Data Science Fondamentaux Et Tudes De Cas Machine Learning Avec Python Et R

Les data sciences en 100 questions/réponsesINIS AtomindexInformation Sur Les Sciences SocialesMemoirs of the Institute of Scientific and Industrial Research, Osaka UniversityCanadian Journal of Earth SciencesDynamics of Curved FrontsNew Serial TitlesData Science Using Python and RA Primer of Conservation GeneticsThe United States ConstitutionHealth Bulletin Trends in the Historiography of Science6èmes Journées d'études sur l'optique dans l'ordinateurBeginning Data Science in RReportCahiers du Centre d'études de recherche opérationnellee-Infrastructure and e-Services for Developing CountriesKey to Economic ScienceDavid HockneyData scienceBulletin Hydrographique InternationalA Brief Practical Guide to Eddy Covariance Flux MeasurementsModeling Discrete Time-to-Event DataEU Competition Law, Data Protection and Online Platforms: Data as Essential FacilityAquaNeural NetworksThe Forestry chronicleThe R BookData Science Projects with PythonRethinking IdiomaticityData science : fondamentaux et études de casApplied Natural Language Processing with PythonInternational Review of Administrative SciencesMastering Machine Learning with Python in Six StepsCanadian Special Publication of Fisheries and Aquatic SciencesInternational Encyclopedia of the Social SciencesPython for Data AnalysisBuild a Career in Data ScienceData Science for BusinessEuro-guide

Les data sciences en 100 questions/réponses

Summary You are going to need more than technical knowledge to succeed as a data scientist. Build a Career in Data Science teaches you what school leaves out, from how to land your first job to the lifecycle of a data science project, and even how to become a manager. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology What are the keys to a data scientist's long-term success? Blending your technical know-how with the right "soft skills" turns out to be a central ingredient of a rewarding career. About the book Build a Career in Data Science is your guide to landing your first data science job and developing into a valued senior employee. By following clear and simple instructions, you'll learn to craft an amazing resume and ace your interviews. In this demanding, rapidly changing field, it can be challenging to keep projects on track, adapt to company needs, and manage tricky stakeholders. You'll love the insights on how to handle expectations, deal with failures, and plan your career path in the stories from seasoned data scientists included in the book. What's inside Creating a portfolio of data science projects Assessing and negotiating an offer Leaving gracefully and moving up the ladder Interviews with professional data scientists About the reader For readers who want to begin or advance a data science career. About the author Emily Robinson is a data scientist at Warby Parker. Jacqueline Nolis is a data science consultant and mentor. Table of Contents: PART 1 - GETTING STARTED WITH DATA SCIENCE 1. What is data science? 2. Data science companies 3. Getting the skills 4. Building a portfolio PART 2 - FINDING YOUR DATA SCIENCE JOB 5. The search: Identifying the right job for you 6. The application: Résumés and cover letters 7. The interview: What to expect and how to handle it 8. The offer: Knowing what to accept PART 3 - SETTLING INTO DATA SCIENCE 9. The first months on the job 10. Making an

effective analysis 11. Deploying a model into production 12. Working with stakeholders PART 4 - GROWING IN YOUR DATA SCIENCE ROLE 13. When your data science project fails 14. Joining the data science community 15. Leaving your job gracefully 16. Moving up the ladder

INIS Atomindex

Information Sur Les Sciences Sociales

Memoirs of the Institute of Scientific and Industrial Research, Osaka University

Learn data science by doing data science! Data Science Using Python and R will get you plugged into the world's two most widespread open-source platforms for data science: Python and R. Data science is hot. Bloomberg called data scientist "the hottest job in America." Python and R are the top two open-source data science tools in the world. In Data Science Using Python and R, you will learn step-by-step how to produce hands-on solutions to real-world business problems, using state-of-the-art techniques. Data Science Using Python and R is written for the general reader with no previous analytics or programming experience. An entire chapter is dedicated to learning the basics of Python and R. Then, each chapter presents step-by-step instructions and walkthroughs for solving data science problems using Python and R. Those with analytics experience will appreciate having a one-stop shop for learning how to do data science using Python and R. Topics covered include data preparation, exploratory data analysis, preparing to model the data, decision trees, model evaluation, misclassification costs, naïve Bayes classification, neural networks, clustering, regression modeling, dimension reduction, and association rules mining. Further, exciting new topics such as random forests and general linear models are also included. The book emphasizes data-driven error costs to enhance profitability, which avoids the common pitfalls that may cost a company millions of dollars. Data Science Using Python and R provides exercises at the end of every chapter, totaling over 500 exercises in the book. Readers will therefore have plenty of opportunity to test their newfound data science skills and expertise. In the Hands-on Analysis exercises, readers are challenged to solve interesting business problems using real-world data sets.

Canadian Journal of Earth Sciences

Dynamics of Curved Fronts

New Serial Titles

Data Science Using Python and R

All are agreed that the digital economy contributes to a dynamic evolution of markets and competition. Nonetheless, concerns are increasingly raised about the market dominance of a few key players. Because these companies hold the power to drive rivals out of business, regulators have begun to seek scope for competition enforcement in cases where companies claim that withholding data is needed to satisfy customers and cut costs. This book is the first focus on how competition law enforcement tools can be applied to refusals of dominant firms to give access data on online platforms such as search engines, social networks, and e-commerce platforms – commonly referred to as the 'gatekeepers' of the Internet. The question arises whether the denial of a dominant firm to grant competitors access to its data could constitute a 'refusal to deal' and lead to competition law liability under the so-called 'essential facilities doctrine', according to which firms need access to shared knowledge in order to be able to compete. A possible duty to share data with rivals also brings to the forefront the interaction of competition law with data protection legislation considering that the required information may include personal data of individuals. Building on the refusal to deal concept, and using a multidisciplinary approach, the analysis covers such issues and topics as the following: – data portability; – interoperability; – data as a competitive advantage or entry barrier in digital markets; – market definition and dominance with respect to data; – disruptive versus sustaining innovation; – role of intellectual property regimes; – economic trade-off in essential facilities cases; – relationship of competition enforcement with data protection law and – data-related competition concerns in merger cases. The author draws on a wealth of relevant material, including EU and US decision-making practice, case law, and policy documents, as well as economic and empirical literature on the link between competition and innovation. The book concludes with a proposed framework for the application of the essential facilities doctrine to potential forms of abuse of dominance relating to data. In addition, it makes suggestions as to how data protection interests can be integrated into competition policy. An invaluable contribution to ongoing academic and policy discussions about how data-related competition concerns should be addressed under competition law, the analysis clearly demonstrates how existing competition tools for market definition and assessment of dominance can be applied to online platforms. It will be of immeasurable value to the many jurists, business persons, and academics concerned with this very timely subject.

A Primer of Conservation Genetics

A union list of serials commencing publication after Dec. 31, 1949.

The United States Constitution

Health Bulletin

Gain hands-on experience with industry-standard data analysis and machine learning tools in Python Key Features Learn techniques to use data to identify the exact problem to be solved Visualize data using different graphs Identify how to select an appropriate algorithm for data extraction Book Description Data Science Projects with Python is designed to give you practical guidance on industry-

standard data analysis and machine learning tools in Python, with the help of realistic data. The book will help you understand how you can use pandas and Matplotlib to critically examine a dataset with summary statistics and graphs, and extract the insights you seek to derive. You will continue to build on your knowledge as you learn how to prepare data and feed it to machine learning algorithms, such as regularized logistic regression and random forest, using the scikit-learn package. You'll discover how to tune the algorithms to provide the best predictions on new and, unseen data. As you delve into later chapters, you'll be able to understand the working and output of these algorithms and gain insight into not only the predictive capabilities of the models but also their reasons for making these predictions. By the end of this book, you will have the skills you need to confidently use various machine learning algorithms to perform detailed data analysis and extract meaningful insights from unstructured data. What you will learn Install the required packages to set up a data science coding environment Load data into a Jupyter Notebook running Python Use Matplotlib to create data visualizations Fit a model using scikit-learn Use lasso and ridge regression to reduce overfitting Fit and tune a random forest model and compare performance with logistic regression Create visuals using the output of the Jupyter Notebook Who this book is for If you are a data analyst, data scientist, or a business analyst who wants to get started with using Python and machine learning techniques to analyze data and predict outcomes, this book is for you. Basic knowledge of computer programming and data analytics is a must. Familiarity with mathematical concepts such as algebra and basic statistics will be useful.

Trends in the Historiography of Science

6èmes Journées d'études sur l'optique dans l'ordinateur

This concise, entry level text provides an introduction to the importance of genetic studies in conservation and presents the essentials of the discipline in an easy-to-follow format, with main points and terms clearly highlighted. The authors assume only a basic knowledge of Mendelian genetics and simple statistics, making the book accessible to those with a limited background in these areas. Connections between conservation genetics and the wider field of conservation biology are interwoven throughout the book. Worked examples are provided throughout to help illustrate key equations and glossary and suggestions for further reading provide additional support for the reader. Many beautiful pen and ink portraits of endangered species are included to enhance the text. Written for short, introductory level courses in genetics, conservation genetics and conservation biology, this book will also be suitable for practising conservation biologists, zoo biologists and wildlife managers.

Beginning Data Science in R

This book was written to familiarize beginners with general theoretical principles, requirements, applications, and processing steps of the Eddy Covariance method. It is intended to assist in further understanding the method, and provides references such as textbooks, network guidelines and journal papers. It is also

intended to help students and researchers in field deployment of instruments used with the Eddy Covariance method, and to promote its use beyond micrometeorology.

Report

Written by renowned data science experts Foster Provost and Tom Fawcett, Data Science for Business introduces the fundamental principles of data science, and walks you through the "data-analytic thinking" necessary for extracting useful knowledge and business value from the data you collect. This guide also helps you understand the many data-mining techniques in use today. Based on an MBA course Provost has taught at New York University over the past ten years, Data Science for Business provides examples of real-world business problems to illustrate these principles. You'll not only learn how to improve communication between business stakeholders and data scientists, but also how participate intelligently in your company's data science projects. You'll also discover how to think data-analytically, and fully appreciate how data science methods can support business decision-making. Understand how data science fits in your organization—and how you can use it for competitive advantage Treat data as a business asset that requires careful investment if you're to gain real value Approach business problems data-analytically, using the data-mining process to gather good data in the most appropriate way Learn general concepts for actually extracting knowledge from data Apply data science principles when interviewing data science job candidates

Cahiers du Centre d'études de recherche opérationnelle

e-Infrastructure and e-Services for Developing Countries

Key to Economic Science

Neural networks represent a powerful data processing technique that has reached maturity and broad application. When clearly understood and appropriately used, they are a mandatory component in the toolbox of any engineer who wants make the best use of the available data, in order to build models, make predictions, mine data, recognize shapes or signals, etc. Ranging from theoretical foundations to real-life applications, this book is intended to provide engineers and researchers with clear methodologies for taking advantage of neural networks in industrial, financial or banking applications, many instances of which are presented in the book. For the benefit of readers wishing to gain deeper knowledge of the topics, the book features appendices that provide theoretical details for greater insight, and algorithmic details for efficient programming and implementation. The chapters have been written by experts and edited to present a coherent and comprehensive, yet not redundant, practically oriented introduction.

David Hockney

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.

Data science

>

Bulletin Hydrographique International

A Brief Practical Guide to Eddy Covariance Flux Measurements

This book focuses on statistical methods for the analysis of discrete failure times. Failure time analysis is one of the most important fields in statistical research, with applications affecting a wide range of disciplines, in particular, demography, econometrics, epidemiology and clinical research. Although there are a large variety of statistical methods for failure time analysis, many techniques are designed for failure times that are measured on a continuous scale. In empirical studies, however, failure times are often discrete, either because they have been measured in intervals (e.g., quarterly or yearly) or because they have been rounded or grouped. The book covers well-established methods like life-table analysis and discrete hazard regression models, but also introduces state-of-the art techniques for model evaluation, nonparametric estimation and variable selection. Throughout, the methods are illustrated by real life applications, and relationships to survival analysis in continuous time are explained. Each section includes a set of exercises on the respective topics. Various functions and tools for the analysis of discrete survival data are collected in the R package discSurv that accompanies the book.

Modeling Discrete Time-to-Event Data

Formerly known as Yearbook of the European Communities, this guide is the only publication to bring together within a single volume nearly all of the intergovernmental bodies working for European integration.EuroGuide describes

the structure, function, and activities of institutions, bodies, and agencies of the European Union and other European intergovernmental organizations. The second part is a directory, giving the names of the persons in charge, with direct telephone numbers and addresses, and organizational charts of European intergovernmental organizations, the diplomatic corps accredited to the European Union, and those establishments offering European higher education courses.Other useful information contained in this guide include websites, national holidays, an overview of the history of the European Union, along with a subject index and alphabetical index.

EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility

Aqua

The articles in this volume have been first presented during an international Conference organised by the Greek Society for the History of Science and Technology in June 1990 at Corfu. The Society was founded in 1989 and planned to hold a series of meetings to impress upon an audience comprised mainly by Greek students and scholars, the point that history of science is an autonomous discipline with its own plurality of approaches developed over the years as a result of long discussions and disputes within the community of historians of science. The Conference took place at a time when more and more people came to realise that the future of the Greek Universities and Research Centres depends not only on the progress of the institutional reforms, but also very crucially on the establishment of new and modern subject areas. Though there have been significant steps towards such a direction in the physical sciences, mathematics and engineering, the situation in the so-called humanities has been, at best, confusing. Political expediencies of the post war years and ideological commitments to a glorious, yet very distant past, paralysed the development of the humanities and constrained them within a framework which could not allow much more than a philological approach.

Neural Networks

The Forestry chronicle

Un livre à la fois théorique et pratique Cet ouvrage a pour ambition de couvrir un large spectre du domaine des data sciences. Il va plus loin qu'un simple tour sur les algorithmes d'apprentissage automatique et s'attaque aux autres aspects, malheureusement négligés mais fondamentaux pour tout data scientist : concepts généraux mais poussés, dont la maîtrise est indispensable ; algorithmes d'apprentissage automatique les plus connus ; aspects liés à l'exploration des données ; mesures de performances et d'autres métriques utilisées par les algorithmes ; différents concepts fondamentaux en mathématiques à connaître pour mieux explorer et comprendre les données ; notions importantes des big data ; études de cas pratiques en langage Python. La première partie de ce livre est

théorique et adopte un format questions/réponses qui présente plusieurs avantages, dont la possibilité pour le lecteur de lire distinctement chacune des questions pour parfaire son savoir. Un autre point fort de cette structure est qu'elle incite à entrer dans un dialogue. Ainsi, grâce aux questions posées, le lecteur est poussé vers une réflexion où il confronte ses réponses à celles données par le livre. La deuxième partie est pratique et propose deux exemples d'implémentation de modèles d'apprentissage automatique. Vous y trouverez des codes écrits en Python et un aperçu de différentes difficultés que peut rencontrer un spécialiste lors de l'exercice de son métier. À qui s'adresse cet ouvrage ? Le présent ouvrage est adapté à toute personne ayant une certaine maîtrise de la data science et du Machine Learning. Il aidera notamment à se rappeler des concepts importants, mais suppose que le lecteur soit initié sur le sujet. Il sera particulièrement utile à ceux qui veulent se préparer pour un concours, un examen ou un entretien.

The R Book

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, natuarl 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.

Data Science Projects with Python

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

Rethinking Idiomaticity

Data science : fondamentaux et études de cas

In recent years, much progress has been made in the understanding of interface dynamics of various systems: hydrodynamics, crystal growth, chemical reactions, and combustion. Dynamics of Curved Fronts is an important contribution to this field and will be an indispensable reference work for researchers and graduate students in physics, applied mathematics, and chemical engineering. The book consist of a 100 page introduction by the editor and 33 seminal articles from various disciplines.

Applied Natural Language Processing with Python

International Review of Administrative Sciences

The high-level language of R is recognized as one of the most powerful and flexible statistical software environments, and is rapidly becoming the standard setting for quantitative analysis, statistics and graphics. R provides free access to unrivalled coverage and cutting-edge applications, enabling the user to apply numerous statistical methods ranging from simple regression to time series or multivariate analysis. Building on the success of the author's bestselling Statistics: An Introduction using R, The R Book is packed with worked examples, providing an all inclusive guide to R, ideal for novice and more accomplished users alike. The book assumes no background in statistics or computing and introduces the advantages of the R environment, detailing its applications in a wide range of disciplines. Provides the first comprehensive reference manual for the R language, including practical guidance and full coverage of the graphics facilities. Introduces all the statistical models covered by R, beginning with simple classical tests such as chi-square and t-test. Proceeds to examine more advance methods, from regression and analysis of variance, through to generalized linear models, generalized mixed models, time series, spatial statistics, multivariate statistics and much more. The R Book is aimed at undergraduates, postgraduates and professionals in science, engineering and medicine. It is also ideal for students and professionals in statistics, economics, geography and the social sciences.

Mastering Machine Learning with Python in Six Steps

This book presents an analysis of the key developments in Hockney's work over the past 30 years, and the chapters advance the critical debate around his work. He is the most widely discussed and well documented artist of his generation.

Canadian Special Publication of Fisheries and Aquatic Sciences

Learn to harness the power of AI for natural language processing, performing tasks such as spell check, text summarization, document classification, and natural language generation. Along the way, you will learn the skills to implement these methods in larger infrastructures to replace existing code or create new algorithms. Applied Natural Language Processing with Python starts with reviewing the necessary machine learning concepts before moving onto discussing various NLP problems. After reading this book, you will have the skills to apply these concepts in your own professional environment. You will: Utilize various machine learning and natural language processing libraries such as TensorFlow, Keras, NLTK, and Gensim Manipulate and preprocess raw text data in formats such as .txt and .pdf Strengthen your skills in data science by learning both the theory and the application of various algorithms.

International Encyclopedia of the Social Sciences

Python for Data Analysis

Nous vivons une époque très excitante, qui ramène l'analyse de données et les méthodes quantitatives au coeur de la société. L'aboutissement de nombreux projets de recherche, la puissance de calcul informatique disponible et des données à profusion permettent aujourd'hui d'incroyables réalisations, grâce au travail des data scientists. Un livre de référence pour les data scientists La data science est l'art de traduire des problèmes industriels, sociaux. scientifiques, ou de toute autre nature, en problèmes de modélisation quantitative, pouvant être résolus par des algorithmes de traitement de données. Cela passe par une réflexion structurée, devant faire en sorte que se rencontrent problèmes humains, outils techniques/informatiques et méthodes statistiques/algorithmiques. Chaque projet de data science est une petite aventure, qui nécessite de partir d'un problème opérationnel souvent flou, à une réponse formelle et précise, qui aura des conséquences réelles sur le quotidien d'un nombre plus ou moins important de personnes. Eric Biernat et Michel Lutz proposent de vous guider dans cette aventure. Ils vous feront visiter les vastes espaces de la data science modeme, de plus en plus présente dans notre société et qui fait tant parler d'elle, parfois par l'intermédiaire d'un sujet qui lui est corollaire. les big data. Des études de cas pour devenir kaggle master Loin des grands discours abstraits, les auteurs vous feront découvrir, claviers à la main, les pratiques de leur métier de data scientist chez OCTO Technology, l'un des leaders français du domaine. Et vous mettrez également la main à la pâte : avec juste ce qu'il faut de théorie pour comprendre ce qu'impliquent les méthodes mathématiques utilisées, mais surtout avec votre ordinateur personnel, quelques logiciels gratuits et puissants, ainsi qu'un peu de réflexion. vous allez participer activement à cette passionnante exploration !

Build a Career in Data Science

This book constitutes the thoroughly refereed proceedings of the 11th International Conference on e-Infrastructure and e-Services for Developing Countries, AFRICOMM 2019, held in Porto-Novo, Benin, in December 2019. The 19 full papers were carefully selected from 46 submissions. The accepted papers provide a wide range of research topics including targeted infrastructures, Internet of Things (IoT), wireless and mobile networks, intelligent transportation systems (ITS), software and network security, cloud and virtualization, data analytics, and machine learning.

Data Science for Business

Euro-guide

Covers each article and amendment of the Constitution in a graphic format designed to be relevant and accessible to modern readers.

ROMANCE ACTION & ADVENTURE MYSTERY & THRILLER BIOGRAPHIES & HISTORY CHILDREN'S YOUNG ADULT FANTASY HISTORICAL FICTION HORROR LITERARY FICTION NON-FICTION SCIENCE FICTION

ROMANCE ACTION & ADVENTURE MYSTERY & THRILLER BIOGRAPHIES & HISTORY CHILDREN'S YOUNG ADULT FANTASY HISTORICAL FICTION HORROR LITERARY FICTION NON-FICTION SCIENCE FICTION