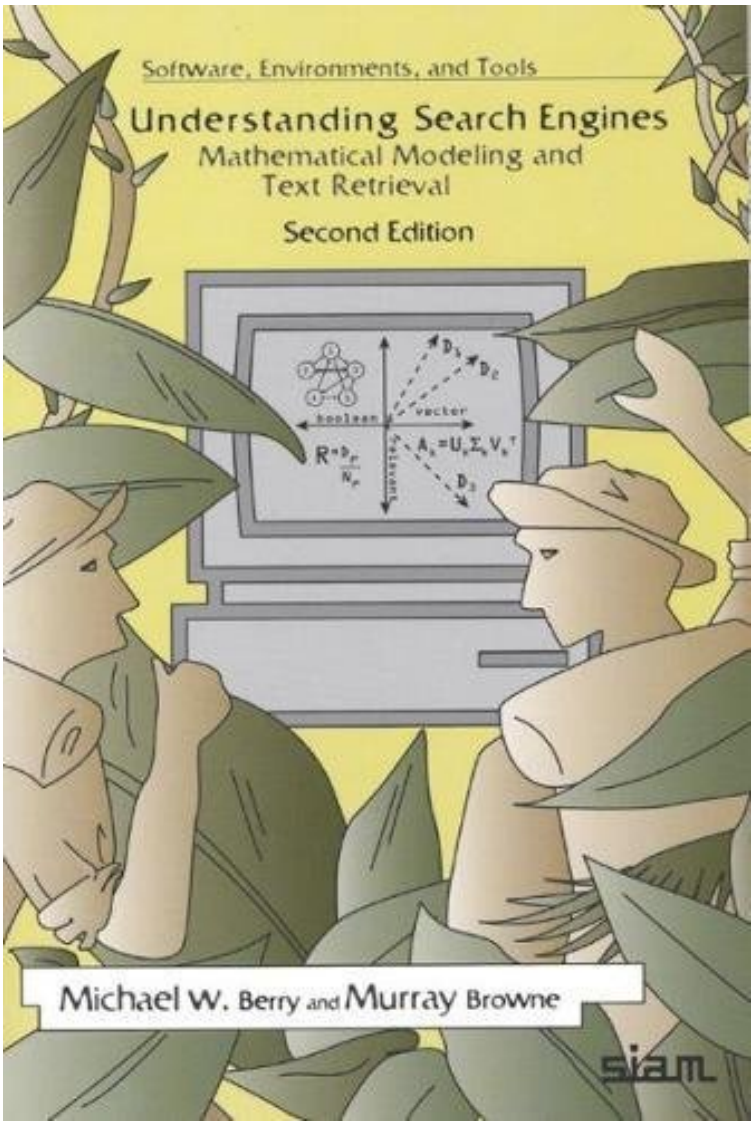
Understanding Search Engines: Mathematical Modeling and Text Retrieval (Software, Environments, Tools), Second Edition

Michael W. Berry, Murray Browne
*ePub | *DOC | audiobook | ebooks | Download PDF*



#1648528 in Books SIAM, Society for Industrial and Applied Mathematics 2005-04-28 2005-04-28Original language:EnglishPDF # 1 8.98 x .35 x 5.98l, .50 #File Name: 0898715814117 pages | File size: 72.Mb

Michael W. Berry, Murray Browne : Understanding Search Engines: Mathematical Modeling and Text Retrieval (Software, Environments, Tools), Second Edition before purchasing it in order to gage whether or not it would be worth my time, and all praised Understanding Search Engines: Mathematical Modeling and Text Retrieval (Software, Environments, Tools), Second Edition:

8 of 8 people found the following review helpful. Good IntroductionBy Geraldo XEXEOThere are better books in the market, and even the author would be the first to recognize it. However, this book is one of the most clear and readable introduction to the subject that you can find.The author fully acomplishes the objective: teach his reader, at undergratuate level, how search engines work. Even some difficult subject, such as LSI, are treated at a level one can easilly understand.One of the most important characteristics of the book is that it does math. Every formula has an example, usually using small matrix that allow the reader to easilly follow them.The book is suitable for an objective introduction to the field. It is not very "academic", in the sense it is rather informal. If it is not a textbook, it could help some bewildered student to grasp the inner workings. It could also help a teacher to find clearer ways for explanations and good examples for classroom.2 of 2 people found the following review helpful. Linear Algebra, Numerical Linear Algebra, and Search EngineBy Man Kam TamOther than showing the readers how to design a search engine, the authors, Michael W. Berry and Murray Browne of "Understanding Search Engines: Mathematical Modeling, and Text Retrieval," intend to fill the gap between applied mathematics and information management. In a latent semantic index (LSI) system, mathematics plays a major role in search engine performance. The term-by-document matrix of the system would be transformed to a lower rank matrix for conceptual indexing. However, nobody knows how low the rank should be for the best performance. The best technique so far for lower rank approximation is called singular value decomposition. In such a system, vectors model both documents and queries. The angle between the document vector and the query vector determines the rank-order of the document. The elements of the vectors are usually the weighted frequency of the term occurrence. Thus the searchers should list as many terms as possible in their queries for better search results. LSI search engine is good for small document system only. Other searching methods such as HITS and PageRank are introduced. For the readers who have the background on linear algebra, numerical linear algebra, and search engine should find this book interesting. Generally speaking, the book is brief. It has 117 pages and 9 chapters. The nine chapters are Introduction, Document File Preparation, Vector Space Models, Matrix Decompositions, Query Management, Ranking and Relevance Feedback, Searching by Link Structure, User Interface Considerations, and Further Reading. Chapter two (Document File Preparation) reminds the readers that the documents of the system needed to be "clean-up" and index. The works may require plenty of manual labor.3 of 3 people found the following review helpful. A mix of good and badBy RayAs others have pointed out, this book is very short. As a consequence, it leaves out a lot of details and forces the reader to refer to another book. This is more noticeable in the sections that do not relate to linear algebra (stemming, performance evaluation, and user interface design). If you want more information about these topics, it is best to look for another book.However, the discussions about latent semantic indexing and querying based on link structure are more detailed in comparison and both topics are mentioned within the context of linear algebra.Don't expect an introduction to QR or SVD matrix decompositions or what an eigenspace is. Also, don't expect a proper definition of what a graph is. For all of this, you will also have to refer to another book. If you do not need such an introduction, then you may not mind.Overall, the book attempts to do too many topics in few pages and suffers from this. However, if you are looking for a "crash course in search engines"-type book, then this might be the one for you. You may end up buying another book afterwards if you want to know implementation details, though.

There is no other information retrieval/search book where the heart is the mathematical foundations. This book is greatly needed to further establish information retrieval as a serious academic, as well as practical and industrial, area." ---Jaime Carbonell, Carnegie Mellon University. Berry and Browne describe most of what you need to know to design your own search engine. Their strength is the description of the solid mathematical underpinnings at a level that is understandable to competent engineering undergraduates, perhaps with a bit of instructor guidance. They discuss the algorithms used by most commercial search engines, so you may find your use of Google and its kind becomes more effective, too. --George Corliss, Marquette University. This book gives a valuable, generally non-technical, insight into how search engines work, how to improve the users' success in Information Retrieval (IR), and an in-depth analysis of a mathematical algorithm for improving a search engine's performance. Written in an informal style, the book is easy to read and is a good introduction on how search engines operate Christopher Dean, Mathematics Today, October 1999. The second edition of Understanding Search Engines: Mathematical Modeling and Text Retrieval follows the basic premise of the first edition by discussing many of the key design issues for building search engines and emphasizing the important role that applied mathematics can play in improving information retrieval. The authors discuss important data structures, algorithms, and software as well as user-centered issues such as interfaces, manual indexing, and document preparation. Readers will find that the second edition includes significant changes that bring the text up to date on current information retrieval methods. For example, the authors have added a completely new chapter on link-structure algorithms used in search engines such as Google, and the chapter on user interface has been rewritten to specifically focus on search engine usability. To reflect updates in the literature on information retrieval, the authors have added new recommendations for further reading and expanded the bibliography. In addition, the index has been updated and streamlined to make it more reader friendly. Instructors will find that the book serves as an excellent companion text for courses in information retrieval, applied linear algebra, and scientific computing.

Because of the authors informal, conversational tone, readers with nonmathematical backgrounds also will appreciate the less technical chapters of the text.

'There is no other information retrieval/search book where the heart is the mathematical foundations. This book is greatly needed to further establish information retrieval as a serious academic, as well as practical and industrial, area.' Jaime Carbonell, Carnegie Mellon University'Berry and Browne describe most of what you need to know to design your own search engine. Their strength is the description of the solid mathematical underpinnings at a level that is understandable to competent engineering undergraduates, perhaps with a bit of instructor guidance. They discuss the algorithms used by most commercial search engines, so you may find your use of Google and its kind becomes more effective, too.' George Corliss, Marquette University.'This book gives a valuable, generally non-technical, insight into how search engines work, how to improve the users' success in Information Retrieval (IR), and an in-depth analysis of a mathematical algorithm for improving a search engine's performance. ...Written in an informal style, the book is easy to read and is a good introduction on how search engines operate...' Christopher Dean, Mathematics Today'Anyone interested in building their own search engine, or looking for a compact and readable introduction to the field of modern information retrieval will find this book to be an excellent first introduction.' Tony Donaldson, MAA sAbout the AuthorMichael W. Berry is a Professor and Interim Department Head in the Department of Computer Science at the University of Tennessee and a faculty member in the Graduate School in Genome Science Technology Program at the University of Tennessee and Oak Ridge National Laboratory. His current research interests include information retrieval, data mining, scientific computing, computational science, numerical linear algebra, and parallel computation. Murray Browne is a Research Associate in the Computer Science Department at the University of Tennessee. He is a member of the American Society for Information Science and Technology and has published numerous essays, book reviews, newspaper articles, and feature stories.