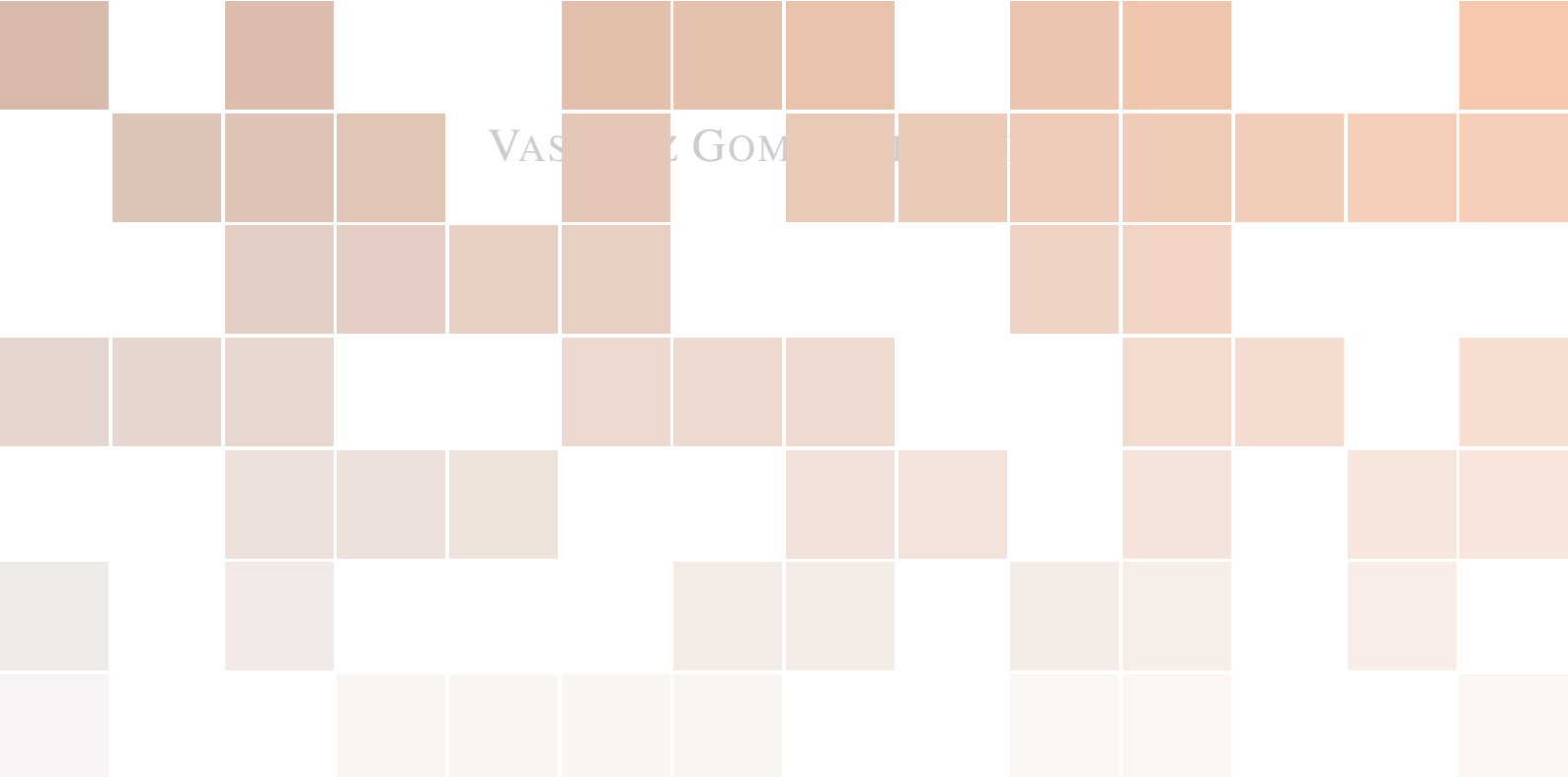


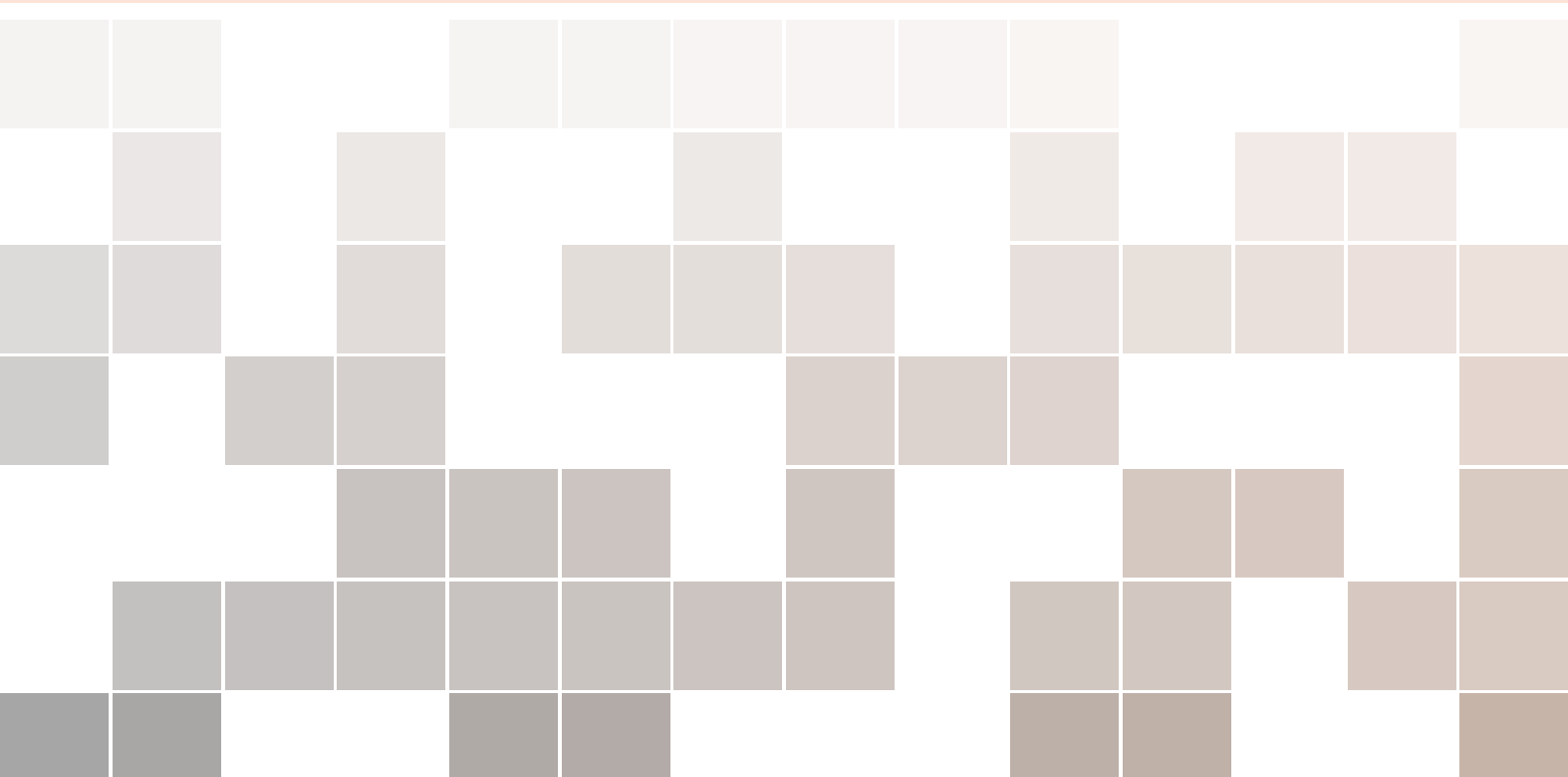


VASQUEZ GOMEZ

Introducción a la redes neuronales y el aprendizaje profundo

Libro de ejercicios

Juan Irving Vasquez Gomez



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	9
1.1	Bienvenida	9
1.2	Acerca del autor	10
1.3	Agradecimientos	10

I

Parte Uno: Fundamentos

2	Preeliminares	15
2.1	Introducción	15
2.2	Programación en Python	15
2.2.1	Numpy	16
2.2.2	Ejercicios	17
2.3	Álgebra lineal	17
2.3.1	Vectores y espacios vectoriales	17
2.3.2	Matrices	18
2.3.3	Ejercicios	20
2.4	Cálculo	20
2.4.1	Calculo diferencial	21
2.4.2	Derivadas parciales	21
2.4.3	Vector gradiente	21
2.4.4	Ejercicios	21
2.4.5	Lecturas recomendadas	22
2.5	Probabilidad y estadística	22
2.5.1	Probabilidad	22

2.5.2	Probabilidad condicional	22
2.5.3	Regla de Bayes	22
2.5.4	Variable aleatoria	22
2.5.5	Distribución de probabilidad	23
2.5.6	Probabilidad marginal	23
2.5.7	Ejercicios	24
2.5.8	Regresión lineal	24
2.6	Aprendizaje automático	24
2.6.1	Tareas a resolver	25
2.6.2	Tipos de aprendizaje	26
2.7	Datos	27
2.8	Optimización	27
2.8.1	Conceptos de optimización	27
2.9	Descenso por gradiente	28
2.9.1	Implementación	29
2.9.2	Ejercicios	30
3	Fundamentos de redes neuronales	33
3.1	Introducción	33
3.2	Modelo McCulloch y Pitts	34
3.2.1	Reglas de operación	34
3.2.2	Parámetros	36
3.2.3	Implementación	36
3.2.4	Limitaciones del modelo	37
3.2.5	Ejercicios	37
3.3	El perceptrón	37
3.3.1	Modelo del perceptrón	38
3.3.2	Selección de parámetros	41
3.3.3	Implementación	42
3.3.4	Ejercicios	42
3.3.5	Limitaciones del perceptrón	44
3.4	Redes neuronales simples	44
3.4.1	Funciones de activación	45
3.4.2	Implementación	49
3.4.3	Ejercicios	49
3.5	Redes neuronales de varias salidas	51
3.5.1	Parámetros	52
3.5.2	Ejemplos	52
3.5.3	Ejercicios	53
3.6	Redes neuronales multicapa	54
3.6.1	Capas ocultas	54
3.6.2	Implementación	57
3.6.3	La base de muchos avances recientes	57
3.6.4	Ejercicios	58
3.7	Acordeón de redes neuronales	61

4	Aprendizaje	63
4.1	Introducción	63
4.2	Regla de aprendizaje del perceptrón	63
4.3	Descenso por gradiente	64
4.3.1	Métricas de error	64
4.3.2	Proceso general de descenso por gradiente	65
4.3.3	Uso en redes neuronales	65
4.3.4	Derivadas de las funciones de activación	67
4.3.5	Actualización para múltiples ejemplos	69
4.3.6	Implementación eficiente	69
4.3.7	Ejercicios	71
4.4	Retropropagación	72
4.4.1	Retropropagación eficiente	74
4.4.2	Ejercicios	76
4.4.3	Algoritmo de retropropagación	77
4.4.4	Implementación de la retropropagación	77
4.4.5	Ejercicios	77
4.5	Inicialización de los pesos	77
4.5.1	Problema del mal condicionamiento	78
4.5.2	Inicializaciones prácticas	79
4.6	Sobre-ajuste	80
4.7	Regularización	80
4.7.1	Paro anticipado	80
4.7.2	Regularización L1 y L2	80
4.7.3	Deserción	81
4.8	Optimización	81
4.8.1	Descenso por gradiente estocástico (SGD)	82
4.8.2	Momento (<i>momentum</i>)	82
4.8.3	Caída de la tasa de aprendizaje <i>Learning rate decay</i>	83
4.9	Normalización	84
4.10	Ejercicios	84
4.11	Acordeón de aprendizaje	85

II

Arquitecturas de redes neuronales

5	Redes neuronales convolucionales	89
5.1	Introducción	89
5.1.1	Operación de convolución	90
5.2	Convolución bidimensional	90
5.2.1	Tamaño de la salida	91
5.2.2	Parámetros	93
5.2.3	Gradiente	93
5.2.4	Implementación	93
5.2.5	Relleno (<i>Padding</i>)	94
5.2.6	Salto <i>Stride</i>	95

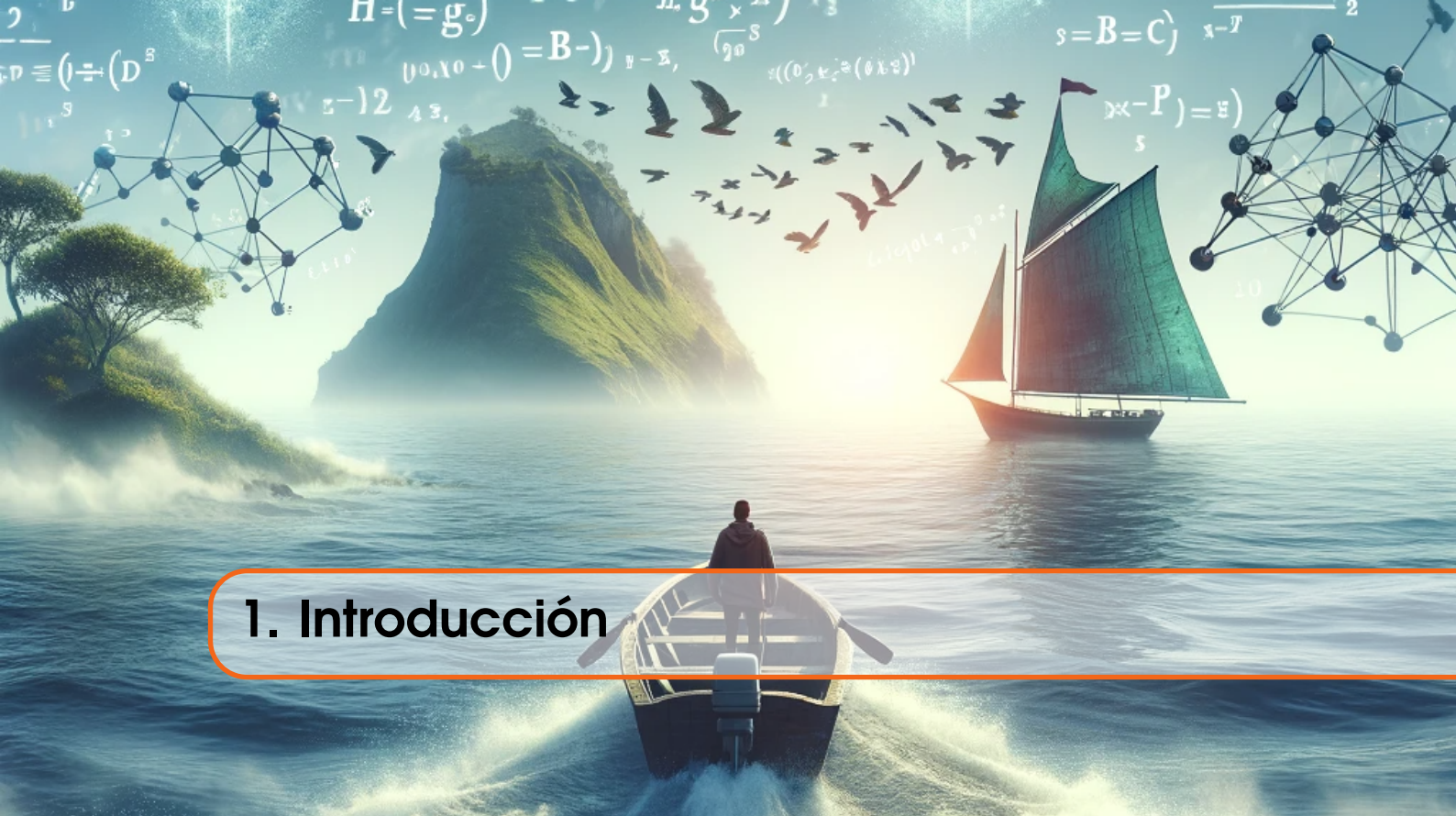
5.3	Submuestreo <i>Pooling</i>	96
5.4	Convolución multicanal	97
5.4.1	Implementación	98
5.5	Apilado	99
5.6	Convoluciones especiales	99
5.6.1	Convolución puntual	100
5.6.2	Convolución separable	100
5.7	Aplanado y regresión	100
5.8	Ejercicios	100
5.9	Acordeon	103
6	Autocodificadores	105
6.1	Introducción	105
6.2	Codificador	105
6.3	Decodificador	105
6.3.1	Convolución transpuesta	105
7	Transformers	107
7.1	Introducción	107
8	In-text Elements	109
8.1	Theorems	109
8.1.1	Several equations	109
8.1.2	Single Line	109
8.2	Definitions	109
8.3	Notations	110
8.4	Remarks	110
8.5	Corollaries	110
8.6	Propositions	110
8.6.1	Several equations	110
8.6.2	Single Line	110
8.7	Problems	110
8.8	Vocabulary	110

III

Aplicación práctica

9	Aplicaciones prácticas	113
9.1	Pre-procesamiento de los datos	113
9.1.1	Adaptación de la base de datos	113
9.2	Regresión	113
9.3	Clasificación	113
9.3.1	Codificación por unos	113
9.3.2	Softmax	114

9.3.3	Entropía cruzada	114
9.4	Entrenamiento	114
9.5	Validación	114
9.6	Evaluación del rendimiento	114
9.6.1	Matriz de confusión	114
9.7	Sobre-ajuste	115
10	Diseño de experimentos	117
10.1	Introducción	117
10.2	Nota sobre sustentabilidad	118
10.3	Una variable a la vez	118
	Index	125



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, sentaremos 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.

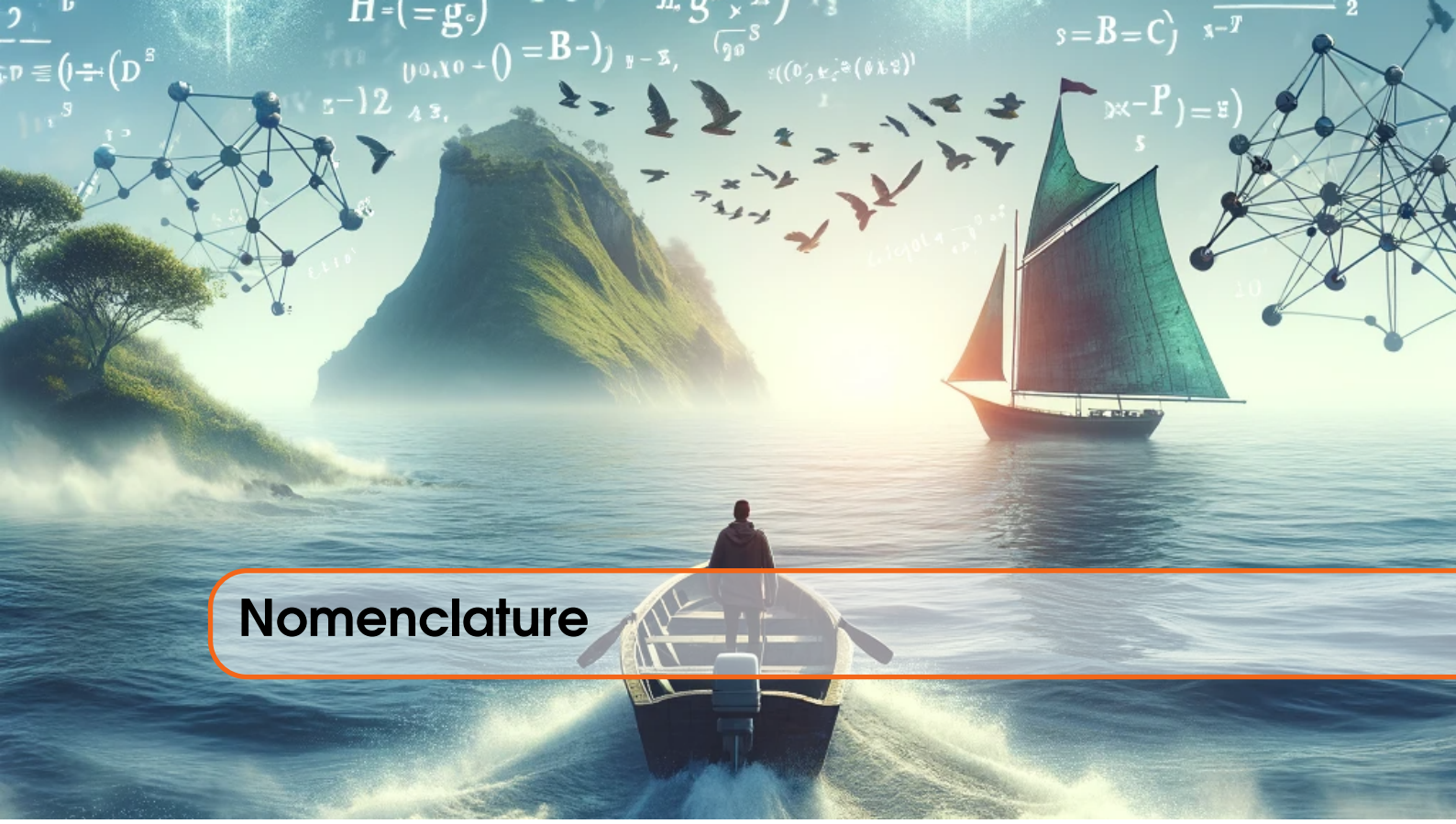
Así que, citando a Constantino Cavafis, "Ten siempre en tu mente a Ítaca. La llegada allí es tu destino. Pero no apresures tu viaje en absoluto" [2]. Con mente abierta y entusiasmo, nuestra meta será el entendimiento de las redes neuronales, pero iremos paso a paso armando un rompecabezas que al principio 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

Agradecimientos a Saulo Gante y Sergio Garrido por participar en la resolución de ejercicios. Fotografías: Clint Adair de Unsplash. Dall-e 3.

A surreal landscape featuring a sailboat with a green sail on the right, a motorboat with a person in the foreground, and a large green mountain in the background. The sky is filled with floating mathematical formulas and molecular structures. The word "Nomenclature" is written in a white box with an orange border.

Nomenclature

A	Matriz
A	Vector renglón
a	Escalar
A^T	Vector transpuesto (columna)
AB^T	Producto punto de los vectores A y B

