Aris Gkoulalas-Divanis Abderrahim Labbi Editors

Large-Scale Data Analytics



Aris Gkoulalas-Divanis and

Large-Scale Data Analytics



This edited book collects state-of-the-art research related to large-scale data analytics that has been accomplished over the last few years. This is among the first books specialized in this important area based on contributions from diverse scientific areas such as for example databases, data mining, supercomputing, equipment architecture, data visualization, figures, and privacy. This requires new distributed architectures for data analysis. There is increasing need for new approaches and technologies that can analyze and synthesize large amounts of data, in the order of petabytes, that are generated by massively distributed data resources. Additionally, the heterogeneity of such resources imposes significant issues for the efficient analysis of the info under many constraints, including consistent data integration, data homogenization and scaling, privacy and security preservation. The authors also broaden reader understanding of emerging real-world applications in domains such as for example client behavior modeling, graph mining, telecommunications, cyber-protection, and social network analysis, which impose extra requirements for large-scale data evaluation.Large-Scale Data Analytics is usually structured in 8 chapters, each providing a survey of a significant direction of large-scale data analytics or specific outcomes of the emerging research in the field. Learners, researchers, experts and practitioners will find this reserve an authoritative and comprehensive resource. The reserve presents key recent study that will help shape the future of large-scale data analytics, leading the way to the look of new approaches and technologies that can analyze and synthesize very large amounts of heterogeneous data.



continue reading

A beautiful map if you want to understand the landscape If you work with expensive enterprise strength data management/analysis items like SAS and Oracle and you want a book that will offer you a map to cover the open up source tools for dealing with "big data" (i.e. One Star By far

the most disheveled and disconnected booklet about technology ever written. From the practical issues in designing internet facing applications to analytic data-sets are protected at the perfect depth for someone who knows a little bit about data and databases. It can an amazingly good work of explaining the utility of the equipment that are used to manage *HUGE* data. Even though you are not a programmer, the writer does a fantastic job of explaining issues from the bottom up without babying the reader (e.g., what are the advantages of using CSV files vs XML vs JSON vs Thrift vs Avro). There are code snippets scattered during that are useful for comparing and contrasting if you know some programming languages (e. A wide book with the right depth to go from 0 to 100 (petabytes) Hive, Hadoop, Shark, Dremel, BigQuery, SciPy, NumPy, Pandas, R, Pig., SQL queries vs HiveQL) but the book does not really attempt to describe the code in great fine detail. So, you wind up with the outline of what a device does without getting bogged down in the gory information. If you would like to go deeper in to the solutions the reserve is filled with references to seminal white papers and various other external references so you can increase on what's covered.So, if you keep hearing approximately things like Hadoop, noSQL, Python, SciPy, Pandas, R and you just want to understand "what is the big offer" or "as to why bother" learning just one more tool, this is actually the perfect book.g. JSON over CSV, with good explanation of why.. whether you are new or a seasoned big data expert, there exists a big and growing universe of keywords to understand. So, if you prefer a very quick summary of what data technology is, that is an easy read and provides you merely that, but if you'd like anything deeper from it, I think this book is somewhat disappointing. I love that he's not afraid to leap and write code, as - when you do it just right - a few lines of code are a lot more illustrative than a picture or block of texts would perform. I think the author has tried to please both more technically inclined and others at the same time, which hasn't actually worked.You'll get valuable knowledge of when to use ex. But if you need to understand the current big data universe, the way the equipment interrelate between each other, and go from data generation to storage to analysis to visualization - this is actually the book. Great "birdperspective" on data Great book for an overview in data, collecting of data, data tools and documents.Wanting to learn whats up and down in the wonderful world of Big Data was achieved by reading this book. If you want to learn Hadoop, purchase a Hadoop publication - or an R book if you want to go deeper for the reason that topic... I recommend this book for those who have little if any experience on how data could be stored, analyzed and visualized. Five Stars very informative Good read Well written In the end I have neither the impression I have a good overview of the various tools available (at least This book has an interesting summary of main technologies in data science, but strikes a slightly odd balance between technical and descriptive -- there are several brief code examples that may get you on the way or that provide you the feeling of the functionality of the particular tool, but it remains very superficial. Ultimately I've neither the impression I've a good summary of the tools obtainable (at least, not really beyond what I already had), nor do I know much at length about each of them. Most items are explained in as well simple vocabulary, using analogies where technical detail could have been more interesting. It's also slightly repetitive sometimes. Totally recommended. In this publication Manoochehri manages to provide a through review on the whys and hows, giving the reader the ideal depth in each topic to understand the motivation for every of these different technology, how they are different to each other, and why you would want to utilize them., Hadoop, Hive, and Pig) get this.

download free Large-Scale Data Analytics e-book

download Large-Scale Data Analytics pdf

download Team Cooperation in a Network of Multi-Vehicle Unmanned Systems: Synthesis of Consensus Algorithms ebook download free Resource Allocation in Decentralized Systems with Strategic Agents: An Implementation Theory Approach (Springer Theses) mobi download free Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints (Systems & Control: Foundations & Applications) txt