Information Theory Inference And Learning Algorithms

Information theory and inference, often taught separately, are here united in one entertaining textbook. These topics lie at the heart of many exciting areas of contemporary science and engineering - communication, signal processing, data mining, machine learning, pattern recognition, computational neuroscience, bioinformatics, and cryptography. This textbook introduces theory in tandem with applications. Information theory is taught alongside practical communication systems, such as arithmetic coding for data compression and sparse-graph codes for error-correction. A toolbox of inference techniques, including message-passing algorithms, Monte Carlo methods, and variational approximations, are developed alongside applications of these tools to clustering, convolutional codes, independent component analysis, and neural networks. The final part of the book describes the state of the art in error-correcting codes, including low-density parity-check codes, turbo codes, and digital fountain codes -- the twenty-first century standards for satellite communications, disk drives, and data broadcast. Richly illustrated, filled with worked examples and over 400 exercises, some with detailed solutions, David MacKay's groundbreaking book is ideal for self-learning and for undergraduate or graduate courses. Interludes on crosswords, evolution, and sex provide entertainment along the way. In sum, this is a textbook on information, communication, and coding for a new generation of students, and an unparalleled entry point into these subjects for professionals in areas as diverse as computational biology, financial engineering, and machine learning.

The standard rules of probability can be interpreted as uniquely valid principles in logic. In this book, E. T. Jaynes dispels the imaginary distinction between 'probability theory' and 'statistical inference', leaving a logical unity and simplicity, which provides greater technical power and flexibility in applications. This book goes beyond the conventional mathematics of probability theory, viewing the subject in a wider context. New results are discussed, along with applications of probability theory to a wide variety of problems in physics, mathematics, economics, chemistry and biology. It contains many exercises and problems, and is suitable for use as a textbook on graduate level courses involving data analysis. The material is aimed at readers who are already familiar with applied mathematics at an advanced undergraduate level or higher. The book will be of interest to scientists working in any area where inference from incomplete information is necessary.

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.

Highly useful text studies logarithmic measures of information and their application to testing statistical hypotheses. Includes numerous worked examples and problems. References. Glossary. Appendix. 1968 2nd, revised edition.

Statistical Inference for Engineers and Data Scientists
Information Theory , Inference And Learning Algorithms
Information Theory, Inference and Learning Algorithms
Quantum Information Theory
Information Theory and Statistics
Information Theory

Machine Learning: A Bayesian and Optimization Perspective, 2nd edition, gives a unified perspective on machine learning by covering both pillars of supervised learning, namely regression and classification. The book starts with the basics, including mean square, least squares and maximum likelihood methods, ridge regression, Bayesian decision theory classification, logistic regression, and decision trees. It then progresses to more recent techniques, covering sparse modelling methods, learning in reproducing kernel Hilbert spaces and support vector machines, Bayesian inference with a focus on the EM algorithm and its approximate inference variational versions, Monte Carlo methods, probabilistic graphical models focusing on Bayesian networks, hidden Markov models and particle filtering. Dimensionality reduction and latent variables modelling are also considered in depth. This palette of techniques concludes with an extended chapter on neural networks and deep learning architectures. The book also covers the fundamentals of statistical parameter estimation, Wiener and Kalman filtering, convexity and convex optimization, including a chapter on stochastic approximation and the gradient descent family of algorithms, presenting related online learning techniques as well as concepts and algorithmic versions for distributed optimization. Focusing on the physical reasoning behind the mathematics, without sacrificing rigor, all the various methods and techniques are explained in depth, supported by examples and problems, giving an invaluable resource to the student and researcher for understanding and applying machine learning concepts. Most of the chapters include typical case studies and computer exercises, both in MATLAB and Python. The chapters are written to be as self-contained as possible, making the text suitable for different courses: pattern recognition, statistical/adaptive signal processing, statistical/Bayesian learning, as well as courses on sparse modeling, deep learning, and probabilistic graphical models. New to this edition: Complete re-write of the chapter on Neural Networks and Deep Learning to reflect the latest advances since the 1st edition. The chapter, starting from the basic perceptron and feed-forward neural networks concepts, now presents an in depth treatment of deep networks, including recent optimization algorithms, batch normalization, regularization techniques such as the dropout method, convolutional neural networks, recurrent neural networks, attention mechanisms, adversarial examples and training, capsule networks and generative architectures, such as restricted Boltzman machines (RBMs), variational autoencoders and generative adversarial networks (GANs). Expanded treatment of Bayesian learning to include nonparametric Bayesian methods, with a focus on the Chinese restaurant and the Indian buffet processes. Presents the physical reasoning, mathematical modeling and algorithmic implementation of each method Updates on the latest trends, including sparsity, convex analysis and optimization, online distributed algorithms, learning in RKH spaces, Bayesian inference, graphical and hidden Markov models, particle filtering, deep learning, dictionary learning and latent variables modeling Provides case studies on a variety of topics, including protein folding prediction, optical character recognition, text authorship identification, fMRI data analysis, change point detection, hyperspectral image unmixing, target localization, and more

This is the first textbook on pattern recognition to present the Bayesian viewpoint. The book presents approximate inference algorithms that permit fast approximate answers in situations where exact answers are not feasible. It uses graphical models to describe probability distributions when no other books apply graphical models to machine learning. No previous knowledge of pattern recognition or machine learning concepts is assumed. Familiarity with multivariate calculus and basic linear algebra is required, and some experience in the use of probabilities would be helpful though not essential as the book includes a self-contained introduction to basic probability theory.

Another title in the reissued Oxford Classic Texts in the Physical Sciences series, Jeffrey's Theory of Probability, first published in 1939, was the first to develop a fundamental theory of scientific inference based on the ideas of Bayesian statistics. His ideas were way ahead of their time and it is only in the past ten years that the subject of Bayes' factors has been significantly developed and extended. Until recently the two schools of statistics (Bayesian and Frequentist) were distinctly different and set apart. Recent work (aided by increased computer power and availability) has changed all that and today's graduate students and researchers all require an understanding of Bayesian ideas. This book is their starting point.

This book is the first cohesive treatment of ITL algorithms to adapt linear or nonlinear learning machines both in supervised and unsupervised paradigms. It compares the performance of ITL algorithms with the second order counterparts in many applications.

Probabilistic Machine Learning
Mathematical Foundation
Elements of Information Theory
Introduction To The Theory Of Neural Computation
Computer Vision
A Bayesian and Optimization Perspective

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.

Originally developed by Claude Shannon in the 1940s, information theory laid the foundations for the digital revolution, and is now an essential tool in telecommunications, genetics, linguistics, brain sciences, and deep space communication. In this richly illustrated book, accessible examples are used to introduce information theory in terms of everyday games like ‘20 questions’ before more advanced topics are explored. Online MatLab and Python computer programs provide hands-on experience of information theory in action, and PowerPoint slides give support for teaching. Written in an informal style, with a comprehensive glossary and tutorial appendices, this text is an ideal primer for novices who wish to learn the essential principles and applications of information theory.

An introduction to a broad range of topics in deep learning, covering mathematical and conceptual background, deep learning techniques used in industry, and research perspectives. “Written by three experts in the field, Deep Learning is the only comprehensive book on the subject.” —Elon Musk, cochair of OpenAI; cofounder and CEO of Tesla and SpaceX Deep learning is a form of machine learning that enables computers to learn from experience and understand the world in terms of a hierarchy of concepts. Because the computer gathers knowledge from experience, there is no need for a human computer operator to formally specify all the knowledge that the computer needs. The hierarchy of concepts allows the computer to learn complicated concepts by building them out of simpler ones; a graph of these hierarchies would be many layers deep. This book introduces a broad range of topics in deep learning. The text offers mathematical and conceptual background, covering relevant concepts in linear algebra, probability theory and information theory, numerical computation, and machine learning. It describes deep learning techniques used by practitioners in industry, including deep feedforward networks, regularization, optimization algorithms, convolutional networks, sequence modeling, and practical methodology; and it surveys such applications as natural language processing, speech recognition, computer vision, online recommendation systems, bioinformatics, and videogames. Finally, the book offers research perspectives, covering such theoretical topics as linear factor models, autoencoders, representation learning, structured probabilistic models, Monte Carlo methods, the partition function, approximate inference, and deep generative models. Deep Learning can be used by undergraduate or graduate students planning careers in either industry or research, and by software engineers who want to begin using deep learning in their products or platforms. A website offers supplementary material for both readers and instructors.

A concise and self-contained introduction to causal inference, increasingly important in data science and machine learning. The mathematization 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.

Information and Inference
The Free Energy Principle in Mind, Brain, and Behavior
The Logic of Science
Variational Bayesian Learning Theory
Perception as Bayesian Inference
Graphical Models, Exponential Families, and Variational Inference

An effective blend of carefully explained theory and practical applications, this text imparts the fundamentals of both information theory and data compression. Although the two topics are related, this unique text allows either topic to be presented independently, and it was specifically designed so that the data compression section requires no prior knowledge of information theory. The treatment of information theory, while theoretical and abstract, is quite elementary, making this text less daunting than many others. After presenting the fundamental definitions and results of the theory, the authors then apply the theory to memoryless, discrete channels with zeroth-order, one-state sources. The chapters on data compression acquaint students with a myriad of lossless compression methods and then introduce two lossy compression methods. Students emerge from this study competent in a wide range of techniques. The authors' presentation is highly practical but includes some important proofs, either in the text or in the exercises, so instructors can, if they choose, place more emphasis on the mathematics. Introduction to Information Theory and Data Compression, Second Edition is ideally suited for an upper-level or graduate course for students in mathematics, engineering, and computer science. Features: Expanded discussion of the historical and theoretical basis of information theory that builds a firm, intuitive grasp of the subject Reorganization of theoretical results along with new exercises, ranging from the routine to the more difficult, that reinforce students' ability to apply the definitions and results in specific situations. Simplified treatment of the algorithm(s) of Gallager and Knuth Discussion of the information rate of a code and the trade-off between error correction and information rate Treatment of probabilistic finite state source automata, including basic results, examples, references, and exercises Octave and MATLAB image compression codes included in an appendix for use with the exercises and projects involving transform methods Supplementary materials, including software, available for download from the authors' Web site at www.dms.auburn.edu/compression

The Minimum Message Length (MML) Principle is an information-theoretic approach to induction, hypothesis testing, model selection, and statistical inference. MML, which provides a formal specification for the implementation of Occam's Razor, asserts that the ‘best’ explanation of observed data is the shortest. Further, an explanation is acceptable (i.e. the induction is justified) only if the explanation is shorter than the original data. This book gives a sound introduction to the Minimum Message Length Principle and its applications, provides the theoretical arguments for the adoption of the principle, and shows the development of certain approximations that assist its practical application. MML appears also to provide both a normative and a descriptive basis for inductive reasoning generally, and scientific induction in particular. The book describes this basis and aims to show its relevance to the Philosophy of Science. Statistical and Inductive Inference by Minimum Message Length will be of special interest to graduate students and researchers in Machine Learning and Data Mining, scientists and analysts in various disciplines wishing to make use of computer techniques for hypothesis discovery, statisticians and econometricians interested in the underlying theory of their discipline, and persons interested in the Philosophy of Science. The book could also be used in a graduate-level course in Machine Learning and Estimation and Model-selection, Econometrics and Data Mining. C.S. Wallace was appointed Foundation Chair of Computer Science at Monash University in 1968, at the age of 35, where he worked until his death in 2004. He received an ACM Fellowship in 1995, and was appointed Professor Emeritus in 1996. Professor Wallace made numerous significant contributions to diverse areas of Computer Science, such as Computer Architecture, Simulation and Machine Learning. His final research focused primarily on the Minimum Message Length Principle.

The latest edition of this classic is updated with new problem sets and material The Second Edition of this fundamental textbook maintains the book's tradition of clear, thought-provoking instruction. Readers are provided once again with an instructive mix of mathematics, physics, statistics, and information theory. All the essential topics in information theory are covered in detail, including entropy, data compression, channel capacity, rate distortion, network information theory, and hypothesis testing. The authors provide readers with a solid understanding of the underlying theory and applications. Problem sets and a telegraphic summary at the end of each chapter further assist readers. The historical notes that follow each chapter recap the main points. The Second Edition features: * Chapters reorganized to improve teaching * 200 new problems * New material on source coding, portfolio theory, and feedback capacity * Updated references Now current and enhanced, the Second Edition of Elements of Information Theory remains the ideal textbook for upper-level undergraduate and graduate courses in electrical engineering, statistics, and telecommunications.

Comprehensive introduction to the neural network models currently under intensive study for computational applications. It also provides coverage of neural network applications in a variety of problems of both theoretical and practical interest.

The Mathematical Theory of Communication
From Statistical Physics to Statistical Inference and Back
Advanced Lectures on Machine Learning
Deep Learning
Understanding Machine Learning
Introduction to Information Theory and Data Compression, Second Edition

The fundamental mathematical tools needed to understand machine learning include linear algebra, analytic geometry, matrix decompositions, vector calculus, optimization, probability and statistics. These topics are traditionally taught in disparate courses, making it hard for data science or computer science students, or professionals, to efficiently learn the mathematics. This self-contained textbook bridges the gap between mathematical and machine learning texts, introducing the mathematical concepts with a minimum of prerequisites. It uses these concepts to derive four central machine learning methods: linear regression, principal component analysis, Gaussian mixture models and support vector machines. For students and others with a mathematical background, these derivations provide a starting point to machine learning texts. For those learning the mathematics for the first time, the methods help build intuition and practical experience with applying mathematical concepts. Every chapter includes worked examples and exercises to test understanding. Programming tutorials are offered on the book's web site.

In the last 25 years, the concept of information has played a crucial role in communication theory, so much so that the terms information theory and communication theory are sometimes used almost interchangeably. It seems to us, however, that the notion of information is also destined to render valuable services to the student of induction and probability, of learning and reinforcement, of semantic meaning and deductive inference, as~well as of scientific method in general. The present volume is an attempt to illustrate some of these uses of information concepts. In 'On Semantic Information' Hintikka summarizes some of his and his associates' recent work on information and induction, and comments briefly on its philosophical suggestions. Jamison surveys from the sub jectivistic point of view some recent results in 'Bayesian Information Usage'. Rosenkrantz analyzes the information obtained by experimen tation from the Bayesian and Neyman-Pearson standpoints, and also from the standpoint of entropy and related concepts. The much-debated principle of total evidence prompts Hilpinen to examine the problem of measuring the information yield of observations in his paper 'On the Information Provided by Observations'. Pietarinen addresses himself to the more general task of evaluating the systematizing ('explanatory') power of hypotheses and theories, a task which quickly leads him to information concepts. Domotor develops a qualitative theory of information and entropy. His paper gives what is probably the first axiomatization of a general qualitative theory of information adequate to guarantee a numerical representation of the standard sort.

A modern treatment focusing on learning and inference, with minimal prerequisites, real-world examples and implementable algorithms.
Information Theory, Inference and Learning AlgorithmsCambridge University Press
The Elements of Statistical Learning
Information-Theoretic Methods in Data Science
The Theory of Probability
Probability Theory
An Introduction
A Tutorial Introduction

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

Developing many of the major, exciting, pre- and post-millennium developments from the ground up, this book is an ideal entry point for graduate students into quantum information theory. Significant attention is given to quantum mechanics for quantum information theory, and careful studies of the important protocols of teleportation, superdense coding, and entanglement distribution are presented. In this new edition, readers can expect to find over 100 pages of new material, including detailed discussions of Bell's theorem, the CHSH game, Tsirelson's theorem, the axiomatic approach to quantum channels, the definition of the diamond norm and its interpretation, and a proof of the Choi – Kraus theorem. Discussion of the importance of the quantum dynamic capacity formula has been completely revised, and many new exercises and references have been added. This new edition will be welcomed by the upcoming generation of quantum information theorists and the already established community of classical information theorists.

A comprehensive introduction to machine learning that uses probabilistic models and inference as a unifying approach. Today's Web-enabled deluge of electronic data calls for automated methods of data analysis. Machine learning provides these, developing methods that can automatically detect patterns in data and then use the uncovered patterns to predict future data. This textbook offers a comprehensive and self-contained introduction to the field of machine learning, based on a unified, probabilistic approach. The coverage combines breadth and depth, offering necessary background material on such topics as probability, optimization, and linear algebra as well as discussion of recent developments in the field, including conditional random fields, L1 regularization, and deep learning. The book is written in an informal, accessible style, complete with pseudo-code for the most important algorithms. All topics are copiously illustrated with color images and worked examples drawn from such application domains as biology, text processing, computer vision, and robotics. Rather than providing a cookbook of different heuristic methods, the book stresses a principled model-based approach, often using the language of graphical models to specify models in a concise and intuitive way. Almost all the models described have been implemented in a MATLAB software package—PMTK (probabilistic modeling toolkit)—that is freely available online. The book is suitable for upper-level undergraduates with an introductory-level college math background and beginning graduate students.

A concise, easy-to-read guide, introducing beginners to the engineering background of modern communication systems, from mobile phones to data storage. Assuming only basic knowledge of high-school mathematics and including many practical examples and exercises to aid understanding, this is ideal for anyone who needs a quick introduction to the subject.

Foundations of Signal Processing

Foundations and Learning Algorithms

Elements of Causal Inference

A Student's Guide to Coding and Information Theory

Models, Learning, and Inference

Data Mining, Inference, and Prediction

This introduction to the theory of variational Bayesian learning summarizes recent developments and suggests practical applications.

This comprehensive and engaging textbook introduces the basic principles and techniques of signal processing, from the fundamental ideas of signals and systems theory to real-world applications. Students are introduced to the powerful foundations of modern signal processing, including the basic geometry of Hilbert space, the mathematics of Fourier transforms, and essentials of sampling, interpolation, approximation and compression The authors discuss real-world issues and hurdles to using these tools, and ways of adapting them to overcome problems of finiteness and localization, the limitations of uncertainty, and computational costs. It includes over 160 homework problems and over 220 worked examples, specifically designed to test and expand students' understanding of the fundamentals of signal processing, and is accompanied by extensive online materials designed to aid learning, including Mathematica® resources and interactive demonstrations.

The first unified treatment of the interface between information theory and emerging topics in data science, written in a clear, tutorial style. Covering topics such as data acquisition, representation, analysis, and communication, it is ideal for graduate students and researchers in information theory, signal processing, and machine learning.

Table of contents

Pattern Recognition and Machine Learning

Fundamentals, Techniques, and Applications

Grammatical Inference

Variational Methods for Machine Learning with Applications to Deep Networks

Learning Automata and Grammars

Knowledge Graphs

This book provides a straightforward look at the concepts, algorithms and advantages of Bayesian Deep Learning and Deep Generative Models. Starting from the model-based approach to Machine Learning, the authors motivate Probabilistic Graphical Models and show how Bayesian inference naturally lends itself to this framework. The authors present detailed explanations of the main modern algorithms on variational approximations for Bayesian inference in neural networks. Each algorithm of this selected set develops a distinct aspect of the theory. The book builds from the ground-up well-known deep generative models, such as Variational Autoencoder and subsequent theoretical developments. By also exposing the main issues of the algorithms together with different methods to mitigate such issues, the book supplies the necessary knowledge on generative models for the reader to handle a wide range of data types: sequential or not, continuous or not, labelled or not. The book is self-contained, promptly covering all necessary theory so that the reader does not have to search for additional information elsewhere. Offers a concise self-contained resource, covering the basic concepts to the algorithms for Bayesian Deep Learning; Presents Statistical Inference concepts, offering a set of elucidative examples, practical aspects, and pseudo-codes; Every chapter includes hands-on examples and exercises and a website features lecture slides, additional examples, and other support material.

The problem of inducing, learning or inferring grammars has been studied for decades, but only in recent years has grammatical inference emerged as an independent field with connections to many scientific disciplines, including bio-informatics, computational linguistics and pattern recognition. This book meets the need for a comprehensive and unified summary of the basic techniques and results, suitable for researchers working in these various areas. In Part I, the objects of use for grammatical inference are studied in detail: strings and their topology, automata and grammars, whether probabilistic or not. Part II carefully explores the main questions in the field: What does learning mean? How can we associate complexity theory with learning? In Part III the author describes a number of techniques and algorithms that allow us to learn from text, from an informant, or through interaction with the environment. These concern automata, grammars, rewriting systems, pattern languages or transducers.

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.

A mathematically accessible textbook introducing all the tools needed to address modern inference problems in engineering and data science.

Renyi's Entropy and Kernel Perspectives

A Probabilistic Perspective

Information Theoretic Learning

Mathematics for Machine Learning

Statistical and Inductive Inference by Minimum Message Length

Machine Learning

Information theory and inference, taught together in this exciting textbook, lie at the heart of many important areas of modern technology - communication, signal processing, data mining, machine learning, pattern recognition, computational neuroscience, bioinformatics and cryptography. The book introduces theory in tandem with applications. Information theory is taught alongside practical communication systems such as arithmetic coding for data compression and sparse-graph codes for error-correction. Inference techniques, including message-passing algorithms, Monte Carlo methods and variational approximations, are developed alongside applications to clustering, convolutional codes, independent component analysis, and neural networks. Uniquely, the book covers state-of-the-art error-correcting codes, including low-density-parity-check codes, turbo codes, and digital fountain codes - the twenty-first-century standards for satellite communications, disk drives, and data broadcast. Richly illustrated, filled with worked examples and over 400 exercises, some with detailed solutions, the book is ideal for self-learning, and for undergraduate or graduate courses. It also provides an unparalleled entry point for professionals in areas as diverse as computational biology, financial engineering and machine learning.

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.

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.

The first comprehensive treatment of active inference, an integrative perspective on brain, cognition, and behavior used across multiple disciplines. Active inference is a way of understanding sentient behavior—a theory that characterizes perception, planning, and action in terms of probabilistic inference. Developed by theoretical neuroscientist Karl Friston over years of groundbreaking research, active inference provides an integrated perspective on brain, cognition, and behavior that is increasingly used across multiple disciplines including neuroscience, psychology, and philosophy. Active inference puts the action into perception. This book offers the first comprehensive treatment of active inference, covering theory, applications, and cognitive domains. Active inference is a "first principles" approach to understanding behavior and the brain, framed in terms of a single imperative to minimize free energy. The book emphasizes the implications of the free energy principle for understanding how the brain works. It first introduces active inference both conceptually and formally, contextualizing it within current theories of cognition. It then provides specific examples of computational models that use active inference to explain such cognitive phenomena as perception, attention, memory, and planning.

From Theory to Algorithms

The Nature of Statistical Learning Theory

ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures

Active Inference

The core of this paper is a general set of variational principles for the problems of computing marginal probabilities and modes, applicable to multivariate statistical models in the exponential family.

This graduate textbook provides a unified view of quantum information theory. Clearly explaining the necessary mathematical basis, it merges key topics from both information-theoretic and quantum- mechanical viewpoints and provides lucid explanations of the basic results. Thanks to this unified approach, it
accessible such advanced topics in quantum communication as quantum teleportation, superdense coding, quantum state transmission (quantum error-correction) and quantum encryption. Since the publication of the preceding book Quantum Information: An Introduction, there have been tremendous strides i
of quantum information. In particular, the following topics – all of which are addressed here – made seen major advances: quantum state discrimination, quantum channel capacity, bipartite and multipartite entanglement, security analysis on quantum communication, reverse Shannon theorem and uncertainty
With regard to the analysis of quantum security, the present book employs an improved method for the evaluation of leaked information and identifies a remarkable relation between quantum security and quantum coherence. Taken together, these two improvements allow a better analysis of quantum state tra
In addition, various types of the newly discovered uncertainty relation are explained. Presenting a wealth of new developments, the book introduces readers to the latest advances and challenges in quantum information. To aid in understanding, each chapter is accompanied by a set of exercises and solutions.

Physicists, when modelling physical systems with a large number of degrees of freedom, and statisticians, when performing data analysis, have developed their own concepts and methods for making the `best' inference. But are these methods equivalent, or not? What is the state of the art in making inference
physicists want answers. More: neural computation demands a clearer understanding of how neural systems make inferences; the theory of chaotic nonlinear systems as applied to time series analysis could profit from the experience already booked by the statisticians; and finally, there is a long-standing conj
some of the puzzles of quantum mechanics are due to our incomplete understanding of how we make inferences. Matter enough to stimulate the writing of such a book as the present one. But other considerations also arise, such as the maximum entropy method and Bayesian inference, information theory an
minimum description length. Finally, it is pointed out that an understanding of human inference may require input from psychologists. This lively debate, which is of acute current interest, is well summarized in the present work.

Bayesian probability theory has emerged not only as a powerful tool for building computational theories of vision, but also as a general paradigm for studying human visual perception. This 1996 book provides an introduction to and critical analysis of the Bayesian paradigm. Leading researchers in computer vi
experimental vision science describe general theoretical frameworks for modelling vision, detailed applications to specific problems and implications for experimental studies of human perception. The book provides a dialogue between different perspectives both within chapters, which draw on insights from ex
and computational work, and between chapters, through commentaries written by the contributors on each others' work. Students and researchers in cognitive and visual science will find much to interest them in this thought-provoking collection.