

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

Un ouvrage de référence pour les (futurs) data scientists Les bibliothèques, les frameworks, les modules et les boîtes à outils sont parfaits pour faire de la data science. Ils sont aussi un bon moyen de plonger dans cette discipline sans la comprendre. Dans cet ouvrage, vous apprendrez comment fonctionnent les outils et algorithmes les plus fondamentaux de la data science, en les réalisant à partir de zéro. Si vous êtes fort en maths et que vous connaissez la programmation, l'auteur, Joel Grus, vous aidera à vous familiariser avec les maths et les statistiques qui sont au cœur de la data science et avec les compétences informatiques indispensables pour démarrer comme data scientist. La profusion des données d'aujourd'hui contient les réponses aux questions que personne n'a encore pensé à poser. Ce livre vous enseigne comment obtenir ces réponses. La deuxième édition, revue et augmentée, de cet ouvrage comporte des codes et exemples associés entièrement réécrits en Python 3.6 et intègre de nouveaux chapitres sur l'apprentissage profond (deep learning), les statistiques et le traitement en langage naturel. Suivez un cours accéléré de Python. Apprenez les fondamentaux de l'algèbre linéaire, des statistiques et des probabilités, et comprenez comment et quand les utiliser en data science. Collectez, explorez, nettoyez, bricollez et manipulez les données. Plongez dans les bases de l'apprentissage automatique. Implémentez des modèles comme les k plus proches voisins, la classification naïve bayésienne, les régressions linéaire ou logistique, les arbres de décision, les réseaux neuronaux et le clustering. Explorez les systèmes de recommandation, le traitement du langage naturel, l'analyse de réseau, MapReduce et les bases de données. À qui s'adresse cet ouvrage ? Aux développeurs, statisticiens, étudiants et chefs de projet ayant à résoudre des problèmes de data science. Aux data scientists, mais aussi à toute personne curieuse d'avoir une vue d'ensemble de l'état de l'art de ce métier du futur.

La bonne qualité des données est aujourd'hui la clé de voûte de toute organisation. La gestion et l'amélioration de cette qualité sont des tâches coûteuses et difficiles, mais néanmoins incontournables. Cet ouvrage propose une étude des différents outils et démarches qui assistent les spécialistes de la qualité et de la gouvernance des données. À travers les expériences de la communauté francophone animée par l'association EXQL (Excellence Qualité, Information), il présente, avec pédagogie et pragmatisme, un panorama des concepts-clés de la gestion de la qualité des données et leurs déclinaisons dans les entreprises (Business Intelligence, Data QualityManagement, Key Performance Indicator, Model Driven Engineering, Master Data Management, etc.). Des solutions théoriques et techniques performantes sont détaillées et de nombreux retours d'expérience permettent d'illustrer les bonnes pratiques à adopter. Mélangant contributions industrielles et académiques, cet ouvrage est un outil de référence en langue française sur la qualité et la gouvernance des données en entreprise.

Une étude sur la légitimité et l'autorité de l'information produite numériquement. Les vingt-cinq dernières années ont vu se remodeler les dispositifs infocommunicationnels visant à concevoir des connaissances, à les organiser et à les rendre publiques, en bouleversant les cadres éditoriaux traditionnels jusqu'à faire primer, dans la société de l'information, la raison computationnelle sur la raison humaine. Cette transformation se caractérise par de nouveaux agencements entre documents, données et connaissances et par l'apparition de nouvelles techniques d'écriture au service de l'éditorialisation. Des pratiques érudites d'écriture par les humanités numériques aux autorités calculées de l'IA, les fondamentaux de l'autorité numérique sont mis en perspective dans le champ des sciences de l'information et de la communication. L'objectif de ce livre est de fournir des clés de compréhension des phénomènes majeurs liés à l'évolution des supports numériques d'inscription. Il propose un modèle d'évaluation de l'autorité dans les dispositifs info-communicationnels. Pour les professionnels de l'information ; pour les chercheurs, enseignants et étudiants en information, communication, documentation et humanités numériques.

Machine Learning with R

Big Data Analysis and Deep Learning Applications

Mathematical Contributions to the Theory of Evolution

A Global Geochemical Database for Environmental and Resource Management

A Cumulative Author List Representing Library of Congress Printed Cards and Titles Reported by Other American Libraries

Data Science : fondamentaux et études de cas

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 puissants, 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 différents rentes difficiles 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.

Contains research project reports arranged by subject with descriptors from the EUDISED Multilingual Thesaurus.

This book helps readers understand the mathematics of machine learning, and apply them in different situations. It is divided into two basic parts, the first of which introduces readers to the theory of linear algebra, probability, and data distributions and it's applications to machine learning. It also includes a detailed introduction to the concepts and constraints of machine learning and what is involved in designing a learning algorithm. This part helps readers understand the mathematical and statistical aspects of machine learning. In turn, the second part discusses the algorithms used in supervised and unsupervised learning. It works out each learning algorithm mathematically and encodes it in R to produce customized learning applications. In the process, it touches upon the specifics of each algorithm and the science behind its formulation. The book includes a wealth of worked-out examples along with R codes. It explains the code for each algorithm, and readers can modify the code to suit their own needs. The book will be of interest to all researchers who intend to use R for machine learning, and those who are interested in the practical aspects of implementing learning algorithms for data analysis. Further, it will be particularly useful and informative for anyone who has struggled to relate the concepts of mathematics and statistics to machine learning.

Money and Banking in Africa

11th EAI International Conference, AFRICOMM 2019, Porto-Novo, Benin, December 3–4, 2019, Proceedings

Learning Data Mining with Python

Machine Learning avec Python et R

International Encyclopedia of the Social Sciences

Information Sur Les Sciences Sociales

Providing code examples in python, this book introduces the concepts of machine learning with mathematical explanations and programming fundamentals. --

Early in the last century the successes of science led a group of French thinkers to apply the principles of science to the study of society. These thinkers purported to have discovered the supposed 'laws' of society and concluded that an elite of social scientists should assume direct control of social life. The Counter-Revolution of Science is Nobel Laureate Friedrich Hayek's forceful attack on this abuse of reason.

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 advanced 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.

Data science

Recommendations for International Geochemical Mapping : Final Report of IGCP Project 259

The Counter-revolution of Science

Machine Learning with Python

Parallel Programming with Python

Les data sciences en 100 questions-réponses

This book introduces machine learning in finance and illustrates how we can use computational tools in numerical finance in real-world context. These computational techniques are particularly useful in financial risk management, corporate bankruptcy prediction, stock price prediction, and portfolio management. The book also offers practical and managerial implications of financial and managerial decision support systems and how these systems capture vast amount of financial data. Business risk and uncertainty are two of the toughest challenges in the financial industry. This book will be a useful guide to the use of machine learning in forecasting, modeling, trading, risk management, economics, credit risk, and portfolio management.

Nous vivons une époque très excitante, qui ramène l'analyse de données et les méthodes quantitatives au cœur 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, évitant faire en sorte que se rencontrent problèmes humains, outils techniques/informatiques et méthodes statistiques/algoritmiques. 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. Éric Biernat et Michel Lutz proposent de vous guider dans cette aventure. Ils vous feront visiter les vastes espaces de la data science moderne, 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 !

À qui s'adresse cet ouvrage ? Aux développeurs, statisticiens, étudiants et chefs de projets ayant à résoudre des problèmes de data science. Aux data scientists, mais aussi à toute personne curieuse d'avoir une vue d'ensemble de l'état de l'art du machine learning.

The Yearbook of International Organizations provides the most extensive coverage of non-profit international organizations currently available. Detailed profiles of international non-governmental (NGO) and intergovernmental organizations (IGO), collected and documented by the Union of International Associations, can be found here. In addition to the history, aims and activities of international organizations, with their events, publications, and contact details, the volumes of the Yearbook include networks between associations, biographies of key people involved and extensive statistical data. Volume 1 (A and B) covers international organizations throughout the world, comprising their aims, activities and events. This includes names (in English, French and, where available, other languages), abbreviations and descriptions of over 34,000 not-for-profit organizations currently active in every field of human endeavor, as well as references to associated organizations, whose goals cross all economic, political and geographical borders, offering an insight into new, productive relationships. Volume 1 also allows quick and easy cross-referencing from volumes 2, 3, 4, and 6.

R pour les data sciences

Reprints

Data science par la pratique

Algorithms, Worked Examples, and Case Studies

Fondamentaux avec Python - Préparez-vous aux métiers du futur !

Beginning Data Science in R

Teaching text developed by U.S. Air Force Academy and designed as a first course emphasizes the universal variable formulation. Develops the basic two-body and n-body equations of motion; orbit determination; classical orbital elements; coordinate transformations; differential correction; more. Includes specialized applications to lunar and interplanetary flight, example problems, exercises. 1971 edition.

A fast, easy-to-follow and clear tutorial to help you develop Parallel computing systems using Python. Along with explaining the fundamentals, the book will also introduce you to slightly advanced concepts and will help you in implementing these techniques in the real world. If you are an experienced Python programmer and are willing to utilize the available computing resources by parallelizing applications in a simple way, then this book is for you. You are required to have a basic knowledge of Python development to get the most of this book.

The novel coronavirus SARS-CoV-2, which causes the disease known as COVID-19, has infected people in 212 countries so far and on every continent except Antarctica. Vast changes to our home lives, social interactions, government functioning and relations between countries have swept the world in a few months and are difficult to hold in one's mind at one time. That is why a collaborative effort such as this edited, multidisciplinary collection is needed. This book confronts the vulnerabilities and interconnectedness made visible by the pandemic and its consequences, along with the legal, ethical and policy responses. These include vulnerabilities for people who have been harmed or will be harmed by the virus directly and those harmed by measures taken to slow its relentless march; vulnerabilities exposed in our institutions, governance and legal structures; and vulnerabilities in other countries and at the global level where persistent injustices harm us all. Hopefully, COVID-19 will force us to deeply reflect on how we govern and our policy priorities; to focus preparedness, precaution, and recovery to include all, not just some. Published in English with some chapters in French.

e-Infrastructure and e-Services for Developing Countries

Studies on the Abuse of Reason

Bulletin Hydrographique International

La qualité et la gouvernance des données : au service de la performance des entreprises

Editorialisation et autorité

Report

The Hands-On, Example-Rich Introduction to Pandas Data Analysis in Python Today, analysts must manage data characterized by extraordinary variety, velocity, and volume. Using the open source Pandas library, you can use Python to rapidly automate and perform virtually any data analysis task, no matter how large or complex. Pandas can help you ensure the veracity of your data, visualize it for effective decision-making, and reliably reproduce analyses across multiple datasets. Pandas for Everyone brings together practical knowledge and insight for solving real problems with Pandas, even if you're new to Python data analysis. Daniel Y. Chen introduces key concepts through simple but practical examples, incrementally building on them to solve more difficult, real-world problems. Chen gives you a jumpstart on using Pandas with a realistic dataset and covers combining datasets, handling missing data, and structuring datasets for easier analysis and visualization. He demonstrates powerful data cleaning techniques, from basic string manipulation to applying functions simultaneously across dataframes. Once your data is ready, Chen guides you through fitting models for prediction, clustering, inference, and exploration. He provides tips on performance and scalability, and introduces you to the wider Python data analysis ecosystem. Work with DataFrames and Series, and import or export data Create plots with matplotlib, seaborn, and pandas Combine datasets and handle missing data Reshape, tidy, and clean datasets so they're easier to work with Convert data types and manipulate text strings Apply functions to scale data manipulations Aggregate, transform, and filter large datasets with groupby Leverage Pandas' advanced date and time capabilities Fit linear models using statsmodels and scikit-learn libraries Use generalized linear modeling to fit models with different response variables Compare multiple models to select the "best" Regularize to overcome overfitting and improve performance Use clustering in unsupervised machine learning

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.

This book stems from the multi-stage International Geochemical Mapping (IGM), an International Geological Correlation Programme (IGCP) project, to set up a global geochemical database on the distribution and quantities present of all 92 chemical elements in the surface of the earth. A comprehensive review and evaluation of methods for regional and national geochemical mapping and providing a recognized, global quantitative base on which local investigations can be built for particular environmental and economic problems concerning various aspects of land use.

The Essentials of Machine Learning in Finance and Accounting

The Forestry chronicle

Eudised R & D Bulletin

Data Analysis, Visualization, and Modelling for the Data Scientist

Cahiers du Centre d'études de recherche opérationnelle

Les data sciences en 100 questions/réponses

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. 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.

This book presents a holistic exploration of the banking systems in Africa. Considering the central role that banks play in most developing countries and the vastly different trends and challenges they face, the book provides a crucial understanding of the specific environments in which banks operate. It addresses specific banking issues relevant to developing countries in general and Africa in particular, and explores the various dynamics of money and banking that separate Africa from the rest of the world. The authors build upon extensive Africa-based research and university teaching, and illustrate each topic with examples and cases from the continent. Written in an accessible style while retaining its practicality and relevance, it is an essential read for professionals, students, and other readers interested in policies affecting the banking sector's development in Africa.

Python Data Analysis

Health Bulletin ...

Vulnerable

Key to Economic Science

Importer, classer, transformer, visualiser et modéliser les données

National Union Catalog

The next step in the information age is to gain insights from the deluge of data coming our way. Data mining provides a way of finding this insight, and Python is one of the most popular languages for data mining, providing both power and flexibility in analysis. This book teaches you to design and develop data mining applications using a variety of datasets, starting with basic classification and affinity analysis. Next, we move on to more complex data types including text, images, and graphs. In every chapter, we create models that solve real-world problems. There is a rich and varied set of libraries available in Python for data mining. This book covers a large number, including the IPython Notebook, pandas, scikit-learn and NLTK. Each chapter of this book introduces you to new algorithms and techniques. By the end of the book, you will gain a large insight into using Python for data mining, with a good knowledge and understanding of the algorithms and implementations.

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

Includes entries for maps and atlases.

Proceedings of the First International Conference on Big Data Analysis and Deep Learning

A Practical Implementation Guide to Predictive Data Analytics Using Python

Mastering Machine Learning with Python in Six Steps

fondamentaux et études de cas : machine learning avec Python et R

Yearbook of International Organizations 2014-2015, Volumes 1a & 1b (Set)

Fundamentals of Astrodynamics

The second edition of a comprehensive introduction to machine learning approaches used in predictive data analytics, covering both theory and practice. Machine learning is often used to build predictive models by extracting patterns from large datasets. These models are used in predictive data analytics applications including price prediction, risk assessment, predicting customer behavior, and document classification. This introductory textbook offers a detailed and focused treatment of the most important machine learning approaches used in predictive data analytics, covering both theoretical concepts and practical applications. Technical and mathematical material is augmented with explanatory worked examples, and case studies illustrate the application of these models in the broader business context. This second edition covers recent developments in machine learning, especially in a new chapter on deep learning, and two new chapters that go beyond predictive analytics to cover unsupervised learning and reinforcement learning.

Les data sciences constituent une discipline fascinante, qui vous servira à transformer des données brutes en idées et connaissances aisément compréhensibles. L'objectif de ce livre est de vous aider à maîtriser les outils essentiels qui vous permettront d'utiliser R dans la pratique des data sciences. Après l'avoir lu, vous pourrez faire face à la plupart des situations que vous rencontrerez dans vos projets, en exploitant au mieux les fonctionnalités de R. Mais avant de devenir un expert en R, vous devrez tout d'abord importer vos données, c'est-à-dire les lire, en général depuis un fichier, une base de données ou une API web, et les charger dans un cadre de données dans R. Si vous ne pouvez pas transférer vos données dans R, vous ne pourrez pas les analyser ! Une fois vos données importées, vous gagnerez beaucoup à les ranger. Une fois vos données rangées, vous passerez bien souvent par une étape de transformation. Une fois vos données rangées avec les variables dont vous avez besoin, la génération de connaissances reposera principalement sur deux moteurs : la visualisation et la modélisation. Leurs forces et faiblesses sont complémentaires, et toute bonne analyse doit tenir compte des deux. Inutile d'être un programmeur expert pour être un bon analyste de données, mais apprendre à programmer This book presents a compilation of selected papers from the first International Conference on Big Data Analysis and Deep Learning Applications (ICBDL 2018), and focuses on novel techniques in the fields of big data analysis, machine learning, system monitoring, image processing, conventional neural networks, communication, industrial information, and their applications. Readers will find insights to help them realize more efficient algorithms and systems used in real-life applications and contexts, making the book an essential reference guide for academic researchers, professionals, software engineers in the industry, and regulators of aviation authorities.

The Law, Policy and Ethics of COVID-19

International Review of Administrative Sciences

+ 2 études de cas commentées

The R Book

The National Union Catalogs, 1963-

Fundamentals of Machine Learning for Predictive Data Analytics, second edition