar los modelos de Deep Learning depende del método a implementar, para ello, al mismo tiempo, existen concepciones de procesos similares que son para los modelos basados en este tipo de aprendizaje.\\

\par En primera instancia, se inicia con el entrenamiento desde cero, tomando un conjunto amplio de datos etiquetados, diseñando la red arquitectónica que aprenda \cite{Pang2017}, de igual modo se determina la transferencia de aprendizaje, donde implica el ajuste detallado de un modelo previamente que ya fue entrenado, como puede ser GoogleLeNet que es un sistema que está relacionado con el Deep Learning. \\

\par En segunda instancia, para trabajar con cantidades masivas de datos se requiere una extracción de características que será un enfoque más especializado, en vista de que las capas con sus tareas asignadas gozan de ciertas características de aprendizaje, estas características se pueden extraer en cualquier momento para darle mayor flexibilidad, clasificación y posible capacidad de regresión, como lo es el ejemplo de las máquinas de vectores de soporte (SVM) \cite{Dong2018}.\\

Por último, lo relacionado respecto al tiempo es fundamental, dado que los principales problemas que enfrenta la tecnología es el tiempo, siendo que hay algoritmos que quizás se pueden demorar hasta semanas con los datos. Para esto, el uso de recursos de cómputo mediante GPU, reduce el tiempo para entrenar una red y acota el tiempo de entrenamiento para resolver un problema, perfeccionando el modelo de Deep Learning \cite{Pang2017}.\\

\doublespacing

### Convolutional Neural Networks (Redes Neuronales Convolucionales)

\spacing{1.5}

El Deep Learning ha demostrado muy buenos resultados para la resolución de problemas, en cambio, las limitaciones, sobre todo en el campo de la imagenología, ha hecho que se elabore un método diferente para que exista un análisis más preciso al momento de analizar la imagen. Este método se llama Convolutional Neural Networks (CNN), que tiene una arquitectura con mejor rendimiento para las tareas complejas relacionales \cite{Pena-Torres}. \\

\par Desde el año 2012, las arquitecturas basadas en CNN para visión artificial han crecido muchísimo; sin embargo, no todas han sido eficientes para ocuparse en tareas de visión artificial (Figueroa Flores, 2021), siendo el Grupo de Geometría Visual (VGG) de la Universidad de Oxford \cite{Simonyan2015} una de las arquitecturas más utilizadas para las tareas de procesamiento de visión artificial. \\

\par Las CNN se forman usando tres tipos de capas, los cuales son las capas convolucionales, capas de pooling y capas totalmente conectadas.\\

\begin{figure}[h]

\centering

\includegraphics[scale=0.7]{img/Marco Teorico/convulcioinales.png}

\caption{Descripción del funcionamiento de una red neuronal convolucional (CNN) }

\end{figure}

\par Para la CNN existen dos arquitecturas básicas de CNN, la cuales son CNN que entrega una salida para toda la imagen, como en la figura anterior y la Fully Convolutional networks que posee un codificador y decodificador, entrega una compresión de la información y su salida es por pixel. En las arquitecturas CNN exiten varios ejemplos, como lo son AlexNet, entrenado con ImageNet, GoogLeNet donde posee 22 capas y sus neuronas son más complejas, entre otras \cite{NIPS2012\_c399862d}.\\

\begin{figure}[h]

\centering

\includegraphics[scale=0.8]{img/Marco Teorico/rnn.PNG}

\caption{Diagrama de grafo ciclico de una red neuronal recurrente (RNN) }

\end{figure}

\doublespacing

### Neural Networks for Scarce Data Domains (Redes neuronales para dominios de datos escasos)

\spacing{1.5}

En algunos casos, los modelos no contemplan con demasiados datos y por ende no se puede ser preciso con total certeza de la predicción de los resultados de la tarea. Por tanto, esto que se apunta a crear modelos que existan pocas muestras etiquetadas disponibles \cite{Carola}. Fei-Fei et al. \cite{Fei-Fei2006} demostraron que es posible aprender nuevas categorías, una o pocas muestras por clase, aprovechando las categorías aprendidas anteriormente. \\

\doublespacing

### Dataset

\spacing{1.5}

\par Los Dataset están relacionados con el Big Data y con el procesamiento de datos, porque son datos tabulados, donde cada columna representa una variable y las filas datos, en consecuencia, el Dataset es una base de datos que se puede obtener de plataformas \cite{Astudillo2021}. Los Datasets, nos ayudan a trabajar con volúmenes de datos más pequeños, ya que la principal función de los Datasets es que hacen referencia a una única base de datos de origen. Estos Datasets se clasifican según su origen y formato, que son utilizados según las necesidades de su modelo, existiendo tipos de archivo, carpeta, bases de datos y web, cada uno con su formato. Se pueden encontrar con grandes voleumnes de información con fuentes privadas como gratuitas, siendo una de ellas Google Dataset Seach.\\

\doublespacing

\section{Aspectos de la salud y la enfermedad}

\spacing{1.5}

\lettrine[lines=4, slope=0.6em, findent=-1em, nindent=0.6em]{L}{a} salud es una de los aspectos más importantes en la vida, muchas veces nuestra salud se ve afectada por factores externos o internos a nuestra persona. “Este concepto involucra un estado completo de bienestar físico, mental y social, y no solamente la ausencia de afecciones o enfermedades” (Wold Health Organization, 1946). Estas palabras son relatadas en la conferencia del 46’ que hacen la referencia a que la persona no solo se compone de un estado físico, sino que hay más aspectos para sentirse saludable, ya que, si uno llegará a fallar, habría un déficit en la persona y poco a poco empezaría a carecer de un estado equilibrado que lo proporciona la salud. \\

\doublespacing

\subsection{Accidente Cerebro Vascular}

\spacing{1.5}

\par Un Accidente Cerebro Vascular es la detención del flujo de sangre a una parte del cerebro. En ocasiones se le llama como “Ataque Cerebral”, al igual, si el flujo sanguíneo se detiene por más de pocos segundos, el cerebro no recibirá nutrientes ni oxígeno, causando una muerte y un daño permanente en la zona afectada \cite{Garcia2019}.

\par Existen dos tipos de ACV, los cuales son Isquémicos y Hemorrágicos.\\

\doublespacing

\subsection{Accidente Cerebro Vascular Isquémico}

\spacing{1.5}

\par Los isquémicos es el más común, generalmente, es causado por un coagulo sanguíneo (masas que se presentan cuando la sangre se endurece, pasando de liquida a sólida) que bloquea el vaso sanguíneo del cerebro, empezando a morir en cuestión de minutos las células cerebrales. Existen los ACV Isquémicos transitorios y los permanentes \cite{Garcia2019}, diferenciándose en que los transitorios solo es cuando la sangre no llega al cerebro por unos instantes y el permanente es cuando ya tiene un tiempo prolongado el ACV, también conocido como Infarto Cerebral.\\

\doublespacing

\subsection{Síntomas ACV Isquémico}

\spacing{1.5}

\par Los síntomas de un ACV Isquémico pueden ser:

\begin{itemize}

\item Entumecimiento o debilidad repentina de la cara, brazo o pierna (especialmente en un lado del cuerpo).

\item Confusión repentina, dificultad para hablar o entender el lenguaje.

\item Dificultad repentina para ver con uno o ambos ojos.

\item Problemas para caminar repentino, mareos, pérdida de equilibrio o coordinación.

\item Problemas para caminar repentino, mareos, pérdida de equilibrio o coordinación.\\

\end{itemize}

\doublespacing

\subsection{Categorías de los ACV}

\spacing{1.5}

\par En los ACV para ayudar a optimizar el tratamiento específico, existen categorías que son identificadas por la escala de TOAST \cite{Adams1993}.

\par La primera categoría es la enfermedad Aterotrombótica aterosclerótica de gran vaso, se basa en la reducción de tejido sanguíneo medio o grande en el cerebro, con ubicación cortical o subcortical, con localización vertebrobasilar o carotídea, donde se encuentra presente una aterosclerosis u obstrucción con estenosis u oclusión de las arterias craneales. También la aterosclerosis sin estenosis con menos factores de riesgo se puede encontrar presente en esta categoría. La segunda categoría es el Cardioembolismo, es una reducción de tejido sanguíneo medio o grande, de localización cortical, en la que existe una cardiopatía embolígena \cite{Molina2018}. La tercera categoría es la enfermedad oclusiva de pequeño vaso infarto lacunar, es una reducción de tejido sanguíneo de tamaño pequeño, en el sector de una arteria de perforante cerebral puede provocar una oclusión en el transporte de nutrientes. La cuarta categoría se debe a otras causas, siendo que son un tamaño o una localización variable que además no están en las tres categorías anteriores, por ello, se pueden producir enfermedades metabólicas, alteraciones de la coagulación, displacía fibromuscular, etc. La quinta categoría hace énfasis a los orígenes desconocidos con estudios incompletos o completos, por más de una etiología \cite{Radu2017}.\\

\doublespacing

\subsection{Tratamiento y Actualizaciones en los ACV}

\spacing{1.5}

\par Para los accidentes cerebro vasculares, los factores que pueden influir son muchísimos y hay algunos que influye hasta los fumadores afectando al sistema nervioso central \cite{Piloto2020}.

\par Las ayudas diagnósticas, nos proveen información sobre el grado de lesión y la identificación de la lesión \cite{Wintermark2013} como las imágenes. La recomendación es que la vía aérea y con asistencia ventilatoria en pacientes con ACV, siendo que se deberían lograr saturaciones de oxígeno mayores al 94\%, también la temperatura > 38 grados Celsius y agregando a todo la anterior se debe monitorizar la hiperglucemia, porque si perdura por más de 24 hrs se asocia a un peor desenlace \cite{Garcia2019}.\\

\doublespacing

# \chapter{TRABAJOS RELACIONADOS}

\spacing{1.5}

\lettrine[lines=4, slope=0.6em, findent=-1em, nindent=0.6em]{E}{l} análisis del estado del Arte contempla una revisión bibliográfica, contextualizando los avances en la investigación acerca de la Inteligencia Artificial y el área de salud. Se agrupará en dos tipos: la primera es sobre programas o métodos de prevención de Accidentes Cerebros Vasculares y la segunda sobre las investigaciones que se han realizado, considerando diferentes variables del problema con Machine Learning.\\

\par Los avances en los hospitales en los países desarrollados han permitido la implementación de la IA en sus sistemas, pero con esto nacen nuevos desafíos \cite{Nagendran2020}, Google es uno de los partidarios en algoritmos con IA que aprenda a partir de la experiencia desde 0 a partir del del Machine Learning, pero en la actualidad el problema es procesar grandes volúmenes de información o a veces los datos pueden estar incompletos debido a la incompatibilidad de los sistemas en que se registran, también la presión de entregar una investigación o producto sin prolijidad y beneficio real para las personas.\\

\par Con el aumento de las tecnologías en el campo medico y avance de las técnicas de Machine Learning, se ha desarrollado un interés por el mundo científico para predecir algún proceso anticipado o secuela después de un ACV. Investigadores en el 2019 \cite{Heo2019}, realizaron estudio utilizando Random Forest, Redes Neuronales y Regresión Logística para la predicción de prognosis de un paciente de ACV Isquémico tres meses después del evento inicial. Los autores deseaban predecir la mortalidad a los 3 años luego de salir de la rehabilitación con un algoritmo basado en Arboles de decisión. El mejor modelo fue Random Forest con la implementación del minority oversampling technique, el cual logró un AUC de 0.928 \cite{Scrutinio2020}. Yu et al. \cite{Yu2020} usaron técnicas de Machine Learning considerando el Árbol de Decisión, siendo el objetivo clasificar la severidad del Accidente Cerebro Vascular Isquémico. El árbol se construyó con 13 variables de 18 propuestas y los datos usados son los del National Institutes of Health Stroke Scale con personas mayores a 65 años. Con esta técnica se logró tener un accuracy del 91.11%. prediciendo el nivel de discapacidad, si poseía discapacidad o no, la información recogida de las primeras 24 horas en el hospital luego de 90 días \cite{Xie2018}. Los predictores incluían información de exámenes de escáner, demografía e información clínica de los pacientes.\\

\par Con la herramienta de las imágenes nace la posibilidad de generar algoritmos para un método más actualizado de la detección de esta enfermedad y futura prevención, ya que existen variadas técnicas \cite{Wintermark2013} como la RM con DWI para la evaluación de presencia y extensión de isquemia posterior y la CTA y DSA para la trombosis de la arteria. Siguiendo con la investigación en imágenes en un estudio en Estados Unidos \cite{Garcia2019} se analizaron 610.000 casos y 185.000 son recurrentes en el 2019, además los signos y síntomas producen lesiones visibles para la imagenología, se pudo llegar a que el manejo médico y prevención secundaria es vital para que el proceso sea exitoso. Tomando en cuenta los beneficios y riesgos para tomar un tipo de terapia especifica y oportuna, es importante educar a la comunidad para reconocer algunos síntomas del ACV y así poder acudir al centro médico más cercano. Finalmente, esta enfermedad llegará a ser un 6,2\% para los países desarrollados, con estos datos mencionados es importante crear un algoritmo para la atención basado en experiencias internacionales como nacionales \cite{Garcia2019} y es igual de importante crear un modelo que nos permita predecir esas secuelas o futuros problemas en el tratamiento para llevar a cabo una prevención exitosa evitando la recurrencia.\\

\par En los diseños la IA puede generar estudios de Aprendizaje Profundo como, estudio de revisión sistemática del diseño, estándares de informes y afirmaciones de los estudios de aprendizaje profundo \cite{Pang2017}, el objetivo siempre es que el modelo con el Aprendizaje Profundo sea efectivo y eficaz, pero se debe ayudar con médicos expertos, ellos evalúan los diagnósticos mediante imágenes y contrastan los resultados obtenidos con los de la IA, teniendo en cuenta que esta investigación cuenta con una gran fuente de datos como Ensayos controlados, datos Medline y ensayos que la Organización Mundial de la Salud poseía desde el 2010 hasta el 2019, en ellos se encontraron registros aleatorios de aprendizaje profundo con bajo nivel de sesgo. En síntesis, la información que es emitida por los modelos de aprendizaje profundo, de igual forma es manipulada por los expertos, que este caso son los médicos, ya que los algoritmos \cite{Nagendran2020} son experimentales y aunque son pioneros, no pueden dejar en total confianza al algoritmo para que determine un diagnóstico más certero.\\

\par Las redes de datos convolucionales demostraron en Corea que el uso de ellas es una herramienta que predice con precisión los cuidados intensivos en servicios médicos (Kang et al., 2020), asumiendo que el modelo predictivo, basado en Deep Learning, es superior a las otras herramientas de predicción y puntuaciones convencionales \cite{Bioetica2022}. Cabe destacar, que el algoritmo de aprendizaje es muy eficiente por la cantidad de capas que puede poseer el modelo, puesto que entre más capas mayor puede ser el aprendizaje. Pese a las evidencias que demuestran de la efectividad y el desempeño del modelo, pueda ser mejor al del humano, existe un miedo por la implementación en los sistemas de salud, lo que puede provocar que el crecimiento del Deep Learning, perezca de una base amplia para su desarrollo \cite{Nagendran2020}.\\

\par En el área de la implementación de un modelo con CNN, encontramos el trabajo de Chunjiao Dong , Chunfu Shao,1,2 Juan Li, and Zhihua Xiong del 2018, que desarrolla específicamente la predicción sobre los accidentes de tránsito \cite{shao2018improved}. Ellos demuestran una técnica novedosa con un modelo de regresión multivariable, que presenta la relación entre lo examinado y los accidentes de tránsitos. Como resultados el módulo identifica las variables de entrada y representaciones de características de salida, aunque se haya reducido su magnitud, se conserva la información original. Además, el modelo propuesto explica mejor los problemas de heterogeneidad en predicción de accidentes de tráfico y puede ser aplicado a casos similares.\\

\begin{figure}[h]

\centering

\includegraphics[scale=0.55]{img/Marco Teorico/predicción.PNG}

\caption{Resultados de la predicción de choque de accidentes}

\end{figure}

\par En este caso el modelo propuesto en contraste con el SVM en la categoría choque con daños menores es significativo (29.961\% versus 61.350\%), así es como la predicción medida por el RMSD se puede mejorar un 84,58\% y un 158,27\% en comparación con el modelo de aprendizaje profundo sin la capa de regresión y el modelo SVM.\\

# 6 -. Capítulo de Tesis: Diseño Metodológico

## 6.1 Método

El diseño metodológico dará una guía con los pasos a seguir a grande escala para la obtención de la finalización del proyecto. Se tendrá en cuenta el tipo, el enfoque, la población y la muestra para iniciar el trabajo. Siguiendo con estrategia, utilizaremos el libro de \emph{Machine Leaning in Action} \cite{Harrington2012}, utilizando el procedimiento de la sección \emph{“Steps in developing a machine learning application”} que consta de 7 pasos para la implementación de una aplicación que utiliza técnicas de Machine Learning.

\begin{enumerate}

\item \textbf{Colección de los datos de entrada:} El primer paso para implementar una aplicación que trabaje utilizando técnicas de Machine Learning es coleccionar los datos que serán analizados, hay muchas formas en las que se pueden obtener grandes cantidades de datos a través de internet, este paso se refiere a esas opciones y al esfuerzo aplicado para obtener dichos datos pudiendo ahorrar algo de tiempo y esfuerzo en adquirir datos públicos.

\item \textbf{Preparación de datos de entrada:} Una vez que ya se obtienen los datos,

es necesario asegurarse que estén en el formato correcto para ser procesados por el algoritmo de Machine Learning seleccionado. El beneficio de tener este formato estándar es que puede mezclar y combinar algoritmos y fuentes de datos.

\par Por lo que este paso involucra, si fuera necesario, formatear los datos para

adaptarlos a la necesidad de cada algoritmo.

\item \textbf{Analizar los datos de entrada:} Este

podría ser tan simple como mirar los datos que ha analizado en un editor de texto para hacer

Asegúrese de que los pasos 1 y 2 realmente funcionen y no tenga un montón de valores vacíos.

También puede mirar los datos para ver si puede reconocer algún patrón o si

hay algo obvio, como algunos puntos de datos que son muy diferentes de

el resto del conjunto.

\item \textbf{Implementar algoritmo:} Con los pasos anteriores realizados y verificados, ahora hace falta implementar el algoritmo con el leguaje de programación que escogimos, asegurándonos que la información de los datos de entrada sea acorde a nuestra técnica empleada y al formato que poseemos. En este paso, debe quedar la mayor cantidad de documentación posible relacionada al desarrollo.

\item \textbf{Entrenamiento del algoritmo:} Aquí es donde tiene lugar el aprendizaje automático. Este paso y el siguiente paso es donde se encuentran los algoritmos "básicos", según el algoritmo. Alimentas el algoritmo con buenos datos limpios de los primeros dos pasos y extraer conocimiento o información. Este conocimiento que a menudo almacena en un formato eso es fácilmente utilizable por una máquina para los siguientes dos pasos.

\item \textbf{Entrenamiento del algoritmo:} Aquí es donde se encuentra la información aprendida en el paso anterior. Cuando esté evaluando un algoritmo, lo probará para ver qué tan bien funciona. En el caso del aprendizaje supervisado, tiene algunos valores conocidos que puede utilizar para evaluar el algoritmo. Si no está satisfecho, puede volver al paso 4, cambiar algunas cosas e intentar probar nuevamente. A menudo la recopilación o preparación de los datos puede haber sido el problema, y ​​usted hay que volver al paso 1.

\item \textbf{Uso del algoritmo:} Una vez se han consumado todos los pasos, no queda

más que usar el algoritmo, esta etapa implica tener que volver a ejecutar los pasos 1, 2, 3 y 5.

\end{enumerate}

\\

\par Sabemos que el proceso de programación podemos desarrollarlo óptimamente en los leguajes más usados para el Machine Learning que son Java, R o Python. Par este proyecto se escogió Python por su simpleza y múltiple documentación para el desarrollo de las técnicas de Machine Learning.\\

## 6.1.2 Enfoque o método de Investigación

La metodología de este trabajo será de enfoque cuantitativa por los resultados que se quieren llegar a obtener, donde primará el análisis matemático de la base de datos y entre los algoritmos a comparar.

## 6.1.2 Tipo de Investigación

El tipo de investigación será de carácter experimental, ya que medirá tendencias en los resultados arrojados por los algoritmos.

## 6.1.3 Población

La población será todas las personas que salen registradas en la base de datos a trabajar. En este caso específico, son los pacientes post Ataque Cerebrovascular Isquémico.

## 6.1.4 Muestra

La muestra serán los pacientes post Ataque Cerebro Vascular Isquémico registrados en la base de datos.

## 6.2 Procedimientos de la Investigación

**6.2.1-. Dominio del problema**

Dentro del análisis de los algoritmos, debemos dominar el problema revisando la bibliografía relacionada con el área que realizaremos la investigación, así como son los enfoques del Machine Learning (informática) y los ACV (salud), en definitiva, estas dos áreas se juntan para dar lugar a los modelos y tecnologías actuales que se implementan en ambas áreas o por separado, en el ámbito nacional o internacional (Objetivo Específico 1).

**6.2.2-. Creación del modelo**

En caso de no existir modelos de Machine Learning con ACV, se investigarían modelos parecidos para contrastar lo introducido en un nuevo modelo a la ciencia. A fin de todo esto habrá que elegir un mecanismo adecuado de Machine Learning (Objetivo Específico 2). En este punto también constará con un modelo como arquitectura, definiendo un marco claro para la interacción con el código fuente del software (Objetivo Específico 3).

**6.2.3-. Implementación: Construcción y eficiencia del Modelo**

Una vez creado el modelo, se baja hasta el cuadro de Implementar y realizaremos lo que son la recolección de datos. Los datos serán de tipo etiquetado, como datos algoritmos, datos estadísticos, entre otros, que nos ayuden al entrenamiento de la red neuronal, en base a la arquitectura planteada. Por consecuencia, la estrategia para la creación de la red neuronal, será aprendida desde el marco teórico para la creación del modelo, siendo que los datos deben ser procesados y ajustados al modelo paralelamente. esos datos. A continuación, una vez finalizadas las tareas antes mencionadas, vendrá la tarea de Visión Data, que consiste en analizar un problema de perdida de información cuando la base de datos no está completa y para eso debemos estandarizar los datos y aplicar una estadística de normalización.

**6.2.4-. Experimentación del Modelo**

Para tener un buen modelo debemos realizar, en primera instancia, la tarea de prueba para ver si funciona o no, con casos clínicos que fueron seguidos durante un transcurso prolongado de tiempo, luego tendríamos la tarea de validación, que sería si logra completar con lo que se solicita para ser utilizado el modelo por otros profesionales.

**6.2.5-. Resultados de la Investigación**

Los Resultados son la parte más importante de la investigación, en donde la tarea será reflejar lo que obtuvimos y deseábamos, o simplemente tener unas métricas desde otro modelo y compararlas con nuestros resultados. Se destaca que el estado del Arte del capítulo 6, nos ayudar a tener una comparativa importante con los resultados de modelos actuales y con los resultados que entrega el modelo creado.

Después de todo lo anterior, las conclusiones serán la interpretación de los resultados y la investigación.

Los estudios experimentales se realizarían en el Hospital Herminda Martin, con pacientes que estén pasando por un ACV.

DESCRIPCIÓN DEL AMBIENTE DE SOFTWARE (SI CORRESPONDE)

## 6.3 Técnicas para la Recolección de Datos

## 6.3.1 Descripción de los Instrumentos

## 6.3.1.1 Representación de los datos

Los datos estarán dispuestos en una base de datos estática tipo hoja de cálculo, siendo que pueden ser transformado a ficheros tipo CVS que son más cercanos a la librería pandas, para poder ser usados por modelos de predicción en Python.

## 6.3.1.2 Elección de variables significativas

-Estimación del ritmo.

- Frecuencia Respiratoria.

## 6.3.1.3 Datos y normalización

Con la base de datos en nuestro poder, ahora es tiempo de procesar los datos en búsqueda de ausencia de datos (valores nulos).

* Normalización por Media y Desviación Estándar
* Normalización por Máximos y Mínimos

## 6.3.2 Validez y Confiabilidad de los Instrumentos

## 6.4 Técnica para el Procesamiento y Análisis de Datos

Entrenamiento

## 6.5 Resguardos Éticos

SI ESTE ANTEPROYECTO HA CONTADO CON LA COLABORACIÓN DE PERSONAS O INSTITUCIONES PARA LAS CUALES, FINALMENTE SERÁ DESARROLLADO EL PROYECTO, POR FAVOR IDENTIFICAR:

# Capítulo 7 Resultados

## 7.1 Predicción del tipo de Accidente Cardiovascular

Aquí se explica que cual de los algoritmos es más efectivo para predecir que es Isquémico.

## 7.2 Predicción de pronóstico

Debería haber datos de monitorización y de acuerdo a eso ver si el paciente vivirá o morirá. Tiene que ver más con los signos.

## 7.3 Predicción de aumento o disminución del total de variables

Aquí se hará un análisis del paciente para saber cuanto tiempo se tardará en recuperarse lo antes posible de todos los síntomas.

## 7.4 Predicción de reincidencia

Aquí se evaluará si el paciente volverá a ser reincidente. Depende de la base de datos a analizar, se podría poner en contraste los síntomas de antes vs los de la segunda vez.

# Capitulo 8 Discusión, Conclusiones y Recomendaciones

FIN :)

Relleno

\par El ACV posee clasificaciones y que se actualizan constantemente, ya que es una enfermedad grave que conlleva alteraciones significativas en la vida del paciente. Además se puede enfatizar que existe un confiable sistema de clasificación, debido a que este comprende un 75,2\% de las muertes y el 81\% de los años de vida ajustados por discapacidad que genera esta enfermedad en los países en desarrollo, asimismo existe un 87\% con el ACV Isquémico \cite{Radu2017}, ) son una clasificación confiable y precisa para la enfermedad a través del TOAST, aunque la clasificación fenotípica ha avanzado a pasos agigantados, aún se necesita más investigación a nivel mundial para lograr un lenguaje más estándar en la investigación Fenotípica.\\