

1. Record Nr.	UNINA9910822283103321
Autore	Torres-Moreno Juan-Manuel
Titolo	Automatic text summarization / / Juan-Manuel, Torres-Moreno
Pubbl/distr/stampa	London, [England] ; ; Hoboken, New Jersey : , : ISTE Limited : , : John Wiley & Sons, , 2014 ©2014
ISBN	1-119-04407-3 1-119-00475-6 1-119-04414-6
Descrizione fisica	1 online resource (376 p.)
Disciplina	025.04
Soggetti	Automatic abstracting
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Cover Page; Half-title Page; Title page; Copyright page; Contents; Foreword by A. Zamora and R. Salvador; Foreword; The need to identify important information; The problem of information storage; Automatic size reduction; The future; Foreword by H. Saggion; Automatic Text Summarization; Juan-Manuel Torres-Moreno; Notation; Introduction; The need to summarize texts; The summarization process; Automatic text summarization; About this book; PART 1: Foundations; 1: Why Summarize Texts?; 1.1. The need for automatic summarization; 1.2. Definitions of text summarization 1.3. Categorizing automatic summaries 1.4. Applications of automatic text summarization; 1.5. About automatic text summarization; 1.6. Conclusion; 2: Automatic Text Summarization: Some Important Concepts; 2.1. Processes before the process; 2.1.1. Sentence-term matrix: the vector space model (VSM) model; 2.2. Extraction, abstraction or compression?; 2.3. Extraction-based summarization; 2.3.1. Surface-level algorithms; 2.3.2. Intermediate-level algorithms; 2.3.3. Deep parsing algorithms; 2.4. Abstract summarization; 2.4.1. FRUMP or the temptation to understand 2.4.2. Information extraction and abstract generation 2.5. Sentence compression and fusion; 2.5.1. Sentence compression; 2.5.2. Multisentence fusion; 2.6. The limits of extraction; 2.6.1. Cohesion and

coherence; 2.6.2. The HexTAC experiment; 2.7. The evolution of automatic text summarization tasks; 2.7.1. Traditional tasks; Single-document summarization; Guided summarization; Multidocument summarization; Multilingual summarization; 2.7.2. Current and future problems; Summarization based on the source of the document; Specialized-domain summarization; Update summarization Sentence compression and multi-sentence fusionSemantic summarization; Opinion summarization; Multi and cross-lingual summarization; Ultra-summarization; Tweet summarization (short texts in microblogs); Multimedia summarization; Abstract generation; 2.8. Evaluating summaries; 2.9. Conclusion; 3: Single-document Summarization; 3.1. Historical approaches; 3.1.1. H.P. Luhn's Automatic Creation of Literature Abstracts; 3.1.2. The Luhn algorithm; 3.1.2.1. Preprocessing; 3.1.2.2. Sentence weighting; 3.1.3. Edmundson's linear combination; Edmundson's algorithm; 3.1.4. Extracts by elimination 3.2. Machine learning approaches3.2.1. Machine learning parameters; 3.3. State-of-the-art approaches; 3.4. Latent semantic analysis; 3.4.1. Singular value decomposition (SVD); 3.4.2. Sentence weighting by SVD; 3.5. Graph-based approaches; 3.5.1. PAGERANK and SNA algorithms; 3.5.2. Graphs and automatic text summarization; 3.5.3. Constructing the graph; 3.5.4. Sentence weighting; 3.5.4.1. LEXRANK; 3.5.4.2. TEXTRANK; 3.6. DIVTEX: a summarizer based on the divergence of probability distribution; 3.7. CORTEX 22; 3.7.1. Frequential measures; 3.7.2. Hamming measures; 3.7.3. Mixed measures 3.7.4. Decision algorithm

Sommario/riassunto

This new textbook examines the motivations and the different algorithms for automatic document summarization (ADS). We performed a recent state of the art. The book shows the main problems of ADS, difficulties and the solutions provided by the community. It presents recent advances in ADS, as well as current applications and trends. The approaches are statistical, linguistic and symbolic. Several examples are included in order to clarify the theoretical concepts. The books currently available in the area of Automatic Document Summarization are not recent. Powerful algorithms have been developed