# *Information Theory Inference And Learning Algorithms*

***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.***

***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 informal introduction provides a fresh perspective on isomorphism theory, which is the branch of ergodic theory that explores the conditions under which two measure preserving systems are essentially equivalent. It contains a primer in basic measure theory, proofs of fundamental ergodic theorems, and material on entropy, martingales, Bernoulli processes, and various varieties of mixing. Original proofs of classic theorems - including the Shannon–McMillan–Breiman theorem, the Krieger finite generator theorem, and the Ornstein isomorphism theorem - are presented by degrees, together with helpful hints that encourage the reader to develop the proofs on their own. Hundreds of exercises and open problems are also included, making this an ideal text for graduate courses. Professionals needing a quick review, or seeking a different perspective on the subject, will also value this book.***

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

***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.*

*This book is devoted to one of the essential functions of modern telecommunications systems: channel coding or error correction coding. Its main topic is iteratively decoded algebraic codes, convolutional codes and concatenated codes.*

*Information Theory, Inference and Learning AlgorithmsCambridge University Press*

*Taken literally, the title "All of Statistics" is an exaggeration. But in spirit, the title is apt, as the book does cover a much broader range of topics than a typical introductory book on mathematical statistics. This book is for people who want to learn probability and statistics quickly. It is suitable for graduate or advanced undergraduate students in computer science, mathematics, statistics, and related disciplines. The book includes modern topics like non-parametric curve estimation, bootstrapping, and classification, topics that are usually relegated to follow-up courses. The reader is presumed to know calculus and a little linear algebra. No previous knowledge of probability and statistics is required. Statistics, data mining, and machine learning are all concerned with collecting and analysing data.*

*Or, The Calculus of Inference, Necessary and Probable*
*A Bayesian and Optimization Perspective*
*ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures*
*Regression, ConvNets, GANs, RNNs, NLP, and more with TensorFlow 2 and the Keras API, 2nd Edition*
*Models, Learning, and Inference*
*A Tutorial Introduction*
*Quantum Information Theory*
*Grammatical Inference*
*Understanding Machine Learning*
*Quantum Information Theory and the Foundations of Quantum Mechanics*
*Information Theory*

**Probability is the bedrock of machine learning. You cannot develop a deep understanding and application of machine learning without it. Cut through the equations, Greek letters, and confusion, and discover the topics in probability that you need to know. Using clear explanations, standard Python libraries, and step-by-step tutorial lessons, you will discover the importance of probability to machine learning, Bayesian probability, entropy, density estimation, maximum likelihood, and much more.**

**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.**

**Christopher G. Timpson provides the first full-length philosophical treatment of quantum information theory and the questions it raises for our understanding of the quantum world. He argues for an ontologically deflationary account of the nature of quantum information, which is grounded in a revisionary analysis of the concepts of information.**

**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.**

**Information Theory and Evolution discusses the phenomenon of life, including its origin and evolution (and also human cultural evolution), against the background of thermodynamics, statistical mechanics, and information theory. Among the central themes is the seeming contradiction between the second law of thermodynamics and the high degree of order and complexity produced by living systems. This paradox has its resolution in the information content of the Gibbs free energy that enters the biosphere from outside sources, as the author will show. The role of information in human cultural evolution is another focus of the book. The first edition of Information Theory and Evolution made a strong impact**

**on thought in the field by bringing together results from many disciplines. The new second edition offers updated results based on reports of important new research in several areas, including exciting new studies of the human mitochondrial and Y-chromosomal DNA. Another extensive discussion featured in the second edition is contained in a new appendix devoted to the relationship of entropy and Gibbs free energy to economics. This appendix includes a review of the ideas of Alfred Lotka, Frederick Soddy, Nicholas Georgiescu-Roegen and Herman E. Daly, and discusses the relevance of these ideas to the current economic crisis. The new edition discusses current research on the origin of life, the distinction between thermodynamic information and cybernetic information, new DNA research and human prehistory, developments in current information technology, and the relationship between entropy and economics.**

**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.**

**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.**

**Every other day we hear about new ways to put deep learning to good use: improved medical imaging, accurate credit card fraud detection, long range weather forecasting, and more. PyTorch puts these superpowers in your hands, providing a comfortable Python experience that gets you started quickly and then grows with you as you—and your deep learning skills—become more sophisticated. Deep Learning with PyTorch will make that journey engaging and fun. Summary Every other day we hear about new ways to put deep learning to good use: improved medical imaging, accurate credit card fraud detection, long range weather forecasting, and more. PyTorch puts these superpowers in your hands, providing a comfortable Python experience that gets you started quickly and then grows with you as you—and your deep learning skills—become more sophisticated. Deep Learning with PyTorch will make that journey engaging and fun. Foreword by Soumith Chintala, Cocreator of PyTorch. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology Although many deep learning tools use Python, the PyTorch library is truly Pythonic. Instantly familiar to anyone who knows PyData tools like NumPy and scikit-learn, PyTorch simplifies deep learning without sacrificing advanced features. It’s excellent for building quick models, and it scales smoothly from laptop to enterprise. Because companies like Apple, Facebook, and JPMorgan Chase rely on PyTorch, it’s a great skill to have as you expand your career options. It’s easy to get started with PyTorch. It minimizes cognitive overhead without sacrificing the access to advanced features, meaning you can focus on what matters the most - building and training the latest and greatest deep learning models and contribute to making a dent in the world. PyTorch is also a snap to scale and extend, and it partners well with other Python tooling. PyTorch has been adopted by hundreds of deep learning practitioners and several first-class players like FAIR, OpenAI, FastAI and Purdue. About the book Deep Learning with PyTorch teaches you to create neural networks and deep learning systems with PyTorch. This practical book quickly gets you to work building a real-world example from scratch: a tumor image classifier. Along the way, it covers best practices for the entire DL pipeline, including the PyTorch Tensor API, loading data in Python, monitoring training, and visualizing results. After covering the basics, the book will take you on a journey through larger projects. The centerpiece of the book is a neural network designed for cancer detection. You'll discover ways for training networks with limited inputs and start processing data to get some results. You'll sift through the unreliable initial results and focus on how to diagnose and fix the problems in your neural network. Finally, you'll look at ways to improve your results by training with augmented data, make improvements to the model architecture, and perform other fine tuning. What's inside Training deep neural networks Implementing modules and loss functions Utilizing pretrained models from PyTorch Hub Exploring code samples in Jupyter Notebooks About the reader For Python programmers with an interest in machine learning. About the author Eli Stevens had roles from software engineer to CTO, and is currently working on machine**

**learning in the self-driving-car industry. Luca Antiga is cofounder of an AI engineering company and an AI tech startup, as well as a former PyTorch contributor. Thomas Viehmann is a PyTorch core developer and machine learning trainer and consultant. consultant based in Munich, Germany and a PyTorch core developer. Table of Contents PART 1 - CORE PYTORCH 1 Introducing deep learning and the PyTorch Library 2 Pretrained networks 3 It starts with a tensor 4 Real-world data representation using tensors 5 The mechanics of learning 6 Using a neural network to fit the data 7 Telling birds from airplanes: Learning from images 8 Using convolutions to generalize PART 2 - LEARNING FROM IMAGES IN THE REAL WORLD: EARLY DETECTION OF LUNG CANCER 9 Using PyTorch to fight cancer 10 Combining data sources into a unified dataset 11 Training a classification model to detect suspected tumors 12 Improving training with metrics and augmentation 13 Using segmentation to find suspected nodules 14 End-to-end nodule analysis, and where to go next PART 3 - DEPLOYMENT 15 Deploying to production**

**Information Theory and Evolution**
**The Mathematical Theory of Communication**
**Information Theory , Inference And Learning Algorithms**
**Statistical and Inductive Inference by Minimum Message Length**
**Information Theory and Statistics**
**Machine Learning**
**Elements of Information Theory**
**All of Statistics**
**Learning Automata and Grammars**
**The Minimum Description Length Principle**
**Deep Learning with TensorFlow 2 and Keras**

**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.**

**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.**

**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.**

**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.**

**Build machine and deep learning systems with the newly released TensorFlow 2 and Keras for the lab, production, and mobile devices Key Features Introduces and then uses TensorFlow 2 and Keras right from the start Teaches key machine and deep learning techniques Understand the fundamentals of deep learning and machine learning through clear explanations and extensive code samples Book Description Deep Learning with TensorFlow 2 and Keras, Second Edition teaches neural networks and deep learning techniques alongside TensorFlow (TF) and Keras. You'll learn how to write deep learning applications in the most powerful, popular, and scalable machine learning stack available. TensorFlow is the machine learning library of choice for professional applications, while Keras offers a simple and powerful Python API for accessing TensorFlow. TensorFlow 2 provides full Keras integration, making advanced machine learning easier and more convenient than ever before. This book also introduces neural networks with TensorFlow, runs through the main applications (regression, ConvNets (CNNs), GANs, RNNs, NLP), covers two working example apps, and then dives into TF in production, TF mobile, and using TensorFlow with AutoML. What you will learn Build machine learning and deep learning systems with TensorFlow 2 and the Keras API Use Regression analysis, the most popular approach to machine learning Understand ConvNets (convolutional neural networks) and how they are essential for deep learning systems such as image classifiers Use GANs (generative adversarial networks) to create new data that fits with existing patterns Discover RNNs (recurrent neural networks) that can process sequences of input intelligently, using one part of a sequence to correctly interpret another Apply deep learning to natural human language and interpret natural language texts to produce an appropriate response Train your models on the cloud and put TF to work in real environments Explore how Google tools can automate simple ML workflows without the need for complex modeling Who this book is for This book is for Python developers and data scientists who want to build machine learning and deep learning systems with TensorFlow. Whether or not you have done machine learning before, this book gives you the theory and practice required to use Keras, TensorFlow 2, and AutoML to build machine learning systems.**

**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.**

**A modern treatment focusing on learning and inference, with minimal prerequisites, real-world examples and implementable algorithms.**

**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.**

**Mathematics for Machine Learning**
**Information Theory, Inference and Learning Algorithms**
**Variational Bayesian Learning Theory**
**Information and Inference**
**Principles of Statistical Inference**
**Sequential Decisions Based on Algorithmic Probability**
**A Probabilistic Perspective**
**A Concise Course in Statistical Inference**
**Deep Learning**
**Elements of Causal Inference**
**Bayesian Reasoning and Machine Learning**

*In this definitive book, D. R. Cox gives a comprehensive and balanced appraisal of statistical inference. He develops the key concepts, describing and comparing the main ideas and controversies over foundational issues that have been keenly argued for more than two-hundred years. Continuing a sixty-year career of major contributions to statistical thought, no one is better placed to give this much-needed* *account of the field. An appendix gives a more personal assessment of the merits of different ideas. The content ranges from the traditional to the contemporary. While specific applications are not treated, the book is strongly motivated by applications across the sciences and associated technologies. The mathematics is kept as elementary as feasible, though previous knowledge of statistics is assumed. The book will be valued by every user or student of statistics who is serious about understanding the uncertainty inherent in conclusions from statistical analyses.*

*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.*

*Did mandatory busing programs in the 1970s increase the school achievement of disadvantaged minority youth? Does obtaining a college degree increase an individual's labor market earnings? Did the use of the butterfly ballot in some Florida counties in the 2000 presidential election cost Al Gore votes? If so, was the number of miscast votes sufficiently large to have altered the election outcome? At their core, these types of questions are simple cause-and-effect questions. Simple cause-and-effect questions are the motivation for much empirical work in the social sciences. This book presents a model and set of methods for causal effect estimation that social scientists can use to address causal questions such as these. The essential features of the counterfactual model of causality for observational data analysis are presented with examples from sociology, political science, and economics.*

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

*This introduction to the MDL Principle provides a reference accessible to graduate students and researchers in statistics, pattern classification, machine learning, and data mining, to philosophers interested in the foundations of statistics, and to researchers in other applied sciences that involve model selection.*

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

*First comprehensive introduction to information theory explores the work of Shannon, McMillan, Feinstein, and Khinchin. Topics include the entropy concept in probability theory, fundamental theorems, and other subjects. 1957 edition.*

*Distills key concepts from linear algebra, geometry, matrices, calculus, optimization, probability and statistics that are used in machine learning.*

*Computer Vision*
*Probability for Machine Learning*
*Foundations and Learning Algorithms*
*From Theory to Algorithms*
*Computer Age Statistical Inference*
*Discover How To Harness Uncertainty With Python*
*Statistical Inference for Engineers and Data Scientists*

*Pattern Recognition and Machine Learning*
*Advanced Lectures on Machine Learning*
*An Outline of Ergodic Theory*

Table of contents

Personal motivation. The dream of creating artificial devices that reach or outperform human inteUigence is an old one. It is also one of the dreams of my youth, which have never left me. What makes this challenge so interesting? A solution would have enormous implications on our society, and there are reasons to believe that the AI problem can be solved in my expected lifetime. So, it's worth sticking to it for a lifetime, even if it takes 30 years or so to reap the benefits. The AI problem. The science of artificial intelligence (AI) may be defined as the construction of intelligent systems and their analysis. A natural definition of a system is anything that has an input and an output stream. Intelligence is more complicated. It can have many faces like creativity, solving prob lems, pattern recognition, classification, learning, induction, deduction, build ing analogies, optimization, surviving in an environment, language processing, and knowledge. A formal definition incorporating every aspect of intelligence, however, seems difficult. Most, if not all known facets of intelligence can be formulated as goal driven or, more precisely, as maximizing some utility func tion. It is, therefore, sufficient to study goal-driven AI; e. g. the (biological) goal of animals and humans is to survive and spread. The goal of AI systems should be to be useful to humans.

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 practical introduction perfect for final-year undergraduate and graduate students without a solid background in linear algebra and calculus.

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.

Take an exhilarating journey through the modern revolution in statistics with two of the ringleaders.

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 vision and 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 experimental 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.

Methods and Principles for Social Research
Codes and turbo codes
Data Mining, Inference, and Prediction
The Elements of Statistical Learning
Perception as Bayesian Inference
Formal Logic
Mathematical Foundations of Information Theory
Deep Learning with PyTorch
The Nature of Statistical Learning Theory
Counterfactuals and Causal Inference
Universal Artificial Intelligence