

JIVG

# Introducción a la redes neuronales y el aprendizaje profundo

Libro de ejercicios

Juan Irving Vasquez Gomez

JIVG.ORG

Copyright © 2023 Juan Irving Vasquez Gomez

PUBLISHED BY PUBLISHER

JIVG.ORG

This work is licensed under a Creative Commons Attribution 4.0 International License.

*First printing, March 2013*



# Contents

<b>1</b>	<b>Introducción</b>	<b>7</b>
1.1	Bienvenida	7
1.2	Acerca del autor	8
1.3	Agradecimientos	8

## I

## Parte Uno: Fundamentos

<b>2</b>	<b>Preeliminares</b>	<b>11</b>
2.1	Introducción	11
2.2	Álgebra lineal	11
2.2.1	Vectores y espacios vectoriales	12
2.2.2	Matrices	13
2.3	Cálculo diferencial	15
2.3.1	Ejercicios	15
2.3.2	Lecturas recomendadas	15
2.4	Probabilidad y estadística	15
2.4.1	Variable aleatoria	15
2.4.2	Distribución de probabilidad	15
2.4.3	Probabilidad marginal	16
<b>3</b>	<b>Fundamentos de redes neuronales</b>	<b>17</b>
3.1	Introducción	17
3.2	Perceptrón	18
3.2.1	El modelo del perceptrón	18

3.2.2	Ejercicios	20
3.2.3	Soluciones a los ejercicios	21
<b>3.3</b>	<b>Redes neuronales simples</b>	<b>21</b>
3.3.1	Funciones de activación	21
3.3.2	Ejercicios	25
<b>3.4</b>	<b>Redes neuronales de varias salidas</b>	<b>25</b>
3.4.1	Ejercicios	25
<b>3.5</b>	<b>Redes neuronales multicapa</b>	<b>26</b>
3.5.1	Pesos	26
3.5.2	Capas ocultas	27
3.5.3	Salida	27
3.5.4	Ejercicios	28
<b>4</b>	<b>Aprendizaje</b>	<b>29</b>
<b>4.1</b>	<b>Introducción</b>	<b>29</b>
<b>4.2</b>	<b>Descenso por gradiente</b>	<b>29</b>
4.2.1	Error	30
4.2.2	Pesos	30
4.2.3	Algoritmo	30
4.2.4	Ejercicios	31
<b>4.3</b>	<b>Retropropagación</b>	<b>31</b>
4.3.1	Algoritmo de retropropagación	32
4.3.2	Ejercicios	32
<b>4.4</b>	<b>Aplicación en clasificación</b>	<b>32</b>
<b>4.5</b>	<b>Sobre ajuste</b>	<b>33</b>
<b>4.6</b>	<b>Regularización</b>	<b>33</b>
4.6.1	Regularización L1 y L2	33
4.6.2	Deserción	34
<b>4.7</b>	<b>Optimizadores</b>	<b>34</b>
4.7.1	Descenso por gradiente estocástico (SGD)	34
<b>4.8</b>	<b>Normalización</b>	<b>35</b>

## II

## Arquitecturas de redes neuronales

<b>5</b>	<b>Redes neuronales convolucionales</b>	<b>39</b>
<b>5.1</b>	<b>Introducción</b>	<b>39</b>
5.1.1	Operación de convolución	40
5.1.2	<i>Stride</i>	40
5.1.3	<i>Padding</i>	40
5.1.4	<i>Pooling</i>	41
5.1.5	Ejercicios	41
<b>5.2</b>	<b>Aritmética de la capa de convolución</b>	<b>42</b>
5.2.1	Sin <i>padding</i> y <i>stride</i> unitario	42
5.2.2	<i>Padding</i> y <i>stride</i> unitario	43
5.2.3	<i>Same padding</i>	43

5.2.4	<i>Padding</i> completo . . . . .	43
5.2.5	Sin <i>padding</i> y <i>stride</i> no unitario . . . . .	43
5.2.6	<i>Padding</i> y <i>stride</i> no unitario . . . . .	43
5.2.7	<i>Pooling</i> . . . . .	44
5.2.8	Ejercicios . . . . .	44
<b>6</b>	<b>Autocodificadores</b> . . . . .	<b>45</b>
<b>6.1</b>	<b>Introducción</b>	<b>45</b>
<b>7</b>	<b>Transformers</b> . . . . .	<b>47</b>
<b>7.1</b>	<b>Introducción</b>	<b>47</b>
<b>8</b>	<b>Text Chapter</b> . . . . .	<b>49</b>
<b>8.1</b>	<b>Paragraphs of Text</b>	<b>49</b>
<b>8.2</b>	<b>Citation</b>	<b>49</b>
<b>8.3</b>	<b>Lists</b>	<b>49</b>
8.3.1	Numbered List . . . . .	49
8.3.2	Bullet Points . . . . .	49
8.3.3	Descriptions and Definitions . . . . .	49
<b>9</b>	<b>In-text Elements</b> . . . . .	<b>51</b>
<b>9.1</b>	<b>Theorems</b>	<b>51</b>
9.1.1	Several equations . . . . .	51
9.1.2	Single Line . . . . .	51
<b>9.2</b>	<b>Definitions</b>	<b>51</b>
<b>9.3</b>	<b>Notations</b>	<b>52</b>
<b>9.4</b>	<b>Remarks</b>	<b>52</b>
<b>9.5</b>	<b>Corollaries</b>	<b>52</b>
<b>9.6</b>	<b>Propositions</b>	<b>52</b>
9.6.1	Several equations . . . . .	52
9.6.2	Single Line . . . . .	52
<b>9.7</b>	<b>Examples</b>	<b>52</b>
9.7.1	Equation and Text . . . . .	52
9.7.2	Paragraph of Text . . . . .	53
<b>9.8</b>	<b>Exercises</b>	<b>53</b>
<b>9.9</b>	<b>Problems</b>	<b>53</b>
<b>9.10</b>	<b>Vocabulary</b>	<b>53</b>
<b>10</b>	<b>Presenting Information</b> . . . . .	<b>55</b>
<b>10.1</b>	<b>Table</b>	<b>55</b>
<b>10.2</b>	<b>Figure</b>	<b>55</b>
	<b>Bibliography</b> . . . . .	<b>57</b>
	<b>Books</b>	<b>57</b>
	<b>Articles</b>	<b>57</b>

<b>Index</b>	<b>59</b>
--------------	-----------



# 1. Introducción

## 1.1 Bienvenida

Bienvenidos a la apasionante travesía por el mundo de las redes neuronales y el aprendizaje profundo. Este libro está diseñado como una guía completa para estudiantes de licenciatura interesados en adentrarse en el campo de las redes neuronales, una de las áreas más dinámicas y revolucionarias de la inteligencia artificial y el aprendizaje automático.

Al adentrarnos en este viaje, exploraremos no solo la teoría que subyace a estas herramientas, sino también su aplicación práctica. A través de ejercicios matemáticos cuidadosamente seleccionados, este libro tiene como objetivo proporcionar una comprensión sólida y práctica del funcionamiento fundamental de las redes neuronales. Estos ejercicios no solo reforzarán los conceptos teóricos, sino que también desarrollarán habilidades prácticas esenciales para aplicar estos conceptos en problemas del mundo real. Encontraremos que los ejercicios matemáticos irán acompañados de su respectiva implementación en el lenguaje de programación Python.

Este libro abarca una variedad de temas clave, comenzando con el perceptrón, la unidad básica y fundamental de las redes neuronales. A través de este simple modelo matemático, sentiremos las bases para comprender cómo las neuronas artificiales pueden imitar, hasta cierto punto, las funciones de las neuronas biológicas en el cerebro humano.

A continuación, nos adentraremos en el mundo de las redes neuronales convolucionales (CNNs), esenciales en el procesamiento y análisis de imágenes. Exploraremos cómo estas redes pueden identificar patrones, objetos y características en imágenes, jugando un papel crucial en aplicaciones que van desde el reconocimiento facial hasta la clasificación de imágenes médicas. Dedicaremos especial atención a la interacción de los Kernels con las entradas para formar los mapas de características.

Los autocodificadores, otra pieza fundamental en nuestro viaje, nos mostrarán cómo las redes neuronales pueden aprender a comprimir y descomprimir datos, una habilidad clave para la reducción de la dimensionalidad y la generación de representaciones más eficientes de los datos. Esto nos ayudará a comprender cómo es que podemos generar nuevos ejemplos a partir de estas representaciones reducidas.

Finalmente, abordaremos los transformers, una innovación relativamente reciente que ha revolu-



cionado el procesamiento del lenguaje natural y la visión computacional generativa. Aprenderemos cómo estas redes pueden manejar secuencias de datos, como textos o series temporales, de una manera más efectiva que las técnicas anteriores.

Este libro está diseñado no solo para proporcionar conocimientos, sino también para inspirar curiosidad y fomentar la experimentación. Los ejercicios prácticos, combinados con explicaciones detalladas, están pensados para estimular el pensamiento crítico y la creatividad, elementos esenciales para cualquier aspirante a científico de datos o ingeniero de inteligencia artificial.

Así que, citando a Constantino Cavafis, "Ten siempre en tu mente a Ítaca. La llegada allí es tu destino. Pero no apresures tu viaje en absoluto". Con mente abierta y entusiasmo, nuestra meta será el entendimiento de las redes neuronales, pero iremos paso a paso armando un rompecabezas que al principio parece abrumador. Finalmente, los invito a sumergirse en el mundo de las redes neuronales, una herramienta poderosa y versátil que continúa redefiniendo lo que es posible en la era de la inteligencia artificial.

## 1.2 Acerca del autor

Juan Irving Vasquez recibió los grados de maestría en ciencias y doctorado en ciencias por el Instituto Nacional de Astrofísica Óptica y Electrónica (INAOE) en 2009 y 2014 respectivamente. El grado de Ingeniero en Sistemas Computacionales lo adquirió por el Instituto Tecnológico de Tehuacán en 2006. De 2016 a 2021 fue investigador del programa cátedras CONACYT. Desde 2021 es profesor de tiempo completo en el Instituto Politécnico Nacional. Su producción científica incluye diversas publicaciones en revistas arbitradas y congresos internacionales, así como desarrollos tecnológicos aplicados a la industria. Sus intereses actuales de investigación incluyen visión computacional basada en aprendizaje, robótica móvil, planificación de movimientos así como sus aplicaciones a vehículos autónomos. Desde 2017 ha sido reconocido como investigador nacional por parte del CONACYT, actualmente en nivel 1.

## 1.3 Agradecimientos

Fotografías: Clint Adair de Unsplash. Dall-e 3.