Foundations Of Machine Learning Mehryar Mohri | 4763faec99c3faade471b2028ff3a00b1

Introduction to Machine Learning, fourth edition

Data Science

This book provides an introduction to the mathematical and algorithmic foundations of data science, including machine learning, high-dimensional geometry, and analysis of large networks. Topics include the counterintuitive nature of data in high dimensions, important linear algebraic techniques such as singular value decomposition, the theory of random walks and Markov chains, the fundamentals of and important algorithms for machine learning, algorithms and analysis for clustering, probabilistic models for large networks, representation learning including topic modelling and non-negative matrix factorization, wavelets and compressed sensing. Important probabilistic techniques are developed including the law of large numbers, tail inequalities, analysis of random projections, generalization guarantees in machine learning, and moment methods for analysis of phase transitions in large random graphs. Additionally, important structural and complexity measures are discussed such as matrix norms and VC-dimension. This book is suitable for both undergraduate and graduate courses in the design and analysis of algorithms for data.

Principles of Neural Design

The term Federated Learning was coined as recently as 2016 to describe a machine learning setting where multiple entities collaborate in solving a machine learning problem, under the coordination of a central server or service provider. Each client’s raw data is stored locally and not exchanged or transferred; instead, focused updates intended for immediate aggregation are used to achieve the learning objective. Since then, the topic has gathered much interest across many different disciplines and the realization that solving many of these interdisciplinary problems likely requires not just machine learning but techniques from distributed optimization, cryptography, security, differential privacy, fairness, compressed sensing, systems, information theory, statistics, and more. This monograph has contributions from leading experts across the disciplines, who describe the latest state-of-the-art from their perspective. These contributions have been carefully curated into a comprehensive treatment that enables the reader to understand the work that has been done and get pointers to where effort is required to solve many of the problems before Federated Learning can become a reality in practical systems. Researchers working in the area of distributed systems will find this monograph an enlightening read that may inspire them to work on the many challenging issues that are outlined. This monograph will get the reader up to speed quickly and easily on what is likely to become an increasingly important topic: Federated Learning.

Perturbation theory for linear operators

The aim of this book is to discuss the fundamental ideas which lie behind the statistical theory of learning and generalization. It considers learning as a general problem of function estimation based on empirical data. Omitting proofs and technical details, the author concentrates on discussing the main results of learning theory and their connections to fundamental problems in statistics. This second edition contains three new chapters devoted to further development of the learning theory and SVM techniques. Written in a readable and concise style, the book is intended for statisticians, mathematicians, physicists, and computer scientists.

Introduction to Natural Language Processing

Apache Mahout is a scalable machine learning library with algorithms for clustering, classification, and recommendations. It empowers users to analyze patterns in large, diverse, and complex datasets faster and more scalable. This book is an all-inclusive guide to analyzing large and complex datasets using Apache Mahout. It explains complicated but very effective machine learning algorithms simply, in relation to real-world practical examples. Starting from the fundamental concepts of machine learning and Apache Mahout, this book guides you through Apache Mahout’s implementations of machine learning techniques including classification, clustering, and recommendations. During this exciting walkthrough, real-world applications, a diverse range of popular algorithms and their implementations, code examples, evaluation strategies, and best practices are given for each technique. Finally, you will learn vdata visualization techniques for Apache Mahout to bring your data to life.

The Hundred-page Machine Learning Book

A concise introduction to the emerging field of data science, explaining its evolution, relation to machine learning, current uses, data infrastructure issues, and ethical challenges. The goal of data science is to improve decision making through the analysis of data. Today data science determines the ads we see online, the books and movies that are recommended to us online, which emails are filtered into our spam folders, and even how much we pay for health insurance. This volume in the MIT Press Essential Knowledge series offers a concise introduction to the emerging field of data science, explaining its evolution, current uses, data infrastructure issues, and ethical challenges. It has never been easier for organizations to gather, store, and process data. Use of data science is driven by the rise of big data and social media, the development of high-performance computing, and the emergence of such powerful methods for data analysis and modeling as deep learning. Data science encompasses a set of principles, problem definitions, algorithms, and processes for extracting non-obvious and useful patterns from large datasets. It is closely related to the fields of data mining and machine learning, but broader in

Page 1/6
Prediction, Learning, and Games

A friendly and systematic introduction to the theory and applications. The book begins with the sums of independent random variables and vectors, with maximal inequalities and sharp estimates on moments, which are later used to develop and interpret decoupling inequalities. Decoupling is first introduced as it applies to randomly stopped processes and unbiased estimation. The authors then proceed with the theory of decoupling in full generality, paying special attention to comparison and interplay between martingale and decoupling theory, and to applications. These include limit theorems, moment and exponential inequalities for martingales and more general dependence structures, biostatistical implications, and moment convergence in Anscombe’s theorem and Wald’s equation for U–statistics. Addressed to researchers in probability and statistics and to graduates, the exposition is at the level of a second graduate probability course, with a good portion of the material fit for use in a first year course.

Interpretable Machine Learning

This is a light-hearted introduction to game theory suitable for advanced undergraduate students or beginning graduate students. It answers three questions. What is game theory? How is game theory applied? Why is game theory right?

A Guide to Convolutional Neural Networks for Computer Vision

This monograph presents a rigorous mathematical introduction to optimal transport as a variational problem, its use in modeling various phenomena, and its connections with partial differential equations. Its main goal is to provide the reader with the techniques necessary to understand the current research in optimal transport and the tools which are most useful for its applications. Full proofs are used to illustrate mathematical concepts and each chapter includes a section that discusses applications of optimal transport to various areas, such as economics, finance, potential games, image processing and fluid dynamics. Several topics are covered that have never been previously in books on this subject, such as the Knothe transport, the properties of functionals on measures, the Dacorogna-Moser flow, the formulation through minimal flows with prescribed divergence formulation, the case of the supremal cost, and the most classical numerical methods. Graduate students and researchers in both pure and applied mathematics interested in the problems and applications of optimal transport will find this to be an invaluable resource.

Playing for Real

Introduces machine learning and its algorithmic paradigms, explaining the principles behind automated learning approaches and the considerations underlying their usage.

The Beauty of Mathematics in Computer Science

An accessible introduction to the artificial intelligence technology that enables computer vision, speech recognition, machine translation, and driverless cars. Deep learning is an artificial intelligence technology that enables computer vision, speech recognition in mobile phones, machine translation, AI games, driverless cars, and other applications. When we use consumer products from Google, Microsoft, Facebook, Apple, or Baidu, we are often interacting with a deep learning system. In this volume in the MIT Press Essential Knowledge series, computer scientist John Kelleher offers an accessible and concise but comprehensive introduction to the fundamental technology at the heart of the artificial intelligence revolution. Kelleher explains that deep learning enables data-driven decisions by identifying and extracting patterns from large datasets; its ability to learn from complex data makes deep learning ideally suited to take advantage of the rapid growth in big data and computational power. Kelleher also explains some of the basic concepts in deep learning, presents a history of advances in the field, and discusses the current state of the art. He describes the most important deep learning architectures, including autoencoders, recurrent neural networks, and long short-term networks, as well as such recent developments as Generative Adversarial Networks and capsule networks. He also provides a comprehensive (and comprehensible) introduction to the two fundamental algorithms in deep learning: gradient descent and backpropagation. Finally, Kelleher considers the future of deep learning—major trends, possible developments, and significant challenges.

Apache Mahout Essentials

State-of-the-art algorithms and theory in a novel domain of machine learning, prediction when the output has structure.

Knowledge Graphs

Lifelong Machine Learning, Second Edition is an introduction to an advanced machine learning paradigm that continuously learns by accumulating past knowledge and then uses it in future learning and prediction. Lifelong learning is the dominant machine learning paradigm learns in isolation: given a training dataset, it runs a machine learning algorithm on the dataset to produce a model that is then used in its intended application. It makes no attempt to retain the learned knowledge and use it in subsequent learning. Unlike this isolated system, humans learn effectively with only a few examples precisely because our learning is very knowledge-driven: the knowledge learned in the past helps us learn new things with little data or effort. Lifelong learning aims to emulate this capability, because without it, an AI system cannot be considered truly intelligent. Research in lifelong learning has developed significantly in the relatively short time since the first edition of this book was published. The purpose of this second edition is to expand the definition of lifelong learning, update the content of several chapters, and add a new chapter about continual learning in deep neural networks—which has been actively researched over the past two or three years. A few chapters have also been reorganized to make each of them more coherent for the reader. Moreover, the authors want to propose a unified framework for the research area. Currently, there are several research topics in machine learning that are closely related to lifelong learning—most notably, multi-task learning, transfer learning, and meta-learning—because they also employ the idea of knowledge sharing and transfer. This book brings all these topics together and discusses their similarities and differences. Its goal is to introduce this emerging machine learning paradigm and present a comprehensive survey and review of the important research results and latest ideas in the area. This book is thus suitable for students, researchers, and practitioners who are interested in machine learning, data mining, natural language processing, or pattern recognition. Lecturers can readily use the book for courses in any of these related fields.
This friendly and accessible guide to AI theory and programming in Python requires no maths or data science background. Key Features Roll up your sleeves and start programming AI models Now math, data science, or machine learning background required Packed with hands-on examples, illustrations, and clear step-by-step instructions 5 hands-on working projects put ideas into action and show step-by-step how to build intelligent software Book Description AI is changing the world—and with this book, anyone can start building intelligent software! Through his best-selling video courses, Hadelin de Ponteves has taught hundreds of thousands of people to write AI software. Now, for the first time, his hands-on, energetic approach is available as a book. Taking a graduated approach that starts with the basics before easing readers into more complicated formulas and notation, Hadelin helps you understand what you really need to build AI systems with reinforcement learning and deep learning. Five full working projects put the ideas into action, showing step-by-step how to build intelligent software using the best and easiest tools for AI programming: Google Colab Python TensorFlow Keras PyTorch AI Crash Course teaches everyone to build an AI to work in their applications. Once you've read this book, you're only limited by your imagination. Who should read this book? Master the key skills of deep learning, reinforcement learning, and deep reinforcement learning. Understand Q-learning and deep Q-learning Learn from friendly, plain English explanations and practical activities Build fun projects, including a virtual-self-driving car Use AI to solve real-world business problems and win classic video games Build an intelligent, virtual robot warehouse worker Who this book is for If you want to add AI to your skillset, this book is for you. It doesn’t require data science or machine learning knowledge. Just maths basics (high school level).

Federated Learning

This book describes how neural networks operate from the mathematical point of view. As a result, neural networks can be interpreted both as function universal approximators and information processors. The book bridges the gap between ideas and concepts of neural networks, which are used nowadays at an intuitive level, and the precise modern mathematical language, presenting the best practices of the former and enjoying the robustness and elegance of the latter. This book can be used in a graduate course in deep learning, with the first few parts being accessible to senior undergraduates. In addition, the book will be of wide interest to machine learning researchers who are interested in a theoretical understanding of the subject.

Networked Life

A new edition of a graduate-level machine learning textbook that focuses on the analysis and theory of algorithms. This book is a general introduction to machine learning that can serve as a textbook for graduate students and a reference for researchers. It covers fundamental modern topics in machine learning while providing the theoretical basis and conceptual tools needed for the discussion and justification of algorithms. It also describes several key aspects of the application of these algorithms. The authors aim to present novel theoretical tools and concepts while giving concise proofs even for relatively advanced topics. Foundations of Machine Learning is unique in its focus on the analysis and theory of algorithms. The first four chapters lay the theoretical foundation for what follows in subsequent chapters and are mostly self-contained. Topics covered include the Probably Approximately Correct (PAC) learning framework; generalization bounds based on Rademacher complexity and VC-dimension; Support Vector Machines (SVMs); kernel methods; boosting; on-line learning; multi-class classification; ranking; algorithmic stability; dimensionality reduction; learning automata and languages; and reinforcement learning. Each chapter ends with a set of exercises. Appendixes provide additional material including concise probability review. This second edition offers three new chapters, on model selection, maximum entropy models, and conditional entropy models. New material in the appendices includes a major section on Fenchel duality, expanded coverage of concentration inequalities, and an entirely new entry on information theory. More than half of the exercises are new to this edition.

Advances and Open Problems in Federated Learning

A survey of computational methods for understanding, generating, and manipulating human language, which offers a synthesis of classical representations and algorithms with contemporary machine learning techniques. This textbook provides a technical perspective on natural language processing—methods for building computer software that understands, generates, and manipulates human language. It emphasizes contemporary data-driven approaches, focusing on techniques from supervised and unsupervised machine learning. The first section establishes a foundation in machine learning by building a set of tools that will be used throughout the book and applying them to word-based textual analysis. The second section introduces some classical constructions of language, including sequences, trees, and graphs. The third section explores different approaches to the representation and analysis of linguistic meaning, ranging from formal logic to neural word embeddings. The final section offers chapter-length treatments of three transformative applications of natural language processing: information extraction, machine translation, and text generation. End-of-chapter exercises include both paper-and-pencil analysis and software implementation. The text synthesizes and distills a broad and diverse research literature, linking contemporary machine learning techniques with the field’s linguistic and computational foundations. It is suitable for use in advanced undergraduate and graduate-level courses and as a reference for software engineers and data scientists. Readers should have a background in computer programming and college-level mathematics. After mastering the material presented, students will have the technical skill to build and analyze novel natural language processing systems and to understand the latest research in the field.

Foundations of Data Science

Two distinguished neuroscientists distill general principles from more than a century of scientific study, “reverse engineering” the brain to describe it, with more than fifty three-dimensional structures applying increasingly advanced methods. A mountain of new facts and mechanisms has emerged. And yet a principled framework to organize this knowledge has been missing. In this book, Peter Sterling and Simon Laughlin, two leading neuroscientists, strive to fill this gap, outlining a set of organizing principles to explain the whys of neural design that allow the brain to compute so efficiently. Setting out to “reverse engineer” the brain—disassembling it to understand it—Sterling and Laughlin first consider why an animal should need a brain, tracing computational abilities from bacteria to protozoa to worm. They examine bigger brains and the advantages of “anticipatory regulation”; identify constraints on neural design and the need to “nanofy”; and demonstrate the routes to efficiency in an integrated molecular system, phototransduction. They show that the principles of neural design at finer scales and lower levels apply at larger scales and higher levels; describe neurodegeneration; and discuss learning as a principle of biological design that includes “the only what is needed.” Sterling and Laughlin avoid speculation about how the brain might work and endeavor to make sense of what is already known. Their distinctive contribution is to gather a coherent set of basic rules and exemplify them across spatial and functional scales.

Statistical Machine Learning

An introduction to the techniques and algorithms of the newest field in robotics. Probabilistic robotics is a new and growing
area in robotics, concerned with perception and control in the face of uncertainty. Building on the field of mathematical statistics, probabilistic robotics endows robots with a new level of robustness in real-world situations. This book introduces the mathematical foundations and algorithms in the field. All algorithms are based on a single overarching mathematical foundation. Each chapter provides example implementations in pseudo code, detailed mathematical derivations, discussions from a practitioner’s perspective, and extensive lists of exercises and class projects. The book’s Web site, www.probabilistic-robotics.org, has additional material. The book is relevant for anyone involved in robotic software development and scientific research. It will also be of interest to applied statisticians and engineers dealing with real-world sensor data.

A Probabilistic Theory of Pattern Recognition

A concise and self-contained introduction to causal inference, increasingly important in data science and machine learning. The matematization of causality is a relatively recent development, and has become increasingly important in data science and machine learning. This book offers a self-contained and concise introduction to causal models and how to learn them from data. After explaining the need for causal models and discussing some of the principles underlying causal inference, the book teaches readers how to use causal models: how to compute intervention distributions, how to infer causal models from observational and interventional data, and how causal ideas could be exploited for classical machine learning problems. All of these topics are discussed first in terms of two variables and then in the more general multivariate case. The bivariate case turns out to be a particularly hard problem for causal learning because there are no conditional independences as used by classical methods for solving multivariate cases. The authors consider analyzing statistical asymmetries between cause and effect to be highly instructive, and they report on their decade of intensive research into this problem. The book is accessible to readers with a background in machine learning or statistics, and can be used in graduate courses or as a reference for researchers. The text includes code snippets that can be copied and pasted, exercises, and an appendix with a summary of the most important technical concepts.

The Nature of Statistical Learning Theory

This accessible textbook demonstrates how to recognize, simplify, model and solve optimization problems - and apply these principles to new projects.

Bandit Algorithms

Multiagent systems combine multiple autonomous entities, each having diverging interests or different information. This overview of the field offers a computer science perspective, but also draws on ideas from game theory, economics, operations research, logic, philosophy and linguistics. It will serve as a reference for researchers in each of these fields, and be used as a text for advanced undergraduate or graduate courses. The authors emphasize foundations to create a broad and rigorous treatment of their subject, with thorough presentations of distributed problem solving, game theory, multiagent communication and learning, social choice, mechanism design, auctions, cooperative game theory, and modal logics of knowledge and belief. For each topic, basic concepts are introduced, examples are given, proofs of key results are offered, and algorithmic considerations are examined. An appendix covers background material in probability theory, classical logic, Markov decision processes and mathematical programming.

Multiagent Systems

A comprehensive and rigorous introduction for graduate students and researchers, with applications in sequential decision-making problems.

Probabilistic Robotics

A Brief Introduction to Machine Learning for Engineers

A rigorous and comprehensive textbook covering the major approaches to knowledge graphs, an active and interdisciplinary area within artificial intelligence. The field of knowledge graphs, which allows us to model, process, and derive insights from complex real-world data, has emerged as an active and interdisciplinary area of artificial intelligence over the last decade, drawing on such fields as natural language processing, data mining, and the semantic web. Current projects involve predicting cyberattacks, recommending products, and even gleaning insights from thousands of papers on COVID-19. This textbook offers rigorous and comprehensive coverage of the field. It focuses systematically on the major approaches, both those that have stood the test of time and the latest deep learning methods.

Predicting Structured Data

Fundamental topics in machine learning are presented along with theoretical and conceptual tools for the discussion and proof of algorithms.

Lifelong Machine Learning

The most comprehensive book on the engineering aspects of building reliable AI systems. “If you intend to use machine learning to solve business problems at scale, I’m delighted you got your hands on this book.” - Cassie Kozyrkov, Chief Decision Scientist at Google “Foundational work about the reality of building machine learning models in production.” - Karolis Urbonas, Head of Machine Learning and Science at Amazon

AI Crash Course

Computer vision has become increasingly important and effective in recent years due to its wide-ranging applications in areas as diverse as smart surveillance and monitoring, health and medicine, sports and recreation, robotics, drones, and self-driving cars. Visual recognition tasks, such as image classification, localization, and detection, are the core building blocks of many of these applications, and recent developments in Convolutional Neural Networks (CNNs) have led to outstanding performance in these state-of-the-art visual recognition tasks and systems. As a result, CNNs now form the crux of deep learning algorithms in computer vision. This self-contained guide will benefit those who seek to both understand the theory behind CNNs and to gain hands-on experience on the application of CNNs in computer vision. It provides a comprehensive introduction to CNNs starting with the essential concepts behind neural networks: training, regularization, and optimization of CNNs. The book also discusses a wide range of loss functions, network layers, and popular CNN architectures, reviews the different techniques for
the evaluation of CNNs, and presents some popular CNN tools and libraries that are commonly used in computer vision. Further, this text describes and discusses case studies that are related to the application of CNN in computer vision, including image classification, object detection, semantic segmentation, scene understanding, and image generation. This book is ideal for undergraduate and graduate students, as no prior background knowledge in the field is required to follow the material, as well as new researchers, developers, engineers, and practitioners who are interested in gaining a quick understanding of CNN models.

**Probabilistic Machine Learning**

A self-contained and coherent account of probabilistic techniques, covering: distance measures, kernel rules, nearest neighbour rules, Vapnik-Chervonenkis theory, parametric classification, and feature extraction. Each chapter concludes with problems and exercises to further the readers understanding. Both research workers and graduate students will benefit from this wide-ranging and up-to-date account of a fast-moving field.

**Machine Learning Engineering**

A comprehensive introduction to Support Vector Machines and related kernel methods. In the 1990s, a new type of learning algorithm was developed, based on results from statistical learning theory: the Support Vector Machine (SVM). This gave rise to a new class of theoretically elegant learning machines that use a central concept of SVMs—kernels—for a number of learning tasks. Kernel machines provide a modular framework that can be adapted to different tasks and domains by the choice of the kernel function and the base algorithm. They are replacing neural networks in a variety of fields, including engineering, information retrieval, and bioinformatics. Learning with Kernels provides an introduction to SVMs and related kernel methods. Although the book begins with the basics, it also includes the latest research. It provides all of the concepts necessary to enable a reader equipped with some basic mathematical knowledge to enter the world of machine learning using theoretically well-founded yet easy-to-use kernel algorithms and to understand and apply the powerful algorithms that have been developed over the last few years.

**Understanding Machine Learning**

This book provides a comprehensive and self-contained introduction to Federated Learning, ranging from the basic knowledge and theories to various key applications, and the privacy and incentive factors are the focus of the whole book. This book is timely needed since Federated Learning is getting popular after the release of the General Data Protection Regulation (GDPR). As Federated Learning aims to enable a machine model to be collaboratively trained without each party exposing private data to others. This setting adheres to regulatory requirements of data privacy protection such as GDPR. This book contains three main parts. First, it introduces different privacy-preserving methods for protecting a Federated Learning model against different types of attacks such as Data Leakage and/or Data Poisoning. Second, the book presents incentive mechanisms which aim to encourage individuals to participate in the Federated Learning ecosystems. Last but not the least, this book also describes how Federated Learning can be applied in industry and business to address data silo and privacy-preserving problems. The book is intended for readers from both academia and industries, who would like to learn federated learning from scratch, practice its implementation, and apply it in their own business. Readers are expected to have some basic understanding of linear algebra, calculus, and neural network. Additionally, domain knowledge in FinTech and marketing are preferred.

**Optimal Transport for Applied Mathematicians**

A substantially revised fourth edition of a comprehensive textbook, including new coverage of recent advances in deep learning and neural networks. The goal of machine learning is to program computers to use example data or past experience to solve a given problem. Machine learning underlies such exciting new technologies as self-driving cars, speech recognition, and translation applications. This substantially revised fourth edition of a comprehensive, widely used machine learning textbook offers new coverage of recent advances in the field in both theory and practice, including developments in deep learning and neural networks. The book covers a broad array of topics not usually included in introductory machine learning texts, including supervised learning, Bayesian decision theory, parametric methods, semiparametric methods, nonparametric methods, multivariate analysis, hidden Markov models, reinforcement learning, kernel machines, graph theoretical methods, Bayesian estimation, and statistical testing. The fourth edition offers a new chapter on deep learning that discusses training, regularizing, and structuring deep neural networks such as convolutional and generative adversarial networks; new material in the chapter on reinforcement learning that covers the use of deep networks, the policy gradient methods, and deep reinforcement learning; new material in the chapter on multilayer perceptrons on autoencoders and the word2vec network; and discussion of a popular method of dimensionality reduction, t-SNE. New appendixes offer background material on linear algebra and optimization. End-of-chapter exercises help readers to apply concepts learned. Introduction to Machine Learning can be used in courses for advanced undergraduate and graduate students and as a reference for professionals.

**Learning with Kernels**

There is a wealth of literature and books available to engineers starting to understand what machine learning is and how it can be used in their everyday work. This presents the problem of where the engineer should start. The answer is often “for a general, but slightly outdated introduction, read this book; for a detailed survey of methods based on probabilistic models, check this reference; to learn about statistical learning, this text is useful” and so on. This monograph provides the starting point to the literature that every engineer new to machine learning needs. It offers a basic and compact reference that describes key ideas and principles in simple terms and within a unified treatment, encompassing recent developments and pointers to the literature for further study. A Brief Introduction to Machine Learning for Engineers is the entry point to machine learning for students, practitioners, and researchers with an engineering background in probability and linear algebra.

**Deep Learning Architectures**

**Foundations of Machine Learning**

A detailed and up-to-date introduction to machine learning, presented through the unifying lens of probabilistic modeling and Bayesian decision theory. This book offers a detailed and up-to-date introduction to machine learning (including deep learning) through the unifying lens of probabilistic modeling and Bayesian decision theory. The book covers mathematical background (including linear algebra and optimization), basic supervised learning (including linear and logistic regression and deep neural networks), as well as more advanced topics (including transfer learning and unsupervised learning). End-of-chapter exercises allow students to apply what they have learned, and an appendix covers notation. Probabilistic Machine
Learning grew out of the author's 2012 book, Machine Learning: A Probabilistic Perspective. More than just a simple update, this is a completely new book that reflects the dramatic developments in the field since 2012, most notably deep learning. In addition, the new book is accompanied by online Python code, using libraries such as scikit-learn, JAX, PyTorch, and TensorFlow, which can be used to reproduce nearly all the figures; this code can be run inside a web browser using cloud-based notebooks, and provides a practical complement to the theoretical topics discussed in the book. This introductory text will be followed by a sequel that covers more advanced topics, taking the same probabilistic approach.

Deep Learning

Emphasizing issues of computational efficiency, Michael Kearns and Umesh Vazirani introduce a number of central topics in computational learning theory for researchers and students in artificial intelligence, neural networks, theoretical computer science, and statistics. Emphasizing issues of computational efficiency, Michael Kearns and Umesh Vazirani introduce a number of central topics in computational learning theory for researchers and students in artificial intelligence, neural networks, theoretical computer science, and statistics. Computational learning theory is a new and rapidly expanding area of research that examines formal models of induction with the goals of discovering the common methods underlying efficient learning algorithms and identifying the computational impediments to learning. Each topic in the book has been chosen to elucidate a general principle, which is explored in a precise formal setting. Intuition has been emphasized in the presentation to make the material accessible to the nontheoretician while still providing precise arguments for the specialist. This balance is the result of new proofs of established theorems, and new presentations of the standard proofs. The topics covered include the motivation, definitions, and fundamental results, both positive and negative, for the widely studied L. G. Valiant model of Probably Approximately Correct Learning: Occam's Razor, which formalizes a relationship between learning and data compression; the Vapnik-Chervonenkis dimension; the equivalence of weak and strong learning; efficient learning in the presence of noise by the method of statistical queries; relationships between learning and cryptography, and the resulting computational limitations on efficient learning; reducibility between learning problems; and algorithms for learning finite automata from active experimentation.

Elements of Causal Inference

This important text and reference for researchers and students in machine learning, game theory, statistics and information theory offers a comprehensive treatment of the problem of predicting individual sequences. Unlike standard statistical approaches to forecasting, prediction of individual sequences does not impose any probabilistic assumption on the data-generating mechanism. Yet, prediction algorithms can be constructed that work well for all possible sequences, in the sense that their performance is always nearly as good as the best forecasting strategy in a given reference class. The central theme is the model of prediction using expert advice, a general framework within which many related problems can be cast and discussed. Repeated game playing, adaptive data compression, sequential investment in the stock market, sequential pattern analysis, and several other problems are viewed as instances of the experts' framework and analyzed from a common nonstochastic standpoint that often reveals new and intriguing connections.

An Introduction to Computational Learning Theory

The Beauty of Mathematics in Computer Science explains the mathematical fundamentals of information technology products and services we use every day, from Google Web Search to GPS Navigation, and from speech recognition to CDMA mobile services. The book was published in Chinese in 2011 and has sold more than 600,000 copies. Readers were surprised to find that many daily-used IT technologies were so tightly tied to mathematical principles. For example, the automatic classification of facial images is the common feature of the cosine similarity, and the book covers many topics related to computer applications and applied mathematics including: Natural language processing Speech recognition and machine translation Statistical language modeling Quantitative measurement of information Graph theory and web crawler Pagerank for web search Matrix operation and document classification Mathematical background of big data Neural networks and Google's deep learning Jun Wu was a staff research scientist in Google who invented Google’s Chinese, Japanese, and Korean Web Search Algorithms and was responsible for many Google machine learning projects. He wrote official blogs introducing Google technologies behind its products in very simple languages for Chinese Internet users from 2006-2010. The blogs had more than 2 million followers. Wu received PhD in computer science from Johns Hopkins University and has been working on speech recognition and natural language processing for more than 20 years. He was one of the earliest engineers of Google, managed many products of the company and was awarded 19 US patents during his 10-year tenure there. Wu became a full-time VC investor and co-founded Amino Capital in Palo Alto in 2014 and is the author of eight books.

Decoupling

How does the internet really work? This book explains the technology behind it all, in simple question and answer format.

Optimization Models

The recent rapid growth in the variety and complexity of new machine learning architectures requires the development of improved methods for designing, analyzing, evaluating, and communicating machine learning technologies. Statistical Machine Learning: A Unified Framework provides students, engineers, and scientists with tools from mathematical statistics and nonlinear optimization theory to become experts in the field of machine learning. In particular, the material in this text directly supports the mathematical analysis and design of old, new, and not-yet-invented nonlinear high-dimensional machine learning algorithms. Features: Unified empirical risk minimization framework supports rigorous mathematical analyses of widely used supervised, unsupervised, and reinforcement machine learning algorithms Matrix calculus methods for supporting machine learning analysis and design applications Explicit conditions for ensuring convergence of adaptive, batch, minibatch, MCEM, and MCMC learning algorithms that minimize both unimodal and multimodal objective functions Explicit conditions for characterizing asymptotic properties of M-estimators and model selection criteria such as AIC and BIC in the presence of possible model misspecification This advanced text is suitable for graduate students or highly motivated undergraduate students in statistics, computer science, electrical engineering, and applied mathematics. The text is self-contained and only assumes knowledge of lower-division linear algebra and upper-division probability theory. Students, professional engineers, and multidisciplinary scientists possessing these minimal prerequisites will find this text challenging yet accessible. About the Author, Dr. Goldgen Golden (B.S.E.E.) is a Recognized Expert in Machine Learning who has been a faculty member in Electrical Engineering at the University of Texas at Dallas for the past five years. Dr. Golden has published articles and given talks at scientific conferences on a wide range of topics in the fields of both statistics and machine learning over the past three decades. His long-term research interests include identifying conditions for the convergence of deterministic and stochastic machine learning algorithms and investigating estimation and inference in the presence of possibly misspecified probability models.