

|                         |   |
|-------------------------|---|
| 1. Record Nr.           | UNINA9910163287003321   |
| Autore                  | Kraft Donald H.   |
| Titolo                  | Fuzzy information retrieval // Donald H. Kraft, Erin Colvin   |
| Pubbl/distr/stampa      | [San Rafael, California] : , : Morgan & Claypool, , 2017  |
| ISBN                    | 1-62705-956-3   |
| Descrizione fisica      | 1 online resource (83 pages) : illustrations  |
| Collana                 | Synthesis lectures on information concepts, retrieval, and services, , 1947-9468 ; ; # 55   |
| Disciplina              | 511.322   |
| Soggetti                | Fuzzy systems<br>Information retrieval - Automation   |
| Lingua di pubblicazione | Inglese   |
| Formato                 | Materiale a stampa  |
| Livello bibliografico   | Monografia  |
| Note generali           | Part of: Synthesis digital library of engineering and computer science.   |
| Nota di bibliografia    | Includes bibliographical references (pages 57-62).  |
| Nota di contenuto       | 1. Introduction to information retrieval -- 1.1 Defining information retrieval -- 1.1.1 Retrieval metric -- 1.2 A bit of IR history -- 1.3 The key notion of relevance -- 1.4 Some interesting applications of information retrieval -- 1.5 Where the fuzz is -- 2. Modeling -- 2.1 Boolean model -- 2.1.1 Deficiencies with Boolean logic -- 2.2 Vector space model -- 2.3 Probability model -- 2.3.1 Language models -- 2.3.2 Alternative probabilistic best match models -- 2.4 Modern concepts of IR -- 2.5 Fuzzy logic and sets -- 2.6 Membership functions -- 2.6.1 Intuition -- 2.6.2 Inference -- 2.6.3 Rank ordering -- 2.6.4 Neural networks -- 2.6.5 Genetic algorithms -- 2.6.6 Inductive reasoning -- 2.6.7 Aggregation -- 2.7 Extended boolean in IR -- 2.8 Fuzzy evaluation metrics -- 2.8.1 Map -- 2.9 Discounted cumulative gain -- 2.10 Summary -- 3. Source of weights -- 3.1 Indexing -- 3.1.1 TF-IDF -- 3.2 Variants -- 3.2.1 Document length normalization -- 3.3 Querying -- 3.4 Summary -- 4. Relevance feedback and query expansion -- 4.1 Defining relevance feedback -- 4.2 Pseudo-relevant feedback -- 4.3 Relevance feedback with the vector space model -- 4.4 Relevance feedback with the probability model -- 4.5 Relevance feedback with the boolean and fuzzy boolean models -- 4.5.1 Genetic algorithms (programming) for Boolean relevance feedback -- 4.6 Query expansion -- 4.6.1 Adding or changing terms -- 4.7 Fuzzy and rough sets for data mining of a controlled vocabulary -- 4.7.1 Fuzzy set notation for rough sets -- 4.8 |

Extensions of the rough set approach to retrieval -- 4.8.1 Rough fuzzy sets -- 4.8.2 Fuzzy rough sets -- 4.8.3 Generalized fuzzy and rough sets -- 4.8.4 Nonequivalence relationships -- 4.8.5 Combining several relationships simultaneously --  
5. Clustering for retrieval -- 5.1 Introduction -- 5.2 Applications -- 5.3 Clustering algorithms -- 5.4 Similarity measures -- 5.5 Fuzzy clustering -- 5.6 The fuzzy C-means algorithm -- 5.7 A testing example -- 5.7.1 Fuzzy rule discovery --  
6. Uses of information retrieval today -- Bibliography -- Author biographies.

---

## Sommario/riassunto

Information retrieval used to mean looking through thousands of strings of texts to find words or symbols that matched a user's query. Today, there are many models that help index and search more effectively so retrieval takes a lot less time. Information retrieval (IR) is often seen as a subfield of computer science and shares some modeling, applications, storage applications and techniques, as do other disciplines like artificial intelligence, database management, and parallel computing. This book introduces the topic of IR and how it differs from other computer science disciplines. A discussion of the history of modern IR is briefly presented, and the notation of IR as used in this book is defined. The complex notation of relevance is discussed. Some applications of IR is noted as well since IR has many practical uses today. Using information retrieval with fuzzy logic to search for software terms can help find software components and ultimately help increase the reuse of software. This is just one practical application of IR that is covered in this book. Some of the classical models of IR is presented as a contrast to extending the Boolean model. This includes a brief mention of the source of weights for the various models. In a typical retrieval environment, answers are either yes or no, i.e., on or off. On the other hand, fuzzy logic can bring in a "degree of " match, vs. a crisp, i.e., strict match. This, too, is looked at and explored in much detail, showing how it can be applied to information retrieval. Fuzzy logic is often times considered a soft computing application and this book explores how IR with fuzzy logic and its membership functions as weights can help indexing, querying, and matching. Since fuzzy set theory and logic is explored in IR systems, the explanation of where the fuzz is ensues. The concept of relevance feedback, including pseudorelevance feedback is explored for the various models of IR. For the extended Boolean model, the use of genetic algorithms for relevance feedback is delved into. The concept of query expansion is explored using rough set theory. Various term relationships is modeled and presented, and the model extended for fuzzy retrieval. An example using the UMLS terms is also presented. The model is also extended for term relationships beyond synonyms. Finally, this book looks at clustering, both crisp and fuzzy, to see how that can improve retrieval performance. An example is presented to illustrate the concepts.

---