DEEP LEARNING

TEORÍA Y APLICACIONES

Jesús Alfonso
López Sotelo

alphaeditorial

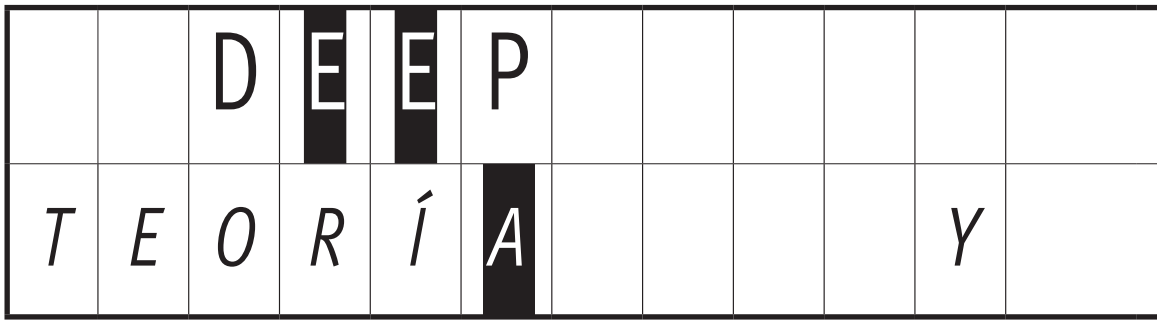
Jesús Alfonso López Sotelo

Ingeniero electricista, magíster en Automática y doctor en Ingeniería.

Tiene más de veinte años de experiencia en docencia y en desarrollo de proyectos relacionados con Inteligencia Artificial. Sus áreas de interés son las redes neuronales artificiales y el aprendizaje profundo (Deep Learning), los sistemas difusos, la computación evolutiva, la enseñanza de la inteligenacia artificial y el impacto que esta tecnología pueda tener en nuestra sociedad.

Es investigador senior del Sistema nacional de ciencia tecnología e innovación en Colombia de Colciencias, es miembro profesional del IEEE (The Institute of Electrical and Electronics Engineers, Inc.) y pertenece a la Sociedad de Inteligencia Computacional de dicho instituto. Actualmente está vinculado a la Universidad Autónoma de Occidente en Cali y pertenece al Grupo de Investigación en Energías, GIEN. Ha publicado diversos artículos, capítulos de libro y libros en las temáticas de redes neuronales artificiales y otras técnicas de inteligencia artificial

alphaeditorial


DEEP
TEORÍA
Y

JESÚS ALFONSO LÓPEZ SOTELO

LEARNING APLICACIONES

AlphaEditorial

Alpha Editorial
Alfaomega Colombiana S.A.

Calle 62 20-46 esquina, Bogotá
Teléfono (57-1) 746 0102
Fax: (57-1) 210 0122
cliente@alfaomegacolombiana.com

www.alpha-editorial.com

Primera edición: Bogotá, enero del 2021

Edición: Sandra Ardila
Portada: Ana Paula Santander

ISBN 978-958-778-686–6
ISBN 978-958-778-687–3 digital

Impreso en Colombia
Printed and made in Colombia

CONTENIDO

3. Superando las dificultades con la propagación inversa en las redes

4. Aprendizaje profundo *(Deep Learning)* o las redes de muchas capas

5. Redes convolucionales

Bibliografía

LISTA DE FIGURAS

LISTA DE TABLAS

LISTA DE VARIABLES

f_i	Activación del filtro convolucional ante *i-ésimo* patrón del minilote
f_{iNor}	Activación del filtro convolucional ante *i-ésimo* patrón del minilote normalizado
f_{iBN}	Activación del filtro convolucional ante *i-ésimo* patrón del minilote luego de aplicar *batch normalization*
$A\Delta\omega(t)$	Acumulación de las variaciones de peso
$AG(t)$	Acumulación del gradiente
b_j^o	Bias de la capa de oculta
b_k^s	Bias de la capa de salida
–	Bias o umbral de la i-ésima neurona
K	Cantidad de filtros en una capa convolucional
n	Cantidad total de pesos en una red neuronal
$x_1, x_2, \ldots x_n$	Componentes del vector de entrada
$d_1, d_2, \ldots d_M$	Componentes del vector de salidas deseadas
$y_1, y_2, \ldots y_M$	Componentes del vector de salidas producidas por la red
X^P	Conjunto de vectores de entrada. Se representa como una matriz
D^P	Conjunto de vectores de salida. Se representa como una matriz
ε	Constante usada para estabilidad numérica (evitar división por cero)
$f'(x)$	Derivada de la función a la que se desea calcular la raíz
f'^o_j	Derivada de la función de activación de la *j-ésima* neurona de la capa oculta
f'^s_k	Derivada de la función de activación de la *k-ésima* neurona de la capa de salida
L'	Derivada de la función de pérdida
σ	Desviación estándar de los datos a normalizar
σ_B	Desviación estándar del minilote
$Neta_j$	Entrada neta de j-ésima neurona
$Neta_{pj}^o$	Entrada neta de la *j-ésima* neurona de la capa oculta
$Neta_{pk}^s$	Entrada neta de la k-*ésima* neurona de la capa salida
e_j^o	Error capa oculta
e	Error producido a la salida de una neurona
α	Factor de aprendizaje
υ	Factor de decremento de la razón de aprendizaje
τ	Factor de incremento de la razón de aprendizaje
λ	Factor para ponderar la sumatoria de los pesos al cuadrado
β	Factor que define la influencia del *momentum*
$f(x)$	Función a la que se desea calcular la raíz
f_j^o	Función de activación de la *j-ésima* unidad oculta
f_k^s	Función de activación de la *j-ésima* unidad oculta

L	Función de pérdida
L_R	Función de pérdida regularizada
L_D	Función de pérdida sin regularizar, solo tiene en cuenta el aprendizaje de la salida deseada
p_i	i-ésimo pixel
H	Matriz Hessiana
W	Matriz o vector de pesos sinápticos
O	Número de neuronas de en la capa oculta
N	Número de neuronas de entrada
M	Número de neuronas en la capa de salida
P	Números de patrones de entrenamiento
p	*p-ésimo* patrón del conjunto de entrenamiento
φ	Parámetro ajustable del *batch normalization*
n	Parámetro ajustable del *batch normalization*
ζ	Parámetro de control en el algoritmo Levenberg-Marquard
κ	Parámetro de la parte exponencial de la función de activación ELU
Pa	Parámetro para definir el *zero-padding*
w_{ji}^{o}	Peso de la capa de oculta
w_{kj}^{s}	Peso de la capa de salida
w_{ji}	Peso sináptico entre la neurona j y la neurona i
μ_B	Promedio del minilote
ρ	Razón de decaimiento
y_j	Salida de j-ésima neurona
Con_Out	Salida de la convolución
y_{pj}^{o}	Salida de la *j-ésima* neurona de la capa oculta
y_{pk}	Salida de la *k-ésima* neurona de la capa de salida
y	Salida de la red neuronal artificial
d	Salida deseada
L''	Segunda derivada de la función de pérdida
S	*Stride* o corrimiento
L_W	Sumatoria de los pesos al cuadrado
W_i^{nl}m	Tamaño de la imagen de entrada
W_{out}^{lm}	Tamaño de la imagen de salida
W_{pol}	Tamaño de la ventana del *polling*
F	Tamaño del filtro en las redes convolucionales
m	Tamaño del minilote
δ^{h}_{pj}	Término de error para la *j-ésima* neurona de la capa oculta *h*
δ^{o}_{pk}	Término de error para la *k-ésima* neurona de la capa de salida
γ	Valor a determinar en el algoritmo de gradiente conjugado
x_{sn}	Valor a sin normalizar
$x_{raiz}(t)$	Valor actual de la raíz

$w(t)$	Valor actual del peso sináptico
$W(t)$	Valor actual del vector o matriz de pesos
$x_{raiz}(t+1)$	Valor actualizado de la raíz
$w(t+1)$	Valor actualizado del peso sináptico
$W(t+1)$	Valor actualizado del vector o matriz de pesos
$x_{máx}$	Valor máximo de los valores a normalizar
$y_{máx}$	Valor máximo de los valores normalizados = +1
$x_{mín}$	Valor mínimo de los valores a normalizar
$y_{mín}$	Valor mínimo de los valores normalizados = -1
y_{nor}	Valor normalizado
μ	Valor promedio de los datos a normalizar
$\Delta w(t)$	Variación del peso sináptico
$\Delta W(t)$	Variación del vector o matriz de pesos
s	Vector conjugado a r
r	Vector conjugado a s
X	Vector de entrada
$Neta$	Vector de entrada neta a una capa de neuronas artificiales
D	Vector de salida deseada
Y	Vector de salida producida por la red
$h(t)$	Vector dirección de actualización en el gradiente conjugado
G	Vector gradiente

A Andrea, Valerie y Emmanuel.

Mi familia, mis “neuronas”,
mi "capa", las "conexiones"
que activan mi vida

PRESENTACIÓN

El aprendizaje profundo o *Deep Learning* es una evolución de las redes neuronales artificiales (RNA). Las RNA constituyen una de las técnicas más relevantes de la inteligencia artificial que trata de emular la manera como trabajan las neuronas del cerebro. Este enfoque se encuentra dentro de la vertiente denominada conexionista, pues se basa en imitar el funcionamiento cerebral por medio de redes formadas por unidades sencillas (neuronas artificiales) interconectadas entre sí. Además, el conocimiento se modifica cambiando la fuerza o el valor de las conexiones existentes entre las diferentes neuronas de la red.

Históricamente, las RNA han tenido varios momentos donde, luego de ser muy utilizadas, pasaron a ser casi que abandonadas. En particular, se han presentado dos periodos de este estilo. El primero comenzó al finalizar la década de los años 60, cuando se hicieron notorias las limitaciones de las primeras RNA como el Perceptron Monocapa. Dichas limitaciones fueron extendidas por algunos investigadores a las redes más complejas de varias capas. De este periodo, las RNA salieron a mediados de los 80, luego de encontrar una solución para entrenar redes multicapa, denominada propagación inversa o *backpropagation*. Infortunadamente, el *backpropagation* funcionaba bien con redes superficiales o de una capa oculta, tratar de entrenar redes neuronales con varias capas ocultas no era viable. Lo anterior produjo el segundo de estos periodos de estancamiento a mediados de los 90. En 2006, se comenzaron a generar resultados que solventarían las dificultades para entrenar redes neuronales de varias capas ocultas, dando origen a lo que se denomina aprendizaje profundo (*Deep Learning,* DL). Actualmente, con el DL se obtienen desempeños comparables a los alcanzados por humanos en tareas de reconocimiento de imágenes y de voz, y superando las otras técnicas de aprendizaje automático en diversas aplicaciones. El DL es usado por los gigantes tecnológicos como Facebook, Google y Microsoft para sus desarrollos, como, por ejemplo, el reconocimiento automático de objetos en una imagen, el etiquetado de imágenes y la traducción simultánea.

Este libro se ha estructurado de tal manera que se realiza un seguimiento por el desarrollo histórico de las RNA, presentando lo principales conceptos asociados a ellas, partiendo desde lo modelos más simples o monocapa, hasta los modelos de muchas capas o basados en *Deep Learning*.

Inicialmente (Capítulo 1) se mostrará una breve historia de las redes neuronales artificiales desde sus primeros modelos, hasta el surgimiento del aprendizaje profundo. También se trabajarán los conceptos básicos de las redes neuronales. Estos conceptos serán fundamentales para lo que se presentará en el resto del libro.

Históricamente hablando, las primeras RNA datan de finales de los años 50 y comienzos de los 60, época en que se propusieron las redes monocapa como el Perceptron y el Adaline; estos modelos se trabajarán en el Capítulo 2. El entendimiento de

cómo se entrenan este tipo de RNA es clave para comprender redes neuronales superficiales y profundas, como las que se presentarán en los siguientes capítulos de este libro.

Posteriormente, se trabajarán las redes neuronales multicapa de una capa oculta (Capítulo 3). Estas redes se denominan redes superficiales. Se revisarán sus algoritmos de entrenamiento y un concepto clave en el desarrollo de las redes neuronales artificiales, como es el de la propagación inversa o *backpropagation*. El desarrollo de este concepto permitió que la RNA salieran de su primer periodo de inactividad y produjera a su vez uno de mucho desarrollo.

En el Capítulo 4 se revisarán los conceptos fundamentales de *Deep Learning* y una de sus arquitecturas más representativas, como son los autocodificadores (*autoencoders*). El *Deep Learning* ha permitido que la RNA se convierta hoy en día en la técnica de inteligencia artificial con más aplicación y proyección a mediano plazo. De hecho, la mayoría de aplicaciones que hoy se publicitan como grandes logros en el desarrollo de la inteligencia artificial se basan en esta técnica.

Si el *Deep Learning* fuera una película, las redes convolucionales (*Convolutional Neural Networks,* CNN) serían los protagonistas. Estas redes se verán en el Capítulo 5. Se presentarán las operaciones que realiza este tipo de redes neuronales, su arquitectura y su entrenamiento. Una de las aplicaciones relevantes de las redes convolucionales es la clasificación de imágenes, por lo que se revisará cómo se puede utilizar este tipo de red para este fin.

Este libro, además de presentar los principales conceptos teóricos de las redes neuronales artificiales y del aprendizaje profundo, se centra en mostrar cómo se puede aplicar esta tecnología para solucionar diversos problemas. Para este propósito, al final de cada capítulo se ha preparado una sección especial. En dicha sección se expondrá cómo implementar redes neuronales artificiales en diversas plataformas, como Matlab y Arduino. Lo anterior con el fin de resolver problemáticas tan diversas como el aprendizaje de funciones de una o más variable, el modelado de sistemas dinámicos, el filtrado adaptativo de señales, el reconocimiento de caracteres alfabéticos y numéricos, clasificación de tipos de flores, el reconocimiento de patrones en imágenes, entre otras.

CAPÍTULO 1

CONCEPTOS SOBRE REDES NEURONALES Y *DEEP LEARNING*

INTRODUCCIÓN

Las redes neuronales artificiales (RNA) son una técnica de inteligencia artificial clasificada dentro del aprendizaje automático o *Machine Learning*. Las RNA siguen una tendencia diferente a los enfoques clásicos de la inteligencia artificial que tratan de modelar la inteligencia humana buscando imitar los procesos de razonamiento que ocurren en nuestro cerebro. El enfoque en que se hallan las RNA se encuentra dentro de la vertiente denominada conexionista, que imita el funcionamiento cerebral por medio de redes formadas por unidades sencillas (neuronas artificiales) interconectadas entre sí. Además, el conocimiento se modifica cambiando la fuerza o el valor de las conexiones existentes entre las diferentes neuronas de la red.

El enfoque de las RNA no es reciente, pero en los últimos años ha resurgido el interés en ellas debido a una técnica que permite entrenar redes de muchas capas ocultas denominada aprendizaje profundo o *Deep Learning*. El aprendizaje profundo es la base en la cual se apoya la mayoría de los últimos avances actuales en inteligencia artificial.

Para entender el aprendizaje profundo, se hace necesario comenzar con los conceptos iniciales y básicos asociados a las RNA, los cuales serán los temas principales de este capítulo.

Objetivos del capítulo

- Comprender la relación entre inteligencia artificial, aprendizaje automático y aprendizaje profundo.
- Estudiar el desarrollo histórico que dio origen al aprendizaje profundo
- Comprender el modelo de neurona artificial y su relación con el funcionamiento de las neuronas biológicas.
- Conocer las diferentes arquitecturas de redes neuronales y los tipos de aprendizaje que se usan en las mismas.
- Presentar algunas aplicaciones del aprendizaje profundo.

Inteligencia artificial, redes neuronales artificiales y Deep Learning

El término de inteligencia artificial se le atribuye a John McCarthy, quien, en 1956, la definió como "la ciencia e ingenio de hacer máquinas inteligentes, especialmente programas de cómputo inteligentes".

De una manera coloquial, la podemos definir como el conjunto de técnicas creadas por el ser humano, que tratan de emular la inteligencia biológica; por razones obvias, el santo grial sería generar un algoritmo que emule la inteligencia humana considerada la más compleja entre la de todos los seres vivos.

Dentro de la inteligencia artificial hay diferentes vertientes, entre ellas una de las más usadas es la del aprendizaje automático *(Machine Learning)*. El aprendizaje automático son una serie de técnicas que son capaces de "aprender" las relaciones intrínsecas que hay en un conjunto de datos. Lo interesante es que no se les define a priori lo que debe aprender, lo aprende de manera automática a partir de la información que se le suministra. Si tengo una imagen y deseo reconocer si hay un gato o un perro, usando un algoritmo convencional, habría que especificar cómo identificar a cada animal, por ejemplo, un gato tiene las orejas puntiagudas, bigotes, etc. Por otro lado, un algoritmo de aprendizaje automático se le da una serie de imágenes de gatos y perros, y de manera automática se espera que aprenda a reconocer estos animales.

Como un subconjunto de las técnicas de aprendizaje automático están las redes neuronales artificiales que, en esencia, son modelos computacionales que tratan de emular el comportamiento de las neuronas del cerebro por medio de la conexión de neuronas artificiales generalmente organizadas en forma de capas. Existen diferentes tipos de redes neuronales como los mapas autoorganizados, el Perceptron, el Adaline, entre otras.

Dentro las redes neuronales artificiales tenemos las redes que se caracterizan por tener muchas capas o profundas. Este conjunto de redes neuronales artificiales y los mecanismos que hacen posible entrenarlas constituyen el denominado aprendizaje profundo o *Deep Learning*.

En la Figura 1.1 vemos cómo se relacionan los campos de la inteligencia artificial, el aprendizaje automático, las redes neuronales artificiales y el aprendizaje profundo.

Uno de los avances que ha permitido el *Deep Learning* respecto a la técnica de aprendizaje automático tradicional es que, debido a su múltiples capas, una red neuronal profunda permite realizar de manera automática una extracción de características de la información que se le suministra como entrada; esta extracción la hace de manera jerárquica, pasando primero por características de bajo nivel, luego encontrando las de nivel intermedio hasta llegar a las de alto nivel. Tomando como ejemplo una imagen, las características de bajo nivel serían bordes y líneas, las de nivel intermedio serían contornos y las de alto nivel podrían ser formas y objetos.

Lo anterior es contrario a lo que las técnicas de aprendizaje automático tradicional realizan, donde primero se realiza una extracción de características generalmente guiada por un usuario. Con las características ya extraídas se procede a realizar la clasificación. En *Deep Learning,* la red profunda, aparte de realizar la extracción de características, puede incluir una capa clasificadora que nos permitirá generar la salida del proceso realizado por la red neuronal.

En la Figura 1.2 mostramos cómo sería el proceso de clasificar una imagen usando aprendizaje automático y *Deep Learning*. Por ejemplo, se desea saber si en la imagen de entrada hay un perro; en el enfoque tradicional de aprendizaje automático

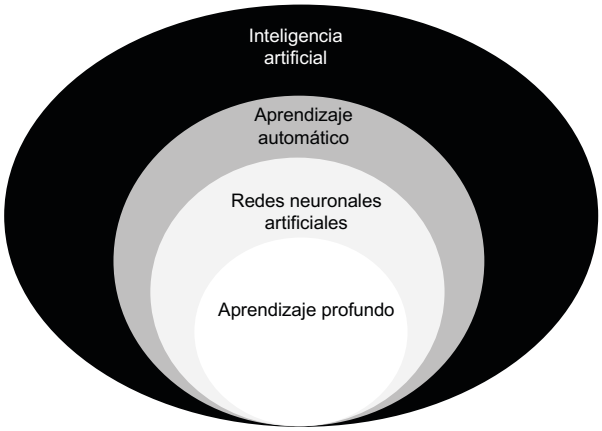


FIGURA 1.1 El aprendizaje profundo como un campo de la inteligencia artificial.

primero se extraen las características de manera manual (para esto hay que resolver la pregunta ¿cómo podemos identificar un perro? Por la forma, por las orejas, etc.) y, con las mismas, se hace una clasificación para determinar si el objeto que hay en la imagen de entrada es un perro o no. Usando *Deep Learning*, a la red profunda se le suministra la imagen de entrada y ella realiza tanto la extracción de características como la clasificación de manera automática, generando como salida nuevamente si hay o no un perro en la imagen de entrada.

Breve reseña histórica

Emular redes neuronales de manera artificial no es un desarrollo reciente. Se hicieron algunos intentos antes del advenimiento de los computadores, pero su verdadero desarrollo tuvo lugar cuando las simulaciones por computador fueron factibles por su capacidad de procesamiento y bajo costo. Luego de un periodo inicial de entusiasmo, las redes neuronales cayeron en una etapa de frustración y desprestigio, ya que el soporte económico era muy limitado, y solo unos pocos investigadores consiguieron logros de algún nivel de importancia. Estos pioneros fueron capaces de desarrollar una tecnología que sobrepasara las limitaciones identificadas en algunas publicaciones de Minsky y Papert en 1969, que sembraron un desencanto y frustración general en la comunidad científica. A este periodo se le puede denominar el primer invierno de las RNA. Luego de superar los inconvenientes para entrenar redes con capas ocultas, llegó un periodo donde se fortaleció el trabajo y la investigación con RNA. Este periodo vino acompañado de un nuevo invierno producido por la incapacidad de solventar los problemas de entrenar RNA con muchas capas ocultas o profundas. Afortunadamente, dichos inconvenientes se solventaron dando origen al denominado aprendizaje profundo, concepto que ha revitalizado el campo de la RNA y han hecho que las mismas sean una de las tecnologías con mejores resultados en aplicaciones de inteligencia artificial.

Acerquémonos a algunas etapas en las que puede resumirse la historia de las redes neuronales artificiales: