

Scientific Data Analysis Using Jython Scripting And Java

Eventually, you will enormously discover a extra experience and success by spending more cash. still when? accomplish you tolerate that you require to acquire those every needs later than having significantly cash? Why dont you try to acquire something basic in the beginning? Thats something that will guide you to comprehend even more re the globe, experience, some places, with history, amusement, and a lot more?

It is your utterly own era to discharge duty reviewing habit. in the course of guides you could enjoy now is **Scientific Data Analysis Using Jython Scripting And Java** below.

Remote Instrumentation for eScience and Related Aspects - Franco Davoli 2011-12-16

This book will focus on new Remote Instrumentation aspects related to middleware architecture, high-speed networking, wireless Grid for acquisition devices and sensor networks, QoS provisioning for real-time control, measurement instrumentation and methodology. Moreover, it will provide knowledge about the

automation of mechanisms oriented to accompanying processes that are usually performed by a human.

Another important point of this book is focusing on the future trends concerning Remote Instrumentation systems development and actions related to standardization of remote instrumentation mechanisms.

Workflows for e-Science - Ian J. Taylor 2007-12-31

This is a timely book

presenting an overview of the current state-of-the-art within established projects, presenting many different aspects of workflow from users to tool builders. It provides an overview of active research, from a number of different perspectives. It includes theoretical aspects of workflow and deals with workflow for e-Science as opposed to e-Commerce. The topics covered will be of interest to a wide range of practitioners.

[PySpark Recipes](#) - Raju Kumar Mishra 2017-12-09

Quickly find solutions to common programming problems encountered while processing big data. Content is presented in the popular problem-solution format. Look up the programming problem that you want to solve. Read the solution. Apply the solution directly in your own code. Problem solved! PySpark Recipes covers Hadoop and its shortcomings. The architecture of Spark, PySpark, and RDD are presented. You will learn to apply RDD to solve day-to-day big data problems. Python and

NumPy are included and make it easy for new learners of PySpark to understand and adopt the model. What You Will Learn Understand the advanced features of PySpark2 and SparkSQL Optimize your code Program SparkSQL with Python Use Spark Streaming and Spark MLlib with Python Perform graph analysis with GraphFrames Who This Book Is For Data analysts, Python programmers, big data enthusiasts

[Handbook of Applied Spatial Analysis](#) - Manfred M. Fischer 2009-12-24

The Handbook is written for academics, researchers, practitioners and advanced graduate students. It has been designed to be read by those new or starting out in the field of spatial analysis as well as by those who are already familiar with the field. The chapters have been written in such a way that readers who are new to the field will gain important overview and insight. At the same time, those readers who are already practitioners in the field will gain through the

advanced and/or updated tools and new materials and state-of-the-art developments included. This volume provides an accounting of the diversity of current and emergent approaches, not available elsewhere despite the many excellent journals and te- books that exist. Most of the chapters are original, some few are reprints from the Journal of Geographical Systems, Geographical Analysis, The Review of Regional Studies and Letters of Spatial and Resource Sciences. We let our contributors - velop, from their particular perspective and insights, their own strategies for m- ping the part of terrain for which they were responsible. As the chapters were submitted, we became the first consumers of the project we had initiated. We gained from depth, breadth and distinctiveness of our contributors' insights and, in particular, the presence of links between them.

Hydroinformatics - Praveen Kumar 2005-11-02
Modern hydrology is more

interdisciplinary than ever. Staggering amounts and varieties of information pour in from GIS and remote sensing systems every day, and this information must be collected, interpreted, and shared efficiently. Hydroinformatics: Data Integrative Approaches in Computation, Analysis, and Modeling introduces the tools, approache

Astronomical Data Analysis Software and Systems XV - Carlos Gabriel 2006

Mastering Java for Data Science - Alexey Grigorev 2017-04-27

Use Java to create a diverse range of Data Science applications and bring Data Science into production About This Book An overview of modern Data Science and Machine Learning libraries available in Java Coverage of a broad set of topics, going from the basics of Machine Learning to Deep Learning and Big Data frameworks. Easy-to-follow illustrations and the running example of building a search engine. Who This Book Is For

This book is intended for software engineers who are comfortable with developing Java applications and are familiar with the basic concepts of data science. Additionally, it will also be useful for data scientists who do not yet know Java but want or need to learn it. If you are willing to build efficient data science applications and bring them in the enterprise environment without changing the existing stack, this book is for you!

What You Will Learn

- Get a solid understanding of the data processing toolbox available in Java
- Explore the data science ecosystem available in Java
- Find out how to approach different machine learning problems with Java
- Process unstructured information such as natural language text or images
- Create your own search engine
- Get state-of-the-art performance with XGBoost
- Learn how to build deep neural networks with DeepLearning4j
- Build applications that scale and process large amounts of data
- Deploy data science models to

production and evaluate their performance

In Detail Java is the most popular programming language, according to the TIOBE index, and it is a typical choice for running production systems in many companies, both in the startup world and among large enterprises. Not surprisingly, it is also a common choice for creating data science applications: it is fast and has a great set of data processing tools, both built-in and external. What is more, choosing Java for data science allows you to easily integrate solutions with existing software, and bring data science into production with less effort. This book will teach you how to create data science applications with Java. First, we will revise the most important things when starting a data science application, and then brush up the basics of Java and machine learning before diving into more advanced topics. We start by going over the existing libraries for data processing and libraries with machine learning algorithms. After that,

we cover topics such as classification and regression, dimensionality reduction and clustering, information retrieval and natural language processing, and deep learning and big data. Finally, we finish the book by talking about the ways to deploy the model and evaluate it in production settings. Style and approach This is a practical guide where all the important concepts such as classification, regression, and dimensionality reduction are explained with the help of examples.

Performance Tuning of Scientific Applications - David H. Bailey 2010-11-23

With contributions from some of the most notable experts in the field, Performance Tuning of Scientific Applications presents current research in performance analysis. The book focuses on the following areas. Performance monitoring: Describes the state of the art in hardware and software tools that are commonly used for monitoring and measuring performance and managing large quantities

of data Performance analysis: Discusses modern approaches to computer performance benchmarking and presents results that offer valuable insight into these studies Performance modeling: Explains how researchers deduce accurate performance models from raw performance data or from other high-level characteristics of a scientific computation Automatic performance tuning: Explores ongoing research into automatic and semi-automatic techniques for optimizing computer programs to achieve superior performance on any computer platform Application tuning: Provides examples that show how the appropriate analysis of performance and some deft changes have resulted in extremely high performance Performance analysis has grown into a full-fledged, sophisticated field of empirical science. Describing useful research in modern performance science and engineering, this book helps real-world users of parallel computer systems to better

understand both the performance vagaries arising in scientific applications and the practical means for improving performance. Read about the book on [HPCwire](#) and [insideHPC](#)

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.

Practical Data Science with Hadoop and Spark -

Ofer Mendelevitch 2016-12-08

The Complete Guide to Data Science with Hadoop—For Technical Professionals, Businesspeople, and Students

Demand is soaring for professionals who can solve real data science problems with Hadoop and Spark. Practical Data Science with Hadoop® and Spark is your complete guide to doing just that. Drawing on immense experience with Hadoop and big data, three leading experts bring together everything you need: high-level concepts, deep-dive techniques, real-world use cases, practical applications, and hands-on tutorials. The authors introduce the essentials of data science and the modern Hadoop ecosystem, explaining how Hadoop and Spark have evolved into an effective platform for solving data science problems at scale. In addition to comprehensive application coverage, the authors also provide useful guidance on the important steps of data ingestion, data munging, and visualization. Once the groundwork is in place, the authors focus on specific applications, including machine learning, predictive modeling for sentiment

analysis, clustering for document analysis, anomaly detection, and natural language processing (NLP). This guide provides a strong technical foundation for those who want to do practical data science, and also presents business-driven guidance on how to apply Hadoop and Spark to optimize ROI of data science initiatives. Learn What data science is, how it has evolved, and how to plan a data science career How data volume, variety, and velocity shape data science use cases Hadoop and its ecosystem, including HDFS, MapReduce, YARN, and Spark Data importation with Hive and Spark Data quality, preprocessing, preparation, and modeling Visualization: surfacing insights from huge data sets Machine learning: classification, regression, clustering, and anomaly detection Algorithms and Hadoop tools for predictive modeling Cluster analysis and similarity functions Large-scale anomaly detection NLP: applying data science to human

language

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 [High Performance Computing in Remote Sensing](#) - Antonio J. Plaza 2007-10-18

Solutions for Time-Critical Remote Sensing Applications The recent use of latest-generation sensors in airborne and satellite platforms is producing a nearly continual stream of high-dimensional data, which, in turn, is creating new processing challenges. To address the computational requirements of time-critical applications, researchers have begun incorporating high performance computing (HPC) models in remote sensing missions. High Performance Computing in Remote Sensing is one of the first volumes to explore state-of-the-art HPC

techniques in the context of remote sensing problems. It focuses on the computational complexity of algorithms that are designed for parallel computing and processing. A Diverse Collection of Parallel Computing Techniques and Architectures The book first addresses key computing concepts and developments in remote sensing. It also covers application areas not necessarily related to remote sensing, such as multimedia and video processing. Each subsequent chapter illustrates a specific parallel computing paradigm, including multiprocessor (cluster-based) systems, large-scale and heterogeneous networks of computers, grid computing platforms, and specialized hardware architectures for remotely sensed data analysis and interpretation. An Interdisciplinary Forum to Encourage Novel Ideas The extensive reviews of current and future developments combined with thoughtful perspectives on the potential challenges of adapting HPC

paradigms to remote sensing problems will undoubtedly foster collaboration and development among many fields.

Beginning Data Science in R

- Thomas Mailund 2017-03-09

Discover best practices for data analysis and software development in R and start on the path to becoming a fully-fledged data scientist. This book teaches you techniques for both data manipulation and visualization and shows you the best way for developing new software packages for R. Beginning Data Science in R details how data science is a combination of statistics, computational science, and machine learning. You'll see how to efficiently structure and mine data to extract useful patterns and build mathematical models. This requires computational methods and programming, and R is an ideal programming language for this. This book is based on a number of lecture notes for classes the author has taught on data science and statistical programming using

the R programming language. Modern data analysis requires computational skills and usually a minimum of programming. What You Will Learn Perform data science and analytics using statistics and the R programming language Visualize and explore data, including working with large data sets found in big data Build an R package Test and check your code Practice version control Profile and optimize your code Who This Book Is For Those with some data science or analytics background, but not necessarily experience with the R programming language.

Neural Information Processing - Tingwen Huang 2012-11-05 The five volume set LNCS 7663, LNCS 7664, LNCS 7665, LNCS 7666 and LNCS 7667 constitutes the proceedings of the 19th International Conference on Neural Information Processing, ICONIP 2012, held in Doha, Qatar, in November 2012. The 423 regular session papers presented were carefully reviewed and selected from

numerous submissions. These papers cover all major topics of theoretical research, empirical study and applications of neural information processing research. The 5 volumes represent 5 topical sections containing articles on theoretical analysis, neural modeling, algorithms, applications, as well as simulation and synthesis.

Massively Multi-Agent Technology - Nadeem Jamali 2008-08-04

This book constitutes the thoroughly refereed post-workshop proceedings of the joint AAMAS 2006

International Workshops on Massively Multi-Agent Systems, MMAS 2006, and on Large scale Multi-Agent Systems, LSMAS 2006, held in Hakodate, Japan, in May 2006, and of the International Workshop on Coordination and Control in Massively Multi-Agent Systems, CCMMS 2007, held in Honolulu, HI, USA, in May 2007 as associated event of AAMAS 2007. The 13 revised full papers presented were carefully selected from the

presentations made at the 3 workshops fall in 4 broad categories, presenting a snapshot of current research. Included are implementation strategies addressing coordination in the space of spatial and temporal distributed systems; approaches to deal with complexity to make decisions such as task allocation and team formation efficiently, by creating implicit or explicit encapsulations; and finally, a diverse range of applications to which these approaches may be applied, from large-scale agent based simulations to managing different types of networks to image segmentation.

Quantitative Quality of Service for Grid Computing: Applications for Heterogeneity, Large-Scale Distribution, and Dynamic Environments - Wang, Lizhe 2009-05-31

"This book provides research into parallel & distributed computing, high performance computing, and Grid computing"--Provided by publisher.

Web Microanalysis of Big Image Data - Peter Bajcsy 2018-01-22

This book looks at the increasing interest in running microscopy processing algorithms on big image data by presenting the theoretical and architectural underpinnings of a web image processing pipeline (WIPP). Software-based methods and infrastructure components for processing big data microscopy experiments are presented to demonstrate how information processing of repetitive, laborious and tedious analysis can be automated with a user-friendly system. Interactions of web system components and their impact on computational scalability, provenance information gathering, interactive display, and computing are explained in a top-down presentation of technical details. Web Microanalysis of Big Image Data includes descriptions of WIPP functionalities, use cases, and components of the web software system (web server and client architecture,

algorithms, and hardware-software dependencies). The book comes with test image collections and a web software system to increase the reader's understanding and to provide practical tools for conducting big image experiments. By providing educational materials and software tools at the intersection of microscopy image analyses and computational science, graduate students, postdoctoral students, and scientists will benefit from the practical experiences, as well as theoretical insights. Furthermore, the book provides software and test data, empowering students and scientists with tools to make discoveries with higher statistical significance. Once they become familiar with the web image processing components, they can extend and re-purpose the existing software to new types of analyses. Each chapter follows a top-down presentation, starting with a short introduction and a classification of related

methods. Next, a description of the specific method used in accompanying software is presented. For several topics, examples of how the specific method is applied to a dataset (parameters, RAM requirements, CPU efficiency) are shown. Some tips are provided as practical suggestions to improve accuracy or computational performance.

Numeric Computation and Statistical Data Analysis on the Java Platform - Sergei V. Chekanov 2016-03-23

Numerical computation, knowledge discovery and statistical data analysis integrated with powerful 2D and 3D graphics for visualization are the key topics of this book. The Python code examples powered by the Java platform can easily be transformed to other programming languages, such as Java, Groovy, Ruby and BeanShell. This book equips the reader with a computational platform which, unlike other statistical programs, is not limited by a

single programming language. The author focuses on practical programming aspects and covers a broad range of topics, from basic introduction to the Python language on the Java platform (Jython), to descriptive statistics, symbolic calculations, neural networks, non-linear regression analysis and many other data-mining topics. He discusses how to find regularities in real-world data, how to classify data, and how to process data for knowledge discoveries. The code snippets are so short that they easily fit into single pages. *Numeric Computation and Statistical Data Analysis on the Java Platform* is a great choice for those who want to learn how statistical data analysis can be done using popular programming languages, who want to integrate data analysis algorithms in full-scale applications, and deploy such calculations on the web pages or computational servers regardless of their operating system. It is an excellent reference for scientific computations to solve real-

world problems using a comprehensive stack of open-source Java libraries included in the DataMelt (DMelt) project and will be appreciated by many data-analysis scientists, engineers and students.

Mathematical Software -- ICMS 2014 - Hoon Hong
2014-08-01

This book constitutes the proceedings of the 4th International Conference on Mathematical Software, ICMS 2014, held in Seoul, South Korea, in August 2014. The 108 papers included in this volume were carefully reviewed and selected from 150 submissions. The papers are organized in topical sections named: invited; exploration; group; coding; topology; algebraic; geometry; surfaces; reasoning; special; Groebner; triangular; parametric; interfaces and general.

Optical, Infrared, and Millimeter Space Telescopes - 2004

Proceedings of SPIE present the original research papers presented at SPIE conferences and other high-quality

conferences in the broad-ranging fields of optics and photonics. These books provide prompt access to the latest innovations in research and technology in their respective fields. Proceedings of SPIE are among the most cited references in patent literature. [Comprehensive Medicinal Chemistry III](#) - 2017-06-03 Comprehensive Medicinal Chemistry III provides a contemporary and forward-looking critical analysis and summary of recent developments, emerging trends, and recently identified new areas where medicinal chemistry is having an impact. The discipline of medicinal chemistry continues to evolve as it adapts to new opportunities and strives to solve new challenges. These include drug targeting, biomolecular therapeutics, development of chemical biology tools, data collection and analysis, in silico models as predictors for biological properties, identification and validation of new targets, approaches to quantify target

engagement, new methods for synthesis of drug candidates such as green chemistry, development of novel scaffolds for drug discovery, and the role of regulatory agencies in drug discovery. Reviews the strategies, technologies, principles, and applications of modern medicinal chemistry Provides a global and current perspective of today's drug discovery process and discusses the major therapeutic classes and targets Includes a unique collection of case studies and personal assays reviewing the discovery and development of key drugs *Machine Learning and Knowledge Discovery in Databases* - Peter A. Flach 2012-09-11

This two-volume set LNAI 7523 and LNAI 7524 constitutes the refereed proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases: ECML PKDD 2012, held in Bristol, UK, in September 2012. The 105 revised research papers presented together with 5

invited talks were carefully reviewed and selected from 443 submissions. The final sections of the proceedings are devoted to Demo and Nectar papers. The Demo track includes 10 papers (from 19 submissions) and the Nectar track includes 4 papers (from 14 submissions). The papers grouped in topical sections on association rules and frequent patterns; Bayesian learning and graphical models; classification; dimensionality reduction, feature selection and extraction; distance-based methods and kernels; ensemble methods; graph and tree mining; large-scale, distributed and parallel mining and learning; multi-relational mining and learning; multi-task learning; natural language processing; online learning and data streams; privacy and security; rankings and recommendations; reinforcement learning and planning; rule mining and subgroup discovery; semi-supervised and transductive learning; sensor data; sequence and string mining;

social network mining; spatial and geographical data mining; statistical methods and evaluation; time series and temporal data mining; and transfer learning.

[Python for Bioinformatics](#) -

Sebastian Bassi 2016-04-19

Programming knowledge is often necessary for finding a solution to a biological problem. Based on the author's experience working for an agricultural biotechnology company, Python for Bioinformatics helps scientists solve their biological problems by helping them understand the basics of programming. Requiring no prior knowledge of programming-related concepts, the book focuses on the easy-to-use, yet powerful, Python computer language. The book begins with a very basic introduction that teaches the principles of programming. It then introduces the Biopython package, which can be useful in solving life science problems. The next section covers sophisticated tools for bioinformatics, including relational database

management systems and XML. The last part illustrates applications with source code, such as sequence manipulation, filtering vector contamination, calculating DNA melting temperature, parsing a genbank file, inferring splicing sites, and more. The appendices provide a wealth of supplementary information, including instructions for installing Python and Biopython and a Python language and style guide. By incorporating examples in biology as well as code fragments throughout, the author places a special emphasis on practice, encouraging readers to experiment with the code. He shows how to use Python and the Biopython package for building web applications, genomic annotation, data manipulation, and countless other applications.

Big Data, Data Mining, and Machine Learning - Jared Dean 2014-05-07

With big data analytics comes big insights into profitability. Big data is big business. But

having the data and the computational power to process it isn't nearly enough to produce meaningful results. **Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners** is a complete resource for technology and marketing executives looking to cut through the hype and produce real results that hit the bottom line. Providing an engaging, thorough overview of the current state of big data analytics and the growing trend toward high performance computing architectures, the book is a detail-driven look into how big data analytics can be leveraged to foster positive change and drive efficiency. With continued exponential growth in data and ever more competitive markets, businesses must adapt quickly to gain every competitive advantage available. Big data analytics can serve as the linchpin for initiatives that drive business, but only if the underlying technology and analysis is fully understood and appreciated by engaged

stakeholders. This book provides a view into the topic that executives, managers, and practitioners require, and includes: A complete overview of big data and its notable characteristics Details on high performance computing architectures for analytics, massively parallel processing (MPP), and in-memory databases Comprehensive coverage of data mining, text analytics, and machine learning algorithms A discussion of explanatory and predictive modeling, and how they can be applied to decision-making processes Big Data, Data Mining, and Machine Learning provides technology and marketing executives with the complete resource that has been notably absent from the veritable libraries of published books on the topic. Take control of your organization's big data analytics to produce real results with a resource that is comprehensive in scope and light on hyperbole.

Scientific Data Analysis using Jython Scripting and Java - Sergei V. Chekanov

2010-08-05

Scientific Data Analysis using Jython Scripting and Java presents practical approaches for data analysis using Java scripting based on Jython, a Java implementation of the Python language. The chapters essentially cover all aspects of data analysis, from arrays and histograms to clustering analysis, curve fitting, metadata and neural networks. A comprehensive coverage of data visualisation tools implemented in Java is also included. Written by the primary developer of the jHepWork data-analysis framework, the book provides a reliable and complete reference source laying the foundation for data-analysis applications using Java scripting. More than 250 code snippets (of around 10-20 lines each) written in Jython and Java, plus several real-life examples help the reader develop a genuine feeling for data analysis techniques and their programming implementation. This is the first data-analysis and data-

mining book which is completely based on the Jython language, and opens doors to scripting using a fully multi-platform and multi-threaded approach. Graduate students and researchers will benefit from the information presented in this book.

Scientific Data Analysis using Jython Scripting and Java - Sergei V. Chekanov
2010-08-20

Scientific Data Analysis using Jython Scripting and Java presents practical approaches for data analysis using Java scripting based on Jython, a Java implementation of the Python language. The chapters essentially cover all aspects of data analysis, from arrays and histograms to clustering analysis, curve fitting, metadata and neural networks. A comprehensive coverage of data visualisation tools implemented in Java is also included. Written by the primary developer of the jHepWork data-analysis framework, the book provides a reliable and complete reference source laying the

foundation for data-analysis applications using Java scripting. More than 250 code snippets (of around 10-20 lines each) written in Jython and Java, plus several real-life examples help the reader develop a genuine feeling for data analysis techniques and their programming implementation. This is the first data-analysis and data-mining book which is completely based on the Jython language, and opens doors to scripting using a fully multi-platform and multi-threaded approach. Graduate students and researchers will benefit from the information presented in this book.

Data Analysis and Visualization Using Python - Dr. Ossama Embarak
2018-11-20

Look at Python from a data science point of view and learn proven techniques for data visualization as used in making critical business decisions. Starting with an introduction to data science with Python, you will take a closer look at the Python environment and get acquainted with editors such as

Jupyter Notebook and Spyder. After going through a primer on Python programming, you will grasp fundamental Python programming techniques used in data science. Moving on to data visualization, you will see how it caters to modern business needs and forms a key factor in decision-making. You will also take a look at some popular data visualization libraries in Python. Shifting focus to data structures, you will learn the various aspects of data structures from a data science perspective. You will then work with file I/O and regular expressions in Python, followed by gathering and cleaning data. Moving on to exploring and analyzing data, you will look at advanced data structures in Python. Then, you will take a deep dive into data visualization techniques, going through a number of plotting systems in Python. In conclusion, you will complete a detailed case study, where you'll get a chance to revisit the concepts you've covered so far. What You Will Learn Use Python programming

techniques for data science
Master data collections in Python
Create engaging visualizations for BI systems
Deploy effective strategies for gathering and cleaning data
Integrate the Seaborn and Matplotlib plotting systems
Who This Book Is For
Developers with basic Python programming knowledge looking to adopt key strategies for data analysis and visualizations using Python.
Astronomical Data Analysis Software and Systems XIV - Patrick L. Shopbell 2005

Web Services Research for Emerging Applications: Discoveries and Trends - Zhang, Liang-Jie 2010-02-28
"This book provides a comprehensive assessment of the latest developments in Web services research, focusing on composing and coordinating Web services, XML security, and service oriented architecture, and presenting new and emerging research in the Web services discipline"--
Provided by publisher.

Python for R Users - Ajay

Ohri 2017-11-13

The definitive guide for statisticians and data scientists who understand the advantages of becoming proficient in both R and Python. The first book of its kind, *Python for R Users: A Data Science Approach* makes it easy for R programmers to code in Python and Python users to program in R. Short on theory and long on actionable analytics, it provides readers with a detailed comparative introduction and overview of both languages and features concise tutorials with command-by-command translations—complete with sample code—of R to Python and Python to R. Following an introduction to both languages, the author cuts to the chase with step-by-step coverage of the full range of pertinent programming features and functions, including data input, data inspection/data quality, data analysis, and data visualization. Statistical modeling, machine learning, and data mining—including supervised and unsupervised

data mining methods—are treated in detail, as are time series forecasting, text mining, and natural language processing. • Features a quick-learning format with concise tutorials and actionable analytics • Provides command-by-command translations of R to Python and vice versa • Incorporates Python and R code throughout to make it easier for readers to compare and contrast features in both languages • Offers numerous comparative examples and applications in both programming languages • Designed for use for practitioners and students that know one language and want to learn the other • Supplies slides useful for teaching and learning either software on a companion website Python for R Users: A Data Science Approach is a valuable working resource for computer scientists and data scientists that know R and would like to learn Python or are familiar with Python and want to learn R. It also functions as textbook for students of computer

science and statistics. A. Ohri is the founder of Decisionstats.com and currently works as a senior data scientist. He has advised multiple startups in analytics off-shoring, analytics services, and analytics education, as well as using social media to enhance buzz for analytics products. Mr. Ohri's research interests include spreading open source analytics, analyzing social media manipulation with mechanism design, simpler interfaces for cloud computing, investigating climate change and knowledge flows. His other books include R for Business Analytics and R for Cloud Computing.

Python Scripting for Computational Science -

Hans Petter Langtangen
2013-03-14

Scripting with Python makes you productive and increases the reliability of your scientific work. Here, the author teaches you how to develop tailored, flexible, and efficient working environments built from small programs (scripts) written in Python. The focus is on

examples and applications of relevance to computational science: gluing existing applications and tools, e.g. for automating simulation, data analysis, and visualization; steering simulations and computational experiments; equipping programs with graphical user interfaces; making computational Web services; creating interactive interfaces with a Maple/Matlab-like syntax to numerical applications in C/C++ or Fortran; and building flexible object-oriented programming interfaces to existing C/C++ or Fortran libraries.

-

Bioimage Data Analysis Workflows -

Kota Miura
2019-10-17

This Open Access textbook provides students and researchers in the life sciences with essential practical information on how to quantitatively analyze data images. It refrains from focusing on theory, and instead uses practical examples and

step-by step protocols to familiarize readers with the most commonly used image processing and analysis platforms such as ImageJ, MatLab and Python. Besides gaining knowhow on algorithm usage, readers will learn how to create an analysis pipeline by scripting language; these skills are important in order to document reproducible image analysis workflows. The textbook is chiefly intended for advanced undergraduates in the life sciences and biomedicine without a theoretical background in data analysis, as well as for postdocs, staff scientists and faculty members who need to perform regular quantitative analyses of microscopy images. *The Handbook On Reasoning-based Intelligent Systems* - Nakamatsu Kazumi 2013-01-18 This book consists of various contributions in conjunction with the keywords “reasoning” and “intelligent systems”, which widely covers theoretical to practical aspects of intelligent systems. Therefore, it is suitable for researchers or

graduate students who want to study intelligent systems generally.

Agile Data Science - Russell Journey 2013-10-15

Mining big data requires a deep investment in people and time. How can you be sure you’re building the right models? With this hands-on book, you’ll learn a flexible toolset and methodology for building effective analytics applications with Hadoop. Using lightweight tools such as Python, Apache Pig, and the D3.js library, your team will create an agile environment for exploring data, starting with an example application to mine your own email inboxes. You’ll learn an iterative approach that enables you to quickly change the kind of analysis you’re doing, depending on what the data is telling you. All example code in this book is available as working Heroku apps. Create analytics applications by using the agile big data development methodology Build value from your data in a series of agile sprints, using the data-value

stack Gain insight by using several data structures to extract multiple features from a single dataset Visualize data with charts, and expose different aspects through interactive reports Use historical data to predict the future, and translate predictions into action Get feedback from users after each sprint to keep your project on track

New Scientist - 2005

The Definitive Guide to

Jython - Josh Juneau

2010-12-28

Jython is an open source implementation of the high-level, dynamic, object-oriented scripting language Python seamlessly integrated with the Java platform. The predecessor to Jython, JPython, is certified as 100% Pure Java. Jython is freely available for both commercial and noncommercial use and is distributed with source code. Jython is complementary to Java. The Definitive Guide to Jython, written by the official Jython team leads, covers

Jython 2.5 (or 2.5.x)—from the basics to more advanced features. This book begins with a brief introduction to the language and then journeys through Jython’s different features and uses. The Definitive Guide to Jython is organized for beginners as well as advanced users of the language. The book provides a general overview of the Jython language itself, but it also includes intermediate and advanced topics regarding database, web, and graphical user interface (GUI) applications; Web services/SOA; and integration, concurrency, and parallelism, to name a few.

Java: Data Science Made

Easy - Richard M. Reese

2017-07-07

Data collection, processing, analysis, and more About This Book Your entry ticket to the world of data science with the stability and power of Java Explore, analyse, and visualize your data effectively using easy-to-follow examples A highly practical course covering a broad set of topics -

from the basics of Machine Learning to Deep Learning and Big Data frameworks. Who This Book Is For This course is meant for Java developers who are comfortable developing applications in Java, and now want to enter the world of data science or wish to build intelligent applications.

Aspiring data scientists with some understanding of the Java programming language will also find this book to be very helpful. If you are willing to build efficient data science applications and bring them in the enterprise environment without changing your existing Java stack, this book is for you!

What You Will Learn

Understand the key concepts of data science Explore the data science ecosystem available in Java Work with the Java APIs and techniques used to perform efficient data analysis Find out how to approach different machine learning problems with Java Process unstructured information such as natural language text or images, and create your own search Learn how to build deep

neural networks with DeepLearning4j Build data science applications that scale and process large amounts of data Deploy data science models to production and evaluate their performance In Detail Data science is concerned with extracting knowledge and insights from a wide variety of data sources to analyse patterns or predict future behaviour. It draws from a wide array of disciplines including statistics, computer science, mathematics, machine learning, and data mining. In this course, we cover the basic as well as advanced data science concepts and how they are implemented using the popular Java tools and libraries. The course starts with an introduction of data science, followed by the basic data science tasks of data collection, data cleaning, data analysis, and data visualization. This is followed by a discussion of statistical techniques and more advanced topics including machine learning, neural networks, and deep learning. You will examine the major

categories of data analysis including text, visual, and audio data, followed by a discussion of resources that support parallel implementation. Throughout this course, the chapters will illustrate a challenging data science problem, and then go on to present a comprehensive, Java-based solution to tackle that problem. You will cover a wide range of topics - from classification and regression, to dimensionality reduction and clustering, deep learning and working with Big Data. Finally, you will see the different ways to deploy the model and evaluate it in production settings. By the end of this course, you will be up and running with various facets of data science using Java, in no time at all. This course contains premium content from two of our recently published popular titles: *Java for Data Science Mastering Java for Data Science Style* and *approach*. This course follows a tutorial approach, providing examples of each of the concepts covered. With a step-

by-step instructional style, this book covers various facets of data science and will get you up and running quickly.

Beginning Programming with Python For Dummies -

John Paul Mueller 2014-09-22

A step-by-step guide to the programming language explains its strength as a Web development tool, covering syntax, data types, modules, lists, classes, and real world examples.

Numeric Computation and Statistical Analysis on the Java Platform -

Sergei V. Chekanov 2018-04-25

Numerical computation, knowledge discovery and statistical data analysis integrated with powerful 2D and 3D graphics for visualization are the key topics of this book. The Python code examples powered by the Java platform can easily be transformed to other programming languages, such as Java, Groovy, Ruby and BeanShell. This book equips the reader with a computational platform which, unlike other statistical

programs, is not limited by a single programming language. The author focuses on practical programming aspects and covers a broad range of topics, from basic introduction to the Python language on the Java platform (Jython), to descriptive statistics, symbolic calculations, neural networks, non-linear regression analysis and many other data-mining topics. He discusses how to find regularities in real-world data, how to classify data, and how to process data for knowledge discoveries. The code snippets are so short that they easily fit into single pages. Numeric Computation and Statistical Data Analysis on the

Java Platform is a great choice for those who want to learn how statistical data analysis can be done using popular programming languages, who want to integrate data analysis algorithms in full-scale applications, and deploy such calculations on the web pages or computational servers regardless of their operating system. It is an excellent reference for scientific computations to solve real-world problems using a comprehensive stack of open-source Java libraries included in the DataMelt (DMelt) project and will be appreciated by many data-analysis scientists, engineers and students.