Introduction To Statistical Learning Theory

Advanced statistical modeling and knowledge representation techniques for a newly emerging area of machine learning and probabilistic reasoning; includes introductory material, tutorials for different proposed approaches, and applications. Handling inherent uncertainty and exploiting compositional structure are fundamental to understanding and designing large-scale systems. Statistical relational learning builds on ideas from probability theory and statistics to address uncertainty while incorporating tools from logic, databases and programming languages to represent structure. In Introduction to Statistical Relational Learning, leading researchers in this emerging area of machine learning describe current formalisms, models, and algorithms that enable effective and robust reasoning about richly structured systems and data. The early chapters provide tutorials for material used in later chapters, offering introductions to representation, inference and learning in graphical models, and logic. The book then describes object-oriented approaches, including probabilistic relational models, relational Markov networks, and probabilistic entity-relationship models as well as logic-based formalisms including Bayesian logic programs, Markov logic, and stochastic logic programs. Later chapters discuss such topics as probabilistic models with unknown objects, relational dependency networks, reinforcement learning in relational domains, and information extraction. By presenting a variety of approaches, the book highlights commonalities and clarifies important differences among proposed approaches and, along the way, identifies important representational and algorithmic issues. Numerous applications are provided throughout.

During the past decade there has been an explosion in computation and information technology. With it have come vast amounts of data in a variety of fields such as medicine, biology, finance, and marketing. The challenge of understanding these data has led to the development of new tools in the field of statistics, and spawned new areas such as data mining, machine learning, and bioinformatics. Many of these tools have common underpinnings but are often expressed with different terminology. This book describes the important ideas in these areas in a common conceptual framework. While the approach is statistical, the emphasis is on concepts rather than mathematics. Many examples are given, with a liberal use of color graphics. It should be a valuable resource for statisticians and anyone interested in data mining in science or industry. The book's coverage is broad, from supervised learning (prediction) to unsupervised learning. The many topics include neural networks, support vector machines, classification trees and boosting---the first comprehensive treatment of this topic in any book. This major new edition features many topics not covered in the original, including graphical models, random forests, ensemble methods, least angle regression & path algorithms for the lasso, non-negative matrix factorization, and spectral clustering. There is also a chapter on methods for “wide” data (p bigger than n), including multiple testing and false discovery rates. Trevor Hastie, Robert Tibshirani, and Jerome Friedman are professors of statistics at Stanford University. They are prominent researchers in this area: Hastie and Tibshirani developed generalized additive models and wrote a popular book of that title. Hastie co-developed much of the statistical modeling software and environment in R/S-PLUS and invented principal curves and surfaces. Tibshirani proposed the lasso and is co-author of the very successful An Introduction to the Bootstrap. Friedman is the co-inventor of many data-mining tools including CART, MARS, projection pursuit and gradient boosting.

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

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; subsequent chapters 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; regression; 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 appendixes 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.

Statistical Foundations of Data Science gives a thorough introduction to commonly used statistical models, contemporary statistical machine learning techniques and algorithms, along with their mathematical insights and statistical theories. It aims to serve as a graduate-level textbook and a research monograph on high-dimensional statistics, sparsity and covariance learning, machine learning, and statistical inference. It includes ample exercises that involve both theoretical studies as well as empirical applications. The book begins with an introduction to the stylized features of big data and their impacts on statistical analysis. It then introduces multiple linear regression and expands the techniques of model building via nonparametric regression and kernel tricks. It provides a comprehensive account on sparsity explorations and model selections for multiple regression, generalized linear models, quantile regression, robust regression, hazards regression, among others. High-dimensional inference is also thoroughly addressed and so is feature screening. The book also provides a comprehensive account on high-dimensional covariance estimation, learning latent factors and hidden structures, as well as their applications to statistical estimation, inference, prediction and machine learning problems. It also introduces thoroughly statistical machine learning theory and methods for classification, clustering, and prediction. These include CART, random forests, boosting, support vector machines, clustering algorithms, sparse PCA, and deep learning.

The implications for philosophy and cognitive science of developments in statistical learning theory. In Reliable Reasoning, Gilbert Harman and Sanjeev Kulkarni—a philosopher and an engineer—argue that philosophy and cognitive science can benefit from statistical learning theory (SLT), the theory that lies behind recent advances in machine learning. The philosophical problem of induction, for example, is in part about the reliability of inductive reasoning, where the reliability of a method is measured by its statistically expected percentage of errors—a central topic in SLT. After discussing philosophical attempts to evade the problem of induction, Harman and Kulkarni provide an admirably clear account of the basic framework of SLT and its implications for inductive reasoning. They explain the Vapnik-Chervonenkis (VC) dimension of a set of hypotheses and distinguish two kinds of inductive reasoning. The authors discuss various topics in machine learning, including nearest-neighbor methods, neural networks, and support vector machines. Finally, they describe transductive reasoning and suggest possible new models of human reasoning suggested by developments in SLT.

Suitable for self study Use real examples and real data sets that will be familiar to the audience Introduction to the bootstrap is included – this is a modern method missing in many other books

A Computational Approach to Statistical Learning gives a novel introduction to predictive modeling by focusing on the algorithmic and numeric motivations behind popular statistical methods. The text contains annotated code to over 80 original reference functions. These functions provide minimal working implementations of common statistical learning algorithms. Every chapter concludes with a fully worked out application that illustrates predictive modeling tasks using a real-world dataset. The text begins with a detailed analysis of linear models and ordinary least squares. Subsequent chapters explore extensions such as ridge regression, generalized linear models, and additive models. The second half focuses on the use of general-purpose algorithms for convex optimization and their application to tasks in statistical learning. Models covered include the elastic net, dense neural networks, convolutional neural networks (CNNs), and spectral clustering. A unifying theme throughout the text is the use of optimization theory in the description of predictive models, with a particular focus on the singular value

decomposition (SVD). Through this theme, the computational approach motivates and clarifies the relationships between various predictive models. Taylor Arnold is an assistant professor of statistics at the University of Richmond. His work at the intersection of computer vision, natural language processing, and digital humanities has been supported by multiple grants from the National Endowment for the Humanities (NEH) and the American Council of Learned Societies (ACLS). His first book, Humanities Data in R, was published in 2015. Michael Kane is an assistant professor of biostatistics at Yale University. He is the recipient of grants from the National Institutes of Health (NIH), DARPA, and the Bill and Melinda Gates Foundation. His R package bigmemory won the Chamber's prize for statistical software in 2010. Bryan Lewis is an applied mathematician and author of many popular R packages, including irlba, doRedis, and threejs.
A practitioner’s tools have a direct impact on the success of his or her work. This book will provide the data scientist with the tools and techniques required to excel with statistical learning methods in the areas of data access, data munging, exploratory data analysis, supervised machine learning, unsupervised machine learning and model evaluation. Machine learning and data science are large disciplines, requiring years of study in order to gain proficiency. This book can be viewed as a set of essential tools we need for a long-

term career in the data science field – recommendations are provided for further study in order to build advanced skills in tackling important data problem domains. The R statistical environment was chosen for use in this book. R is a growing phenomenon worldwide, with many data scientists using it exclusively for their project work. All of the code examples for the book are written in R. In addition, many popular R packages and data sets will be used.

One of the primary motivations for clinical trials and observational studies of humans is to infer cause and effect. Disentangling causation from confounding is of utmost importance. Fundamentals of Causal Inference explains and relates different methods of confounding adjustment in terms of potential outcomes and graphical models, including standardization, difference-in-differences estimation, the front-door method, instrumental variables estimation, and propensity score methods. It also covers effect-measure modification, precision variables, mediation analyses, and time-dependent confounding. Several real data examples, simulation studies, and analyses using R motivate the methods throughout. The book assumes familiarity with basic statistics and probability, regression, and R and is suitable for seniors or graduate students in statistics, biostatistics, and data science as well as PhD students in a wide variety of other disciplines, including epidemiology, pharmacy, the health sciences, education, and the social, economic, and behavioral sciences.

Beginning with a brief history and a review of essential elements of probability and statistics, a unique feature of

the book is its focus on real and simulated datasets with all binary variables to reduce complex methods down to their fundamentals. Calculus is not required, but a willingness to tackle mathematical notation, difficult concepts, and intricate logical arguments is essential. While many real data examples are included, the book also features the Double What-If Study, based on simulated data with known causal mechanisms, in the belief that the methods are best understood in circumstances where they are known to either succeed or fail. Datasets, R code, and solutions to odd-numbered exercises are available at www.routledge.com.
Applied Predictive Modeling covers the overall predictive modeling process, beginning with the crucial steps of data preprocessing, data splitting and foundations of model tuning. The text then provides intuitive explanations of numerous common and modern regression and classification techniques, always with an emphasis on illustrating and solving real data problems. The text illustrates all parts of the modeling process through many hands-on, real-life examples, and every chapter contains extensive R code for each step of the process. This multi-purpose text can be used as an introduction to predictive models and the overall modeling process, a practitioner's reference handbook, or as a text for advanced undergraduate or graduate level predictive modeling courses. To that end, each chapter contains problem sets to help solidify the covered concepts and uses data available in the book's R package. This text is intended for a broad audience as both an introduction to predictive models as well as a

guide to applying them. Non-mathematical readers will appreciate the intuitive explanations of the techniques while an emphasis on problem-solving with real data across a wide variety of applications will aid practitioners who wish to extend their expertise. Readers should have knowledge of basic statistical ideas, such as correlation and linear regression analysis. While the text is biased against complex equations, a mathematical background is needed for advanced topics.

An Introduction to Statistical Learningwith Applications in RSpringer Science & Business Media

'It is not often I can use "accessible" and "phenomenology" in the same sentence, but reading the new book, Interpretative Phenomenological Analysis...certainly provides me the occasion to do so. I can say this because these authors provide an engaging and clear introduction to a relatively new analytical approach' - The Weekly Qualitative Report Interpretative phenomenological analysis (IPA) is an increasingly popular approach to qualitative inquiry. This handy text covers its theoretical foundations and provides a detailed guide to conducting IPA research. Extended worked examples from the authors' own studies in health, sexuality, psychological distress and identity illustrate the breadth and depth of IPA research. Each of the chapters also offers a guide to other good exemplars of IPA research in the designated area. The final section of the book considers how IPA connects with other contemporary qualitative approaches like discourse and narrative analysis and how it addresses issues to do with validity. The book is written in an accessible style and will be extremely useful to students and researchers in psychology and related disciplines in the health and social sciences.

The goal of machine learning is to program computers to use example data or past experience to solve a given problem. Many successful applications of machine learning exist already, including systems that analyze past sales data to predict customer behavior, optimize robot behavior so that a task can be completed using minimum resources, and extract knowledge from bioinformatics data. Introduction to Machine Learning is a comprehensive textbook on the subject, covering a broad array of topics not usually included in introductory machine learning texts. Subjects include supervised learning; Bayesian decision theory; parametric, semi-parametric, and nonparametric methods; multivariate analysis; hidden Markov models; reinforcement learning; kernel machines; graphical models; Bayesian estimation; and statistical testing.Machine learning is rapidly becoming a skill that computer science students must master before graduation. The third edition of Introduction to Machine Learning reflects this shift, with added support for beginners, including selected solutions for exercises and additional example data sets (with code available online). Other substantial changes include discussions of outlier detection; ranking algorithms for perceptrons and support vector machines; matrix decomposition and spectral methods; distance estimation; new kernel algorithms; deep learning in multilayered perceptrons; and the nonparametric approach to Bayesian methods. All learning algorithms are explained so that students can easily move from the equations in the book to a computer program. The book can be used by both advanced undergraduates and graduate students. It will also be of interest to professionals who are concerned with the application of machine learning methods.

Machine learning allows computers to learn and discern patterns without actually being programmed. When Statistical techniques and machine learning are combined together they are a powerful tool for analysing various kinds of data in many computer science/engineering areas including, image processing, speech processing, natural language processing, robot control, as well as in fundamental sciences such as biology, medicine, astronomy, physics, and materials. Introduction to Statistical Machine Learning provides a general introduction to machine learning that covers a wide range of topics concisely and will help you bridge the gap between theory and practice. Part I discusses the fundamental concepts of statistics and probability that are used in describing machine learning algorithms. Part II and Part III explain the two major approaches of machine learning techniques; generative methods and discriminative methods. While Part III provides an in-depth look at advanced topics that play essential roles in making machine learning algorithms more useful in practice. The accompanying MATLAB/Octave programs provide you with the necessary practical skills needed to accomplish a wide range of data analysis tasks. Provides the necessary background material to understand machine learning such as statistics, probability, linear algebra, and calculus. Complete coverage of the generative approach to statistical pattern recognition and the discriminative approach to statistical machine learning. Includes MATLAB/Octave programs so that readers can test the algorithms numerically and acquire both mathematical and practical skills in a wide range of data analysis tasks Discusses a wide range of applications in machine learning and statistics and provides examples drawn from image processing, speech processing, natural language processing, robot control, as well as biology, medicine, astronomy, physics, and materials.

Statistical learning theory is aimed at analyzing complex data with necessarily approximate models. This book is intended for an audience with a graduate background in probability theory and statistics. It will be useful to any reader wondering why it may be a good idea, to use as is often done in practice a notoriously "wrong'' (i.e. over-simplified) model to predict, estimate or classify. This point of view takes its roots in three fields: information theory, statistical mechanics, and PAC-Bayesian theorems. Results on the large deviations of trajectories of Markov chains with rare transitions are also included. They are meant to provide a better understanding of stochastic optimization algorithms of common use in computing estimators. The author focuses on non-asymptotic bounds of the statistical risk, allowing one to choose adaptively between rich and structured families of models and corresponding estimators. Two mathematical objects pervade the book: entropy and Gibbs measures. The goal is to show how to turn them into versatile and efficient technical tools, that will stimulate further studies and results.

Data analysis is changing fast. Driven by a vast range of application domains and affordable tools, machine learning has become mainstream. Unsupervised data analysis, including cluster analysis, factor analysis, and low dimensionality mapping methods continually being updated, have reached new heights of achievement in the incredibly rich data wor

Offering accessible and nuanced coverage, Richard W. Hamming discusses theories of probability with unique clarity and depth. Topics covered include the basic philosophical assumptions, the nature of stochastic methods, and Shannon entropy. One of the best introductions to the topic, The Art of Probability is filled with unique insights and tricks worth knowing.

AN INTRODUCTION TO MACHINE LEARNING THAT INCLUDES THE FUNDAMENTAL TECHNIQUES, METHODS, AND APPLICATIONS Machine Learning: a Concise Introduction offers a comprehensive introduction to the core concepts, approaches, and applications of machine learning. The author—an expert in the field—presents fundamental ideas, terminology, and techniques for solving applied problems in classification, regression, clustering, density estimation, and dimension reduction. The design principles behind the techniques are emphasized, including the bias-variance trade-off and its influence on the design of ensemble methods. Understanding these principles leads to more flexible and successful applications. Machine Learning: a Concise Introduction also includes methods for optimization, risk estimation, and model selection—essential elements of most applied projects. This important resource: Illustrates many classification methods with a single, running example, highlighting similarities and differences between methods Presents R source code which shows how to apply and interpret many of the techniques covered Includes many thoughtful exercises as an integral part of the text, with an appendix of selected solutions Contains useful information for effectively communicating with clients A volume in the popular Wiley Series in Probability and Statistics, Machine Learning: a Concise Introduction offers the practical information needed for an understanding of the methods and application of machine learning. STEVEN W. KNOX holds a Ph.D. in Mathematics from the University of Illinois and an M.S. in Statistics from Carnegie Mellon University. He has over twenty years' experience in using Machine Learning, Statistics, and Mathematics to solve real-world problems. He currently serves as Technical Director of Mathematics Research and Senior Advocate for Data Science at the National Security Agency.

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. This book presents the Statistical Learning Theory in a detailed and easy to understand way, by using practical examples, algorithms and source codes. It can be used as a textbook in graduation or undergraduation courses, for self-learners, or as reference with respect to the main theoretical concepts of Machine Learning. Fundamental concepts of Linear Algebra and Optimization applied to Machine Learning are provided, as well as source codes in R, making the book as self-contained as possible. It starts with an introduction to Machine Learning concepts and algorithms such as the Perceptron, Multilayer Perceptron and the Distance-Weighted Nearest Neighbors with examples, in order to provide the necessary foundation so the reader is able to understand the Bias-Variance Dilemma, which is the central point of the Statistical Learning Theory. Afterwards, we introduce all assumptions and formalize the Statistical Learning Theory, allowing the practical study of different classification algorithms. Then, we proceed with concentration inequalities until arriving to the Generalization and the Large-Margin bounds, providing the main motivations for the Support Vector Machines. From that, we introduce all necessary optimization concepts related to the implementation of Support Vector Machines. To provide a next stage of development, the book finishes with a discussion on SVM kernels as a way and motivation to study data spaces and improve classification results.

Aimed at advanced undergraduate and beginning graduate students, this book covers the theory of foreign policy analysis. Beginning with an overview, it then tackles theory and research at multiple levels of analysis, ending with an examination of the areas in which the next generation of foreign policy analysts can make important contributions.

In nonparametric and high-dimensional statistical models, the classical Gauss–Fisher–Le Cam theory of the optimality of maximum likelihood estimators and Bayesian posterior inference does not apply, and new foundations and ideas have been developed in the past several decades. This book gives a coherent account of the statistical theory in infinite-dimensional parameter spaces. The mathematical foundations include self-contained 'mini-courses' on the theory of Gaussian and empirical processes, approximation and wavelet theory, and the basic theory of function spaces. The theory of statistical inference in such models - hypothesis testing, estimation and confidence sets - is presented within the minimax paradigm of decision theory. This includes the basic theory of convolution kernel and projection estimation, but also Bayesian nonparametrics and nonparametric maximum likelihood estimation. In a final chapter the theory of adaptive inference in nonparametric models is developed, including Lepski's method, wavelet thresholding, and adaptive inference for self-similar functions. Winner of the 2017 PROSE Award for Mathematics.

Extensive treatment of the most up-to-date topics Provides the theory and concepts behind popular and

emerging methods Range of topics drawn from Statistics, Computer Science, and Electrical Engineering
This book provides a broad yet detailed introduction to neural networks and machine learning in a statistical framework. A single, comprehensive resource for study and further research, it explores the major popular neural network models and statistical learning approaches with examples and exercises and allows readers to gain a practical working understanding of the content. This updated new edition presents recently published results and includes six new chapters that correspond to the recent advances in computational learning theory, sparse coding, deep learning, big data and cloud computing. Each chapter features state-of-the-art descriptions and significant research findings. The topics covered include: • multilayer perceptron; • the Hopfield network; • associative memory models;• clustering models and algorithms; • t he radial basis function network; • recurrent neural networks; • nonnegative matrix factorization; • independent component analysis; •probabilistic and Bayesian networks; and • fuzzy sets and logic. Focusing on the prominent accomplishments and their practical aspects, this book provides academic and technical staff, as well as graduate students and researchers with a solid foundation and comprehensive reference on the fields of neural networks, pattern recognition, signal processing, and machine 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.

An Introduction to Statistical Learning provides an accessible overview of the field of statistical learning, an essential toolset for making sense of the vast and complex data sets that have emerged in fields ranging from biology to finance to marketing to astrophysics in the past twenty years. This book presents some of the most important modeling and prediction techniques, along with relevant applications. Topics include linear regression, classification, resampling methods, shrinkage approaches, tree-based methods, support vector machines, clustering, and more. Color graphics and real-world examples are used to illustrate the methods presented. Since the goal of this textbook is to facilitate the use of these statistical learning techniques by practitioners in science, industry, and other fields, each chapter contains a tutorial on implementing the analyses and methods presented in R, an extremely popular open source statistical software platform. Two of the authors co-wrote The Elements of Statistical Learning (Hastie, Tibshirani and Friedman, 2nd edition 2009), a popular reference book for statistics and machine learning researchers. An Introduction to Statistical Learning covers many of the same topics, but at a level accessible to a much broader audience. This book is targeted at statisticians and non-statisticians alike who wish to use cutting-edge statistical learning techniques to analyze their data. The text assumes only a previous course in linear regression and no knowledge of matrix algebra.

Discover New Methods for Dealing with High-Dimensional Data A sparse statistical model has only a small number of nonzero parameters or weights; therefore, it is much easier to estimate and interpret than a dense model. Statistical Learning with Sparsity: The Lasso and Generalizations presents methods that exploit sparsity to help recover the underlying signal in a set of data. Top experts in this rapidly evolving field, the authors describe the lasso for linear regression and a simple coordinate descent algorithm for its computation. They discuss the application of l1 penalties to generalized linear models and support vector machines, cover generalized penalties such as the elastic net and group lasso, and review numerical methods for optimization. They also present statistical inference methods for fitted (lasso) models, including the bootstrap, Bayesian methods, and recently developed approaches. In addition, the book examines matrix decomposition, sparse multivariate analysis, graphical models, and compressed sensing. It concludes with a survey of theoretical results for the lasso. In this age of big data, the number of features measured on a person or object can be large and might be larger than the number of observations. This book shows how the sparsity assumption allows us to tackle these problems and extract useful and reproducible patterns from big datasets. Data analysts, computer scientists, and theorists will appreciate this thorough and up-to-date treatment of sparse statistical modeling.

Praise for the first edition: "[This book] succeeds singularly at providing a structured introduction to this active field of research. ... it is arguably the most accessible overview yet published of the mathematical ideas and principles that one needs to master to enter the field of high-dimensional statistics. ... recommended to anyone interested in the main results of current research in high-dimensional statistics as well as anyone interested in acquiring the core mathematical skills to enter this area of research." —Journal of the American Statistical Association Introduction to High-Dimensional Statistics, Second Edition preserves the philosophy of the first edition: to be a concise guide for students and researchers discovering the area and interested in the mathematics involved. The main concepts and ideas are presented in simple settings, avoiding thereby unessential technicalities. High-dimensional statistics is a fast-evolving field, and much progress has been made on a large variety of topics, providing new insights and methods. Offering a succinct presentation of the mathematical foundations of high-dimensional statistics, this new edition: Offers revised chapters from the previous edition, with the inclusion of many additional materials on some important topics, including compress sensing, estimation with convex constraints, the slope estimator, simultaneously low-rank and row-sparse linear regression, or aggregation of a continuous set of estimators. Introduces three new chapters on iterative algorithms, clustering, and minimax lower bounds. Provides enhanced appendices, minimax lower-bounds mainly with the addition of the Davis-Kahan perturbation bound and of two simple versions of the Hanson-Wright concentration inequality. Covers cutting-edge statistical methods including model selection, sparsity and the Lasso, iterative hard thresholding, aggregation, support vector machines, and learning theory. Provides detailed exercises at the end of every chapter with collaborative solutions on a wiki site. Illustrates concepts with simple but clear practical examples. A thought-provoking look at statistical learning theory and its role in understanding human learning and inductive reasoning A joint endeavor from leading researchers in the fields of philosophy and electrical engineering, An Elementary Introduction to Statistical Learning Theory is a comprehensive and accessible primer on the rapidly evolving fields of statistical pattern recognition and statistical learning theory. Explaining these areas at a level and in a way that is not often found in other books on the topic, the authors present the basic theory behind contemporary machine learning and uniquely utilize its foundations as a framework for philosophical thinking about inductive inference. Promoting the fundamental goal of statistical learning, knowing what is achievable and what is not, this book demonstrates the value of a systematic methodology when used along with the needed techniques for evaluating the performance of a learning system. First, an introduction to machine learning is presented that includes brief discussions of applications such as image recognition, speech recognition, medical diagnostics, and statistical arbitrage. To enhance accessibility, two chapters on relevant aspects of probability theory are provided. Subsequent chapters feature coverage of topics such as the pattern recognition problem, optimal Bayes decision rule, the nearest neighbor rule, kernel rules, neural networks, support vector machines, and boosting. Appendices throughout the book explore the relationship between the discussed material and related topics from mathematics, philosophy, psychology, and statistics, drawing insightful connections between problems in these areas and statistical learning theory. All chapters conclude with a summary section, a set of practice questions, and a reference sections that supplies historical notes and additional resources for further study. An Elementary Introduction to Statistical Learning Theory is an excellent book for courses on statistical learning theory, pattern recognition, and machine learning at the upper-undergraduate and graduate levels. It also serves as an introductory reference for researchers and practitioners in the fields of engineering, computer science, philosophy, and cognitive science that would like to further their knowledge of the topic.

This textbook considers statistical learning applications when interest centers on the conditional distribution of a response variable, given a set of predictors, and in the absence of a credible model that can be specified before the data analysis begins. Consistent with modern data analytics, it emphasizes that a proper statistical learning data analysis depends in an integrated fashion on sound data collection, intelligent data management, appropriate statistical procedures, and an accessible interpretation of results. The unifying theme is that supervised learning properly can be seen as a form of regression analysis. Key concepts and procedures are illustrated with a large number of real applications and their associated code in R, with an eye toward practical implications. The growing integration of computer science and statistics is well represented including the occasional, but salient, tensions that result. Throughout, there are links to the big picture. The third edition considers significant advances in recent years, among which are: the development of overarching, conceptual frameworks for statistical learning; the impact of “big data” on statistical learning; the nature and consequences of post-model selection statistical inference; deep learning in various forms; the special challenges to statistical inference posed by statistical learning; the fundamental connections between data collection and data analysis; interdisciplinary ethical and political issues surrounding the application of algorithmic methods in a wide variety of fields, each linked to concerns about transparency, fairness, and accuracy. This edition features new sections on accuracy, transparency, and fairness, as well as a new chapter on deep learning. Precursors to deep learning get an expanded treatment. The connections between fitting and forecasting are considered in greater depth. Discussion of the estimation targets for algorithmic methods is revised and expanded throughout to reflect the latest research. Resampling procedures are emphasized. The material is written for upper undergraduate and graduate students in the social, psychological and life sciences and for researchers who want to apply statistical learning procedures to scientific and policy problems.

This interdisciplinary text offers theoretical and practical results of information theoretic methods used in statistical learning. It presents a comprehensive overview of the many different methods that have been developed in numerous contexts.

The goal of learning theory is to approximate a function from sample values. To attain this goal learning theory draws on a variety of diverse subjects, specifically statistics, approximation theory, and algorithmics. Ideas from all these areas blended to form a subject whose many successful applications have triggered a rapid growth during the last two decades. This is the first book to give a general overview of the theoretical foundations of the subject emphasizing the approximation theory, while still giving a balanced overview. It is based on courses taught by the authors, and is reasonably self-contained so will appeal to a broad spectrum of researchers in learning theory and adjacent fields. It will also serve as an introduction for graduate students and others entering the field, who wish to see how the problems raised in learning theory relate to other disciplines.

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: Richard M. Golden (Ph.D., M.S.E.E., B.S.E.E.) is Professor of Cognitive Science and Participating Faculty Member in Electrical Engineering at the University of Texas at Dallas. 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.
Table of contents

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.

This book is for anyone who has biomedical data and needs to identify variables that predict an outcome, for two-group outcomes such as tumor/not-tumor, survival/death, or response from treatment. Statistical learning machines are ideally suited to these types of prediction problems, especially if the variables being studied may not meet the assumptions of traditional techniques. Learning machines come from the world of probability and computer science but are not yet widely used in biomedical research. This introduction brings learning machine techniques to the biomedical world in an accessible way, explaining the underlying principles in nontechnical language and using extensive examples and figures. The authors connect these new methods to familiar techniques by showing how to use the learning machine models to generate smaller, more easily interpretable traditional models. Coverage includes single decision trees, multiple-tree techniques such as Random ForestsTM, neural nets, support vector machines, nearest neighbors and boosting.

Machine Learning has become a key enabling technology for many engineering applications, investigating scientific questions and theoretical problems alike. To stimulate discussions and to disseminate new results, a summer school series was started in February 2002, the documentation of which is published as LNAI 2600. This book presents revised lectures of two subsequent summer schools held in 2003 in Canberra, Australia, and in Tübingen, Germany. The tutorial lectures included are devoted to statistical learning theory, unsupervised learning, Bayesian inference, and applications in pattern recognition; they provide in-depth overviews of exciting new developments and contain a large number of references. Graduate students, lecturers, researchers and professionals alike will find this book a useful resource in learning and teaching machine learning.