



BENEMERITA UNIVERSIDAD AUTONOMA DE  
PUEBLA

---

---

FACULTAD DE CIENCIAS  
FÍSICO MATEMÁTICAS

Comparación entre niños obesos y con  
normopeso vía análisis del hipotálamo con  
redes neuronales.

T E S I S

Para obtener el título de:

Lic. en física

presenta:

María Lucero Sánchez Ramírez

Asesores:

Dr. Benito De Celis Alonso  
Dr. Javier M. Hernández López.

Ciudad Universitaria, PUEBLA 2022

*A mi familia*

# Agradecimientos

En primer lugar quiero dar gracias a mi mamá Eva, gracias mamá porque sin tu sacrificio al irte a otro país y sin tu amor no hubiera podido llegar hasta aquí.

En segundo lugar quiero agradecer a mi familia, porque gracias a sus presiones constantes pude ponerle más empeño a mi trabajo. A mi hermana Ariana porque sin su dulzura, ternura y una que otra travesura cuando estaba estresada no hubiera logrado calmarme y a mi perrita Luna quien estuvo conmigo en todas esas noches de desvelo, acompañándome hasta que me fuera a dormir.

En tercer lugar, quiero agradecer a mis amigos Adrián y Arianna, son a quienes conocí en la facultad y con quienes pasé mis años de licenciatura, infinitas horas de estudio en las palapas, cubículos o en la biblioteca, gracias a ellos sentí más amena la dificultad de algunas materias, sobre todo en esta era del COVID. También a mi amigo Fer, por adentrarme al mundo de las redes neuronales, sin sus enseñanzas no hubiese logrado entender los conceptos de mis redes.

Por último y no por ello menos importante a mis asesores, el doctor Benito de Celis Alonso y el doctor Javier Miguel Hernández, gracias por darme la oportunidad de trabajar en un proyecto así, por responderme aun cuando yo me demoraba más del tiempo establecido para las correcciones, por ayudarme cuando no sabía que rumbo darle a este trabajo.

Quiero agregar una mención especial a la persona más importante en mi vida, mi madre Adriana, una mención tan importante como ella en mi vida porque sin sus palabras de aliento, apoyo emocional y su infinito amor no hubiera logrado terminar esta tesis.

# Índice general

AGRADECIMIENTOS .....	II
RESUMEN.....	1
CAPÍTULO 1 INTRODUCCIÓN .....	2
1.1. JUSTIFICACIÓN DEL PROBLEMA.....	2
1.2. HIPÓTESIS .....	3
1.3. OBJETIVOS PRIMARIOS Y SECUNDARIOS DEL TRABAJO.....	3
1.4. RESUMEN SOBRE ESTRUCTURA DE LA TESIS.....	4
CAPÍTULO 2 FUNDAMENTO TEÓRICO .....	5
2.1. ANATOMÍA Y FISIOLÓGIA .....	5
2.1.1. <i>Anatomía y fisiología general del cerebro</i> .....	5
2.2. OBESIDAD.....	16
2.2.1. <i>Obesidad infantil</i> .....	21
2.2.2. <i>Alteraciones</i> .....	23
2.3. RESONANCIA MAGNÉTICA .....	26
2.3.1. <i>Fundamentos de resonancia magnética</i> .....	29
2.3.2. <i>BOLD</i> .....	32
2.4. INTELIGENCIA ARTIFICIAL .....	34
2.4.1. <i>Redes Neuronales</i> .....	36
Aplicación de Redes Neuronales.....	38
2.4.2. <i>Machine Learning</i> .....	40
Aprendizaje supervisado .....	44
Aprendizaje no supervisado .....	46
Deep Learning o Aprendizaje profundo.....	48
2.4.3. <i>Redes Neuronales Convolucionales</i> .....	49
Capas convolutivas .....	51
Capas de pooling.....	53
2.4.4. <i>Clasificación de imágenes</i> .....	54
Arquitectura.....	54
LeNet.....	54
AlexNet.....	55
VGGNet.....	56
GoogLeNet.....	58
ResNet.....	59
DenseNet.....	60
CAPÍTULO 3 METODOS.....	62
3.1. POBLACIÓN PARA ESTUDIAR.....	62
3.2. HARDWARE Y SOFTWARE.....	62
3.3. OBTENCIÓN Y PROCESAMIENTO DE IMÁGENES.....	64
3.4. DISEÑO O ESTRUCTURA DE LA RED NEURONAL.....	65
3.5. ESTADÍSTICA.....	69
CAPÍTULO 4 RESULTADOS.....	74

<b>CAPÍTULO 5 DISCUSIÓN Y CONCLUSIONES.....</b>	<b>80</b>
<b>5.1. RESUMEN DE LOS RESULTADOS.....</b>	<b>80</b>
<b>5.2. LIMITACIONES DEL STUDIO .....</b>	<b>82</b>
<b>5.3. TRABAJO FUTURO .....</b>	<b>83</b>
<b>CAPÍTULO 6 BIBLIOGRAFÍA .....</b>	<b>84</b>
<b>CAPÍTULO 7 APÉNDICE I. CÓDIGO DE LA RED NEURONAL .....</b>	<b>90</b>

# Índice de figuras

FIGURA 2.1: RELACION GENERAL DEL SNC Y SNP. MODIFICADO DE REFERENCIA (1). .....	5
FIGURA 2.2: DIVISIONES PRINCIPALES DEL SNC. MODIFICADO DE REFERENCIA (1). .....	6
FIGURA 2.3: ANATOMÍA MACROSCÓPICA DEL SNC. MODIFICADO DE REFERENCIA (3). .....	7
FIGURA 2.4: NEURONA TÍPICA. ....	12
FIGURA 2.5: LÓBULOS DE LA CORTEZA CEREBRAL. MODIFICADO DE REFERENCIA (3). .....	15
FIGURA 2.6: RELACIÓN ENTRE LOS PLANOS DE IMAGEN Y EL ENCÉFALO. EL DIAGRAMA MUESTRA LA ORIENTACIÓN HABITUAL DE UN PACIENTE EN UN EQUIPO DE RM Y LOS PLANOS DE ADQUISICIÓN DE LAS CUATRO EXPLORACIONES (POTENCIADAS EN T1). A Y B, IMÁGENES CORONALES. C Y D, IMÁGENES AXIALES. MODIFICADO DE REFERENCIA (1). .....	26
FIGURA 2.7: REPRESENTACIÓN DE LA DISTRIBUCIÓN DE LA ORIENTACIÓN DE MAGNETIZACIÓN DE ESPÍN EN UNA EXPLORACIÓN DE CUERPO ENTERO CON RESPECTO AL CAMPO MAGNÉTICO EXTERNO. MODIFICADO DE REFERENCIA (15). .....	27
FIGURA 2.8: IMAGEN POTENCIADA EN T1 EN FUNCIÓN DE LOS TIEMPOS DE REPETICIÓN Y DE ECO. MODIFICADO DE REFERENCIA (16). .....	28
FIGURA 2.9: GIRO DE PROTONES. MODIFICADO DE REFERENCIA (15). .....	30
FIGURA 2.10: GRAFICA SEÑAL/TIEMPO DE LA RELAJACIÓN LONGITUDINAL Y TRANSVERSAL DE DOS TEJIDOS. MODIFICADO DE REFERENCIA (16). .....	31
FIGURA 2.11: BLOQUE F EN LA RED. MODIFICADO DE REFERENCIA (24). .....	37
FIGURA 2.12: ENCADENACION DE BLOQUES F. MODIFICADO DE REFERENCIA (24). .....	37
FIGURA 2.13: CLASIFICADOR LINEAL. MODIFICADO DE REFERENCIA (23). .....	45
FIGURA 2.14: ALGORITMO DE AGRUPAMIENTO. MODIFICADO DE REFERENCIA (23). .....	47
FIGURA 2.15: DEEP LEARNING. MODIFICADO DE REFERENCIA (23). .....	48
FIGURA 2.16: ARQUITECTURA DE LENET. MODIFICADO DE REFERENCIA (27). .....	55
FIGURA 2.17: ARQUITECTURA DE ALEXNET. MODIFICADO DE REFERENCIA (28). .....	56
FIGURA 2.18: ARQUITECTURA DE VGG. MODIFICADO DE REFERENCIA (28). .....	57
FIGURA 2.19: ARQUITECTURA DE VGG-16. MODIFICADO DE REFERENCIA (29). .....	58
FIGURA 2.20: ARQUITECTURA DE VGG-19. MODIFICADO DE REFERENCIA (30). .....	58
FIGURA 2.21: ARQUITECTURA COMPLETA DE GOOGLNET. MODIFICADO DE REFERENCIA (28). .....	59
FIGURA 2.22: ARQUITECTURA DE RESNET. MODIFICADO DE REFERENCIA (27). .....	60
FIGURA 2.23: ARQUITECTURA DE DENSENET. MODIFICADO DE REFERENCIA (28). .....	61
FIGURA 3.1: DIAGRAMA DE LA ESTRUCTURA DE UNA RED NEURONAL CONVOLUCIONAL. MODIFICADO DE REFERENCIA (34). .....	69
FIGURA 4.1: GRAFICA MODELO LOSS VS PACIENTES. ....	75
FIGURA 4.2: GRAFICA MODELO ACCURACY VS PACIENTES. ....	76

# Resumen

Las redes neuronales convolucionales son una de las ramas más importantes del mundo de la informática, esto gracias a que los algoritmos de machine learning son capaces de entrenar a computadoras que simulan el comportamiento humano. El presente trabajo se utiliza para crear una red neuronal convolucional que sea capaz de clasificar a pacientes obesos y normopesos, utilizando tensorflow y dentro de este la librería keras. El modelo establecido es una clasificación utilizando los valores 1 para obeso y 0 para normopeso. La solución es la creación de un código de entrenamiento y otro de validación de imágenes.

# Capítulo 1

## Introducción

### 1.1. Justificación del problema

Este proyecto de investigación se centra en el análisis de imágenes anatómicas segmentadas del hipotálamo de niños de entre 7 y 9 años, ello con el fin de estudiar las diferencias entre personas con normopeso y personas con obesidad. El estudio se realizó mediante redes neuronales convolucionales.

Las redes neuronales son la base de la inteligencia artificial, se han convertido en la familia de algoritmos de Machine learning más populares a lo largo del tiempo. Son usadas para múltiples cosas, entre ellas el pronóstico de enfermedades, en este caso fueron utilizadas para la clasificación de imágenes. Las redes neuronales están compuestas por múltiples neuronas agrupadas. Existen varios tipos de redes neuronales pero la usada aquí fue la red neuronal convolucional para el procesamiento y clasificación de imágenes. Consta de capas que se encargan de hacer convoluciones y maxpooling, es decir, al dar una imagen al algoritmo las convoluciones y el maxpooling reducen el tamaño de la imagen conforme avance en la estructura, cuando se reduce el tamaño de la imagen cada capa identifica los elementos más importantes y cada una de estas consta con un nivel de abstracción más elaborados.

Las redes neuronales cuentan con capas de entradas las cuales reciben los pixeles de la imagen, después pasa toda la información a las capas ocultas y son ahí donde se encuentran las convoluciones y el pooling. Siendo el pooling una operación para reducir los tamaños de las imágenes y las convoluciones para pasar filtros a la imagen para detectar patrones que serán importantes dentro de esta misma.

Las redes neuronales convolucionales son usadas tanto para texto como para imágenes o video. Al contar con las imágenes de pacientes se usó este tipo de red para clasificarlas. La clasificación de imagen fue de obesos y normopesos. Esto fue posible una vez segmentados los hipotálamos de los pacientes.

## **1.2. Hipótesis**

Es posible crear una red neuronal que pueda clasificar los hipotálamos de pacientes con normopeso y obesidad con una exactitud cercana al 80%.

## **1.3. Objetivos primarios y secundarios del trabajo**

Estudiar el volumen y estructura del hipotálamo segmentado de personas con normopeso y obesos en imágenes anatómicas por resonancia magnética. Se pretende usar redes neuronales convolucionales para la clasificación de aspectos de interés (biomarcadores). En este caso se espera que la red sea capaz de diferenciar el hipotálamo de cualquiera de estos dos grupos de voluntarios.

- Segmentación de distintos tejidos del cerebro en imágenes de resonancia magnética.
- Cálculo de volúmenes tisulares en imágenes de resonancia magnética.

- Desarrollar una red neuronal convolucional capaz de realizar clasificaciones lo más precisas posible.

## **1.4. Resumen sobre estructura de la tesis**

En el capítulo 1 se menciona la justificación del problema, la hipótesis, los objetivos primarios y secundarios, también una descripción breve de cada capítulo de la tesis.

En el capítulo 2 se encuentra el fundamento teórico, se describe la anatomía y fisiología del cerebro, la obesidad, resonancia magnética, inteligencia artificial, donde en este último se centra en redes neuronales convolucionales.

En el capítulo 3 se describen los métodos, se explican las herramientas utilizadas para construir la red neuronal convolucional, iniciando con los pacientes a estudiar, especificaciones del hardware y software usados. Se continúa con una explicación de la estructura de la red neuronal y se finaliza con la estadística.

En el capítulo 4 se explican los resultados, argumentando sobre la función de pérdida y la exactitud, además de la comprobación de la hipótesis.

En el capítulo 5 se discuten los resultados obtenidos, las limitaciones de este estudio y el trabajo futuro.

En el capítulo 6 se muestra la bibliografía de los libros, artículos y revistas consultados para esta tesis.

En el apéndice I se muestra y describe el código de la red neuronal usado.

# Capítulo

# 2

## Fundamento teórico

### 2.1. Anatomía y Fisiología

#### 2.1.1. Anatomía y fisiología general del cerebro

El sistema nervioso humano se divide en dos; sistema nervioso central (SNC) y sistema nervioso periférico (SNP). El SNC está constituido por el encéfalo y la médula espinal, estas son las estructuras más protegidas del cuerpo ya que están en el interior del cráneo y de la columna vertebral. El SNP está compuesto por nervios que conectan en encéfalo y la médula espinal como se muestra en la figura 2.1 [1].

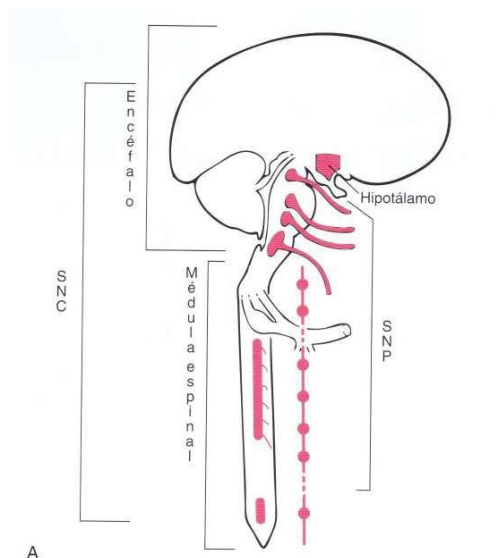


Figura 2.1: Relación general del SNC y SNP. Modificado de referencia (1).

El sistema nervioso central está formado por el encéfalo y la médula espinal. El encéfalo es la parte del sistema nervioso central contenido en el cráneo y el cual comprende el cerebro, el cerebelo y el tronco encefálico. En la figura 2.2 se observa que la médula espinal es la parte del sistema nervioso central situada en el interior del canal vertebral y se conecta con el encéfalo a través del agujero occipital del cráneo. El SNC recibe, integra y correlaciona distintos tipos de información sensorial así como también es la fuente de nuestros pensamientos, emociones y recuerdos. Integran esta información a través de funciones motoras que viajan por nervios del SNP y este ejecuta una respuesta adecuada. [2]

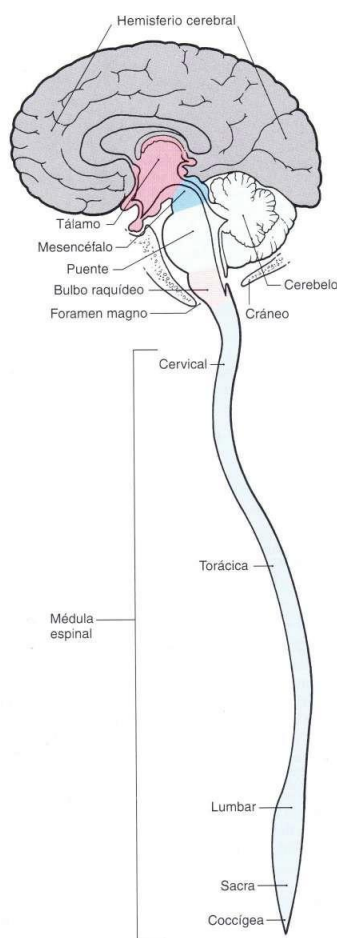


Figura 2.2: Divisiones principales del SNC. Modificado de referencia (1).

El sistema nervioso periférico está formado por nervios que conectan el encéfalo y

la médula espinal con otras partes del cuerpo los nervios que se originan en el encéfalo son llamados nervios craneales y a su vez los que se originan en la médula espinal son llamados nervios raquídeos y espinales. [2]

Se puede ver al SNC de manera macroscópica en la figura 2.3 [3]:

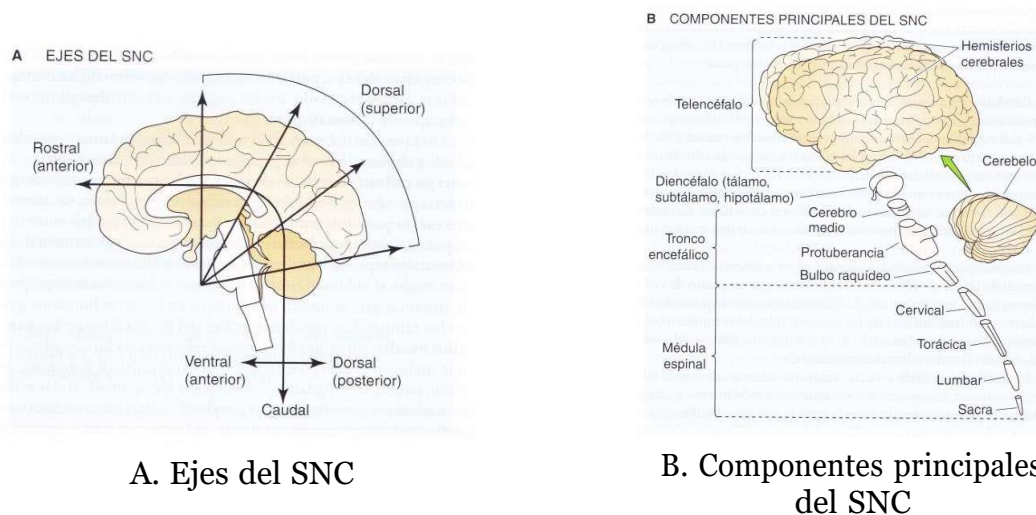


Figura 2.3: Anatomía macroscópica del SNC. Modificado de referencia (3).

El SNC bien puede dividirse en:

- Telencéfalo

La corteza cerebral humana es la porción más sobresaliente de los pares de hemisferio cerebrales. La corteza cerebral tiene un superficie aproximadamente  $2.200 \text{ cm}^2$  y se calcula que contiene  $1.5-2 \times 10^{10}$  neuronas. Es la capacidad de procesamiento de la información de esta maquinaria neural abarca una amplia gama de funciones: razonamiento, aprendizaje, memoria y consciencia. [3]

La corteza está organizada de dos formas. Primera, áreas de la corteza mediante funciones concretas. Segundo, dentro de una porción de la corteza que gestiona una función específica las regiones del cuerpo, están cartografiadas de forma

ordenada. [3]

Otra porción del telencéfalo es la gran masa de axones que salen y entran a la corteza y se conectan con otras regiones. El volumen de axones necesarios para interconectar las neuronas corticales aumenta exponencialmente en función del volumen cortical, de este modo, el volumen relativo de sustancia blanca con respecto a la sustancia gris es 5 veces mayor en el humano que en los ratones. [3]

La porción final del encéfalo son los ganglios basales, un grupo de neuronas relacionadas funcionalmente que incluyen el cuerpo estriado, el globo pálido, la amígdala y la formación hipocampal. Los ganglios basales tienen conexiones indirectas con porciones motoras de la corteza cerebral y están implicadas en el control motor. La amígdala participa en la expresión de la emoción. La formación hipocampal es crucial en la formación de nuevos recuerdos. [3]

- Cerebelo

Representa solamente un 10% del volumen del SNC, contiene aproximadamente el 50% de todas las neuronas del SNC. Muchas conexiones transportan información al cerebelo entre éstas la información visual y auditiva. [3]

El cerebelo puede dividirse en tres partes: el vestibulocerebelo, el espinocerebelo y el cerebrocerebelo. El vestibulocerebelo ayuda a mantener el equilibrio del cuerpo, relacionado con el sistema vestibular, sus sensores se sitúan en el oído interno y las estaciones de paso se localizan en la protuberancia y el bulbo raquídeo. El espinocerebelo ayuda a regular el tono muscular, ya que recibe un gran número de aferencias procedentes de los receptores de estiramiento musculares a través de conexiones en la médula y el tronco encefálico. El cerebrocerebelo coordina el comportamiento motor, este recibe un gran número de proyecciones desde las porciones sensoriomotoras de la corteza cerebral a

través de neuronas en la protuberancia. [3]

Así mismo el puente y el cerebelo están originados en el mismo segmento del tubo neural en desarrollo, en los adultos el puente forma parte del tronco del encéfalo y el cerebelo es una estructura supra segmentaria situada posterior al tronco del encéfalo. El puente contiene muchas somas neurales, algunos organizados en núcleos de los nervios craneales, atravesados por vías ascendentes y descendentes. [4]

El cerebelo está conectado con varias regiones del SNC y es considerado parte del sistema motor, es encargado de coordinar las actividades de los grupos musculares individuales que producen movimientos voluntarios finos y sinérgicos.

- **Diencéfalo**

Esta región consta del tálamo, el subtálamo y el hipotálamo. El tálamo es la estación integradora principal para la formación sensorial hacia la corteza cerebral, donde alcanza el nivel de percepción consciente, así como ciertos aspectos de la memoria. Junto con el subtálamo, el tálamo recibe proyecciones desde ganglios basales importantes para la función motora. Las aferencias al tálamo desde el cerebrocerebelo son importantes para el control motor normal. El hipotálamo es la estructura del SNC, realiza esta función a través de conexiones directas con núcleos Autónomos del tronco encefálico y la médula espinal. También actúan sintetizando hormonas y transportándolas distalmente a través de sus axones hasta la hipófisis posterior donde se secretan a la sangre, también, sintetizan hormonas liberadoras y las secretan hasta el sistema portal que las transporta a células en la hipófisis anterior y allí estas hormonas liberadoras estimulan las células para secretar hormonas hacia el torrente sanguíneo. [3]

Por muy pequeño que parezca el hipotálamo participa en la conducta sexual, la

alimentación, en la secreción de hormonas hipofisarias, en la regulación de la temperatura corporal y en funciones vegetativas. [4]

- Tronco encefálico (consta del mesencéfalo, la protuberancia y el bulbo raquídeo)

El tronco encefálico descansa por encima de la médula espinal, el mesencéfalo, la protuberancia y el bulbo raquídeo recibe información sensorial aferente y envían señales motoras eferentes y aferentes a través de nervios llamados pares craneales. [3]

El mesencéfalo posee motoneuronas que controlan el movimiento ocular. También contiene grupos de neuronas que están implicadas en transmitir señales relacionadas con la audición y la visión. [3]

Al mesencéfalo se le sitúa la protuberancia que contiene motoneuronas que controlan la masticación, el movimiento ocular y los músculos faciales. Recibe información sensorial desde la cara, el cuero cabelludo, la boca y la nariz. Así como El proceso de la información con la audición y el equilibrio. [3]

En esta parte del mesencéfalo continua el puente, esta parte del encéfalo es el enlace entre el tronco del encéfalo y el prosencéfalo. Las vías ascendentes y descendentes del prosencéfalo deben atravesar el mesencéfalo ya que en este se encuentran los núcleos de los nervios craneales oculomotor y troclear, también participan en las vías visuales y auditivas reflejas, en la función motora, en la transmisión del dolor y en funciones viscerales. [4]

El bulbo raquídeo con motoneuronas que inervan a los músculos del cuello y la lengua. Éste junto con la protuberancia se encarga del control de la presión arterial, la frecuencia cardíaca, la respiración y la digestión. [3]

- Médula espinal

La médula espinal es la continuación del bulbo raquídeo, consta de 31 segmentos, cada uno de ellos con raíz nerviosa motora y sensorial. Estas raíces se combinan para formar 31 pares de nervios raquídeos bilateralmente simétricos. [3]

El sistema nervioso se compone de neuronas y células gliales. Las neuronas son unidades básicas estructurales y funcionales del sistema nervioso, se encargan de recibir información, transmitir impulsos eléctricos e influir en otras neuronas o tejidos. Cada información llega a la neurona a través de las prolongaciones llamadas axones, que terminan en la neurona en uniones especializadas denominadas sinapsis, dichas sinapsis pueden ser localizadas en las prolongaciones neuronales llamadas dendritas o en el cuerpo celular neuronal denominado soma o pericarion. [4]

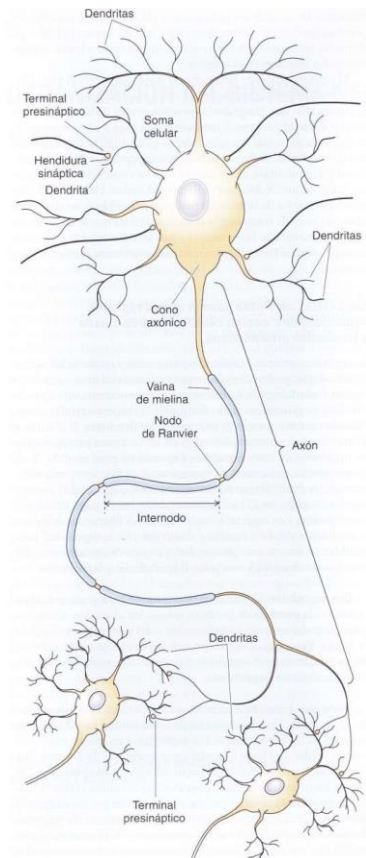
Las células gliales proporcionan a las neuronas soporte estructural y mantienen el microambiente apropiado esencial para la función neuronal. Constituyen la mayoría de las células del sistema nervioso y es requerida por la función cerebral normal. Los tipos principales de células gliales en el SNC son los astrocitos y los oligodendrocitos. Los tipos celulares análogos en el SNP son las células satélites, las células Schwann y los macrófagos. [4]

Se denomina soma o pericarion al cuerpo celular de una neurona. El pericarion es el centro metabólico de la célula nerviosa. Tiene un alto consumo de energía de las células y es por eso que tiene abundantes mitocondrias. [4]

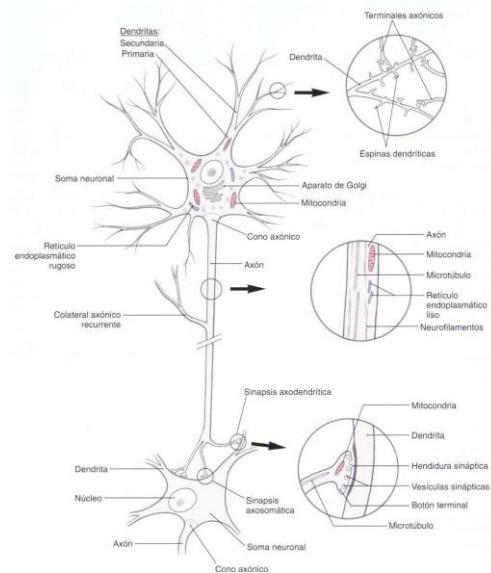
El axón es originado del soma, de una elevación llamada como axónico. Es una proyección que surge del cuerpo celular al igual que las dendritas pero su punto de origen es una región de forma cónica conocida como cono axónico. [4]

El tipo más común de sinapsis en el SNC engloba un terminal axónico de una neurona situado en aposición a una dendrita.

Las dendritas reciben señales desde otras neuronas a través de contactos axonales (sinapsis) que están formados en las superficies como se observa en la figura 2.4. Se ramifican extensamente en la proximidad del soma, provocando una apariencia de arbusto o árbol. Sus ramas aumentan de grosor a medida que se acercan al soma. Las dendritas finales distales son estructuras de las localizaciones de la síntesis proteica y están asociadas a la formación de la memoria [4].



A. Morfología de una neurona típica. Modificado de referencia (3).



B. Representación esquemática de una neurona multipolar típica. Modificado de referencia (4).

Figura 2.4: Neurona típica.

Los astrocitos se encuentran en todo el SNC. Son células ramificadas con prolongaciones que contactan con la superficie de las dendritas y somas neurales. [4]

Los oligodendrocitos aparecen en la sustancia gris y en la sustancia blanca. Su función es la mielinización, es decir, la formación de una vaina electroquímicamente aislante alrededor de los axones, excepto los de la sustancia blanca. Una vaina mielina es una envoltura membranosa alrededor de un axón. [4]

Las neuronas son muy complejas por eso presentan una morfología y una fisiología elaborada. Una neurona recibe aferencias de miles de otras neuronas y combina estas señales por una membrana celular para transformarlos en un mensaje a través de su axón.

Cada región del cerebro posee distintas clases de neuronas y cada una tiene una fisiología adaptada para llevar a cabo funciones específicas. [5]

El sistema nervioso regula algunos mecanismos fisiológicos de manera independiente del control voluntario. El control de la temperatura corporal es un ejemplo de un proceso que la mayoría de los individuos no puede regular conscientemente, así como la presión arterial y la frecuencia cardíaca. [4]

El sistema nervioso autónomo (SNA) se divide en tres partes: simpático, parasimpático y entérico. Las divisiones simpática y parasimpática tienen porciones en el SNC y en el SNP. La división entérica depende del SNP. [4]

Todas las divisiones tienen conexiones aferentes y eferentes, aunque las eferentes

suelen recalcar en el SNA.

Otra parte del sistema nervioso, es el córtex cerebral el cual es el órgano del pensamiento, constituye las funciones intelectuales que nos caracterizan como humanos y que nos hacen individuos únicos, estas funciones incluyen la capacidad para emplear el lenguaje y la lógica, así como ejercitar la imaginación y el juicio. [6]

El córtex cerebral las primeras células que migran abandonan a las células gliales radiales y ocupan posiciones cercanas a las superficies ventricular. Las sucesivas de neuronas inmaduras, se abren paso a través de las capas de células diferenciadas para tomar posiciones más cerca de la superficie pial. Estas filas de células forman la placa cortical a medida que los axones de las células previamente establecidas crecen hacia sus dianas. [7]

El córtex cerebral se forma a partir de la expansión de la parte superficial de la zona intermedia, la subplaca y la placa cortical. Los axones en desarrollo que se originan en regiones como el tálamo que inervan el córtex, envían y forman contactos sinápticos transitorios en la subplaca. Los somas neuronales abandonan el área entre la subplaca y la superficie ventricular; donde la mayoría de los somas celulares restantes son células gliales y esta región forma la sustancia blanca del sistema nervioso adulto.

El córtex cerebral es una densa agregación de somas neuronales que va de 2 y 4 mm de espesor, cuya área total es de aproximadamente  $2.5 \text{ cm}^2$ . Las neuronas de este órgano reciben información a través del tálamo y también de otras regiones por medio de fibras de asociación. Está dividido en diferentes áreas funcionales, algunas están dedicadas al procesamiento de la información sensitiva, otras a la organización de la actividad motora y a las funciones intelectuales superiores. Entre las funciones están la memoria, juicio, planificación de actividades complejas, procesamiento del lenguaje, cálculos

matemáticos y la construcción de una imagen interna. [6]

El córtex cerebral se divide en dos hemisferios: el hemisferio derecho y el hemisferio izquierdo, a su vez cada hemisferio está dividido en cuatro lóbulos. Estos lóbulos están delimitados por cuatro cisuras cerebrales: cisura central, cisura lateral, cisura parieto-occipital y la cisura cingular. Las áreas de corteza cerebral que quedan entre estas cisuras son lo que conocemos como lóbulos y se aprecian en la figura 2.5 [8]:

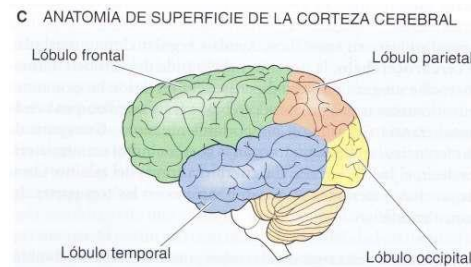


Figura 2.5: Lóbulos de la corteza cerebral. Modificado de referencia (3).

- **Lóbulo frontal:**

Es el lóbulo cerebral más grande de la corteza cerebral. Está ubicado en la parte delantera, justo detrás de la frente. Se extiende desde la parte más anterior hasta la cisura central. Es el centro de mando y control del cerebro humano: "el gran director de orquesta". Está íntimamente relacionado con las funciones ejecutivas (Miller, 2000; Miller & Cohen, 2001), por lo que se involucra en la planificación, el razonamiento y la resolución de problemas, el juicio y el control de impulsos, y también en la regulación de emociones, como la empatía y la generosidad, y la conducta

- **Lóbulo temporal:**

Queda separado de los lóbulos frontal y parietal por la cisura lateral y los límites del lóbulo occipital. Se encuentra situado en los laterales derecho e izquierdo del cerebro. Interviene en el procesamiento auditivo y del lenguaje. También está implicado en las funciones de memoria y en la gestión de emociones.

- Lóbulo parietal:

Ocupa el área comprendida entre la cisura central y la parte superior de la cisura parietooccipital. Está encargado de la integración de la información sensorial. Contribuye al procesamiento del dolor y del tacto entre otras importantes funciones.

- Lóbulo occipital:

Está delimitado por los extremos posteriores de los lóbulos parietal y temporal (en la parte trasera del cerebro). Se ocupa principalmente de la visión (aunque delega el procesamiento de algunos subprocesos de la visión a otros lóbulos) (Kosslyn, 1994). Analiza aspectos como la forma, el color y el movimiento para interpretar y sacar conclusiones de las imágenes visuales.

## 2.2. Obesidad

Se define a la obesidad como la acumulación excesiva de la grasa corporal. Esto está regido por múltiples factores entre los que se incluye: edad, metabolismo, educación y el sexo. De manera eficaz se usa un parámetro que relacione la altura y el peso para cuantificar la obesidad. Es el parámetro que es conocido como índice de masa corporal. Está expresado mediante la siguiente ecuación [9].

$$IMC = \frac{\text{peso (kg)}}{\text{estatura}^2(\text{m}^2)} \quad (2.1)$$

El peso que se considera saludable y normal en una persona es lo que se llama normopeso, es el parámetro de IMC ideal, lo cual implica que el metabolismo quema adecuadamente la grasa y es utilizada para un buen funcionamiento del cuerpo. En niños recién nacidos el IMC es de aproximadamente  $13.9 \text{ kg/m}^2$ , alcanza los  $15.9 \text{ kg/m}^2$

a los 6 años y de ahí hasta los 18 años logra estabilizarse.

El normopeso o también llamado peso ideal es el peso del cuerpo humano más adecuado para la salud. Es el peso que se considera saludable para una persona.

El rango según la Organización Mundial de la Salud (OMS) en adultos es considerado normal si es de 18.5 a 24.9  $kg/m^2$ , sobrepeso de 25 a 29.9  $kg/m^2$  y obesidad es mayor a 30  $kg/m^2$ , existen tres tipos de obesidad, para el I es de 30 a 34.9  $kg/m^2$ , para el II de 35 a 39.9  $kg/m^2$  y el III  $\geq 40 kg/m^2$ .

La causa fundamental del sobrepeso y la obesidad es un desequilibrio energético entre calorías consumidas y calorías gastadas, universalmente tienen una mayor ingesta de alimentos ricos en grasa, sal y azúcares pero pobres en vitaminas, minerales y otros micronutrientes.

Otro de los aspectos es la disminución de la actividad física, producto del estilo de vida sedentario debido a las actividades laborales, métodos de transporte y de la vida urbana.

La grasa corporal es un componente fundamental del organismo, necesaria para un completo funcionamiento, entre las cuales se pueden destacar:

- Proteger los órganos vitales
- Almacenar vitaminas esenciales
- Amortiguar las articulaciones
- Mantener y regular la temperatura corporal

Sin embargo una acumulación de grasa superior a la necesaria puede ocasionar problemas de salud en diferentes maneras, no toda la grasa corporal es igual ya que se distribuye de las siguientes maneras [10]:

- Grasa dura

Ubicada entre los órganos en la zona del abdomen, su función principal es amortiguar y proteger los órganos vitales pero por otra parte la acumulación excesiva de esta grasa es una de las más peligrosas ya que cuesta eliminarla y tiene una relación directa con problemas cardiovasculares o diabetes tipo II, es más común en hombres.

- Grasa blanda

Situada bajo la piel, es más sencilla de eliminar y mucho menos peligros por estar lejos de los órganos vitales, este tipo de acumulación es más común en mujeres en las zonas de la cadera, muslos y glúteos.

- Grasa intramuscular

Localizada en las fibras musculares para protegerlas y se incrementa con el paso de los años, este tipo de grasa es igual de común tanto como en mujeres como hombres.

- Grasa parda

Llamada así porque su color varía del dorado al marrón rojizo, se caracteriza por presentar células grasas. La grasa parda ejerce una importante función en el feto y recién nacido, llegando a representar el 2-5 % del peso corporal. Este tipo de grasa se encuentra localizada entre las escápulas, en las axilas, en la nuca, y

alrededor de los grandes vasos del tronco. La función principal es termogénica durante el primer año de vida, al final del cual se creía prácticamente desaparecida en su totalidad, transformándose en grasa blanca o amarilla. Sin embargo, hoy ya se acepta su existencia también en la edad adulta.

- Grasa blanca

La grasa blanca recibe esta denominación por contraposición a la grasa de color pardo o marrón. Las funciones de la grasa blanca pueden resumirse en cuatro principales: sintetizar lípidos a partir de excedentes de hidratos de carbono o proteínas; responder a estímulos hormonales y nerviosos; secretar sus propias hormonas y la más clásica de todas, actuar como reservorio de energía, formando, almacenando y descomponiendo ácidos grasos en equilibrio con la concentración correspondiente en el torrente sanguíneo, aunque recientemente está surgiendo como fuente importantísima de células madre adultas .

- Grasa beige

Es un tipo de grasa situada dentro de la grasa blanca.

Existen diferentes tipos de obesidad, primeramente se dividen en dos tipos: obesidad tipo androide (manzana) y obesidad tipo ginoide (pera). La primera está localizada por encima de la cintura y está relacionada con la diabetes, triglicéridos y colesterol. El otro tipo de obesidad está localizada bajo la cintura, este tipo es menos dañino, sin embargo, está relacionada con el problema de circulación en las piernas.

Sumado a esto, existen seis subtipos de obesidad [11]:

- **Obesidad alimentaria**

Localizada en la parte superior del cuerpo: cara, cuello, etc. Causada principalmente por el consumo de alcohol, tabaco y mala alimentación.

- **Obesidad abdominal nerviosa**

Localizada en el abdomen y cintura, es causada principalmente por el origen de la depresión, nervios y la ansiedad.

- **Obesidad glútea**

Localizada en mujeres, en la zona de la cadera y los muslos. Generalmente es desarrollada en la pubertad, sin embargo, también es causada en el embarazo o cuando ocurre un desajuste hormonal, pero si se da en el último caso los problemas son más serios.

- **Obesidad metabólica aterógena**

Es más comúnmente generada por la genética, el vientre se inflama y puede ocasionar problemas respiratorios.

- **Obesidad circulatoria venosa**

Localizada en extremidades inferiores y superiores (piernas y brazos), comienza en forma de celulitis en jóvenes y suele agravarse con el paso del tiempo, además de que provoca problemas circulatorios.

- Obesidad por sedentarismo

Aparece cuando no se tiene ningún tipo de actividad física, si sumado a esto no se tiene una buena alimentación la situación tiende a empeorar.

### **2.2.1. Obesidad infantil**

La obesidad infantil es un problema mundial que afecta a países de medianos y bajos recursos. Es muy probable que los niños obesos, sigan siendo obesos en una edad adulta. Esto llevaría asociado un mayor riesgo de enfermedad cardiovascular, diabetes mellitus tipo 2 y otras complicaciones metabólicas [12].

En niños obesos definir el concepto de obesidad es un tanto difícil ya que en ellos hay un aumento de masa libre de grasa. Muchos de los niños que presentan un índice de masa corporal dentro de los límites normales tienen valores de masa grasa determinada por petismo grafía por desplazamiento de aire en rasgos de obesidad. El índice de masa corporal presenta además una serie de limitaciones como la necesidad de encontrar la población de referencia ideal los cambios seculares que está puede sufrir. [13]

Es difícil encontrar una clasificación para los niños probablemente la única distinción importante en este tema es que se puede diferenciar aquellos niños que tienen un patrón patológico debido a enfermedades genéticas o endocrinas de otros cuyo obesidad está ligada etiopatogénica mente a la ingesta de dietas hipercalóricas o a la escasa actividad física y quizá también a una predisposición genética para conservar y almacenar energía esto es también llamado obesidad simple llega a marcar hasta un 95 % de los casos. [13]

Se calcula que la prevalencia de la obesidad se ha duplicado desde 1972 a 1990

aproximadamente el 80 % de los niños y adolescentes obesos seguirán presentando esta patología en la edad adulta al analizar una población de adultos efectos de obesidad se observó que tres cuartas partes de los adultos que tenían un exceso de peso del 160 % en relación con su peso ideal habían sido obesos desde la infancia. [13]

Existen múltiples factores por los cuales hay niños con obesidad entre estos están los factores genéticos ambientales, neuroendocrinos, metabólicos conductuales y de estilo de vida. [13]

Aproximadamente del 25 al 35 % de los casos de obesidad ocurren en familias en las cuales el peso de los padres es normal Aunque el riesgo es mayor si los padres son obesos. El riesgo de obesidad de un niño es cuatro veces mayor si uno de los padres es obeso y ocho veces mayor si ambos lo son, así mismo el patrón de distribución de la grasa corporal sigue también la misma tendencia que lo observado en los padres por tanto el riesgo de ser obeso puede estar atribuido al seguimiento de hábitos similares de alimentación en la familia genéticamente predisuestas. [13]

### Factores ambientales

Uno de los principales factores que ha provocado este aumento de la obesidad en nuestra sociedad es el número de horas que se ve la televisión ya que tiene una relación directa con el riesgo de padecer esta enfermedad. La televisión es uno de los factores más importantes de obesidad en adolescentes y muestra un efecto de dosis-respuesta ya que puede verse implicado el hecho de que la televisión anuncia alimentos con alto contenido calórico, los personajes de la televisión general muestran unos hábitos alimentarios malos, los niños tienen más posibilidades de tomar aperitivos mientras ven la televisión y además de que está reemplaza otras actividades al aire libre que consumen más energía, entre estos están sus juegos y deportes [13].

## Factores neuroendocrinos

Los cambios hormonales en pacientes pediátricos obesos que muestran que los niveles séricos son significativamente superiores en niños obesos que en niños no obesos, estos niveles séricos de los que hablamos pueden estar más elevados si el niño tiene obesidad y la mayor parte de estas alteraciones hormonales pueden ser reversibles si el niño pierde peso. [13]

## Factores relacionados con el gasto energético

Se considera balance energético cuando la ingesta energética es igual al gasto, si la Ingesta excede al gasto, Ésta se vuelve en forma de grasa, Estos excesos relativamente pequeños en la ingesta energética pero que se mantienen durante un largo tiempo van a producir unos aumentos significativos en la grasa corporal, aunque, no está definido si la obesidad está causada por una ingesta excesiva o por un gasto reducido. La energía ingerida a través de los alimentos pierde un 5 % con las heces, orina y sudor, y el resto en energía metabolizable. [13]

### **2.2.2. Alteraciones**

Los niños y adolescentes como efectos de la obesidad pueden presentar alteraciones del desarrollo puberal como las que se mencionan a continuación [13]:

- Pubertad adelantada: un gran porcentaje de estos niños tienen crecimiento lineal avanzado, edad ósea igualmente adelantada y en ellos son precoces la maduración sexual y el estirón puberal.

- Pseudohipogenitalismo en varones prepúberes es un problema muy frecuente ya que la grasa suprapúbica oculta la base del pene disminuyendo su tamaño real. [13]
- Ginecomastia este también es un problema frecuente en varones, debido al acumulamiento de grasa en la región mamaria sin existir un aumento real de tejido glandular mamario.
- Algunas niños con obesidad severa también presentan problemas dermatológicos como estrías, infecciones por hongos en las zonas de pliegues y acantosis nigricans, está aparece en relación a la alteración en la sensibilidad a la insulina.
- En cuanto es la función pulmonar de los niños pueden presentar disnea de esfuerzo ante el ejercicio físico moderado e incluso insuficiencia respiratoria caracterizada por el síndrome de pickwick.
- En cuanto a la patología osteoarticular secundaria ya que tienen que soportar un exceso de peso las articulaciones en la infancia, esta es una sobrecarga para el aparato locomotor por eso es frecuente encontrar en los niños obesos algunos trastornos ortopédicos como el Genu valgum y la epifisiolisis de la cabeza femoral, pie plano, escoliosis, coxa vara y enfermedad de Perthes. [13]

Se estima que la obesidad es responsable del 44 % de diabetes, 23 % de cardiopatías isquémicas y entre el 7 y 41 % de algún tipo de cáncer. Sin embargo entre los problemas más comunes que pueden desarrollar se encuentran:

- Glucosa alta en la sangre o diabetes
- Presión arterial alta (hipertensión)
- Diabetes mellitus tipo 2
- Colesterol alto

- Triglicéridos en la sangre
- Cardiopatías
- Accidente cerebrovascular
- Osteoartritis
- Síndrome de ovario poliquístico
- Embolia pulmonar
- Depresión
- Trastornos de la alimentación
- Problemas óseos
- Apnea del sueño
- Cálculos biliares
- Problemas del hígado
- Resistencia a la insulina
- Diferentes tipos de cáncer (útero, cuello uterino, endometrio, ovarios, mama, colon, recto, esófago, hígado, vesícula, páncreas, riñón, etc.)
- Problemas digestivos (acidez estomacal, enfermedades en la vesícula)
- Infertilidad

- Periodos menstruales irregulares en las mujeres.
- Disfunción eréctil en los hombres
- Alteraciones neurológicas
- Riesgo de desarrollar síntomas de gravedad de COVID-19

## 2.3. Resonancia Magnética

Este trabajo está centrado en el desarrollo de algoritmos de redes neuronales, por lo tanto solo se describirán brevemente las bases físicas de la resonancia magnética.

La resonancia magnética (RM) es una técnica de diagnóstico por imágenes médicas. Para realizar un estudio se necesita disponer de un campo magnético estático potente, generado por imanes superconductoras para alinear los spines. También se precisan otro tipo de imanes que generen gradientes de campo magnético, esto para realizar una codificación del volumen a explorar, se observa en la figura 2.6 [14].

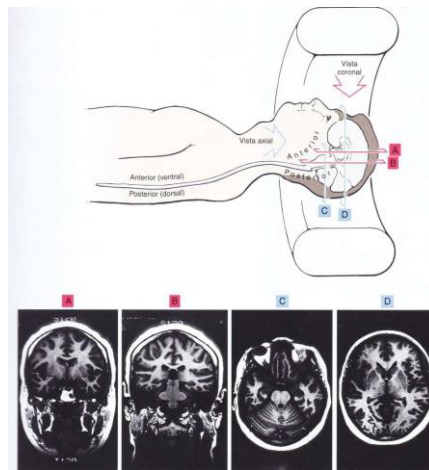


Figura 2.6: Relación entre los planos de imagen y el encéfalo. El diagrama muestra la orientación habitual de un paciente en un equipo de RM y los planos de adquisición de las cuatro exploraciones (potenciadas en T1). A y B, Imágenes coronales. C y D, Imágenes axiales. Modificado de referencia (1).

Los protones hidrógeno constituyen una gran porción de los tejidos corporales. Estos átomos poseen un núcleo y una capa de electrones dos polos norte y sur y giran alrededor de un eje inclinado como pequeños planetas. Mientras los electrones se mueven y rotan con el átomo inducen a una corriente eléctrica que crea un campo magnético, se alinean de forma aleatoria debido a las cambiantes interacciones magnéticas que ejercen unos sobre otros. Cuando los protones se exponen a un campo magnético potente se alinean en paralelo al campo magnético externo a diferentes niveles de energía. Si el campo magnético externo es muy potente la frecuencia de giro para este ángulo será más rápida ya que al someterse a una exploración de resonancia magnética el paciente se convierte en un imán con todos los protones alineados a lo largo del campo magnético externo y girando en un ángulo y a una frecuencia determinada, se aprecia en la figura 2.7 [1].

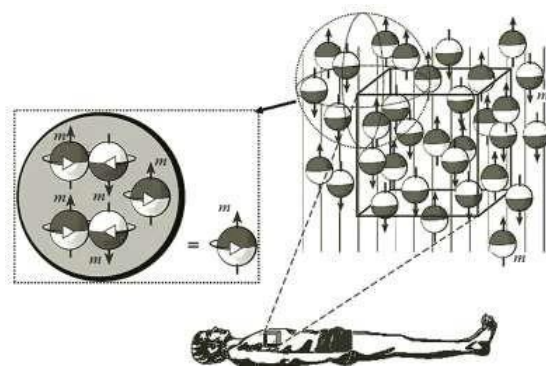


Figura 2.7: Representación de la distribución de la orientación de magnetización de espín en una exploración de cuerpo entero con respecto al campo magnético externo. Modificado de referencia (15).

Cuando una onda de radio o también llamada onda electromagnética se envía en forma de una breve ráfaga hacia el imán que contiene el paciente estamos hablando de un pulso de radiofrecuencia (RF) este pulso de RF tiene diferentes niveles de frecuencia. Cuando un protón gira a una determinada inclinación sólo absorbe la energía de la onda de radio cuando la frecuencia del pulso de RF coincide con la frecuencia del propio protón, esto es a lo que llamamos resonancia, de ahí que recibe el nombre de imagen de resonancia magnética esta resonancia tiene dos opciones: anula los efectos magnéticos de ciertos protones e incrementa los niveles de energía y los efectos magnéticos de otro grupo de protones y cuando la onda de radio cesa los protones estos regresan a su estado

y fuerza de magnetización originales, este fenómeno también es conocido como relajación y está definido por una constante de tiempo denominada  $T_1$ . Los protones alineados en un mayor nivel de energía y magnetización también pierden energía (relajación) y esto se denomina por una constante de tiempo  $T_2$ . El tiempo de relajación  $T_1$  es más prolongado que  $T_2$ . [1]

Los protones relajados liberan su energía en forma de un eco de ondas de radio. Una antena absorbe la información y un ordenador determina las características de las ondas de radio emitidas por los puntos contenidos en la sección del cuerpo examinadas y entonces la imagen de RM es generada y transferida a un monitor en una placa fotográfica, es aquí donde las imágenes potenciadas en  $T_1$  y  $T_2$  pueden obtenerse modificando los tiempos de recepción de los ecos (TE) como se ve en la figura 2.8 [1].

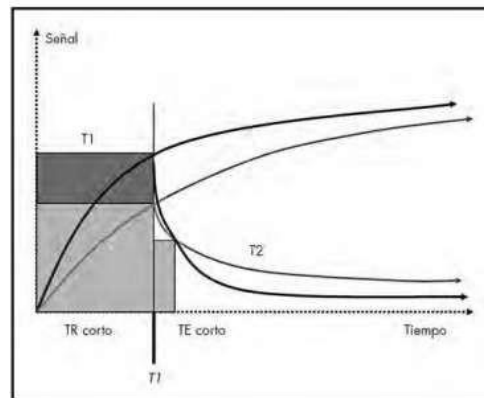


Figura 2.8: Imagen potenciada en  $T_1$  en función de los tiempos de repetición y de eco. Modificado de referencia (16).

La secuencia de Spin eco generan imágenes que pueden ser potenciadas en  $T_1$  o  $T_2$  dependiendo del intervalo de tiempo en milisegundos entre cada onda de radio excitadora, este intervalo se denomina tiempo de repetición (TR). [1]

El tiempo de eco (TE) es el intervalo de tiempo en milisegundos necesario para recoger estas ondas de radio de los protones que se relajan. De acuerdo al pulso Spin-

eco Cuanto más corto sea el TR y el TE más potenciada en  $T_1$  está considerada la imagen y por otro lado Cuanto más largo sea el TR y el TE más potenciada en  $T_2$  es considerada la imagen. [1]

### 2.3.1. Fundamentos de resonancia magnética

En el diagnostico medico utiliza los núcleos del átomo  $^1H$ , que contienen un solo protón. Los protones contienen una propiedad cuántica intrínseca denominada spin asociado al momento angular. [17]

El átomo  $^1H$  se encuentra en las moléculas del agua  $H_2O$ . Todos los órganos y tejidos del cuerpo contienen agua en mayor o menor medida. De esta manera cuando se realiza una resonancia magnética a algún paciente lo que ocurre es que el campo magnético realinea las moléculas de agua y las ondas de radio hacen que estos átomos produzcan señales.

Los protones son partículas compuestas subatómicas con carga positiva que se encuentra en el núcleo de los átomos, formado por tres partículas elementales de espín  $\frac{1}{2}$ , un quark abajo y dos quarks arriba.

El espín es una propiedad de las partículas, su valor solamente adquiere ciertos valores, ya sean discretos o semienteros. Únicamente se puede explicar mediante la mecánica cuántica ya que de manera clásica no se puede explicar la existencia de este [12]. Tiene valores de  $0, \frac{1}{2}, 1, \frac{3}{2}$  y  $2$ , pero solo las partículas con espín distinto de cero son susceptibles a la resonancia magnética.

El protón, como consecuencia de su carga positiva y el spin, se comporta como un

pequeño dipolo magnético  $\mu$ . El momento magnético de los protones se orienta aleatoriamente en todas las direcciones del espacio, cancelándose entre ellos. Cuando se introduce un paciente dentro de un imán de RM los protones interactúan con el campo magnético del imán y les suceden dos cosas simultáneamente: la alineación y la precesión. [17]

Los protones en el núcleo de los átomos tienen una rotación intrínseca alrededor de un eje central, esta rotación se asemeja al giro como se ilustra en la figura 2.9. El ángulo y la velocidad angular de la rotación,  $\omega$ , son intrínsecos, propiedades de cada partícula. El campo magnético  $m$  también es específico para cada partícula o elemento. Esto significa que una carga en movimiento, como un protón, produce su propia carga magnética, que es una función de la carga y la velocidad del movimiento. [15]

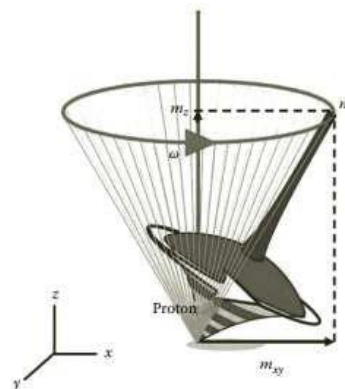


Figura 2.9: Giro de protones. Modificado de referencia (15).

La frecuencia de precesión  $\omega_0$  depende de la intensidad del campo magnético  $B_0$  la cual es de 0.012 y 2 Teslas, con base a este aumenta o disminuye de manera proporcional, está expresado mediante la fórmula de Larmor, donde  $\gamma$  es la constante giromagnética, diferente para cada átomo [17]:

$$\omega_0 = \gamma B_0 \quad (2.2)$$

Los protones están en estado de magnetización y relajación cuando predominan los protones de orientación paralela o de baja energía, formando un vector de

magnetización neto orientado en dirección al campo magnético [18].

La relajación se define como el conjunto de procesos por los cuales la magnetización obtiene un equilibrio tras un pulso de un campo magnético. Es decir, decaimientos exponenciales. Existen dos tipos de relajación: longitudinal ( $T_1$ ) y transversal ( $T_2$ ). Estas dos son constantes de tiempo y  $T_1$  siempre es mayor que  $T_2$ .

Relajación longitudinal ( $T_1$ ) o relajación de espín-retícula es el componente del vector de magnetización ( $M$ ) paralelo al campo magnético ( $B_0$ ) llamado magnetización longitudinal ( $M_z$ ), representa la convergencia de  $M_z$  a  $M_0$  causado por la interacción de los espines con su entorno [19]. Necesario para alinear los protones en un campo magnético estático.

Relajación transversal ( $T_2$ ) o relajación espín-espín es el componente del vector de magnetización ( $M$ ) perpendicular al campo magnético ( $B_0$ ) llamado magnetización transversal ( $M_{xy}$ ), representa una pérdida de coherencia intrínseca e irreversible de los protones [19].

Estos dos tipos de relajación se muestran en la figura 2.10.

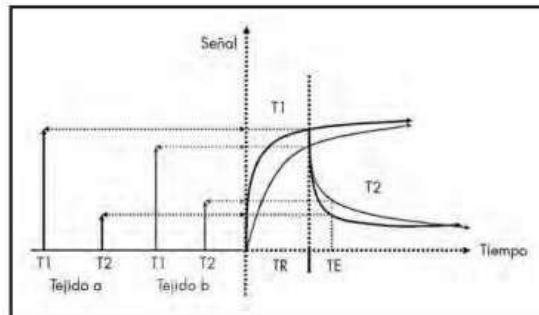


Figura 2.10: Grafica señal/tiempo de la relajación longitudinal y transversal de dos tejidos. Modificado de referencia (16).

La relajación de la RM es descrita matemáticamente mediante ecuaciones diferenciales de primer orden, conocidas como las ecuaciones de Bloch. Después de un pulso de ángulo de punta arbitrario  $\alpha$ , el valor de  $M_z$  en un tiempo  $t$  es dado por [20]:

$$M_z(t) = M_0 \cos \alpha + (M_0 - M_0 \cos \alpha)(1 - e^{-t/T_1}) \quad (2.3)$$

Los diferentes tejidos del cuerpo tienen diferentes valores de  $T_2$  y se utilizan para diferenciar entre tejidos sanos y enfermos en imágenes clínicas [20].

### 2.3.2. BOLD

En 1992 la universidad de Minnesota y del Medical College of Wisconsin, mostraron un registro de resonancia magnética en humanos con técnicas no invasivas basadas en las propiedades magnéticas de la hemoglobina.

Mediante experimentación con ratas, Ogawa probó que la desoxihemoglobina en la sangre podría ser usada como un medio de contraste intrínseco en imágenes obtenidas por RM, fenómeno que recibió la denominación de contraste- bold (blood oxygenation level dependent) o técnica dependiente de nivel de oxigenación sanguínea. [21]

Para hacer una resonancia magnética funcional mediante la técnica bold se requiere de un equipo multidisciplinario de varias áreas. La resonancia magnética funcional es una técnica de diagnóstico por imagen sensible a los cambios hemodinámicos que acontecen durante la actividad cerebral. La relación entre el flujo sanguíneo cerebral (FSC) y la intensidad de señal (IS) de la secuencia bold es una secuencia compleja. [21]

La desoxihemoglobina es la hemoglobina sin moléculas unidas de oxígeno y la

oxihemoglobina es la hemoglobina unida al oxígeno.

La secuencia bold se fundamenta en los efectos de la hemoglobina sobre el campo magnético circundante en función de su porcentaje de oxigenación: la oxihemoglobina tiene un efecto despreciable sobre el campo magnético, mientras que la desoxihemoglobina causa heterogeneidad sobre el campo magnético. En consecuencia, la desoxihemoglobina causa una pérdida de señal  $T_2^*$  más rápida y, en consecuencia, una disminución en la intensidad de señal  $T_2^*$  de alrededor del 1 -5 % en equipos mayores de 1.5 T. [21]

El  $T_2^*$  local es determinado por el balance de la desoxihemoglobina a hemoglobina oxigenada dentro de un voxel. Al producirse actividad neuronal, se incrementa el flujo de la sangre oxigenada cerebral en una región focal determinada, lo cual aumenta la intensidad de señal dentro del voxel en  $T_2^*$  y es a lo que se llama función de autoregulación arterial local. [22]

La intensidad de señal producida en un voxel en el cerebro, puede disminuir significativamente si hay más sangre desoxihemoglobina en una región y aumenta si hay más oxihemoglobina dentro de una región, distribuirlo a través de la red veno-capilar. [22]

La pérdida de señal se considera proporcional a su concentración, la cuál depende del hematocrito, el flujo y el volumen sanguíneo cerebral, así como el estado de oxigenación. Estas variables cambian cuando la activación cerebral produce un cambio en el metabolismo neuronal. [21]

## 2.4. Inteligencia artificial

Según la definición clásica de Elaine Rich, la inteligencia artificial es el estudio de cómo hacer que los ordenadores hagan cosas que, por ahora, los humanos hacemos mejor. Los orígenes de la inteligencia artificial se remontan en el año 1956, donde investigadores como John McCarthy, Marvin Minsky, Nathaniel Rochester, Claude Shannon, Arthur Samuel, Herbert Simon y Allen Newell tuvieron un evento donde propusieron que “cualquier aspecto del aprendizaje o de cualquier otra característica de la inteligencia puede describirse de manera tan precisa que se puede construir una máquina para simularlo”. [23]

Entre 1955-1956 Allen Newell, Herbert Simon y Cliff Shaw escribieron el primer programa con Inteligencia artificial el cual era capaz de imitar las habilidades de resolución de problemas de un ser humano a la hora de demostrar teoremas, dicho programa fue capaz de demostrar 38 de los 52 primeros teoremas de dos Principia Mathematica de Alfred Whitehead y Bertrand Russell. Para el año 1959 crearon un programa más sofisticado de demostrar teoremas llamado GPS (General Problem Solver). Cada uno de ellos tres tenía diferentes optimismo en cuanto a lo que creían que sería en un futuro las máquinas sin embargo no tardó mucho en que se dieran cuenta de que lo que esperaban no sería un total éxito ya que los arrojaba a un callejón sin salida un ejemplo claro es la traducción automática que resultó no ser nada fácil ya que si traducimos del inglés al ruso una frase y luego este resultado del ruso al inglés nos van a robar una frase completamente diferente ya que se enfrentan a las innumerables matices del lenguaje natural.

En los años setenta se diseñaron los primeros sistemas basados en el conocimiento: los sistemas expertos, fueron llamados así porque pretendía contemplar o sustituir las habilidades de expertos en dominios específicos que requerían un conocimiento experto que no era de fácil acceso, entre estos se encontraban [23]:

- Sistemas para identificar compuestos orgánicos a partir de la información proporcionada por un espectrógrafo de masas (DENDRAL).
- Diagnóstico y tratamiento de enfermedades infecciosas (MYCIN).
- Exploración de posibles yacimientos minerales (PROSPECTOR).
- Elaboración de la lista de componentes necesarios para la configuración de un sistema informático (R1/XCON).

Una de las técnicas en donde la Inteligencia artificial despierta un gran interés es el de una de las capacidades humanas que nos diferencian de la mayor parte de los animales [23]:

- El juego de las damas: Arthur Samuel Programa un ordenador que fuese capaz de aprender a jugar a las damas a finales de los años cincuenta y traza algunas horas de práctica Este programa era capaz de jugar mejor que su programador internamente utilizaba un algoritmo llamado poda alfabeta el cual permite analizar las consecuencias de los movimientos de cada jugador sin explorar todas las ramas del árbol de búsqueda asociado al juego. Este juego fue el primer programa de ordenador capaz de ganarle a un campeón mundial.
- El juego de ajedrez estaba programado en C y se ejecutaba bajo el sistema operativo AIX el cuál era capaz de evaluar 200 millones de posiciones por segundo.
- El juego de backgammon, el programa de ordenador BKG 9.8 se basaba en el reconocimiento De patrones: La posición resultante de cada movimiento legal se evaluaba directamente, sin realizar ninguna búsqueda, por medio de funciones SNAC (Smooth, Nonlinearity and Application Coefficients) que capturaban las

características clave de cada configuración del tablero.

- Juego de cartas como pókar combinan el azar, el ordenador se encuentra con dificultades al explorar el espacio de búsqueda asociado a un problema con información asimétrica ya que se ve obligado a evaluar un número enorme de posibilidades. El nombre del programa hace referencia tanto al uso de técnicas del deep learning como a la jerga del poker.
- El juego de comecocos o también conocido como Pac-Man.
- Los juegos de carrera, esta técnica es una forma sencilla de adaptar la dificultad de un videojuego a la capacidad de un jugador. Después de superar a nuestroponente virtual éste se va a limitar a seguirnos a una distancia prudente y en la espera de un fallo en nuestro volver a darle algo de aliciente a la carrera, además el motor del videojuego se ahorra algunos ciclos de reloj para determinar la estrategia de nuestro perseguidor con lo que puede aprovechar el cpu para otros menesteres.

Es la encargada de estudiar diferentes modelos de computación para identificar o crear sistemas inteligentes, apegados cada vez más a los de un ser humano. A pesar de ser ocupada para un sin fin de aplicaciones. Estos desarrollos siempre siguen un patrón; primero se intenta que una maquina tenga nuestro mismo razonamiento, dicho de otra manera, que analice los datos que le damos. Segundo, que sean racionales al momento de identificación y todo esto para poder emular resultados que nos sean satisfactorios. Estos van desde reconocer un texto, patrón, rostro hasta diagnósticos médicos. [23]

### **2.4.1. Redes Neuronales**

Para dar una definición de que son las redes neuronales, se parte del concepto de qué es una neurona; una neurona es una unidad de procesamiento que consta de valores de entrada. Estos valores reciben estímulos externos para que la neurona pueda procesar los cálculos y arrojarlos un valor. Los valores de entrada, en su camino a la

neurona constan de capas ocultas en las que nos dan el peso y esto es de suma importancia ya que nos modificará la neurona a nuestra conveniencia. Entonces, una vez definido esto, las redes neuronales son neuronas conectadas entre sí para trabajar con el mismo fin y generar de mejor manera lo que se requiera procesar.

De manera matemática, una Red Neuronal o Neural Network (NN) es una función  $g$  que asigna datos  $x$ , sean estos una imagen, un vector de salida  $y$ , la etiqueta de una imagen, etc. La función  $g = f_L \dots f_1$  es la composición de una secuencia de funciones más simples, llamados capas o bloques computacionales [24].

En la figura 2.11 se considera un bloque  $f$  en la red. Representa una caja que recibe datos  $x$  y parámetros  $w$  como entradas y produce datos  $y$  como salida [24]:

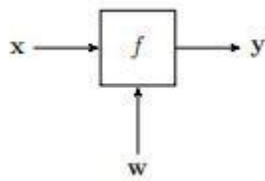


Figura 2.11: Bloque  $f$  en la red. Modificado de referencia (24).

De manera encadenada se puede ver a la figura 2.12.

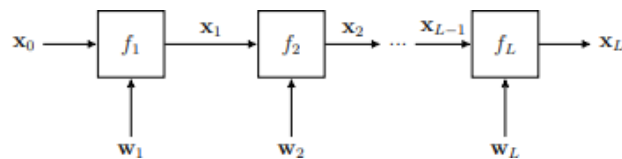


Figura 2.12: Encadenación de bloques  $f$ . Modificado de referencia (24).

Todas las redes neuronales constan de capas de entrada, ocultas y de salida. Sin embargo, lo que las hace diferentes son sus capas ocultas, ya que se estructuran de manera diferente en cada caso.

## Aplicación de Redes Neuronales

Las aplicaciones de las redes neuronales son las siguientes [25]:

- Control de procesos
- Reconocimiento de imagen
- Diagnósticos médicos
- Análisis y procesamiento de señales
- Procesamiento del lenguaje
- Robótica
- Conversión de texto a voz

Consiste en cambiar los símbolos gráficos de un texto a lenguaje hablado. Presentado por Sejnowski y Rosemberg, el sistema llamado NetTalk, convierte texto en fonemas y con la ayuda de un sintetizador de voz (Dectalk) genera voz a partir de un texto escrito.

- Compresión de imágenes

Es la transformación de datos de una imagen a una representación diferente con menos memoria o que se pueda reconstruir una imagen imperceptible. Cottrel, Munro y Zisper fueron los que diseñaron este sistema de compresión de imágenes utilizando una red neuronal con un factor de compresión de 8:1.

- Reconocimiento de caracteres

Es un proceso de interpretación visual y de clasificación de símbolos desarrollado por los investigadores de Néstor, Inc, crearon un sistema de computación neuronal que después del entrenamiento con un conjunto de caracteres de letras sea capaz de interpretar un tipo de carácter o letra que no haya visto con anterioridad.

- Reconocimiento de patrones en imágenes

Basada en Backpropagation donde una de sus típicas aplicaciones es la clasificación de objetos detectados por un sonar, su comportamiento tiene cierta similitud con la de los operadores humanos.

- Procesamiento de la señal

Se divide en tres clases que son:

- Predicción: Se conoce su comportamiento a través de una serie temporal de datos o valores.
  - Modelado de sistema: Caracterizados por la función de transferencia Qué es una expresión analítica entre la variable de salida y una variable independiente y sus derivadas.
  - Filtro de ruido: Eliminan el ruido de una señal y son capaces de mantener un alto grado de las estructuras y valores de los filtros tradicionales.
- Modelos económicos y financieros

Es una de las aplicaciones más importantes para los precios de existencias, la producción de cosechas, el interés de las cuentas, el volumen de las ventas, etcétera.

## 2.4.2. Machine Learning

El Machine Learning o también llamado aprendizaje automático es una rama dentro del campo de la inteligencia artificial, cuyo objetivo es dotar a los ordenadores de la capacidad de aprender sin ser programados. Es una disciplina científica en la inteligencia artificial, que como se menciona antes crea sistemas que sean capaces de aprender de manera automática, es decir, identifica patrones entre cientos de datos. De esta manera el algoritmo es capaz de revisar los datos y posteriormente predecir el comportamiento futuro que tendrán, esto implica que al predecir el comportamiento futuro estos mismos sistemas van a mejorar de forma automática sin la mínima intervención humana.

Es utilizado desde los 50's pero ha ido tomando gran relevancia por el aumento de la capacidad de computación como también los volúmenes de datos que manejan las empresas, en cuyos algoritmos catalogados en machine learning están clasificados en supervise learning y unsupervised learning, donde en el primero se dispone de un conocimiento anterior que ayuda a entender los datos cuando estos llegan para así tomar decisiones o hacer predicciones, como por ejemplo el sistema de control de spam, en un principio el usuario indica cuales correos son inadecuados y de esta manera el machine learning después podrá identificarlos de manera automática, también es utilizado para clasificar enfermedades, ya que el sistema logra aprender cuales son los síntomas y de esta manera logra identificar las enfermedades, es por eso que el unsepervised learning está orientado a la búsqueda de patrones, se clasifican los análisis de datos científicos ya que puede ayudar a clasificar patrones, así como para la segmentación de clientes, incluso sirve para el análisis de redes sociales analizando los patrones de comportamiento, ya que las técnicas de machine learning abarcan un amplio campo y son parte del big data.

El aprendizaje automático proporcionan mecanismos mediante los cuales el

ordenador es capaz de aprender por sí mismo a resolver un problema el programador diseña un algoritmo de aprendizaje adecuado para resolver el problema que se pretenda pero es el ordenador el que resuelve dicho problema en cierto modo el ordenador es capaz de programarse asimismo. [23]

Arthur Samuel define el aprendizaje automático como el campo de estudio que dota a los ordenadores de la capacidad de aprender a resolver problemas para los que no han sido explícitamente programados en Inteligencia artificial el aprendizaje se entiende como un proceso por el cual un ordenador es capaz de mejorar su habilidad en la resolución de un problema a través de la adquisición de conocimiento conocimiento que obtiene a través de experiencia. [23]

El aprendizaje de nota cambios en un sistema que son adaptativos en el sentido de que permiten al sistema hacer la misma tarea a partir de la misma posición de un modo más efectivo.

Al hablar de aprendizaje automático estamos hablando de dos aspectos los cuales son el refinamiento de las habilidades en la resolución de problemas lo que permite al ordenador ser cada vez más efectivo y la adquisición de conocimientos por medio de la experiencia.

Simón en 1969 propuso el uso del ordenador como fuente de conocimiento que ayudan a comprender mejor el comportamiento humano. Los ordenadores sólo trabajan con datos símbolos mediante los cuales se representan formalmente hechos conceptos o instrucciones de forma adecuada para su comunicación interpretación y procesamiento por seres humanos y otros medios automáticos es sólo el ser humano quienes atribuye un significado a esos datos estos datos son información en potencia colecciones de símbolos en crudo que deben ser procesados para que sean significativos

la información la obtenemos nosotros asociando datos y dándole significado a un contexto determinado.[23]

Roger pressman reconoció a mediados de los años noventa la construcción de sistemas extraían conocimiento de los datos estos eran unos de los desafíos más importantes de la informática y este es el objetivo final del aprendizaje automático es entonces que el aprendizaje automático se basa en el uso de datos estos datos de entrenamiento son utilizados para obtener conclusiones más o menos generales partiendo de casos particulares el razonamiento inductivo usado en el aprendizaje automático va de lo particular a lo general que a diferencia del razonamiento deductivo el cual va de lo general a lo particular este razonamiento establecer reglas que permiten establecer de forma precisa cuándo un razonamiento es válido partiendo de las premisas y la conclusión se infiere necesariamente sin embargo en el razonamiento inductivo no existe ningún tipo de garantía de que nos indica cuando un argumento es válido. [23]

El aprendizaje automático es lo que permite que un ordenador sea creativo en vez de programar un ordenador para resolver un problema se programa para que sea capaz de encontrar el mismo la forma de resolverlo en ocasiones hasta puede llegar a sorprendernos por sugerir soluciones que difícilmente se nos podrían haber ocurrido nosotros en cierto modo los algoritmos de aprendizaje automático son algoritmos que constituyen otros algoritmos de forma que el ordenador se programa así mismo y nosotros no tenemos que hacerlo.

Los problemas de reconocimiento de patrones son un ejemplo de este tipo de problemas para los cuales el aprendizaje automático resulta indicado también los sistemas de reconocimiento de voz o de identificación de objetos en imágenes requieren tener en cuenta múltiples factores variables tono entonación velocidad y ruido de fondo de una señal de voz; solapamientos oclusiones sombras condiciones de iluminación y

cambios de perspectiva en el caso de las imágenes. [23]

Es gracias al aprendizaje automático el que el día de hoy tengamos asistentes digitales que reconoce nuestra voz que almacenan el spam de nuestro buzón de correo porque éstos tienen un filtro que elimina automáticamente el correo no deseado así como nos recomiendan libros en amazon o películas de netflix también existen sistemas de traducción automática de un idioma a otro disponemos de aplicaciones para el móvil que nos dan el título y el intérprete de una canción que nos guste sólo con grabar un fragmento tardamos sólo unos instantes en realizar búsquedas en internet etcétera.

Hay diferentes formas de clasificar las técnicas del aprendizaje automático entre las cuales están [23]:

- Aprendizaje supervisado
- Aprendizaje no supervisado
- Deep Learning o Aprendizaje profundo

Estas técnicas son descritas más adelante.

Algunas de las aplicaciones del Machine Learning son:

- Predicción de fallos en equipos
- Detección de fraudes
- Diagnósticos médicos basados en síntomas
- Comportamiento de alguna aplicación móvil

- Potenciar evaluación de créditos
- Ciber vigilancia
- Identificar perfiles de alto riesgo
- Análisis de datos de sensores
- Minimizar el robo de identidad
- Recomendar artículos con base en compras anteriores
- Encontrar nuevas fuentes de energía
- Análisis de minerales del suelo
- Predicción de fallos de sensores de refinerías
- Optimización de la distribución de petróleo
- Rutas de transporte más eficientes

## **Aprendizaje supervisado**

El aprendizaje supervisado está asociado a los problemas de clasificación los cuales son uno de los problemas más estudiados en Inteligencia artificial donde el objetivo de cualquier algoritmo de aprendizaje supervisado es construir un modelo de clasificación a partir de un conjunto de datos de entrada denominado conjunto de entrenamiento que contiene algunos ejemplos de cada una de las clases que pretendemos modelar el objetivo del aprendizaje supervisado es la obtención de una descripción precisa para cada clase utilizando los atributos incluidos en el conjunto de entrenamiento, el modelo que se obtiene se puede usar para clasificar nuevos ejemplos casos cuyas clases se

desconozcan o simplemente para comprender mejor los datos de los que disponemos si nuestro modelo de clasificación es interpretable [23].

En la figura 2.13 se observa como el clasificador lineal es capaz de separar dos clases diferentes si las clases son linealmente separables. La línea recta define la frontera de division entre las dos clases [23].

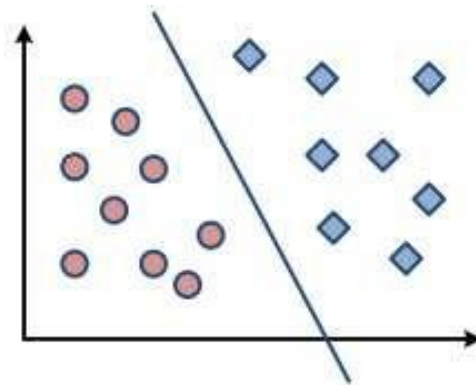


Figura 2.13: Clasificador lineal. Modificado de referencia (23).

Se puede definir de la siguiente manera: Si suponemos que todos los ejemplos que el modelo construido ha de reconocer son elementos potenciales de  $K$  clases distintas denotadas  $y_k$ .

$$Y = \{y_k | 1 \leq k \leq K\} \quad (2.4)$$

De la ecuacion anterior, se llama  $Y$  al conjunto de las clases de nuestro problema de clasificación. En determinadas ocasiones, extenderemos el conjunto de clases  $Y^*$  con una clase de rechazo  $y_o$  que utilizaremos cuando el clasificador no sepa muy bien qué clase asignarle a un ejemplo dado [23].

Al conjunto extendido de clases se le denota como:

$$Y^* = Y \cup \{y_o\} \quad (2.5)$$

Un clasificador tradicional, o regla de clasificación, es una función  $f : X \rightarrow Y^*$  definida sobre el conjunto de posibles ejemplos  $X$  tal que para todo ejemplo  $x \in X$ , el clasificador es capaz de asignarle una clase  $f(x) = y \in Y^*$ . Para construir ese clasificador, utilizaremos un conjunto de ejemplos de entrenamiento de la forma  $(x, y)$ , en los que  $x$  es un vector de características e  $y \in Y$  es la clase a la que pertenece el ejemplo de entrenamiento. Una vez entrenado el clasificador, podremos utilizarlo para tratar de inferir la clase asociada a ejemplos  $x$  para los que no conozcamos su clase:  $\hat{y} = f(x)$ . [23]

Si lo que se quiere es realizar una predicción concreta, simplemente nos fijamos en el máximo de la función  $f$  para cada una de las clases del problema dado un ejemplo  $x$  [23]:

$$\hat{y} = \operatorname{argmax}_{y_k} \hat{p}(y_k|x) \quad (2.6)$$

Algunas de las aplicaciones de este aprendizaje son [23]:

- Múltiples clases, jerarquías de conceptos y ontologías.
- Clases poco balanceadas y detección de anomalías.

### **Aprendizaje no supervisado**

En este tipo de aprendizaje se estimamos la probabilidad  $\hat{p}(y_k|x)$  de que un ejemplo  $x$  pertenezca a cada una de las clases de nuestro problema. El clasificador es un modelo de la distribución de probabilidad condicionada  $p(Y|X)$ . El objetivo es encontrar patrones en los datos de entrada  $x \in X$  que nos permitan construir un modelo de la distribución de probabilidad  $p(X)$ . [23]

Puede servir como herramienta de análisis exploratorio de datos y para preprocesar los datos antes de utilizar una técnica supervisada. El análisis exploratorio de datos son las técnicas de aprendizaje no supervisado que nos pueden servir para analizar la estructura de los datos cuando uno sabemos muy bien sus características y con la ayuda de las técnicas no supervisadas podemos descubrir los patrones que se repiten. Como se agrupan los datos de forma natural y si existen anomalías.

En la figura 2.14 un algoritmo de agrupamiento será capaz de identificar las agrupaciones naturales presentes en un conjunto de datos, denominadas clusters [23].

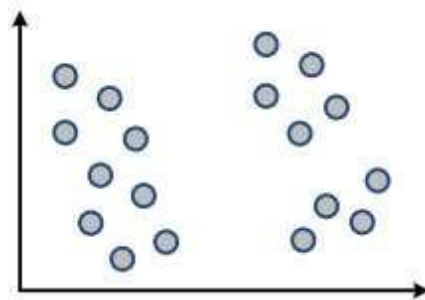


Figura 2.14: Algoritmo de agrupamiento. Modificado de referencia (23).

También con el apoyo de técnicas de visualización de datos podemos inspeccionar visualmente los datos disponibles. Las técnicas no supervisadas nos ayudan en la construcción de mapas temáticos que representen la distribución geográfica de las variables que sean de nuestro interés ya que los datos van asociados a coordenadas en un mapa y entonces la visualización es inmediata. Este tipo de técnicas nos pueden servir también en marketing, en recuperación de información, en el análisis de web logs, entre otras. [23]

En el preprocesamiento de datos utilizamos las técnicas no supervisadas como paso previo a la resolución de otros problemas de minería de datos, generalmente de tipo

supervisado.

Utilizamos las técnicas de discretización en distintas distintas ocasiones esto para convertir un atributo numérico continuo en un atributo categórico discreto. [23]

### **Deep Learning o Aprendizaje profundo**

Es una rama de la inteligencia artificial, un subconjunto del Machine Learning, un tipo de algoritmo con una arquitectura en capa, basada en redes neuronales profundas. Es un tipo de algoritmo de aprendizaje profundo, el cual utiliza varias capas para poder ir procesando cada una de estas y poder procesar las siguientes de manera eficaz. Este tipo de aprendizaje también es conocido como redes neuronales, donde los conexionistas se inspiran en el cerebro humano y construyen modelos informáticos formados por múltiples unidades relativamente simples a las que denominan neuronas. Estas neuronas, o elementos de procesamiento, se conectan entre sí para formar redes neuronales artificiales. Mediante la manipulación de las conexiones entre las neuronas de la red se consigue que la red neuronal tenga el comportamiento deseado. [23]

Se observa en la figura 2.15 el Deep Learning como un subconjunto de técnicas de aprendizaje automático englobadas dentro del campo de la inteligencia artificial [23].

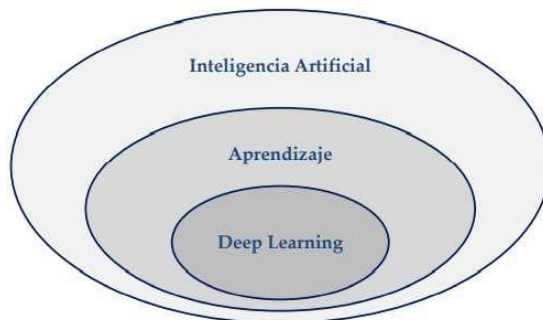


Figura 2.15: Deep learning. Modificado de referencia (23).

Las técnicas de deep learning son capaces de descubrir características a partir de

los datos. Cómo las técnicas de aprendizaje de representaciones que también son capaces de crear nuevas características a partir de otras características de esta forma su función no es solo identificar características que ayuden a discriminar en función a la situación que se encuentre sino también aumenta el nivel de abstracción mediante creación de jerarquías con varios niveles de características y entre las nuevas encuentra la forma automática. En cuanto a las redes neuronales que están empleadas se forman a partir de múltiples capas de neuronas como el córtex cerebral, el cerebelo o la retina humana. Cada una de estas capas hace que sea posible construir características a partir de las ya identificadas anteriormente lo cual resulta útil en el análisis de datos no estructurados como señales de audio o imágenes. Aplicaciones de los algoritmos del deep learning [23]:

- Recopilar un conjunto de datos asociado al problema.
- Diseñar una función de coste apropiada para el problema, lo que comúnmente se conoce como función de pérdida (loss function).
- Seleccionar un modelo de red neuronal y establecer sus hiperparámetros.
- Aplicar un algoritmo de optimización para minimizar la función de coste ajustando los parámetros de la red.

La principal diferencia entre este tipo de aprendizaje y otras técnicas es su capacidad de abstracción, la de formar jerarquías utilizando niveles de representación. la abstracción es empleada para aislar un componente de su entorno y ocultar los detalles de implementación. Y esto permite resolver problemas complejos, Otra manera de describir la abstracción es la forma en la que los humanos tratan la complejidad del mundo que los rodea.[23]

### **2.4.3. Redes Neuronales Convolucionales**

Se describe a la convolución como una función que es la suma o integral de dos

funciones y mide la cantidad de superposición ya que una función se desplaza. Matemáticamente se indica mediante el operador \* [22].

Si  $f$  y  $g$  son funciones en  $t$ , entonces la convolución de estas en un intervalo infinito es una integral dada por [26]:

$$f * g \equiv \int_{-\infty}^{\infty} f(\tau)g(t - \tau) d\tau \quad (2.7)$$

Si la convolución se realiza en un rango finito  $[0, t]$ , la convolución es [26]:

$$[f * g](t) \equiv \int_0^t f(\tau)g(t - \tau) d\tau \quad (2.7)$$

Se puede entender como un conjunto de capas de procesamiento que puede verse como un diagrama secuencial de bloques [21]. Basadas en el uso de convoluciones, haciendo referencia a este tipo de redes como CNN (Convolutional Neural Network) o ConvNets.

Consiste en resolver múltiples problemas prácticos que requieren procesar imágenes (clasificación de imágenes), detectar tipos de objetos que aparecen en la imagen correspondiente (detección de objetos), incluso como parte de un sistema que genere una descripción textual del contenido de una imagen. Recibiendo como entrada un vector ( $1D$ ), matriz ( $2D$ ) o tensor ( $>2D$ ) [23].

Es comúnmente utilizada para el procesamiento de texto e imágenes. Las capas ocultas son las que hacen convoluciones, es decir, dada una imagen la primera capa oculta va a seleccionar la parte que quiera y la va a mandar a otra capa oculta donde de igual forma va a ir reduciendo la imagen hasta que se obtenga la parte que importe identificar. Ya teniendo algo en concreto se llega a la capa de salida.

## Capas convolutivas

Las capas convolutivas sustituyen la multiplicación de pesos por entradas. Recurre a técnicas ad hoc que permiten incorporar conocimiento acerca del problema en el diseño de una red neuronal, dicho conocimiento proviene de características típicas de señales en las que se aplican redes convolutivas [23]:

- Invarianza en el tiempo para señales temporales (audio).
- Invarianza en el espacio para señales espaciales (imagen).
- Ambas señales espacio-temporales (video).

La entrada de una capa convolutiva es una señal, unidimensional en el caso de sonidos, bidimensional en el caso de imágenes. Dicha señal se procesa realizando una convolución con una máscara, también llamada kernel. Los pesos correspondientes a la máscara de convolución son los parámetros de la capa convolutiva que se entrenan. [23]

Matemáticamente, una función es equivariante si, cuando la entrada cambia, la salida cambia de la misma forma:  $f(g(x)) = g(f(x))$ . Las capas convolutivas son, por tanto, equivariantes frente a traslaciones. Características de estas capas [23]:

- Uso de conexiones locales que restringen la conectividad de la red.
- Uso de pesos compartidos por varias neuronas.

Al diseñar una capa convolutiva, se establecen valores para cuatro hiperparámetros [23]:

- Número de filtros K

Determina la profundidad, el número de características que es capaz de detectar. El conjunto de neuronas que detecta las K características en una región particular de la entrada se suele denominar columna o fibra.

- Conectividad local

Las neuronas se conectan a una región local de la entrada de la capa. La extensión de esa región es un hiperparámetro F al que se le denomina tamaño de filtro o del campo receptivo de la neurona, es habitual que esta extensión sea la misma en horizontal y en vertical, se emplean filtros de convolución de  $3 \times 3$  ( $F=3$ ),  $5 \times 5$  ( $F=5$ ) o  $7 \times 7$  ( $F=7$ ).

- Rellenado con ceros

Empleado para conseguir que tanto las imágenes de entrada como las imágenes de salida tengan exactamente el mismo número de píxeles.

- Paso, salto o zancada

Al diseñar una capa convolutiva, se puede establecer un tamaño S de paso, salto o zancada que indique cada cuántas muestras de la entrada quiere obtener una muestra de salida. Cuando  $S = 1$ , realizamos la convolución normalmente. Cuando  $S = 2$ , se salta una de cada dos muestras. Se puede incluso establecer tamaños de salto diferentes para cada dimensión.

Una capa convolutiva no es más que una capa convencional con una conectividad restringida y un conjunto de parámetros reducido, compartido entre todas las neuronas del mismo mapa de características. Los parámetros de la capa convolutiva corresponden a un conjunto de filtros, un conjunto de máscaras de convolución que se aplican localmente sobre diferentes regiones de la entrada. Por ejemplo, en una imagen en color, con tres canales, los filtros pueden ser de tamaño  $5 \times 5 \times 3$  ( $5 \times 5$

correspondientes a la dimensión espacial de la convolución, 3 por la profundidad de la entrada).[23]

## **Capas de pooling**

Las redes convolutivas suelen incluir un segundo componente clave: las capas de pooling o submuestreo. Esta capa reduce el tamaño de su entrada para que las capas posteriores puedan trabajar con datos de entrada reducidos esto puede hacerse para reducir la dimensión espacial de los datos o para reducir su profundidad. [23]

En una red neuronal convolutiva, las capas de pooling se suelen utilizar intercaladas entre capas convolutivas. Objetivos [23]:

- Reducir las dimensiones de las entradas que reciben las capas posteriores de una red neuronal para ser capaces de calcular las características distintas en paralelo, al trabajar con entradas más pequeñas.
- Proporcionar una cantidad pequeña de invarianza frente a traslaciones en la señal de entrada. Esto hace que las señales con las que trabajan las capas posteriores a una capa de pooling sean aproximadamente invariantes frente a pequeñas traslaciones en la entrada.

Uno de los problemas con el uso de este tipo de capas es que después de varios niveles se pierde la información acerca de la posición exacta de las cosas en la imagen de la entrada ya que combina múltiples capas y con esto pierde por completo la capacidad de localizar los objetos que es capaz de clasificar; lo que hace es que reduce la resolución de las imágenes de entrada.

Algunas redes están diseñadas para que utilicen estas capas en ciertos canales esto

con el objetivo de conseguir tanto invarianza de traslación en algunas características como la precisión a la hora de procesar otras características. Sin embargo si el pooling se realiza sobre las salidas de diferentes filtros es posible conseguir que la red aprenda ante qué tipo de transformaciones debe ser invariante. [23]

Las capas de pooling contienen menos unidades que las capas convolutivas que las preceden. El uso de estas capas cuenta con dos hiperparametros [23]:

- Ventana

Se especifica su tamaño para obtener cada salida de la capa de pooling, es decir, el tamaño  $F$  del filtro que se aplica a la imagen de entrada.

- Paso, salto o zancada

Se especifica el tamaño  $S$  del salto sobre la entrada para obtener salidas consecutivas. Se utiliza un paso de tamaño igual al tamaño del filtro de pooling.

#### **2.4.4. Clasificación de imágenes**

Arquitectura

##### **LeNet**

Es una de las primeras CNN publicadas. En la figura 3.16 muestra la arquitectura del modelo que fue introducido en 1989 por Yann LeCun, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W y Jackel LD con propósito de reconocer dígitos escritos a mano en imágenes. Este estudio fue desarrollado para entrenar las CNN a través de la retro propagación que finalmente se adaptó para reconocer dígitos para procesar depósitos en cajeros automáticos, incluso hasta en año actual algunos cajeros siguen usando este código. Para 1990 los mismos autores describieron las redes de retro

propagación en el reconocimiento de dígitos escritos a mano y para 1998 Yann LeCun, Bottou L, Bengio Y Haffner P tomaron su anterior artículo para comparar los puntos de referencia estándar, utilizando el aprendizaje basado en gradientes aplicado al reconocimiento de documentos.

LeNet consta de:

- Codificador convolucional de dos capas convolucionales.
- Bloque denso de tres capas conectadas.

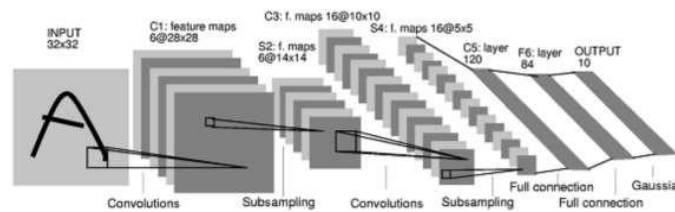


Figura 2.16: Arquitectura de LeNet. Modificado de referencia (27).

Fue diseñado para bajas resoluciones de reconocimiento de objetos en blanco y negro, con alrededor de 32 entradas por imágenes de 32 bits y esas imágenes de convolucionaron para convertirse en 6 canales de 28 x 28 píxeles que más tarde se redujeron agrupándolos en 14 x 14, después es seguido por otra convolucion que logró reducirlo a 10 x10 imágenes con 16 canales, en este caso la resolución que se obtiene otra vez se reduce por una agrupación promedio, dejándola de 5 x 5, entonces, se tiene 120 unidades conectadas, 84 unidades conectadas y en este caso se está tratando de una red RBF gaussiana para mapear cosas en 10 clases.

## AlexNet

Diseñada por Alex Krizhevsky, Ilya Sutskever y Geoffrey Hinton, propuesta en 2012, es un tipo de arquitectura similar a la anterior como se observa en la figura 2.17 pero la

principal diferencia es que esta es más profunda, tiene más filtros por capa y estas capas convolucionales están apiladas. Utiliza la función de activación ReLU, la cual provoca una aceleración de hasta seis veces con la misma precisión, teniendo un mejor rendimiento de entrenamiento sobre tanh y sigmoide, así como también utiliza dropout para tratar el overfitting, doblando el tiempo de entrenamiento a 0.5 de dropout.

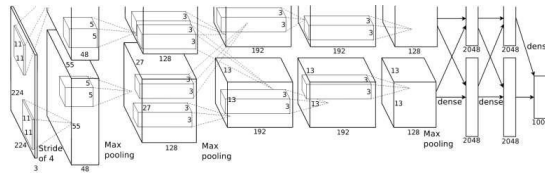


Figura 2.17: Arquitectura de AlexNet. Modificado de referencia (28).

Este tipo de arquitectura participó en una competencia llamada ImageNet donde resultó ganadora considerablemente. la computación dió un paso adelante tras la primera implementación de redes profundas, ocurriendo en GPU con esta arquitectura.

Con entradas de imágenes de tamaño 224x224 bits, las cuales se convulsionaron para convertirse en 96 canales de 54x54 pixeles, después por una agrupación fueron reducidas a 26x26, con el tamaño de canal igual, seguido por otra convolución que logró reducirlo a 26x26 imágenes con 256 canales, seguido a esto continua a reducir la resolución a la mitad por medio de la agrupación promedio, obteniendo 12x12, con 256 canales, continua otra convolución, donde quedan del mismo tamaño pero con 384 canales, seguidos de otra con la misma cantidad en pixeles y canales y se agrega otra convolución que arroja imágenes de 12x12 con 256 canales, por último se da una agrupación donde quedan de 5x5 píxeles con 256 canales y finalmente 4096 unidades completamente conectadas, luego 4096 unidades completamente conectadas tratándose se una red ReLu para mapearlas en 1000 clases.

## VGGNet

Propuesto por K. Simonyan y A. Zisserman de la Universidad de Oxford en 2014,

observaron el progreso en la visión por computadora y así este tipo de arquitectura que se muestra en la figura 2.18 fue una mejora de la anterior.

Se implementan agrupaciones por bloque, ya que después de muchas agrupaciones separadas una por una esto resulta muy tedioso y entonces se convierte en bloques repetidos de parámetros que se usan para visión por computadora. Primeramente se resuelve si deben usar menos convoluciones o más estrechas e hicieron un análisis que mostró que más capas de las convoluciones estrechas son más poderosas que un número más pequeño de convoluciones blancas. Una mayor cantidad de composiciones de funciones simples es más expresiva y se adapta a modelos significativos más que a una cantidad de modelos menos profundos.

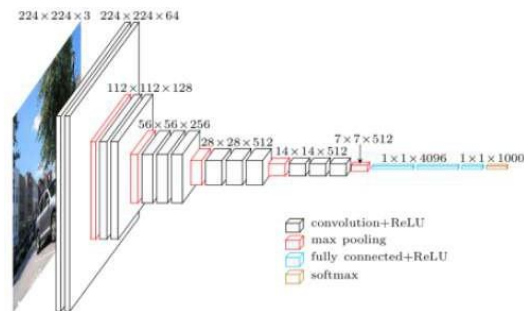


Figura 2.18: Arquitectura de VGG. Modificado de referencia (28).

Cuenta con convoluciones de 3x3, teniendo una agrupación máxima de 2x2 y si se afilan, al final mientras tiene las capas densas que son idénticas a las otras arquitecturas y si se combinan se puede obtener VGG16 y VGG19.

- VGG-16

La figura 2.19 está compuesta por 16 capas, 13 de convolución, 2 conectadas y una capa de softmax.

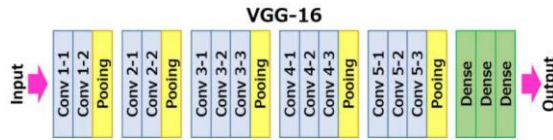


Figura 2.19: Arquitectura de VGG-16. Modificado de referencia (29).

- VGG-19

La figura 2.20 está compuesta por 19 capas, 16 de convolución, 2 conectadas y una capa de softmax.

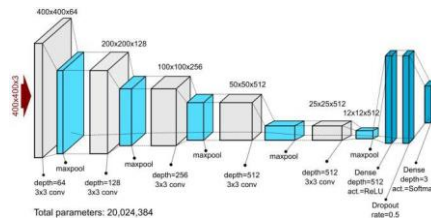


Figura 2.20: Arquitectura de VGG-19. Modificado de referencia (30).

## GoogLeNet

La figura 2.21 está compuesta por 22 capas (27 capas, incluidas las capas de agrupación), y parte de estas capas son un total de 9 módulos de inicio, a diferencia de la arquitectura anterior esta fue diseñada con una mayor eficiencia computacional, lo cual la hace mas rapida y tiene una precisión mayor, logra esto mediante la reducción de la imagen de entrada, al tiempo que retiene información espacial importante. [31]

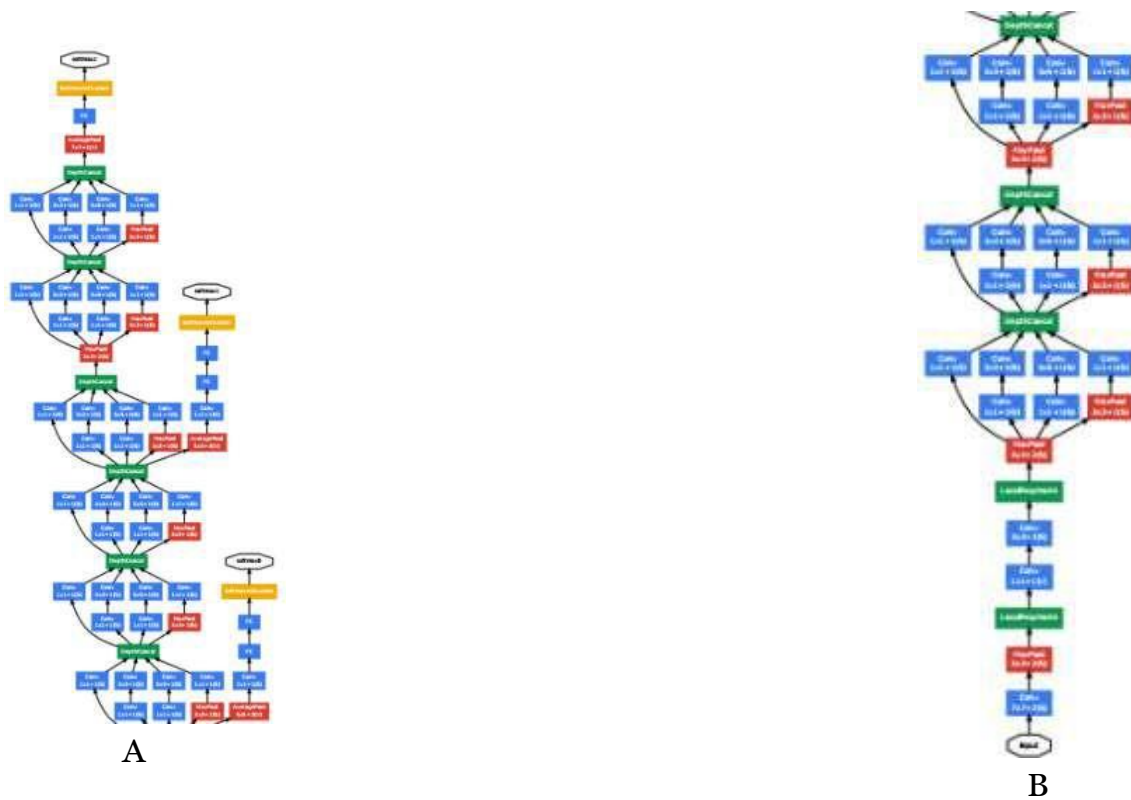


Figura 2.21: Arquitectura completa de GoogLeNet. Modificado de referencia (28).

El propósito principal de esta capa es reducir inmediatamente la imagen de entrada, pero no perder información espacial al utilizar filtros de gran tamaño. Consta de dos capas de agrupación donde el propósito de estas es reducir la muestra de la entrada a medida que avanza a través de la red. Esto se logra mediante la reducción de la altura y el ancho de los datos de entrada. Se utiliza una capa de abandono (40 %) justo antes de la capa lineal. La capa de abandono es una técnica de regularización que se utiliza durante el entrenamiento para evitar el sobreajuste de la red. [31]

## ResNet

También llamada redes residuales, fue propuesta por Kaiming He, Shaqing Ren y Jian Sun. Entrena a miles de capas logrando un rendimiento convincente, mejora el rendimiento de muchas aplicaciones de visión por computadora distintas de la clasificación de imágenes, como detección de objetos y reconocimiento facial. A medida que se toma una arquitectura y se extiende en muchas capas como en la figura 2.22 el gradiente desaparece, sin embargo para que no suceda el caso donde el gradiente de

cero proponen retropropagar a través de la función identidad y entonces simplemente se va a multiplicar por 1 el gradiente, para esto apilan los bloques residuales donde ahí se usa la función de identidad para preservar el gradiente.

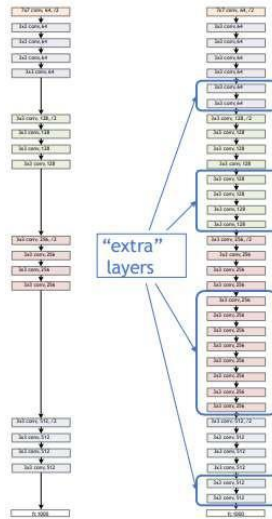


Figura 2.22: Arquitectura de ResNet. Modificado de referencia (27).

Con esto la probabilidad de que el gradiente sea cero es casi nula y simplemente se propaga el gradiente hacia atrás y este se distribuye uniformemente en sumas en un gráfico de cálculo y permite observar el degradado a medida que retrocede.

## DenseNet

La figura 2.23 muestra la arquitectura más reciente, también son llamadas como redes de convolución densas, introdujo un nuevo bloque llamado Dense Block y apilo los bloques uno encima de otro, con capas en medio para construir una red profunda.[28]

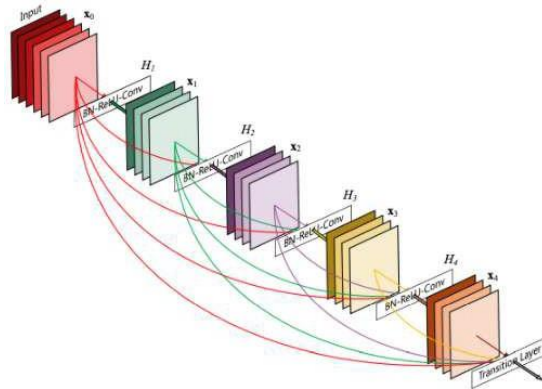


Figura 2.23: Arquitectura de DenseNet. Modificado de referencia (28).

Para un bloque denso, consideramos todos los bloques densos anteriores como entrada y producimos una salida que alimentamos a todos los bloques densos posteriores. Para que las capas sean compatibles entre sí, aplicamos convoluciones y normalizaciones por lotes, esto por la reutilización de funciones. Se resuelven problemas del gradiente que desaparece y se tienen menos parámetros.[28]

# Capítulo 3 Metodos

## 3.1. Población para estudiar

Se estudiaron a 30 pacientes, divididos en dos grupos de la siguiente manera: 15 niños con normopeso y 15 con obesidad. Los datos de los niños voluntarios fueron obtenidos del estudio METCOG para evaluar la obesidad infantil en México. Se trata de niños varones con edades comprendidas entre 7 y 9 años. Todos zurdos y sin antecedentes de enfermedades psicológicas. A cada niño se le hicieron toda una serie de pruebas cognitivas, antropométricas y de resonancia magnética. Para este estudio se usaron solo los resultados de dos de los estudios de resonancia: imágenes anatómicas y estados de reposo. Cada niño tiene un archivo BOLD y un T1. En el T1 se encuentran las imágenes anatómicas del cerebro y en el BOLD las funcionales (estados de reposo, adquisición de 240 imágenes del cerebro una tras otra sin realizar acción alguna dentro del escáner).

## 3.2. Hardware y Software

Las características del Hardware usadas son:

Procesador: AMD A12-9720P RADEON R7, 12 COMPUTE CORES 4C+8G 2.70 GHz

RAM instalada: 12.0 GB (11.5 GB utilizable)

Tipo de sistema: Sistema operativo de 64 bits, procesador x64

Las herramientas usadas para el Software son descritas por partes.

Para la segmentacion de los hipotalamos y el calculo de sus volumenes, se usa el

programa 3D Slicer, el cual está escrito en C++, Python y QT, es un paquete de software gratuito que cuenta con código abierto para análisis de imágenes, se escogió por su variedad en aplicaciones médicas.

Para la construcción de las redes se utiliza el lenguaje de programación de Python. La principal razón para usar este lenguaje es que tengo conocimientos previos, sin embargo esta no es la única razón. Python es un lenguaje orientado a objetos, es decir, la manera de programar trata de encontrar una solución a problemas. Este tipo de lenguaje es parecido a la sintaxis humana, de ahí que es muy sencillo de aprender. Las librerías utilizadas fueron:

- Librerías para visualización: Matplotlib: es la biblioteca de visualización de Python 2D más utilizada. Está equipado con gran variedad de comandos e interfaces para producir gráficos con calidad de publicación a partir de sus datos [32].
- Librerías para cálculo numérico y análisis de datos: NumPy: Abreviatura de Numerical Python, NumPy es el paquete fundamental requerido para la computación científica de alto rendimiento y el análisis de datos en el ecosistema de Python [32].
- Librerías para deep learning: TensorFlow: Es una plataforma para construir y entrenar redes neuronales, que permiten detectar y descifrar patrones y correlaciones, análogos al aprendizaje y razonamiento usados por los humanos [33].
- Keras: Es una librería de aprendizaje profundo que facilita la creación de redes neuronales convolucionales, para poder usar de manera adecuada esta librería debe de descargarse después de instalar python, así mismo al descargarlo se incluyen las siguientes funciones:
  - Sys y os: usadas para poder moverse en carpetas dentro del sistema operativo.

- **ImageDataGenerator:** procesa las imágenes que se le dan a entrenar. Su función es modificar las imágenes que recibe para que el algoritmo las vea todas diferentes, es decir las rota, acerca, hace zoom, etc. esto para que el algoritmo aprenda que no todas las imágenes se van a ver igual o van a estar completas.
- **Optimizers:** optimizador para entrenar el algoritmo, en este caso el usado es Adam porque calcula el promedio exponencial del gradiente y el gradiente al cuadrado.
- **Sequential:** permite hacer redes neuronales secuenciales, es decir, tiene varias capas secuenciadas una por una entre ellas.
- **Conv2D y MaxPooling2D:** capas donde se realizan las convoluciones, el 2D es porque se trabaja con imágenes, 1D es para datos y 3D para videos.
- **Flatten:** es usada para que la imagen profunda y pequeña se convierte a plana para tener una dimensión que contenga toda la información.
- **Dense:** Es la función usada para conectar las capas finales, guarda los parámetros para el backpropagation.

Por simplicidad en esta tesis no se hace una descripción del hardware de resonancia magnética con el que se tomaron imágenes ni de los parámetros de las secuencias que se usaron.

### **3.3. Obtención y Procesamiento de imágenes**

En el fichero T1 se tomó una de las variaciones de la imagen anatómica del cerebro. Se tomaron las que empiezan con mwcocs1 que cuenta con un cerebro sin cráneo (cráneo segmentado y eliminado de la imagen) y normalizado.

En primer lugar se segmentaron en el programa 3D Slicer los hipotálamos de los 30 pacientes. Se hizo un corte para poder dividirlos en los dos hemisferios izquierdo y derecho esto para poder apreciar la diferencia entre las personas con obesidad y normopeso. Una vez obtenidas estas imágenes se guardaron en una carpeta.

### **3.4. Diseño o estructura de la Red Neuronal**

La red neuronal construida fue de tipo: red neuronal convolucional.

Se hizo un programa el cual ayudó a clasificar imágenes. Se entrenó con un set de datos de obesos y normopesos, para que cuando leyera una imagen el sistema dijera si es una persona obesa o una con normopeso.

Primero se tuvo un archivo para entrenar a la red y otro para predecirla, se llamaron entrenar y predecir. El archivo de entrenar tiene el código donde está la red neuronal, la cual se entrenó sobre el mismo archivo. El archivo predecir cargó el modelo ya antes generado e hizo una predicción de alguna imagen dada.

Después se creó una carpeta llamada tesis, la cual tenía dentro otra carpeta llamada datos y dentro de esta estaban las carpetas de entrenamiento y validación. En las carpetas de entrenamiento y validación contenían cada una las carpetas de los pacientes con normopeso y las carpetas de los pacientes con obesidad.

El código creado fue con uso de tensorflow y dentro de este se llamó a keras. Para el primer archivo, el de entrenar, primeramente se importaron todas las librerías que se necesitaban para el algoritmo, para pre procesar las imágenes, optimizar el algoritmo, poder hacer redes neuronales secuenciales, es decir tener cada capa en orden, capas con las que se hicieron convoluciones y maxpooling.

Se pusieron dos variables, una de entrenamiento y otra de validación. En las variables se llamó a los directorios donde están guardadas las imágenes.

Parámetros utilizados:

- Numero de épocas = 30. Se refiere al número de veces que se va a iterar sobre todo el set de datos.
- Altura y longitud = 150, 150. Es el tamaño para procesar las imágenes.
- Batch size = 32. Sirve para dar el número de imágenes que va a procesar en cada paso.
- Pasos = 1. E el número de veces que la información va a ser procesada en cada época.
- Números de filtros aplicados a cada convolucion, en la primera convolucion se usan 32 y en la segunda 64.
- Tamaño de los filtros con altura y longitud, para el primer filtro es de (3,3) y para el segundo de (2,2).
- Tamaño de pool = (2,2). Es el tamaño del filtro usado en maxpooling.
- Clases = 2. Son las carpetas de obesos y normopeso.
- Tamaño de ajuste para acercarse a una solución es de 0.0004.

Se hizo el pre procesamiento de imágenes para después dárselo a la red neuronal. A la red se le dieron las imágenes pre procesadas. Se usó:

- ImageDataGenerator para reescalar las imágenes a 1./255. Las imágenes estaban

de 1 a 255 y las puso de 0 a 1.

- Shear range = 0.3 para generar las imágenes inclinadas. Para que la red aprendiera a que no todas las imágenes iban a estar de manera perfecta.
- Zoom range = 0.3 porque aplicó zoom a algunas imágenes. Para que la red aprendiera que no todas las imágenes iban a estar completas.
- Horizontal flip para invertir las imágenes y hacer que la red supiera distinguir la dirección en la que se encontraban.

Las imágenes guardadas fueron creadas en un programa en 3D por lo que no todas tenían la misma direccionalidad. Para el caso de validación se hizo lo mismo que en entrenar pero ya solo se reescalaron las imágenes ya que se le dieron las imágenes tal y como eran.

Después se hizo uso de las mismas variables para generar las imágenes que se iban a usar en la red neuronal. Entre las variables estaban: target size, batch size y class model, para que entraran al directorio de entrenamiento, procesaran las carpetas a la altura y longitud dadas y las clasificaran de manera categórica y lo mismo se hizo para la validación.

Una vez que se tuvo todo esto se creó la red neuronal convolucional. Se puso que la red que se quería era secuencial y se procedió a crear los filtros de la convolucion 1 con todos los datos ya antes dados: capa con 32 filtros, con cierta altura y longitud. Se añadió una capa de maxpooling donde solo se puso el filtro 1. Se añadió la siguiente capa convolucional, capa con 64 filtros y se añadió otra capa de maxpooling.

Para la clasificación se usó `cnn.add (Flatten)` para que la imagen profunda se convirtiera en plana. Después de aplanarlas se mandó a una capa que estaba conectada

a una capa anterior. Se procedió a apagar la mitad de las neuronas usando Dropout para que el algoritmo aprendiera caminos alternos para clasificar. La última capa fue Dense para saber la probabilidad de ser obeso o normopeso y tomar el valor más alto para saber cuál de estas dos era la correcta.

Se compilaron los parámetros para optimizar el algoritmo, dando la función de pérdida, optimizador y la métrica de que tan bien aprendió la red.

Se guardó todo esto en una carpeta llamada modelo, creada por la misma red, donde guardo la estructura y los pesos de las capas.

Para entender este análisis se consideró la manera en la que podía empezar a verse de manera encadenada como en la figura 3.1 y se procedió a utilizar la ecuación 2.4 para el conjunto de clases de clasificación. Se extendió a una clase de rechazo  $y_0$  para cuando el clasificador no supiera que clase asignarle al ejemplo dado, tal como en la ecuación 2.5. Para el caso de las predicciones concretas, es decir, cuando el clasificador supiera a qué clase pertenece, se utilizó la ecuación 2.6 por el máximo de la función en cada una de las clases del problema dado.

Se creó el archivo predecir, usando la misma altura y longitud. Se le dijo dónde estaba el modelo y donde estaban los pesos. Load model se usó para cargar el modelo en la red y lo mismo para cargar los pesos en la red. Después se cargó la imagen que se quería predecir. Se cargó en un arreglo la imagen, se le dio una dimensión extra para procesarla sin ningún problema. El algoritmo las clasificó en 0 y 1, donde 0 correspondía a normopeso y 1 a obeso.

La figura muestra un diagrama de la estructura de este tipo de red neuronal.

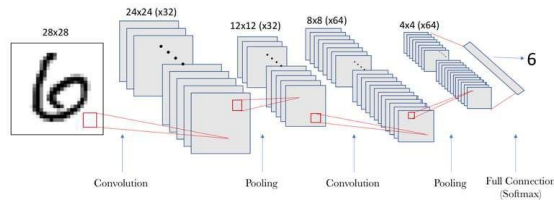


Figura 3.1: Diagrama de la estructura de una red neuronal convolucional. Modificado de referencia (34).

La red se encuentra descrita por códigos en el apéndice I.

### 3.5. Estadística

Para saber si los resultados son correctos se usa una prueba de hipótesis para probarlo. Las pruebas de hipótesis buscan mediante cálculos estadísticos descartar una hipótesis nula de una alternativa.

Ajustando los distintos pesos de la red neuronal se reduce al mínimo la desviación entre el valor predicho y el valor real para la observación dada. Este error de aprendizaje también es llamado error local y se calcula obteniendo la diferencia entre el valor real que hay que predecir y el valor predicho por la neurona artificial. Primeramente se obtiene la media de la muestra ( $Z$ ), dividida por una desviación estándar de la muestra.

Si la media de la muestra es mayor que 2 significa que la desviación excede 2 veces la desviación estándar y viene dada por el teorema del límite central.

También se obtiene el valor  $P$  que representa la probabilidad de apoyar la hipótesis nula. Si  $P < 0.05$  entonces se rechaza la hipótesis nula y aceptar la alternativa, que significa que los valores cumplen con el estándar.

Otra herramienta para evaluar este tipo de modelo es la matriz de confusión, es usada para entender como se comporta el modelo y puede mostrar cuando la red se confunde con otra de las clases. Tiene la siguiente estructura:

		Positives (1)	Negatives (0)
Predicted	Positives (1)	<i>TP</i>	<i>FP</i>
	Negatives (0)	<i>FN</i>	<i>TN</i>

La matriz de confusión es una de las métricas que permite resumir el rendimiento de un algoritmo de clasificación. En este caso se tiene un problema de clasificación en el que predecimos si una persona es obesa o no, esto a partir de la segmentación de su hipotálamo. Se asigna la etiqueta a la variable: 1 cuando una persona es obesa y 0 cuando una persona tiene normopeso. Una vez identificado el problema, la matriz de confusión, es una tabla con dos dimensiones (Actual y Predicción) y conjuntos de clases en ambas dimensiones. Las clasificaciones actuales son columnas y las pronosticadas son filas.

La matriz de confusión tiene cuatro posibles resultados donde dos son posibles valores reales y los otros dos posibles valores de predicción:

- True Positives o Verdadero positivo (TP): Son los casos en los que el valor real es positivo y la prueba predicen positivo.
- True Negatives o Verdadero negativo (TN): Son dos casos en los que el valor real es negativo y la prueba predice negativo.
- False Positive o Falso positivo (FP): Son los casos en los que el valor real es negativo y la prueba predice un positivo.

- False Negative o Falso negativo (FN): Son los casos en el que el valor real es positivo y la prueba predice un negativo.

La exactitud en los problemas de clasificación es el número de predicciones correctas sobre el total de predicciones, la exactitud se refiere a lo cerca que el resultado está de una medición del valor verdadero, está relacionada con el sesgo de una estimación:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

La precisión es el valor más cercano al resultado de la medición del valor verdadero, es decir da información respecto a los falsos positivos, es una proporción entre los positivos reales predichos por el algoritmo y los casos positivos, la precisión se refiere a la dispersión del conjunto de valores obtenidos a partir de mediciones repetidas de una magnitud:

$$Precisión = \frac{TP}{TP + FP}$$

La sensibilidad y la especificidad son dos valores que indican la capacidad del estimador para discriminar los casos positivos de los negativos. El recall, también llamado sensibilidad, la sensibilidad es conocida como tasa de verdaderos positivos y se refiere a la proporción de casos positivos que fueron correctamente identificadas por el algoritmo:

$$Recall = \frac{TP}{TP + FN}$$

La especificidad es conocida como la tasa de verdaderos negativos son los casos negativos que el algoritmo ha clasificado correctamente:

$$Specificity = \frac{TN}{TN + FP}$$

La tasa de falsos positivos es la probabilidad de que se produzca una falsa alarma

que se dé un resultado positivo cuando el valor verdadero sea negativo:

$$TFP = \frac{FP}{FP + TN}$$

La tasa de falsos negativos o también llamada tasa de error es la probabilidad de que la prueba pase por alto un verdadero positivo:

$$TFN = \frac{FN}{FN + TP}$$

La tasa de verdadero positivo o también llamada sensibilidad es la probabilidad de que un resultado positivo real de positivo:

$$TTP = \frac{TP}{TP + FN}$$

La tasa de verdadero negativo también llamada especificidad es la probabilidad de que un resultado negativo real de un resultado negativo:

$$TTN = \frac{TN}{TN + FP}$$

El valor predictivo positivo es la probabilidad de que sí se ha obtenido un resultado positivo realmente tenga la enfermedad:

$$VPP = \frac{TP}{TP + FP}$$

El valor predictivo negativo por el contrario es la probabilidad de que sí se ha obtenido un resultado negativo en la prueba en realidad no tenga la enfermedad:

$$VPN = \frac{TN}{TN + FN}$$

Como ultima herramienta que usan las redes neuronales es la function de costo de la entropia cruzada, para dos variables aleatorias discretas p y q, la entropía cruzada se

define mediante la formula:

$$H(p, q) = - \sum_x p(x) \log q(x)$$

Esta definición no es simétrica. P está pensada como una distribución *verdadera*, solo parcialmente observada, mientras que Q está pensada como una distribución *innatural* obtenida de un modelo estadístico construido.

En la teoría de la información, la entropía cruzada puede interpretarse como la longitud esperada en bits para la codificación de mensajes, cuando se utiliza Q en lugar de P. Esta perspectiva excede el objetivo de la presente discusión y no es de importancia primaria desde el punto de vista de la cadena de suministro.

En la práctica, debido a que P no se conoce, la entropía cruzada se calcula empíricamente a partir de las observaciones, suponiendo simplemente que todas las observaciones recopiladas son igualmente probables; es decir, que  $p(x)=1/N$ , donde N es el número de observaciones.

$$H(q) = - \frac{1}{N} \sum_x \log q(x)$$

# Capítulo

# 4

## Resultados

Las gráficas presentadas aquí son de los cálculos que hizo la red neuronal. En el apéndice I se explica cómo se pre-procesan las imágenes y una vez así se alimenta la red neuronal. Al final, para saber si la red convolucional es capaz de diferenciar a pacientes con normopeso u obesos, se cargan las imágenes de estos mismos pacientes. Las imágenes cargadas son tal y como se obtuvieron al momento de segmentar los hipotálamos, antes de ser pre-procesadas por la red.

El preprocesamiento fue con el único de fin de que sin importar si las imágenes estuvieran en cualquier dirección, con o sin zoom, horizontales o verticales, la red fuera capaz de leerla. Es por eso que para la parte de predecir ya no importa que las imágenes cargadas sean de igual manera pre-procesadas porque la red ya sabía leerlas.

Se presentan dos gráficas. Una sobre la exactitud (accuracy) y otra sobre la función de costo (Loss).

Se tiene una una pareja de datos  $[x,y]$  donde  $x$  son las imágenes y  $y$  la clasificación. Al iniciar con pesos arbitrarios hay una diferencia entre la predicción de la red y la relación  $[x,y]$ , esta diferencia es la cantidad a minimizar, eso se hace cambiando los pesos hasta que la diferencia entre lo que la red predice y los datos de entrenamiento alcancen mas del 90%.

En la figura 4.1 se muestra la función de costo también llamada modelo loss. Esta función evalúa la desviación entre las predicciones que realiza la red neuronal y los valores reales

observados durante el aprendizaje. Es usada para determinar la pérdida entre la salida y el valor dado. La función de costo es un método para evaluar qué tan bien logra modelar un conjunto de datos dado un algoritmo, esta función se minimiza utilizando mínimos cuadrados o la función de la entropía cruzada, siendo la segunda la más usada en la actualidad. Si la función de costo tiene un número mayor a 0.9 o menor a 0.01 significa que las predicciones hechas son equivocadas. Si tiene un número entre 0.01 y 0.9 entonces hizo una buena predicción.

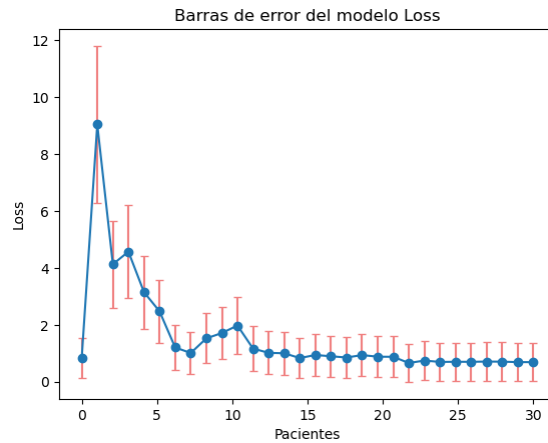


Figura 4.1: Gráfica modelo Loss vs pacientes.

La gráfica 4.1 representa el porcentaje de pérdida de cada paciente. Loss de manera vertical y pacientes de manera horizontal. En la red neuronal se cargaron las imágenes de un total de 30 pacientes, entonces, pacientes se refiere a la imagen del hipotálamo segmentado de cada uno de estos. Estos datos son arrojados cuando la red prueba imagen tras imagen y en cada una da la pérdida y la exactitud.

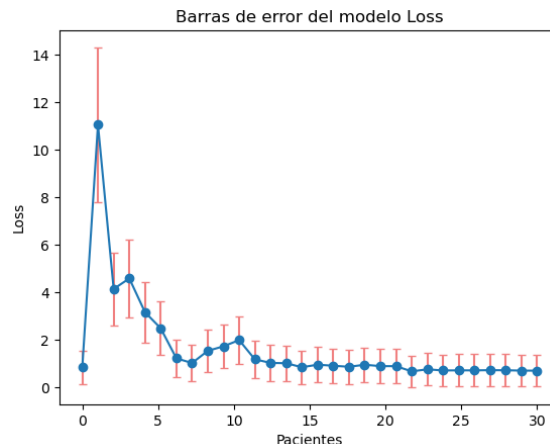


Figura 4.2: Gráfica barras de error del modelo Loss vs pacientes.

La grafica 4.2 representa las barras de error, las cuales son utilizadas para indicar el error que se estima en las medidas del modelo Loss.

La siguiente gráfica que se presenta en la Figura 4.2. Es la gráfica de accuracy o modelo de exactitud. Esta gráfica se refiere al número de predicciones correctas sobre el total de predicciones.

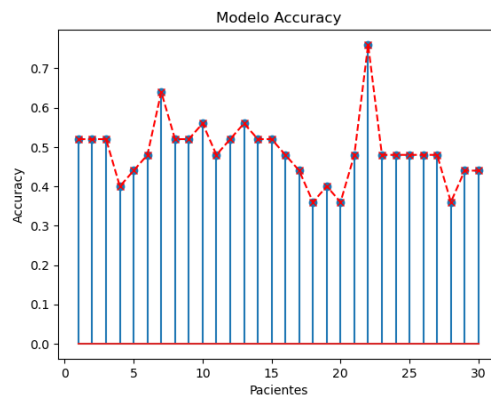


Figura 4.3: Grafica modelo Accuracy vs pacientes.

De la misma manera que la gráfica anterior esta representa el porcentaje de exactitud de cada paciente, siendo accuracy de manera vertical y pacientes de manera horizontal.

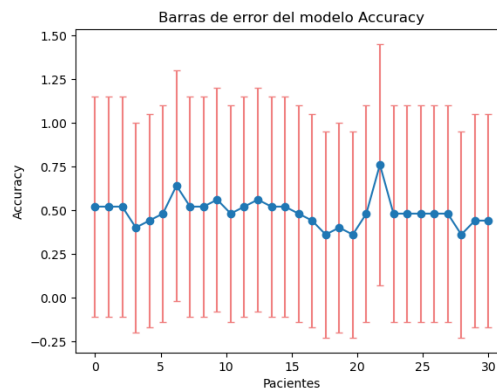


Figura 4.4: Grafica barras de error del modelo Loss vs pacientes.

La grafica 4.2 representa las barras de error, las cuales son utilizadas para indicar el error que se estima en las medidas del modelo Accuracy.

La siguiente tabla muestra los valores de cada paciente que se aprecian en las dos graficas anteriores.

Normopeso		y	Obesos	
Pacientes	Loss		Pacientes	Accurracy
1	0.8286		1	0.52
2	11.0457		2	0.52
3	4.1317		3	0.52
4	4.575		4	0.4
5	3.15		5	0.44
6	2.4744		6	0.48
7	1.2077		7	0.64
8	1.0108		8	0.52
9	1.5252		9	0.52
10	1.7147		10	0.52
11	1.9795		11	0.48
12	1.1671		12	0.52
13	1.0171		13	0.56
14	1.003		14	0.52
15	0.8335		15	0.52
16	0.9416		16	0.48
17	0.8969		17	0.44
18	0.8532		18	0.36

19	0.9387		19	0.4
20	0.8812		20	0.36
21	0.8785		21	0.48
22	0.6597		22	0.76
23	0.7459		23	0.48
24	0.6977		24	0.48
25	0.7029		25	0.48
26	0.7017		26	0.48
27	0.7107		27	0.48
28	0.7068		28	0.36
29	0.6917		29	0.44
30	0.6901		30	0.44

Para comprobar los promedios de los grupos de imágenes, se realizó una prueba de hipótesis. Los resultados de la prueba son los siguientes:

	Desviación estándar	Media	Estadística Z	Valor P
Accuracy	0.08	0.48	-2.558	0.0106
Loss	2.03	0.9178	1.005	0.315

La desviación estándar es una medida de un conjunto de datos que indica cuánto pueden alejarse los valores con respecto a la media, mientras mayor sea esta desviación mayor es la dispersión de los datos. La media es cuando sumamos todos los valores y lo

dividimos entre el número total de observaciones o en este caso, se le llamaría observaciones al número total de pacientes. El valor P representa la probabilidad de apoyar la hipótesis nula.

# Capítulo 5 Discusión

## y Conclusiones

### 5.1. Resumen de los resultados

Se diseñó y programó una red neuronal, en la cual se encuentra que la red neuronal convolucional creada no es del todo capaz de diferenciar las imágenes de los pacientes con normopeso y la de los pacientes con obesidad. Sin embargo al ser pocas imágenes la precisión máxima alcanzada fue de 60 %, ninguna bajó del 50 %.

Si bien en los resultados de la red no se apreciaron muchas diferencias, no fue el caso en el volumen, ya que en estos se encontraron los siguientes datos:

*mm3*

	Normopeso		Obesos	
	Izquierdo	Derecho	Izquierdo	Derecho
1	2000	2000	1600	1520
2	2120	1900	1570	1310
3	2010	2100	1780	1260
4	2300	1600	1800	1670
5	2130	1960	1630	1290
6	2320	1800	1590	1140
7	2000	1890	1720	1900
8	2210	1740	1640	1070
9	2000	1860	1800	1400
10	2200	1990	1500	1600
11	2340	2100	1640	1400

12	2070	2040	1710	1290
13	2100	1920	1690	1500
14	2200	1980	1800	1480
15	2000	2050	1530	1460
<i>Promedio</i>	2133.33	1928.67	1666.67	1419.33
<i>Desviación estándar (±)</i>	122.979	137.730	100.119	210.458

A partir de esta tabla se observa que el volumen promedio del hipotálamo de una persona sana ronda entre los  $4\text{cm}^3$  mientras que el de una persona con obesidad tiene un promedio de  $3.08\text{cm}^3$ . Tanto de una persona obesa como de normopeso el volumen de su hipotálamo izquierdo es mayor que el del derecho.

A estos mismos datos se le hizo una prueba tipo T, a los volúmenes del hipotálamo izquierdo y derecho. Donde Normopeso izquierdo corresponde al volumen del hipotálamo izquierdo de un paciente con normopeso y lo mismo para el resto.

<b>Normopeso izquierdo, Obeso derecho</b>	<b>4.9737E-12</b>
<b>Normopeso derecho, Obeso derecho</b>	1.5259E-08
<b>Normopeso izquierdo, Obeso izquierdo</b>	4.9737E-12
<b>Normopeso derecho, Obeso izquierdo</b>	2.0417E-06

El objetivo principal de este estudio fue: “Estudiar el volumen y estructura del hipotálamo segmentado de personas con normo- peso y obesos en imágenes anatómicas por resonancia magnética. Se pretende usar redes neuronales convolucionales para la clasificación de aspectos de interés (biomarcadores). Eneste caso se espera que la red sea capaz de diferenciar el hipotálamo de cualquiera de estos dos grupos de voluntarios.”

Todos aspectos se completaron con la excepción de la efectividad de la red creada.

Como objetivos secundarios de esta tesis se tenía:

- Segmentación de distintos tejidos del cerebro en imágenes de resonancia magnética.
- Cálculo de volúmenes tisulares en imágenes de resonancia magnética.
- Desarrollar una red neuronal convolucional capaz de realizar clasificaciones lo más precisas posible.

Todos estos objetivos también se cumplieron con excepción del último. En cuanto a las técnicas de segmentación se aprendió a usar las existentes y se aplicaron a las imágenes de este estudio.

Con el fin de mejorar los resultados deberemos de agregar más imágenes, ello permitiría a la red aprender de mejor manera. Originalmente se pensaba utilizar los volúmenes de los hipotálamos junto con la red. Al tener pocas imágenes lo más recomendable es usar otros clasificadores como SVM o regresión logística, es decir, codificar la imagen con descriptores o clasificarlos via volúmenes.

## **5.2. Limitaciones del estudio**

Entre las limitaciones de este estudio se encuentra que se cuenta con un número muy bajo de pacientes. Esto hace que la red no sea capaz de dar una exactitud y precisión adecuada. La red neuronal convolucional funciona si solo se cargan los pesos de los pacientes una vez, sin embargo, si se llegaran a volver a cargar estos pesos provocaría que la red guarde ambos pesos y tenga el doble que de imágenes. Al mezclarse los datos de pesos guardados en una carpeta hace que la red confunda los resultados cuando se le pide que diga a que grupo pertenece la imagen que se requiere comprobar. Esto ocurre a causa de que la red no conoce con exactitud las imágenes y puede pasar que incluso algunas las clasifique mal. La manera más sencilla para que los resultados no estén

en torno al 50 % es agregar más imágenes, de esta manera aumentaría el número de pasos para que la información pueda ser más veces procesada en el número de épocas. El aumento de pasos hace que la red revise más de una vez imagen tras imagen, esto cuantas veces se requiera y se pueda. Sin embargo no es posible poner un número de pasos que al ser multiplicado por el número de épocas este sea mayor a la cantidad total de imágenes.

### **5.3. Trabajo futuro**

En cuanto a las líneas de trabajo futuras, sería ideal poder obtener imágenes en tres dimensiones, para que al aplicar el clasificador a cualquier paciente las máquinas de cualquier hospital sean capaces de apreciar este tipo de pruebas. Además de para lo que han sido aplicadas en este caso tal vez podrían mejorarse para que en un futuro puedan detectar los diferentes tipos de enfermedades o deterioro de la misma que pueda sufrir algún paciente, así como también medir el avance de pacientes con radiación o por el contrario el deterioro en pacientes. Claro que se necesitaría un mayor número de imágenes y muchas más clasificaciones, pero a la larga daría mejores resultado.

# Capítulo

# 6

## Bibliografía

[1] D.E. Haines y A.C. Terrell. (2019). Introducción a la estructura e imagen del sistema nervioso central. En Principios de neurociencia(3-14). España: Elsevier.

[2] Ludwing V Romero F. (2015). Anatomía y Fisiología del Sistema Nervioso. Madrid: XinXii.

[3] Bruce R. Ransom. (2017). Organización del sistema nervioso. En Fisiología médica (254-274). España: elsevier.

[4] D.E. Haines y G.A. Mihailoff. (2018). Principios de neurociencia. Barcelona, España: Elsevier.

[5] Walter F. B. y Emile L. B. (2017). Fisiología médica. España: ELSEVIER.

[6] W. A. Grow. (2019). Córtex cerebral. En Principios de Neurociencia(468-479). España: Elsevier.

[7] W. A. Grow. (2019). Desarrollo del sistema nervioso. En Principios de Neurociencia(72-89). España: Elsevier.

[8] Shlomo Breznitz . (2021). Partes del cerebro. Anatomía del cerebro. 10 de septiembre del 2021, de CogniFit research Sitio web: <https://www.cognifit.com/es/partes-del-cerebro>.

[9] Susan L. McElroy, David B. Allison and George A. Bray. (2006). Obesity and Mental Disorders. New York, NY: Taylor & Francis Group.

[10] M. J. Pérez Miguelsanz, W. Cabrera Parra, G. Varela Moreiras y M. Garaulet. (2010). Distribución regional de la grasa corporal. Uso de técnicas de imagen como herramienta de diagnóstico nutricional. Nutrición Hospitalaria, 2, 17.

[11] Narbona M., León R., Gómez L. (2021). ¿Qué es la obesidad y cuántos tipos existen? 13 de septiembre de 2021, de Obymed Sitio web: <https://www.obymed.es/tipos-de-obesidad/>.

[12] Peter G. Kopelman, Ian D. Caterson and William H. Dietz. (2010). Clinical Obesity in Adults and Children. London, UK: Wiley-Blackwell.

[13] M. Chueca, C. Azcona, M. Oyarzábal. (2002). Obesidad infantil. 1 de agosto de 2021, de ANALES Sis San Navarra Sitio web: <http://5495-Texto%20del%20artículo-8372-1-10-20090220.pdf>.

[14] Laura O. Zufiría y Javier L. Martínez. (2006). Monografía SERAM Aprendiendo

los fundamentos de la resonancia magnética. Madrid, España: Médica Panamericana.

[15] Kayvan Najarian y Robert Splinter. (2012). Biomedical Signal and Image Processing. Boca Raton: CRC Press.

[16] L. Oleaga y J. Lafuente. (2007). Monografía SERAM Aprendiendo los fundamentos de la Resonancia Magnética. España: Médica Panamericana.

[17] Javier Lafuente, Miguel Ángel García García-Tuñón y Antonio Hernando. (Enero-marzo 2020). La imagen por resonancia magnética: sus fundamentos. Revista Española de Física, vol. 34-1, pp 17-24.

[18] Javier Lafuente Martínez y Luis Hernández Moreno. (2008). Capítulo 1: Técnica de la Imagen por Resonancia Magnética. En Cerezal Pesquera L, García Valtuille R, Abascal Abascal F, Corral Sampedro JF, editores. Resonancia Magnética del sistema músculo esquelético (9-28). Madrid, España: SERAM.

[19] WA Worthoff, SD Yun y NJ Shah. (2018). Capítulo 1: Introducción a Imágenes Resonancia Magnética. En Hybrid MR-PET Imaging: Sistemas, métodos y aplicaciones (1-44). Canadá: Royal Society of Chemistry.

[20] N. Smith and A. Webb. (2011). Introduction to Medical Imaging. Cambridge: Cambridge University Press.

[21] Jorge Paz Gutiérrez, Perla Salgado, Sergio Gómez LLata. (2007). Utilidad de la técnica bold de resonancia magnética funcional en los tumores intracraneales de

pacientes prequirúrgicos. 18 de noviembre de 2021, de Medigraphic Sitio web: <https://www.medigraphic.com/pdfs/arcneu/ane-2007/ane073d.pdf>.

[22] Perla Salgado-Lujambio, Rafael Rojas-Jasso, Rosa Delgado-Hernández, Julián Sánchez-Cortazar, Gerardo Reynoso-Padilla y Fernando A. Barrios. (2002). Aplicaciones Clínicas de la Resonancia Magnética Funcional. 19 de noviembre de 2021, de Anales de Radiología México Sitio web: <https://www.medigraphic.com/pdfs/anaradmex/arm-2002/arm022g.pdf>.

[23] Fernando Berzal. (2018). Redes Neuronales & Deep Learning. Granada, España: EUG, Editorial Universidad de Granada.

[24] A. Vedaldi, K. Lenc, A. Gupta. (2015) MatConvNet Convolutional Neural Networks for MATLAB. Oxford: University Oxford.

[25] Xabier Basogain Olabe. (2008). Redes Neuronales Artificiales y sus Aplicaciones . Bilbao, España: EHU .

[26] K. Sung, C. Riley. (2013). Applications of Convolution in Image Processing with MATLAB. Washington: University of Washington.

[27] Mauricio Delbracio, José Lezama, Guillermo Carbajal. (2017). Aprendizaje profundo para visión artificial. 24 de septiembre de 2021, de Universidad de la República Sitio web: <https://iie.fing.edu.uy/~mdelbra/DL2017/slides/c11.pdf>

[28] Mohit Deshpande. (2017). Aprendizaje automático de Python. 16 de septiembre

de 2021, de ZENVA Sitio web: <https://pythonmachinelearning.pro/understanding-advanced-convolutional-neural-networks/>.

[29] Muneeb ul Hassan. (2018). VGG16 – Convolutional Network for Classification and Detection. 24 de septiembre de 2021, de Neurohive Sitio web: [https://neurohive.io/en/popular-networks/vgg16/?fbclid=IwAR1\\_8qTXKK2Jo4\\_WpXdoGnbK46g2NpiUorJeFUf5BxezUVt3AhfF4F-FTuY/](https://neurohive.io/en/popular-networks/vgg16/?fbclid=IwAR1_8qTXKK2Jo4_WpXdoGnbK46g2NpiUorJeFUf5BxezUVt3AhfF4F-FTuY/).

[30] Joanna Jaworek-Korjakowska, Pawel Kleczek, Marek Gorgon. (17 de junio de 2019). Melanoma Thickness Prediction Based on Convolutional Neural Network with VGG-19 Model Transfer Learning. IEEE Xplore, 1, 9.

[31] Richmond Alake. (2020). Aprendizaje profundo: explicación de GoogLeNet. 16 de septiembre de 2021, de towards data science Sitio web: <https://towardsdatascience.com/deep-learning-googlenet-explained-de8861c82765>

[32] Tirthajyoti Sarkar . (2018). Machine learning with Python: Essential hacks and tricks. 17 de septiembre de 2021, de opensource Sitio web: [BasesNN-python.tar/BasesNN-python/BasesNN-python/Machine%20learning%20with%20Python.html](https://basesnn-python.tar/BasesNN-python/BasesNN-python/Machine%20learning%20with%20Python.html).

[33] Javier Buhigas. (2018). Todo lo que necesitas saber sobre TensorFlow, la plataforma para Inteligencia Artificial de Google. 20 de septiembre de 2021, de Puentes Digitales Sitio web: <https://puentesdigitales.com/2018/02/14/todo-lo-que-necesitas-saber-sobre-tensorflow-la-plataforma-para-inteligencia-artificial-de-google/>.

[34] Jordi Torres. (2018). DEEP LEARNING Introducción práctica con Keras. Barcelona: Independently published.

# Capítulo 7

## Apéndice I. Código de la red neuronal

Se utiliza tensorflow para crear el código y dentro de este a keras, sys y os son librerías para poder moverse en carpetas dentro del sistema operativo, después ImageDataGenerator ayuda a preprocesar las imágenes que le daremos a entrenar, el optimizers es el optimizador para entrenar el algoritmo, Sequential es la librería que permite hacer redes neuronales secuenciales, dicho de otra manera pone en orden las capas de la red y por último Convolution 2D y MaxPooling 2D son las capas donde se realizarán las convoluciones.

```
import sys
import os
from tensorflow.python.keras.preprocessing.image import ImageDataGenerator
from tensorflow.python.keras import optimizers
from tensorflow.python.keras.models import Sequential
from tensorflow.python.keras.layers import Dropout, Flatten, Dense, Activation
from tensorflow.python.keras.layers import Convolution2D, MaxPooling2D
from tensorflow.python.keras import backend as K
from keras.optimizers import Adam

K.clear_session()
```

Almacena el directorio del lugar donde se encuentran las imágenes que se tiene para entrenar y validar.

```
K.clear_session()

data_entrenamiento = r"C:\Users\Luz\Documents\Tesis\entrenamiento"
data_validacion = r"C:\Users\Luz\Documents\Tesis\validacion"
```

## Parámetros

Épocas es el número de veces que vamos a iterar sobre todos los datos durante el entrenamiento, longitud y altura es el tamaño al cual se procesan las imágenes, *batch\_size* es el número de imágenes que le vamos a mandar a procesar en cada paso, pasos es el número de veces que la información se va a procesar en cada época, *validation\_steps* corre los pasos con el set de datos de validación para que el algoritmo aprenda, *filtrosConv1* y *filtrosConv2* es el número de filtros que se aplican a cada convolucion, *tamano\_pool* es el tamaño del filtro en el *MaxPooling*, clases son las carpetas con las imágenes, en este caso son Obesos y Normopeso, *lr* es para hacer los ajustes en nuestra red para acercarse a una solución.

```
epocas=30
longitud, altura = 150, 150
batch_size = 32
pasos = 1
validation_steps = 300
filtrosConv1 = 32
filtrosConv2 = 64
tamano_filtro1 = (3, 3)
tamano_filtro2 = (2, 2)
tamano_pool = (2, 2)
clases = 2
lr = 0.0004
```

Pre procesamiento de imágenes para dárselo a la red neuronal, se crea un generador el cual dice cómo se va a pre procesar la información y después se hace la transformación de las imágenes. Rescale es donde se reescalan las imágenes, es decir cada pixel tiene un rango de 0 a 255 y al hacer el rescale se quedan con un rango de 0 a 1, *shear\_range* genera las imágenes pero las inclina para que el algoritmo aprenda que no todas las imágenes van a estar horizontales o verticales, pueden estar un poco inclinadas, *zoom\_range* toma las imágenes que tenemos y a algunas les hace zoom para que aprenda que no todos los hipotálamos van a aparecer completos, en este caso si aparecen así pero si se deseara poner más imágenes y estas un poco recortadas el algoritmo sería capaz de analizarlo, *horizontal\_flip* toma las imágenes y las invierte para que la red pueda aprender a distinguir su dirección. Se hace lo mismo para validación, solo que en este caso solo van a ser re escaladas ya que se le da a la red las imágenes tal cual son.

```

entrenamiento_datagen = ImageDataGenerator(
    rescale=1. / 255,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True)

validacion_datagen = ImageDataGenerator(
    rescale=1. / 255)

```

Entra al directorio de entrenamiento, abre las carpetas y las procesa a cierta altura y longitud específica, las procesa en un *batch\_size* y *class\_mode* es como las va a clasificar, es categórica porque tenemos dos clases: obesos y normopeso. Se hace lo mismo para la validación.

```

entrenamiento_generador = entrenamiento_datagen.flow_from_directory(
    data_entrenamiento,
    target_size=(altura, longitud),
    batch_size=batch_size,
    class_mode='categorical')

validacion_generador = validacion_datagen.flow_from_directory(
    data_validacion,
    target_size=(altura, longitud),
    batch_size=batch_size,
    class_mode='categorical')

```

```

Found 30 images belonging to 2 classes.
Found 30 images belonging to 2 classes.

```

## Crear la red neuronal convolucional.

La red está sobre una variable *cnn*, *Sequential* es para decir que la red es secuencial, es decir, tiene varias capas secuenciadas entre ellas. Después se crea la primera capa donde ña primera capa va a hacer una convolucion donde el número de filtros y el tamaño ya están definidos anteriormente, posteriormente se añade una capa de *MaxPooling*, después se añade otra capa convolucional donde los tamaños de filtros también ya se han definido y añadimos otra capa de *MaxPooling*.

```

cnn = Sequential()
cnn.add(Convolution2D(filtrosConv1, tamano_filtro1, padding="same", input_shape=(longitud, altura, 3), activation='relu'))
cnn.add(MaxPooling2D(pool_size=tamano_pool))

cnn.add(Convolution2D(filtrosConv2, tamano_filtro2, padding="same"))
cnn.add(MaxPooling2D(pool_size=tamano_pool))

cnn.add(Flatten())
cnn.add(Dense(256, activation='relu'))
cnn.add(Dropout(0.5))
cnn.add(Dense(clases, activation='softmax'))

cnn.compile(loss='categorical_crossentropy',
            optimizer=Adam(lr=lr),
            metrics=['accuracy'])

cnn.fit_generator(
    entrenamiento_generador,
    steps_per_epoch=pasos,
    epochs=epoches,
    validation_data=validacion_generador,
    validation_steps=validacion_steps)

target_dir = './modelo/'
if not os.path.exists(target_dir):
    os.mkdir(target_dir)
cnn.save('./modelo/modelo.h5')
cnn.save_weights('./modelo/pesos.h5')

```

Comenzando la parte de clasificación usamos *Flatten* para que la imagen profunda y pequeña se convierte a plana para tener una dimensión que contenga toda la información, después se manda a una capa normal donde todas las neuronas se conectan a las neuronas de la anterior capa, posteriormente *Dropout* es usado durante el entrenamiento se le apaga la mitad de las neuronas en cada paso para que la red sea capaz de crear caminos alternos a la hora de clasificar, de no hacer este paso la red solo tomará un camino y no será capaz de clasificar, agregamos otra capa densa que contiene las clases y *softmax*, este *softmax* va a decir qué porcentaje de probabilidad tiene de ser Obeso y Normopeso, por eso el porcentaje más alto es el que considerará correcto.

Agregamos *Compile* para los parámetros usados para optimizar el algoritmo para que durante el entrenamiento la función de pérdida sea *categorical\_crossentropy*, se optimice con *Adam* e intentará mejorar el *accuracy*.

Después se entrena a la red con imágenes de entrenamiento, corre cada época en los pasos datos, luego darle qué tantos pasos de validación va a tener después de cada época, por lo tanto correrá pasos en cada época y cuando la época termine correrá pasos de validación y después pasará a la siguiente época. Una vez entrenado el modelo se genera una carpeta donde guarde la estructura del modelo y los pesos de cada capa entrenada.

```
cnn.save_weights('./modelo/pesos.h5')
-----
1/1 [=====] - 1s 1s/step - loss: 0.7493 - accuracy: 0.4667
Epoch 22/30
1/1 [=====] - 1s 1s/step - loss: 0.7005 - accuracy: 0.5333
Epoch 23/30
1/1 [=====] - 1s 1s/step - loss: 0.7937 - accuracy: 0.3000
Epoch 24/30
1/1 [=====] - 1s 1s/step - loss: 0.6712 - accuracy: 0.5667
Epoch 25/30
1/1 [=====] - 1s 1s/step - loss: 0.6787 - accuracy: 0.6000
Epoch 26/30
1/1 [=====] - 1s 1s/step - loss: 0.7532 - accuracy: 0.4333
Epoch 27/30
1/1 [=====] - 1s 1s/step - loss: 0.7047 - accuracy: 0.4667
Epoch 28/30
1/1 [=====] - 1s 1s/step - loss: 0.7200 - accuracy: 0.4667
Epoch 29/30
1/1 [=====] - 1s 1s/step - loss: 0.6781 - accuracy: 0.6000
Epoch 30/30
1/1 [=====] - 1s 1s/step - loss: 0.7040 - accuracy: 0.5000
```

## Predecir

Se importan las librerías requeridas.

```
import numpy as np
from keras.preprocessing.image import load_img, img_to_array
from keras.models import load_model
```

Se cargan los archivos que ya se generaron para poder utilizar otra vez la red neuronal y al modelo se le cargan los pesos que ya se obtuvieron y están guardados en el archivo pesos.

```
longitud, altura = 150, 150
modelo = './modelo/modelo.h5'
pesos_modelo = './modelo/pesos.h5'
cnn = load_model(modelo)
cnn.load_weights(pesos_modelo)
```

Con la función *predict* va a decir si la imagen pertenece a Obesos o Normopeso, para esto se convierte en un arreglo de valores la imagen dándole una dimensión extra para poder procesar la información, después se llama a la red y se le dice que se quiere hacer una predicción sobre tal imagen dada y debe dar el arreglo de dos dimensiones el cual con el arreglo dado va a darles un numero 0 y 1, de esta manera si pertenece a la carpeta Normopeso va a arrojar un 0 y si por el contrario pertenece a Obeso dará un 1, la respuesta es dada porque arroja el índice más alto en el arreglo dado.


```
def predict(file):
    x = load_img(file, target_size=(longitud, altura))
    x = img_to_array(x)
    x = np.expand_dims(x, axis=0)
    array = cnn.predict(x)
    result = array[0]
    answer = np.argmax(result)
    if answer == 0:
        print("pred: Normopeso")
    elif answer == 1:
        print("pred: Obeso")
    return answer
```

Finalmente, solo se llama a la función que va a predecir y se le pide que entre las

imágenes aleatorias que se colocan diga a que pertenece, se prueba con dos imágenes de distintas carpetas y la red es capaz de diferenciarlas.

```
from matplotlib.pyplot import imshow
from PIL import Image
from IPython.display import Image
from IPython.display import display

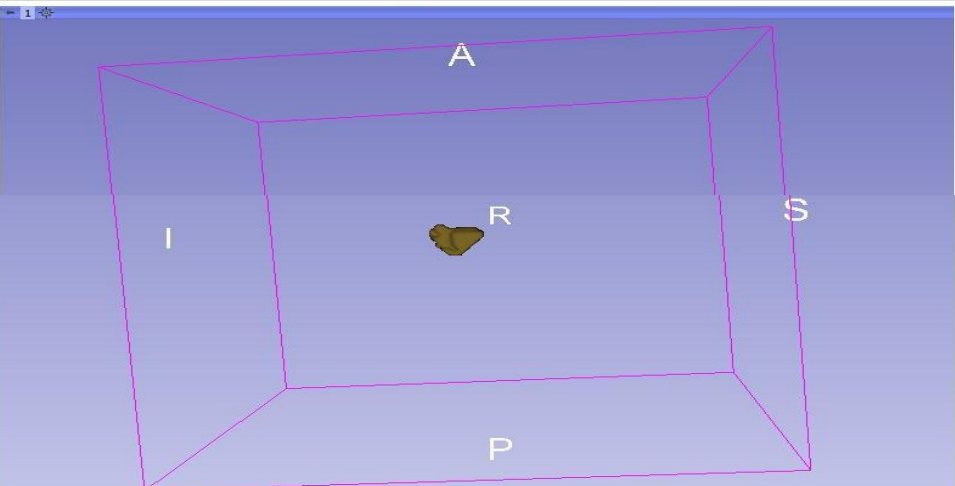
pil_im = display(Image(filename=r'C:\Users\Luz\Pictures\Normopeso - 12.jpg'))
path="/path/to/Normopeso - 12.jpg"
predict(r'C:\Users\Luz\Pictures\Normopeso - 12.jpg')
```



pred: Normopeso  
0

```
from matplotlib.pyplot import imshow
from PIL import Image
from IPython.display import Image
from IPython.display import display

pil_im = display(Image(filename=r'C:\Users\Luz\Pictures\Obesos - 30.jpg'))
path="/path/to/Obesos - 30.jpg"
predict(r'C:\Users\Luz\Pictures\Obesos - 30.jpg')
```



pred: Obeso  
1