Machine Learning

Zhi-Hua Zhou

Machine Learning

Springer

Zhi-Hua Zhou
Nanjing University
Nanjing, Jiangsu, China

Translated by
Shaowu Liu
University of Technology Sydney
Ultimo, NSW, Australia

ISBN 978-981-15-1966-6 ISBN 978-981-15-1967-3 (eBook)
https://doi.org/10.1007/978-981-15-1967-3

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

This is an introductory-level machine learning textbook. To make the content accessible to a wider readership, the author has tried to reduce the use of mathematics. However, to gain a decent understanding of machine learning, basic knowledge of probability, statistics, algebra, optimization, and logic seems unavoidable. Therefore, this book is more appropriate for advanced undergraduate or graduate students in science and engineering, as well as practitioners and researchers with equivalent background knowledge.

The book has 16 chapters that can be roughly divided into three parts. The first part includes Chapters 1–3, which introduces the basics of machine learning. The second part includes Chapters 4–10, which presents some classic and popular machine learning methods. The third part includes Chapters 11–16, which covers advanced topics. As a textbook, Chapters 1–9 and 10 can be taught in one semester at the undergraduate level, while the whole book could be used for the graduate level.

This introductory textbook aims to cover the core topics of machine learning in one semester, and hence is unable to provide detailed discussions on many important frontier research works. The author believes that, for readers new to this field, it is more important to have a broad view than drill down into the very details. Hence, in-depth discussions are left to advanced courses. However, readers who wish to explore the topics of interest are encouraged to follow the further reading section at the end of each chapter.

The book was originally published in Chinese and had a wide readership in the Chinese community. The author would like to thank Dr. Shaowu Liu for his great effort of translating the book into English and thank Springer for the publication.

Zhi-Hua Zhou
Nanjing, China

Contents

Symbols

x	Scalar
$\boldsymbol{x}$	Vector
$\mathbf{x}$	Variable set
$\mathbf{A}$	Matrix
$\mathbf{I}$	Identity matrix
$\mathcal{X}$	Sample space or state space
$\mathcal{D}$	Probability distribution
D	Data set
$\mathcal{H}$	Hypothesis space
H	Hypothesis set
$\mathfrak{L}$	Learning algorithm
$(\cdot, \cdot, \cdot)$	Row vector
$(\cdot; \cdot; \cdot)$	Column vector
$(\cdot)^{\mathrm{T}}$	Transpose of vector or matrix
$\{\cdots\}$	Set
$\lvert\{\cdots\}\rvert$	Number of elements in set $\{\cdots\}$
$\lVert \cdot \rVert_p$	L_p norm; L_2 norm when p is absent
$P(\cdot), P(\cdot\lvert\cdot)$	Probability mass function, conditional probability mass function
$p(\cdot), p(\cdot\lvert\cdot)$	Probability density function, conditional probability density function
$\mathbb{E}_{\cdot\sim\mathcal{D}}[f(\cdot)]$	Expectation of function $f(\cdot)$ with respect to $\cdot$ over distribution $\mathcal{D}$; $\mathcal{D}$ and/or $\cdot$ are omitted when context is clear
$\sup(\cdot)$	Supremum
$\mathbb{I}(\cdot)$	Indicator function, returns 1 if $\cdot$ is true, and 0 if $\cdot$ is false
$\mathrm{sign}(\cdot)$	Sign function, returns -1 if $\cdot < 0$, 0 if $\cdot = 0$, and 1 if $\cdot > 0$