

Data Science Mit Python Das Handbuch Fur Den Eins

When somebody should go to the book stores, search launch by shop, shelf by shelf, it is in point of fact problematic. This is why we allow the books compilations in this website. It will certainly ease you to see guide **data science mit python das handbuch fur den eins** as you such as.

By searching the title, publisher, or authors of guide you in reality want, you can discover them rapidly. In the house, workplace, or perhaps in your method can be every best area within net connections. If you point to download and install the data science mit python das handbuch fur den eins, it is utterly easy then, past currently we extend the connect to purchase and make bargains to download and install data science mit python das handbuch fur den eins appropriately simple!

Datenvisualisierung mit Tableau Alexander Loth 2021-06-29 • Visuelle Datenanalyse leicht gemacht: Von den ersten Balkendiagrammen über Cluster und Trendlinien bis zu geografischen Analysen auf Landkarten • Erhalten Sie aussagefähige Prognosen durch vorausschauende Zukunftsanalysen • Erstellen und teilen Sie interaktive Dashboards und übersichtliche Infografiken Alexander Loth zeigt Ihnen in diesem Buch, wie Sie Ihre Daten ganz einfach visuell darstellen und analysieren. So können Sie selbst komplexe Datenstrukturen besser verstehen und daraus gewonnene Erkenntnisse effektiv kommunizieren. Der Autor erläutert Schritt für Schritt die grundlegenden Funktionen von Tableau. Anhand von Fallbeispielen lernen Sie praxisnah, welche Visualisierungsmöglichkeiten wann sinnvoll sind. Ferner zeigt er Anwendungen, die weit über gängige Standardanalysen hinausreichen, und geht auf Funktionen ein, die selbst erfahrenen Nutzern oft nicht hinlänglich bekannt sind. Sie erhalten außerdem zahlreiche Hinweise und Tipps, die Ihnen das Arbeiten mit Tableau merklich erleichtern. So können Sie zukünftig Ihre eigenen Daten bestmöglich visualisieren und analysieren. Das Buch richtet sich an: • alle, die Zugang zu Daten haben und diese verstehen möchten, • Führungskräfte, die Entscheidungen auf der Grundlage von Daten treffen, • Analysten und Entwickler, die Visualisierungen und Dashboards erstellen, • angehende Data Scientists Sie brauchen weder Tableau-Kenntnisse noch besondere mathematische Fähigkeiten oder Programmiererfahrung, um mit diesem Buch effektiv arbeiten zu können. Es eignet sich daher auch für Einsteiger und Anwender, die sich dem Thema Datenvisualisierung und -analyse praxisbezogen nähern möchten. Aus dem Inhalt: • Einführung und erste Schritte in Tableau • Datenquellen in Tableau anlegen • Visualisierungen erstellen • Aggregationen, Berechnungen und Parameter • Tabellenberechnungen und Detailgenauigkeitsausdrücke • Mit Karten zu weitreichenden Erkenntnissen • Tiefgehende Analysen mit Trends, Prognosen, Clustern und Verteilungen • Interaktive Dashboards • Teilen Sie Ihre Analysen

mit Ihrem Unternehmen oder der ganzen Welt • Daten integrieren und vorbereiten mit Tableau Prep Builder

Machine Learning mit Python und Keras, TensorFlow 2 und Scikit-learn Sebastian Raschka / Vahid Mirjalili 2021-03-03 • Datenanalyse mit ausgereiften statistischen Modellen des Machine Learnings • Anwendung der wichtigsten Algorithmen und Python-Bibliotheken wie NumPy, SciPy, Scikit-learn, Keras, TensorFlow 2, Pandas und Matplotlib • Best Practices zur Optimierung Ihrer Machine-Learning-Algorithmen Mit diesem Buch erhalten Sie eine umfassende Einführung in die Grundlagen und den effektiven Einsatz von Machine-Learning- und Deep-Learning-Algorithmen und wenden diese anhand zahlreicher Beispiele praktisch an. Dafür setzen Sie ein breites Spektrum leistungsfähiger Python-Bibliotheken ein, insbesondere Keras, TensorFlow 2 und Scikit-learn. Auch die für die praktische Anwendung unverzichtbaren mathematischen Konzepte werden verständlich und anhand zahlreicher Diagramme anschaulich erläutert. Die dritte Auflage dieses Buchs wurde für TensorFlow 2 komplett aktualisiert und berücksichtigt die jüngsten Entwicklungen und Technologien, die für Machine Learning, Neuronale Netze und Deep Learning wichtig sind. Dazu zählen insbesondere die neuen Features der Keras-API, das Synthetisieren neuer Daten mit Generative Adversarial Networks (GANs) sowie die Entscheidungsfindung per Reinforcement Learning. Ein sicherer Umgang mit Python wird vorausgesetzt.

Python Data Analytics Fabio Nelli 2015-08-25 Python Data Analytics will help you tackle the world of data acquisition and analysis using the power of the Python language. At the heart of this book lies the coverage of pandas, an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language. Author Fabio Nelli expertly shows the strength of the Python programming language when applied to processing, managing and retrieving information. Inside, you will see how intuitive and flexible it is to discover and communicate meaningful patterns of data using Python scripts, reporting systems, and data export. This book examines how to go about obtaining, processing, storing, managing and analyzing data using the Python programming language. You will use Python and other open source tools to wrangle data and tease out interesting and important trends in that data that will allow you to predict future patterns. Whether you are dealing with sales data, investment data (stocks, bonds, etc.), medical data, web page usage, or any other type of data set, Python can be used to interpret, analyze, and glean information from a pile of numbers and statistics. This book is an invaluable reference with its examples of storing and accessing data in a database; it walks you through the process of report generation; it provides three real world case studies or examples that you can take with you for your everyday analysis needs.

Python for Data Analysis Wes McKinney 2017-09-25 Get complete instructions for manipulating, processing, cleaning, and crunching datasets in Python. Updated for Python 3.6, the second edition of this hands-on guide is packed with practical case studies that show you how to solve a broad set of data analysis problems effectively. You'll learn the latest versions of pandas, NumPy,

IPython, and Jupyter in the process. Written by Wes McKinney, the creator of the Python pandas project, this book is a practical, modern introduction to data science tools in Python. It's ideal for analysts new to Python and for Python programmers new to data science and scientific computing. Data files and related material are available on GitHub. Use the IPython shell and Jupyter notebook for exploratory computing Learn basic and advanced features in NumPy (Numerical Python) Get started with data analysis tools in the pandas library Use flexible tools to load, clean, transform, merge, and reshape data Create informative visualizations with matplotlib Apply the pandas groupby facility to slice, dice, and summarize datasets Analyze and manipulate regular and irregular time series data Learn how to solve real-world data analysis problems with thorough, detailed examples

Advances in Data Science and Information Engineering Robert Stahlbock 2021-10-29 The book presents the proceedings of two conferences: the 16th International Conference on Data Science (ICDATA 2020) and the 19th International Conference on Information & Knowledge Engineering (IKE 2020), which took place in Las Vegas, NV, USA, July 27-30, 2020. The conferences are part of the larger 2020 World Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE'20), which features 20 major tracks. Papers cover all aspects of Data Science, Data Mining, Machine Learning, Artificial and Computational Intelligence (ICDATA) and Information Retrieval Systems, Information & Knowledge Engineering, Management and Cyber-Learning (IKE). Authors include academics, researchers, professionals, and students. Presents the proceedings of the 16th International Conference on Data Science (ICDATA 2020) and the 19th International Conference on Information & Knowledge Engineering (IKE 2020); Includes papers on topics from data mining to machine learning to informational retrieval systems; Authors include academics, researchers, professionals and students.

Foundations of Data Science Avrim Blum 2020-01-23 This book provides an introduction to the mathematical and algorithmic foundations of data science, including machine learning, high-dimensional geometry, and analysis of large networks. Topics include the counterintuitive nature of data in high dimensions, important linear algebraic techniques such as singular value decomposition, the theory of random walks and Markov chains, the fundamentals of and important algorithms for machine learning, algorithms and analysis for clustering, probabilistic models for large networks, representation learning including topic modelling and non-negative matrix factorization, wavelets and compressed sensing. Important probabilistic techniques are developed including the law of large numbers, tail inequalities, analysis of random projections, generalization guarantees in machine learning, and moment methods for analysis of phase transitions in large random graphs. Additionally, important structural and complexity measures are discussed such as matrix norms and VC-dimension. This book is suitable for both undergraduate and graduate courses in the design and analysis of algorithms for data.

Feature Engineering for Machine Learning Alice Zheng 2018-03-23 Feature

Downloaded from avenza-dev.avenza.com
on October 3, 2022 by guest

engineering is a crucial step in the machine-learning pipeline, yet this topic is rarely examined on its own. With this practical book, you'll learn techniques for extracting and transforming features—the numeric representations of raw data—into formats for machine-learning models. Each chapter guides you through a single data problem, such as how to represent text or image data. Together, these examples illustrate the main principles of feature engineering. Rather than simply teach these principles, authors Alice Zheng and Amanda Casari focus on practical application with exercises throughout the book. The closing chapter brings everything together by tackling a real-world, structured dataset with several feature-engineering techniques. Python packages including numpy, Pandas, Scikit-learn, and Matplotlib are used in code examples. You'll examine: Feature engineering for numeric data: filtering, binning, scaling, log transforms, and power transforms Natural text techniques: bag-of-words, n-grams, and phrase detection Frequency-based filtering and feature scaling for eliminating uninformative features Encoding techniques of categorical variables, including feature hashing and bin-counting Model-based feature engineering with principal component analysis The concept of model stacking, using k-means as a featurization technique Image feature extraction with manual and deep-learning techniques

Python: Deeper Insights into Machine Learning Sebastian Raschka 2016-08-31
Leverage benefits of machine learning techniques using Python About This Book
Improve and optimise machine learning systems using effective strategies.
Develop a strategy to deal with a large amount of data. Use of Python code for implementing a range of machine learning algorithms and techniques. Who This Book Is For This title is for data scientist and researchers who are already into the field of data science and want to see machine learning in action and explore its real-world application. Prior knowledge of Python programming and mathematics is must with basic knowledge of machine learning concepts. What You Will Learn Learn to write clean and elegant Python code that will optimize the strength of your algorithms Uncover hidden patterns and structures in data with clustering Improve accuracy and consistency of results using powerful feature engineering techniques Gain practical and theoretical understanding of cutting-edge deep learning algorithms Solve unique tasks by building models Get grips on the machine learning design process In Detail Machine learning and predictive analytics are becoming one of the key strategies for unlocking growth in a challenging contemporary marketplace. It is one of the fastest growing trends in modern computing, and everyone wants to get into the field of machine learning. In order to obtain sufficient recognition in this field, one must be able to understand and design a machine learning system that serves the needs of a project. The idea is to prepare a learning path that will help you to tackle the real-world complexities of modern machine learning with innovative and cutting-edge techniques. Also, it will give you a solid foundation in the machine learning design process, and enable you to build customized machine learning models to solve unique problems. The course begins with getting your Python fundamentals nailed down. It focuses on answering the right questions that cover a wide range of powerful Python libraries, including scikit-learn Theano and Keras. After getting familiar with Python core concepts,

it's time to dive into the field of data science. You will further gain a solid foundation on the machine learning design and also learn to customize models for solving problems. At a later stage, you will get a grip on more advanced techniques and acquire a broad set of powerful skills in the area of feature selection and feature engineering. Style and approach This course includes all the resources that will help you jump into the data science field with Python. The aim is to walk through the elements of Python covering powerful machine learning libraries. This course will explain important machine learning models in a step-by-step manner. Each topic is well explained with real-world applications with detailed guidance. Through this comprehensive guide, you will be able to explore machine learning techniques.

Hands-On Image Processing with Python Sandipan Dey 2018-11-30 Explore the mathematical computations and algorithms for image processing using popular Python tools and frameworks. Key Features Practical coverage of every image processing task with popular Python libraries Includes topics such as pseudo-coloring, noise smoothing, computing image descriptors Covers popular machine learning and deep learning techniques for complex image processing tasks Book Description Image processing plays an important role in our daily lives with various applications such as in social media (face detection), medical imaging (X-ray, CT-scan), security (fingerprint recognition) to robotics & space. This book will touch the core of image processing, from concepts to code using Python. The book will start from the classical image processing techniques and explore the evolution of image processing algorithms up to the recent advances in image processing or computer vision with deep learning. We will learn how to use image processing libraries such as PIL, scikit-mage, and scipy ndimage in Python. This book will enable us to write code snippets in Python 3 and quickly implement complex image processing algorithms such as image enhancement, filtering, segmentation, object detection, and classification. We will be able to use machine learning models using the scikit-learn library and later explore deep CNN, such as VGG-19 with Keras, and we will also use an end-to-end deep learning model called YOLO for object detection. We will also cover a few advanced problems, such as image inpainting, gradient blending, variational denoising, seam carving, quilting, and morphing. By the end of this book, we will have learned to implement various algorithms for efficient image processing. What you will learn Perform basic data pre-processing tasks such as image denoising and spatial filtering in Python Implement Fast Fourier Transform (FFT) and Frequency domain filters (e.g., Weiner) in Python Do morphological image processing and segment images with different algorithms Learn techniques to extract features from images and match images Write Python code to implement supervised / unsupervised machine learning algorithms for image processing Use deep learning models for image classification, segmentation, object detection and style transfer Who this book is for This book is for Computer Vision Engineers, and machine learning developers who are good with Python programming and want to explore details and complexities of image processing. No prior knowledge of the image processing techniques is expected.

Applied Data Science in Tourism Roman Egger 2022 Access to large data sets has

Downloaded from avenza-dev.avenza.com
on October 3, 2022 by guest

led to a paradigm shift in the tourism research landscape. Big data is enabling a new form of knowledge gain, while at the same time shaking the epistemological foundations and requiring new methods and analysis approaches. It allows for interdisciplinary cooperation between computer sciences and social and economic sciences, and complements the traditional research approaches. This book provides a broad basis for the practical application of data science approaches such as machine learning, text mining, social network analysis, and many more, which are essential for interdisciplinary tourism research. Each method is presented in principle, viewed analytically, and its advantages and disadvantages are weighed up and typical fields of application are presented. The correct methodical application is presented with a "how-to" approach, together with code examples, allowing a wider reader base including researchers, practitioners, and students entering the field. The book is a very well-structured introduction to data science not only in tourism and its methodological foundations, accompanied by well-chosen practical cases. It underlines an important insight: data are only representations of reality, you need methodological skills and domain background to derive knowledge from them.

- Hannes Werthner, Vienna University of Technology. Roman Egger has accomplished a difficult but necessary task: make clear how data science can practically support and foster travel and tourism research and applications. The book offers a well-taught collection of chapters giving a comprehensive and deep account of AI and data science for tourism.

- Francesco Ricci, Free University of Bozen-Bolzano. This well-structured and easy-to-read book provides a comprehensive overview of data science in tourism. It contributes largely to the methodological repository beyond traditional methods.

- Rob Law, University of Macau.

Graph Algorithms Mark Needham 2019-05-16 Discover how graph algorithms can help you leverage the relationships within your data to develop more intelligent solutions and enhance your machine learning models. You'll learn how graph analytics are uniquely suited to unfold complex structures and reveal difficult-to-find patterns lurking in your data. Whether you are trying to build dynamic network models or forecast real-world behavior, this book illustrates how graph algorithms deliver value—from finding vulnerabilities and bottlenecks to detecting communities and improving machine learning predictions. This practical book walks you through hands-on examples of how to use graph algorithms in Apache Spark and Neo4j—two of the most common choices for graph analytics. Also included: sample code and tips for over 20 practical graph algorithms that cover optimal pathfinding, importance through centrality, and community detection. Learn how graph analytics vary from conventional statistical analysis Understand how classic graph algorithms work, and how they are applied Get guidance on which algorithms to use for different types of questions Explore algorithm examples with working code and sample datasets from Spark and Neo4j See how connected feature extraction can increase machine learning accuracy and precision Walk through creating an ML workflow for link prediction combining Neo4j and Spark

[Introduction to Machine Learning with Python](#) Andreas C. Müller 2016-09-26

Downloaded from avenza-dev.avenza.com
on October 3, 2022 by guest

Machine learning has become an integral part of many commercial applications and research projects, but this field is not exclusive to large companies with extensive research teams. If you use Python, even as a beginner, this book will teach you practical ways to build your own machine learning solutions. With all the data available today, machine learning applications are limited only by your imagination. You'll learn the steps necessary to create a successful machine-learning application with Python and the scikit-learn library. Authors Andreas Müller and Sarah Guido focus on the practical aspects of using machine learning algorithms, rather than the math behind them. Familiarity with the NumPy and matplotlib libraries will help you get even more from this book. With this book, you'll learn:

- Fundamental concepts and applications of machine learning
- Advantages and shortcomings of widely used machine learning algorithms
- How to represent data processed by machine learning, including which data aspects to focus on
- Advanced methods for model evaluation and parameter tuning
- The concept of pipelines for chaining models and encapsulating your workflow
- Methods for working with text data, including text-specific processing techniques
- Suggestions for improving your machine learning and data science skills

Deep Learning Ian Goodfellow 2016-11-10 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.

Applied Data Science with Python and Jupyter Alex Galea 2018-10-31 Become the

Downloaded from avenza-dev.avenza.com
on October 3, 2022 by guest

master player of data exploration by creating reproducible data processing pipelines, visualizations, and prediction models for your applications. Key FeaturesGet up and running with the Jupyter ecosystem and some example datasetsLearn about key machine learning concepts such as SVM, KNN classifiers, and Random ForestsDiscover how you can use web scraping to gather and parse your own bespoke datasetsBook Description Getting started with data science doesn't have to be an uphill battle. Applied Data Science with Python and Jupyter is a step-by-step guide ideal for beginners who know a little Python and are looking for a quick, fast-paced introduction to these concepts. In this book, you'll learn every aspect of the standard data workflow process, including collecting, cleaning, investigating, visualizing, and modeling data. You'll start with the basics of Jupyter, which will be the backbone of the book. After familiarizing ourselves with its standard features, you'll look at an example of it in practice with our first analysis. In the next lesson, you dive right into predictive analytics, where multiple classification algorithms are implemented. Finally, the book ends by looking at data collection techniques. You'll see how web data can be acquired with scraping techniques and via APIs, and then briefly explore interactive visualizations. What you will learnGet up and running with the Jupyter ecosystemIdentify potential areas of investigation and perform exploratory data analysisPlan a machine learning classification strategy and train classification modelsUse validation curves and dimensionality reduction to tune and enhance your modelsScrape tabular data from web pages and transform it into Pandas DataFramesCreate interactive, web-friendly visualizations to clearly communicate your findingsWho this book is for Applied Data Science with Python and Jupyter is ideal for professionals with a variety of job descriptions across a large range of industries, given the rising popularity and accessibility of data science. You'll need some prior experience with Python, with any prior work with libraries such as Pandas, Matplotlib, and Pandas providing you a useful head start.

Python Data Science Handbook Jacob T. Vanderplas 2016 For many researchers, Python is a first-class tool mainly because of its libraries for storing, manipulating, and gaining insight from data. Several resources exist for individual pieces of this data science stack, but only with the Python Data Science Handbook do you get them all--IPython, NumPy, Pandas, Matplotlib, Scikit-Learn, and other related tools. Working scientists and data crunchers familiar with reading and writing Python code will find this comprehensive desk reference ideal for tackling day-to-day issues: manipulating, transforming, and cleaning data; visualizing different types of data; and using data to build statistical or machine learning models. Quite simply, this is the must-have reference for scientific computing in Python. With this handbook, you'll learn how to use: IPython and Jupyter: provide computational environments for data scientists using Python NumPy: includes the ndarray for efficient storage and manipulation of dense data arrays in Python Pandas: features the DataFrame for efficient storage and manipulation of labeled/columnar data in Python Matplotlib: includes capabilities for a flexible range of data visualizations in Python Scikit-Learn: for efficient and clean Python implementations of the most important and established machine learning algorithms

Data Science mit Python Jacob T. Vanderplas 2018

Software Technologies Marten van Sinderen 2021-07-20 This book constitutes the thoroughly refereed proceedings of the 15th International Conference on Software Technologies, ICSOFT 2020, which was held virtually due to the Covid-19 pandemic. The 12 revised full papers were carefully reviewed and selected from 95 submissions. The papers deal with the following topics: business process modelling; IT service management; interoperability and service-oriented architecture; project management software; scheduling and estimating; software metrics; requirements elicitation and specification; software and systems integration among others.

Humanities Data Analysis Folgert Karsdorp 2021-01-12 A practical guide to data-intensive humanities research using the Python programming language The use of quantitative methods in the humanities and related social sciences has increased considerably in recent years, allowing researchers to discover patterns in a vast range of source materials. Despite this growth, there are few resources addressed to students and scholars who wish to take advantage of these powerful tools. *Humanities Data Analysis* offers the first intermediate-level guide to quantitative data analysis for humanities students and scholars using the Python programming language. This practical textbook, which assumes a basic knowledge of Python, teaches readers the necessary skills for conducting humanities research in the rapidly developing digital environment. The book begins with an overview of the place of data science in the humanities, and proceeds to cover data carpentry: the essential techniques for gathering, cleaning, representing, and transforming textual and tabular data. Then, drawing from real-world, publicly available data sets that cover a variety of scholarly domains, the book delves into detailed case studies. Focusing on textual data analysis, the authors explore such diverse topics as network analysis, genre theory, onomastics, literacy, author attribution, mapping, stylometry, topic modeling, and time series analysis. Exercises and resources for further reading are provided at the end of each chapter. An ideal resource for humanities students and scholars aiming to take their Python skills to the next level, *Humanities Data Analysis* illustrates the benefits that quantitative methods can bring to complex research questions. Appropriate for advanced undergraduates, graduate students, and scholars with a basic knowledge of Python Applicable to many humanities disciplines, including history, literature, and sociology Offers real-world case studies using publicly available data sets Provides exercises at the end of each chapter for students to test acquired skills Emphasizes visual storytelling via data visualizations

Beginning Data Science with Python and Jupyter Alex Galea 2018-06-05 Getting started with data science doesn't have to be an uphill battle. This step-by-step guide is ideal for beginners who know a little Python and are looking for a quick, fast-paced introduction. Key Features Get up and running with the Jupyter ecosystem and some example datasets Learn about key machine learning concepts like SVM, KNN classifiers and Random Forests Discover how you can use web scraping to gather and parse your own bespoke datasets Book Description Get

to grips with the skills you need for entry-level data science in this hands-on Python and Jupyter course. You'll learn about some of the most commonly used libraries that are part of the Anaconda distribution, and then explore machine learning models with real datasets to give you the skills and exposure you need for the real world. We'll finish up by showing you how easy it can be to scrape and gather your own data from the open web, so that you can apply your new skills in an actionable context. What you will learn Get up and running with the Jupyter ecosystem and some example datasets Learn about key machine learning concepts like SVM, KNN classifiers, and Random Forests Plan a machine learning classification strategy and train classification, models Use validation curves and dimensionality reduction to tune and enhance your models Discover how you can use web scraping to gather and parse your own bespoke datasets Scrape tabular data from web pages and transform them into Pandas DataFrames Create interactive, web-friendly visualizations to clearly communicate your findings Who this book is for This book is ideal for professionals with a variety of job descriptions across large range of industries, given the rising popularity and accessibility of data science. You'll need some prior experience with Python, with any prior work with libraries like Pandas, Matplotlib and Pandas providing you a useful head start.

Python Data Science Handbook Jake VanderPlas 2016-11-21 For many researchers, Python is a first-class tool mainly because of its libraries for storing, manipulating, and gaining insight from data. Several resources exist for individual pieces of this data science stack, but only with the Python Data Science Handbook do you get them all—IPython, NumPy, Pandas, Matplotlib, Scikit-Learn, and other related tools. Working scientists and data crunchers familiar with reading and writing Python code will find this comprehensive desk reference ideal for tackling day-to-day issues: manipulating, transforming, and cleaning data; visualizing different types of data; and using data to build statistical or machine learning models. Quite simply, this is the must-have reference for scientific computing in Python. With this handbook, you'll learn how to use: IPython and Jupyter: provide computational environments for data scientists using Python NumPy: includes the ndarray for efficient storage and manipulation of dense data arrays in Python Pandas: features the DataFrame for efficient storage and manipulation of labeled/columnar data in Python Matplotlib: includes capabilities for a flexible range of data visualizations in Python Scikit-Learn: for efficient and clean Python implementations of the most important and established machine learning algorithms

Balancing the Tension between Digital Technologies and Learning Sciences Dirk Ifenthaler 2021-02-16 This volume focuses on the implications of digital technologies for educators and educational decision makers that is not widely represented in the literature. While there are many volumes on how one might integrate a particular technology, there are no volumes on how digital technologies can or should be exploited to address the needs and propel the benefits of large-scale teaching, learning and assessment.

Getting Started with Data Science Murtaza Haider 2015-12-14 Master Data

Analytics Hands-On by Solving Fascinating Problems You'll Actually Enjoy! Harvard Business Review recently called data science "The Sexiest Job of the 21st Century." It's not just sexy: For millions of managers, analysts, and students who need to solve real business problems, it's indispensable. Unfortunately, there's been nothing easy about learning data science—until now. Getting Started with Data Science takes its inspiration from worldwide best-sellers like Freakonomics and Malcolm Gladwell's Outliers: It teaches through a powerful narrative packed with unforgettable stories. Murtaza Haider offers informative, jargon-free coverage of basic theory and technique, backed with plenty of vivid examples and hands-on practice opportunities. Everything's software and platform agnostic, so you can learn data science whether you work with R, Stata, SPSS, or SAS. Best of all, Haider teaches a crucial skillset most data science books ignore: how to tell powerful stories using graphics and tables. Every chapter is built around real research challenges, so you'll always know why you're doing what you're doing. You'll master data science by answering fascinating questions, such as: • Are religious individuals more or less likely to have extramarital affairs? • Do attractive professors get better teaching evaluations? • Does the higher price of cigarettes deter smoking? • What determines housing prices more: lot size or the number of bedrooms? • How do teenagers and older people differ in the way they use social media? • Who is more likely to use online dating services? • Why do some purchase iPhones and others Blackberry devices? • Does the presence of children influence a family's spending on alcohol? For each problem, you'll walk through defining your question and the answers you'll need; exploring how others have approached similar challenges; selecting your data and methods; generating your statistics; organizing your report; and telling your story. Throughout, the focus is squarely on what matters most: transforming data into insights that are clear, accurate, and can be acted upon.

Data Science mit Python Jake VanderPlas 2017-11-24 Die wichtigsten Tools für die Datenanalyse und-bearbeitung im praktischen Einsatz Python effizient für datenintensive Berechnungen einsetzen mit IPython und Jupyter Laden, Speichern und Bearbeiten von Daten und numerischen Arrays mit NumPy und Pandas Visualisierung von Daten mit Matplotlib Python ist für viele die erste Wahl für Data Science, weil eine Vielzahl von Ressourcen und Bibliotheken zum Speichern, Bearbeiten und Auswerten von Daten verfügbar ist. In diesem Buch erläutert der Autor den Einsatz der wichtigsten Tools. Für Datenanalytiker und Wissenschaftler ist dieses umfassende Handbuch von unschätzbarem Wert für jede Art von Berechnung mit Python sowie bei der Erledigung alltäglicher Aufgaben. Dazu gehören das Bearbeiten, Umwandeln und Bereinigen von Daten, die Visualisierung verschiedener Datentypen und die Nutzung von Daten zum Erstellen von Statistiken oder Machine-Learning-Modellen. Dieses Handbuch erläutert die Verwendung der folgenden Tools: ● IPython und Jupyter für datenintensive Berechnungen ● NumPy und Pandas zum effizienten Speichern und Bearbeiten von Daten und Datenarrays in Python ● Matplotlib für vielfältige Möglichkeiten der Visualisierung von Daten ● Scikit-Learn zur effizienten und sauberen Implementierung der wichtigsten und am meisten verbreiteten Algorithmen des Machine Learnings Der Autor zeigt Ihnen, wie Sie die zum Betreiben von Data

Science verfügbaren Pakete nutzen, um Daten effektiv zu speichern, zu handhaben und Einblick in diese Daten zu gewinnen. Grundlegende Kenntnisse in Python werden dabei vorausgesetzt. Leserstimme zum Buch: »Wenn Sie Data Science mit Python betreiben möchten, ist dieses Buch ein hervorragender Ausgangspunkt. Ich habe es sehr erfolgreich beim Unterrichten von Informatik- und Statistikstudenten eingesetzt. Jake geht weit über die Grundlagen der Open-Source-Tools hinaus und erläutert die grundlegenden Konzepte, Vorgehensweisen und Abstraktionen in klarer Sprache und mit verständlichen Erklärungen.« – Brian Granger, Physikprofessor, California Polytechnic State University, Mitbegründer des Jupyter-Projekts

Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow Aurélien Géron 2019-09-05 Through a series of recent breakthroughs, deep learning has boosted the entire field of machine learning. Now, even programmers who know close to nothing about this technology can use simple, efficient tools to implement programs capable of learning from data. This practical book shows you how. By using concrete examples, minimal theory, and two production-ready Python frameworks—Scikit-Learn and TensorFlow—author Aurélien Géron helps you gain an intuitive understanding of the concepts and tools for building intelligent systems. You'll learn a range of techniques, starting with simple linear regression and progressing to deep neural networks. With exercises in each chapter to help you apply what you've learned, all you need is programming experience to get started. Explore the machine learning landscape, particularly neural nets Use Scikit-Learn to track an example machine-learning project end-to-end Explore several training models, including support vector machines, decision trees, random forests, and ensemble methods Use the TensorFlow library to build and train neural nets Dive into neural net architectures, including convolutional nets, recurrent nets, and deep reinforcement learning Learn techniques for training and scaling deep neural nets

Python Machine Learning Sebastian Raschka 2015-09-23 Unlock deeper insights into Machine Learning with this vital guide to cutting-edge predictive analytics About This Book Leverage Python's most powerful open-source libraries for deep learning, data wrangling, and data visualization Learn effective strategies and best practices to improve and optimize machine learning systems and algorithms Ask – and answer – tough questions of your data with robust statistical models, built for a range of datasets Who This Book Is For If you want to find out how to use Python to start answering critical questions of your data, pick up Python Machine Learning – whether you want to get started from scratch or want to extend your data science knowledge, this is an essential and unmissable resource. What You Will Learn Explore how to use different machine learning models to ask different questions of your data Learn how to build neural networks using Keras and Theano Find out how to write clean and elegant Python code that will optimize the strength of your algorithms Discover how to embed your machine learning model in a web application for increased accessibility Predict continuous target outcomes using regression analysis Uncover hidden patterns and structures in data with clustering Organize data using effective pre-processing techniques Get to grips with sentiment analysis to delve deeper

into textual and social media data In Detail Machine learning and predictive analytics are transforming the way businesses and other organizations operate. Being able to understand trends and patterns in complex data is critical to success, becoming one of the key strategies for unlocking growth in a challenging contemporary marketplace. Python can help you deliver key insights into your data – its unique capabilities as a language let you build sophisticated algorithms and statistical models that can reveal new perspectives and answer key questions that are vital for success. Python Machine Learning gives you access to the world of predictive analytics and demonstrates why Python is one of the world's leading data science languages. If you want to ask better questions of data, or need to improve and extend the capabilities of your machine learning systems, this practical data science book is invaluable. Covering a wide range of powerful Python libraries, including scikit-learn, Theano, and Keras, and featuring guidance and tips on everything from sentiment analysis to neural networks, you'll soon be able to answer some of the most important questions facing you and your organization. Style and approach Python Machine Learning connects the fundamental theoretical principles behind machine learning to their practical application in a way that focuses you on asking and answering the right questions. It walks you through the key elements of Python and its powerful machine learning libraries, while demonstrating how to get to grips with a range of statistical models.

Developing Analytic Talent Vincent Granville 2014-03-24 Learn what it takes to succeed in the the most in-demand tech job Harvard Business Review calls it the sexiest tech job of the 21st century. Data scientists are in demand, and this unique book shows you exactly what employers want and the skill set that separates the quality data scientist from other talented IT professionals. Data science involves extracting, creating, and processing data to turn it into business value. With over 15 years of big data, predictive modeling, and business analytics experience, author Vincent Granville is no stranger to data science. In this one-of-a-kind guide, he provides insight into the essential data science skills, such as statistics and visualization techniques, and covers everything from analytical recipes and data science tricks to common job interview questions, sample resumes, and source code. The applications are endless and varied: automatically detecting spam and plagiarism, optimizing bid prices in keyword advertising, identifying new molecules to fight cancer, assessing the risk of meteorite impact. Complete with case studies, this book is a must, whether you're looking to become a data scientist or to hire one. Explains the finer points of data science, the required skills, and how to acquire them, including analytical recipes, standard rules, source code, and a dictionary of terms Shows what companies are looking for and how the growing importance of big data has increased the demand for data scientists Features job interview questions, sample resumes, salary surveys, and examples of job ads Case studies explore how data science is used on Wall Street, in botnet detection, for online advertising, and in many other business-critical situations Developing Analytic Talent: Becoming a Data Scientist is essential reading for those aspiring to this hot career choice and for employers seeking the best candidates.

Machine Learning mit Python und Scikit-Learn und TensorFlow Sebastian Raschka
2017-12-22

Mastering Machine Learning with Python in Six Steps Manohar Swamynathan
2019-10-01 Explore fundamental to advanced Python 3 topics in six steps, all designed to make you a worthy practitioner. This updated version's approach is based on the "six degrees of separation" theory, which states that everyone and everything is a maximum of six steps away and presents each topic in two parts: theoretical concepts and practical implementation using suitable Python 3 packages. You'll start with the fundamentals of Python 3 programming language, machine learning history, evolution, and the system development frameworks. Key data mining/analysis concepts, such as exploratory analysis, feature dimension reduction, regressions, time series forecasting and their efficient implementation in Scikit-learn are covered as well. You'll also learn commonly used model diagnostic and tuning techniques. These include optimal probability cutoff point for class creation, variance, bias, bagging, boosting, ensemble voting, grid search, random search, Bayesian optimization, and the noise reduction technique for IoT data. Finally, you'll review advanced text mining techniques, recommender systems, neural networks, deep learning, reinforcement learning techniques and their implementation. All the code presented in the book will be available in the form of iPython notebooks to enable you to try out these examples and extend them to your advantage. What You'll Learn
Understand machine learning development and frameworks
Assess model diagnosis and tuning in machine learning
Examine text mining, natural language processing (NLP), and recommender systems
Review reinforcement learning and CNN
Who This Book Is For Python developers, data engineers, and machine learning engineers looking to expand their knowledge or career into machine learning area.

Machine Learning mit Python und Keras, TensorFlow2 und Scikit-learn Sebastian Raschka 2021-03-31

Datenvisualisierung mit Power BI Alexander Loth / Peter Vogel 2021-11-22 •
Visuelle Datenanalyse leicht gemacht: Von den ersten Balkendiagrammen über Cluster und Trendlinien zu geografischen Analysen auf Landkarten • Erhalten Sie aussagefähige Prognosen durch vorausschauende Zukunftsanalysen • Erstellen und teilen Sie interaktive Dashboards und übersichtliche Infografiken Visualisieren Sie Ihre Daten schnell und ausdrucksstark mit Power BI, um praktisch umsetzbare Ergebnisse zu erhalten. Alexander Loth und Peter Vogel zeigen Ihnen Schritt für Schritt, wie Sie ganz einfach visuelle Analysen erstellen und so selbst komplexe Datenstrukturen verstehen sowie gewonnene Erkenntnisse effektiv kommunizieren können. Das Buch richtet sich an die folgenden Zielgruppen: • Alle, die Zugang zu Daten haben und diese verstehen möchten • Führungskräfte, die Entscheidungen auf Grundlage von Daten treffen • Analysten und Entwickler, die Visualisierungen und Dashboards erstellen • angehende Data Scientists
Zum Verständnis dieses Buches und dem Erwerb von Power BI-Kenntnissen sind weder besondere mathematische Fähigkeiten noch Programmiererfahrung nötig. Es eignet sich daher auch für Einsteiger und Anwender, die sich dem Thema Datenvisualisierung und -analyse praxisbezogen nähern möchten, ohne

ausschweifende theoretische Abhandlungen. Die grundlegenden Funktionen von Power BI werden Schritt für Schritt erläutert und Sie lernen, welche Visualisierungsmöglichkeiten wann sinnvoll sind. Die Autoren zeigen Fallbeispiele auf, die weit über eine »Standardanalyse« hinausreichen und gehen auf Funktionen ein, die selbst erfahrenen Nutzern oft nicht hinlänglich bekannt sind. Sie geben Ihnen außerdem wertvolle Hinweise und Tipps, die das Arbeiten mit Power BI merklich erleichtern. So können Sie zukünftig Ihre eigenen Daten bestmöglich visualisieren und analysieren. Aus dem Inhalt: • Einführung und erste Schritte in Power BI • Datenquellen in Power BI anlegen • Visualisierungen erstellen • Aggregationen, Berechnungen und Parameter • Dimensionsübergreifende Berechnungen mit DAX-Ausdrücken • Mit Karten zu weitreichenden Erkenntnissen • Tiefgehende Analysen mit Trends, Prognosen, Clustern und Verteilung • Interaktive Dashboards • Analysen teilen

Hands-On Data Analysis with NumPy and Pandas Miller Miller 2018-06-29 Get to grips with the most popular Python packages that make data analysis possible Key Features Explore the tools you need to become a data analyst Discover practical examples to help you grasp data processing concepts Walk through hierarchical indexing and grouping for data analysis Book Description Python, a multi-paradigm programming language, has become the language of choice for data scientists for visualization, data analysis, and machine learning. Hands-On Data Analysis with NumPy and Pandas starts by guiding you in setting up the right environment for data analysis with Python, along with helping you install the correct Python distribution. In addition to this, you will work with the Jupyter notebook and set up a database. Once you have covered Jupyter, you will dig deep into Python's NumPy package, a powerful extension with advanced mathematical functions. You will then move on to creating NumPy arrays and employing different array methods and functions. You will explore Python's pandas extension which will help you get to grips with data mining and learn to subset your data. Last but not the least you will grasp how to manage your datasets by sorting and ranking them. By the end of this book, you will have learned to index and group your data for sophisticated data analysis and manipulation. What you will learn Understand how to install and manage Anaconda Read, sort, and map data using NumPy and pandas Find out how to create and slice data arrays using NumPy Discover how to subset your DataFrames using pandas Handle missing data in a pandas DataFrame Explore hierarchical indexing and plotting with pandas Who this book is for Hands-On Data Analysis with NumPy and Pandas is for you if you are a Python developer and want to take your first steps into the world of data analysis. No previous experience of data analysis is required to enjoy this book.

Deep Reinforcement Learning. Das umfassende Praxis-Handbuch Maxim Lapan 2020-06-18 • Alle wichtigen Methoden und Algorithmen praxisnah erläutert mit Codebeispielen in Python • Selbstständig lernende Agenten programmieren für die Steuerung von Robotern, NLP in interaktiven Spielen, Chatbots und mehr • Deep Q-Networks, Wertiteration, Policy Gradients, Trust Region Policy Optimization (TRPO), genetische Algorithmen, moderne Explorationsverfahren u.v.m. Reinforcement Learning ist ein Teilgebiet des Machine Learnings. Hierbei werden

selbstständig lernende Agenten programmiert, deren Lernvorgang ausschließlich durch ein Belohnungssystem und die Beobachtung der Umgebung gesteuert wird. In diesem umfassenden Praxis-Handbuch zeigt Ihnen Maxim Lapan, wie Sie diese zukunftsweisende Technologie in der Praxis einsetzen. Sie lernen, wie Sie passende RL-Methoden für Ihre Problemstellung auswählen und mithilfe von Deep-Learning-Methoden Agenten für verschiedene Aufgaben trainieren wie zum Beispiel für das Lösen eines Zauberwürfels, für Natural Language Processing in Microsofts TextWorld-Umgebung oder zur Realisierung moderner Chatbots. Alle Beispiele sind so gewählt, dass sie leicht verständlich sind und Sie diese auch ohne Zugang zu sehr großer Rechenleistung umsetzen können. Unter Einsatz von Python und der Bibliothek PyTorch ermöglicht Ihnen der Autor so einen einfachen und praktischen Einstieg in die Konzepte und Methoden des Reinforcement Learnings wie Deep Q-Networks, Wertiteration, Policy Gradients, Trust Region Policy Optimization (TRPO), genetische Algorithmen und viele mehr. Es werden grundlegende Kenntnisse in Machine Learning und Deep Learning sowie ein sicherer Umgang mit Python vorausgesetzt.

Practical Data Analysis Using Jupyter Notebook Marc Wintjen 2020-06-19

Understand data analysis concepts to make accurate decisions based on data using Python programming and Jupyter Notebook Key Features Find out how to use Python code to extract insights from data using real-world examples Work with structured data and free text sources to answer questions and add value using data Perform data analysis from scratch with the help of clear explanations for cleaning, transforming, and visualizing data Book Description Data literacy is the ability to read, analyze, work with, and argue using data. Data analysis is the process of cleaning and modeling your data to discover useful information. This book combines these two concepts by sharing proven techniques and hands-on examples so that you can learn how to communicate effectively using data. After introducing you to the basics of data analysis using Jupyter Notebook and Python, the book will take you through the fundamentals of data. Packed with practical examples, this guide will teach you how to clean, wrangle, analyze, and visualize data to gain useful insights, and you'll discover how to answer questions using data with easy-to-follow steps. Later chapters teach you about storytelling with data using charts, such as histograms and scatter plots. As you advance, you'll understand how to work with unstructured data using natural language processing (NLP) techniques to perform sentiment analysis. All the knowledge you gain will help you discover key patterns and trends in data using real-world examples. In addition to this, you will learn how to handle data of varying complexity to perform efficient data analysis using modern Python libraries. By the end of this book, you'll have gained the practical skills you need to analyze data with confidence. What you will learn Understand the importance of data literacy and how to communicate effectively using data Find out how to use Python packages such as NumPy, pandas, Matplotlib, and the Natural Language Toolkit (NLTK) for data analysis Wrangle data and create DataFrames using pandas Produce charts and data visualizations using time-series datasets Discover relationships and how to join data together using SQL Use NLP techniques to work with unstructured data to create sentiment analysis models Discover patterns in real-world datasets that provide accurate

insightsWho this book is for This book is for aspiring data analysts and data scientists looking for hands-on tutorials and real-world examples to understand data analysis concepts using SQL, Python, and Jupyter Notebook. Anyone looking to evolve their skills to become data-driven personally and professionally will also find this book useful. No prior knowledge of data analysis or programming is required to get started with this book.

Python Data Analytics Fabio Nelli 2018-09-27 Explore the latest Python tools and techniques to help you tackle the world of data acquisition and analysis. You'll review scientific computing with NumPy, visualization with matplotlib, and machine learning with scikit-learn. This revision is fully updated with new content on social media data analysis, image analysis with OpenCV, and deep learning libraries. Each chapter includes multiple examples demonstrating how to work with each library. At its heart lies the coverage of pandas, for high-performance, easy-to-use data structures and tools for data manipulation Author Fabio Nelli expertly demonstrates using Python for data processing, management, and information retrieval. Later chapters apply what you've learned to handwriting recognition and extending graphical capabilities with the JavaScript D3 library. Whether you are dealing with sales data, investment data, medical data, web page usage, or other data sets, Python Data Analytics, Second Edition is an invaluable reference with its examples of storing, accessing, and analyzing data. What You'll LearnUnderstand the core concepts of data analysis and the Python ecosystem Go in depth with pandas for reading, writing, and processing data Use tools and techniques for data visualization and image analysis Examine popular deep learning libraries Keras, Theano, TensorFlow, and PyTorch Who This Book Is For Experienced Python developers who need to learn about Pythonic tools for data analysis

Mathematics for Machine Learning Marc Peter Deisenroth 2020-04-23 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.

Data Science and Machine Learning Dirk P. Kroese 2019-11-20 "This textbook is a well-rounded, rigorous, and informative work presenting the mathematics behind modern machine learning techniques. It hits all the right notes: the choice of

topics is up-to-date and perfect for a course on data science for mathematics students at the advanced undergraduate or early graduate level. This book fills a sorely-needed gap in the existing literature by not sacrificing depth for breadth, presenting proofs of major theorems and subsequent derivations, as well as providing a copious amount of Python code. I only wish a book like this had been around when I first began my journey!" -Nicholas Hoell, University of Toronto "This is a well-written book that provides a deeper dive into data-scientific methods than many introductory texts. The writing is clear, and the text logically builds up regularization, classification, and decision trees. Compared to its probable competitors, it carves out a unique niche. -Adam Loy, Carleton College The purpose of Data Science and Machine Learning: Mathematical and Statistical Methods is to provide an accessible, yet comprehensive textbook intended for students interested in gaining a better understanding of the mathematics and statistics that underpin the rich variety of ideas and machine learning algorithms in data science. Key Features: Focuses on mathematical understanding. Presentation is self-contained, accessible, and comprehensive. Extensive list of exercises and worked-out examples. Many concrete algorithms with Python code. Full color throughout. The Authors: Dirk P. Kroese, PhD, is a Professor of Mathematics and Statistics at The University of Queensland. He has published over 120 articles and five books in a wide range of areas in mathematics, statistics, data science, machine learning, and Monte Carlo methods. He is a pioneer of the well-known Cross-Entropy method—an adaptive Monte Carlo technique, which is being used around the world to help solve difficult estimation and optimization problems in science, engineering, and finance. Zdravko Botev, PhD, is an Australian Mathematical Science Institute Lecturer in Data Science and Machine Learning with an appointment at the University of New South Wales in Sydney, Australia. He is the recipient of the 2018 Christopher Heyde Medal of the Australian Academy of Science for distinguished research in the Mathematical Sciences. Thomas Taimre, PhD, is a Senior Lecturer of Mathematics and Statistics at The University of Queensland. His research interests range from applied probability and Monte Carlo methods to applied physics and the remarkably universal self-mixing effect in lasers. He has published over 100 articles, holds a patent, and is the coauthor of Handbook of Monte Carlo Methods (Wiley). Radislav Vaisman, PhD, is a Lecturer of Mathematics and Statistics at The University of Queensland. His research interests lie at the intersection of applied probability, machine learning, and computer science. He has published over 20 articles and two books.

Interpretable Machine Learning Christoph Molnar 2020 This book is about making machine learning models and their decisions interpretable. After exploring the concepts of interpretability, you will learn about simple, interpretable models such as decision trees, decision rules and linear regression. Later chapters focus on general model-agnostic methods for interpreting black box models like feature importance and accumulated local effects and explaining individual predictions with Shapley values and LIME. All interpretation methods are explained in depth and discussed critically. How do they work under the hood? What are their strengths and weaknesses? How can their outputs be interpreted? This book will enable you to select and correctly apply the interpretation

method that is most suitable for your machine learning project.

The Data Science Design Manual Steven S. Skiena 2017-07-01 This engaging and clearly written textbook/reference provides a must-have introduction to the rapidly emerging interdisciplinary field of data science. It focuses on the principles fundamental to becoming a good data scientist and the key skills needed to build systems for collecting, analyzing, and interpreting data. The Data Science Design Manual is a source of practical insights that highlights what really matters in analyzing data, and provides an intuitive understanding of how these core concepts can be used. The book does not emphasize any particular programming language or suite of data-analysis tools, focusing instead on high-level discussion of important design principles. This easy-to-read text ideally serves the needs of undergraduate and early graduate students embarking on an "Introduction to Data Science" course. It reveals how this discipline sits at the intersection of statistics, computer science, and machine learning, with a distinct heft and character of its own. Practitioners in these and related fields will find this book perfect for self-study as well. Additional learning tools: Contains "War Stories," offering perspectives on how data science applies in the real world Includes "Homework Problems," providing a wide range of exercises and projects for self-study Provides a complete set of lecture slides and online video lectures at www.data-manual.com Provides "Take-Home Lessons," emphasizing the big-picture concepts to learn from each chapter Recommends exciting "Kaggle Challenges" from the online platform Kaggle Highlights "False Starts," revealing the subtle reasons why certain approaches fail Offers examples taken from the data science television show "The Quant Shop" (www.quant-shop.com)

R for Data Science Hadley Wickham 2016-12-12 Learn how to use R to turn raw data into insight, knowledge, and understanding. This book introduces you to R, RStudio, and the tidyverse, a collection of R packages designed to work together to make data science fast, fluent, and fun. Suitable for readers with no previous programming experience, R for Data Science is designed to get you doing data science as quickly as possible. Authors Hadley Wickham and Garrett Grolemund guide you through the steps of importing, wrangling, exploring, and modeling your data and communicating the results. You'll get a complete, big-picture understanding of the data science cycle, along with basic tools you need to manage the details. Each section of the book is paired with exercises to help you practice what you've learned along the way. You'll learn how to: Wrangle—transform your datasets into a form convenient for analysis Program—learn powerful R tools for solving data problems with greater clarity and ease Explore—examine your data, generate hypotheses, and quickly test them Model—provide a low-dimensional summary that captures true "signals" in your dataset Communicate—learn R Markdown for integrating prose, code, and results

MACHINE LEARNING MIT PYTHON; DAS PRAXIS-HANDBUCH FÜR DATA SCIENCE, PREDICTIVE ANALYTICS UND DEEP LEARNING. SEBASTIAN RASCHKA.

