

JIVG.ORG

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

1	Introducción	7
1.1	Bienvenida	7
1.2	Acerca del autor	8
1.3	Agradecimientos	8

I

Parte Uno: Fundamentos

2	Preeliminares	11
2.1	Introducción	11
2.2	Álgebra lineal	11
2.2.1	Vectores y espacios vectoriales	12
2.2.2	Matrices	13
2.2.3	Ejercicios	14
2.3	Cálculo	15
2.3.1	Cálculo diferencial	15
2.3.2	Derivadas parciales	15
2.3.3	Vector gradiente	15
2.3.4	Ejercicios	15
2.3.5	Lecturas recomendadas	16
2.4	Probabilidad y estadística	16
2.4.1	Probabilidad	16
2.4.2	Probabilidad condicional	16
2.4.3	Regla de Bayes	16
2.4.4	Variable aleatoria	16
2.4.5	Distribución de probabilidad	17

2.4.6	Probabilidad marginal	17
2.4.7	Ejercicios	17
2.5	Programación en Python	18
2.5.1	Numpy	18
3	Fundamentos de redes neuronales	21
3.1	Introducción	21
3.2	Modelo McCulloch y Pitts	22
3.2.1	Reglas de operación	22
3.2.2	Implementación	23
3.2.3	Ejercicios	23
3.3	Perceptrón	24
3.3.1	Modelo del perceptrón	24
3.3.2	Implementación	26
3.3.3	Ejercicios	27
3.4	Redes neuronales simples	28
3.4.1	Funciones de activación	28
3.4.2	Implementación	31
3.4.3	Ejercicios	31
3.5	Redes neuronales de varias salidas	33
3.5.1	Ejercicios	35
3.6	Redes neuronales multicapa	35
3.6.1	Capas ocultas	36
3.6.2	La base de muchos avances recientes	37
3.6.3	Ejercicios	37
4	Aprendizaje	39
4.1	Introducción	39
4.2	Descenso por gradiente	39
4.2.1	Métricas de error	39
4.2.2	Proceso general de descenso por gradiente	40
4.2.3	Uso en redes neuronales	40
4.2.4	Implementación	42
4.3	Retropropagación	42
4.3.1	Algoritmo de retropropagación	44
4.4	Inicialización de los pesos	44
4.4.1	Inicializaciones prácticas	44
4.5	Sobre ajuste	45
4.6	Regularización	45
4.6.1	Regularización L1 y L2	45
4.6.2	Deserción	46
4.7	Optimizadores	46
4.7.1	Descenso por gradiente estocástico (SGD)	46
4.7.2	Momento (<i>momentum</i>)	47
4.7.3	Caída de la tasa de aprendizaje <i>Learning rate decay</i>	47

4.8	Normalización	48
4.9	Ejercicios	48
4.10	Acordeón	49

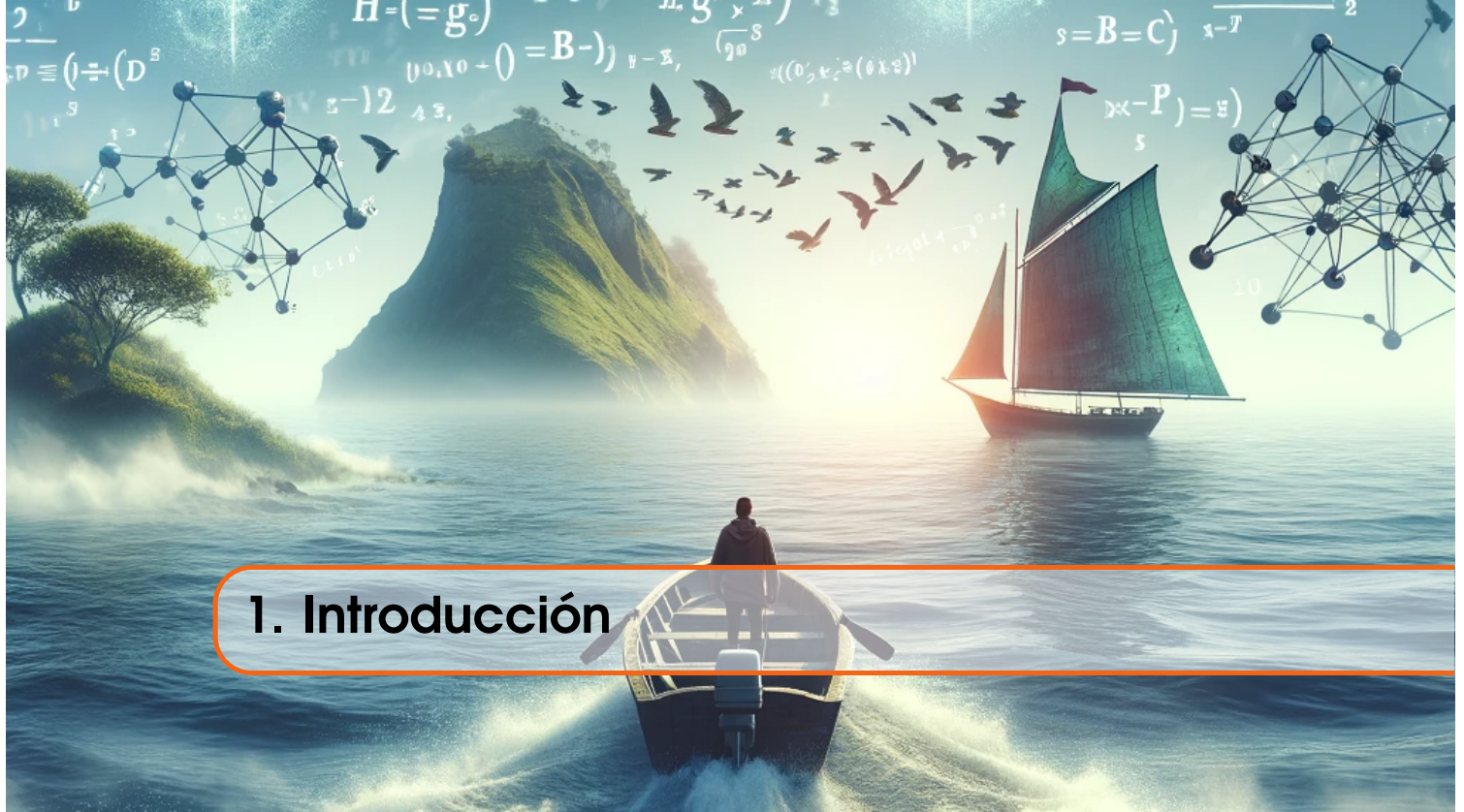
5	Aplicaciones prácticas	51
5.1	Aplicación en clasificación	51
5.2	Entrenamiento	51
5.2.1	Validación	51
5.2.2	Rendimiento	51
5.2.3	Sobre-ajuste	51

II

Arquitecturas de redes neuronales

6	Redes neuronales convolucionales	55
6.1	Introducción	55
6.1.1	Operación de convolución	56
6.1.2	Relleno (<i>Padding</i>)	58
6.1.3	Salto <i>Stride</i>	58
6.1.4	Submuestreo <i>Pooling</i>	59
6.2	Convolución multicanal	61
6.3	Apilación de mapas de características	61
6.4	Convoluciones especiales	61
6.5	Aplanado y regresión	61
6.6	Ejercicios	61
7	Autocodificadores	65
7.1	Introducción	65
7.2	Codificador	65
7.3	Decodificador	65
7.3.1	Convolución transpuesta	65
8	Transformers	67
8.1	Introducción	67
9	In-text Elements	69
9.1	Theorems	69
9.1.1	Several equations	69
9.1.2	Single Line	69
9.2	Definitions	69
9.3	Notations	70
9.4	Remarks	70
9.5	Corollaries	70

9.6	Propositions	70
9.6.1	Several equations	70
9.6.2	Single Line	70
9.7	Examples	70
9.7.1	Paragraph of Text	70
9.8	Problems	71
9.9	Vocabulary	71
10	Presenting Information	73
10.1	Table	73
10.2	Figure	73
	Bibliography	75
	Books	75
	Articles	77
	Index	79



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.