

1. Record Nr.	UNINA9910830069703321
Autore	De Ville Barry
Titolo	Text as data : computational methods of understanding written expression using SAS // Barry deVille, Gurpreet Singh Bawa
Pubbl/distr/stampa	Hoboken, NJ : , : John Wiley & Sons, Inc., , [2022] ©2022
ISBN	1-119-48714-5 1-119-48715-3
Descrizione fisica	1 online resource (235 pages)
Collana	Wiley and SAS business series
Disciplina	006.312
Soggetti	Text data mining
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Cover -- Title Page -- Copyright Page -- Contents -- Preface -- Acknowledgments -- About the Authors -- Introduction -- Chapter 1 Text Mining and Text Analytics -- Background and Terminology -- Text Analytics: What Is It? -- Brief History of Text -- Writing Systems of the World -- Meaning and Ambiguity -- Notes -- Chapter 2 Text Analytics Process Overview -- Text Analytics Processing -- Process Building Blocks -- Preparation -- Utilization -- Process Description -- Text Mining Data Sources -- Capture -- Linguistic Processing -- Parsing and Parse Products -- Internal Representation and Text Products -- Representation -- Notes -- Chapter 3 Text Data Source Capture -- Text Mining Data Source Assembly -- Use Case: Accessing Text from SAS Conference Proceedings -- Text Data Capture Process -- Consuming Linguistics Text Products -- Notes -- Chapter 4 Document Content and Characterization -- Authorship Analytics: Early Text Indicators and Measures -- Function Words as Indicators -- Beyond Function Words -- Words and Word Forms as Psychological Artifacts -- A Case Study in Gender Detection -- Data Product Example -- Analysis Results -- Summarization and Discourse Analysis -- Elementary Operations as Building Blocks to Results -- Fact Extraction -- Sentiment Extraction -- Conditional Inference -- Deployment -- Summarization -- Conclusion -- Notes -- Chapter 5 Textual Abstraction: Latent Structure, Dimension Reduction -- Text Mining

Data Source Assembly -- Latent Structure and Dimensional Reduction -- Singular Value Decomposition as Dimension Reduction -- Latent Semantic Analysis -- Clustering Approach to Document Classification -- SVD Approach to Document Indexing -- Rough Meaning - Approximation for Singular Value Dimensions -- Semantic Indexing: Assigning Category Based on Singular Value Dimensional Scores -- Identifying Topics Using Latent Structure. Latent Structure: Tracking Topic Term Variability Across Semantic Fields -- Conclusion -- Notes -- Chapter 6 Classification and Prediction -- Use Case Scenario -- Composite Document Construction -- Model Development -- Ensemble or Multiagent Models -- Identifying Drivers of Textual Consumer Feedback Using Distance-Based Clustering and Matrix Factorization -- Use Case Scenario: Retailer Reliability Ecommerce -- Discussion -- Notes -- Chapter 7 Boolean Methods of Classification and Prediction -- Rule-Based Text Classification and Prediction -- Method Description -- Characteristics of Boolean Rule Methods -- Example of Boolean Rules Applied to Text Mining Vaccine Data -- An Example Analysis -- Summary -- Notes -- Chapter 8 Speech to Text -- Introduction -- Processing Audio Feedback -- Business Problem -- Process Components -- Further Analysis: Sentiment and Latent Topics -- Conclusion -- Notes -- Appendix A Mood State Identification in Text -- Origins of Mood State Identification -- An Approach to Mood State Developed at SAS -- Background and Discussion -- An Example Mood State Process Flow -- Notes -- Appendix B A Design Approach to Characterizing Users Based on AudioInteractions on a Conversational AI Platform -- Audio-Based User Interaction Inference -- Recommendation Perspective vs. Conventional -- Sole Dependency on Text-Based Bots -- Implementation Scenario: Voice-Based Conversational AI Platform -- Component Process Flow -- Constructed Interaction -- Note -- Appendix C SAS Patents in Text Analytics -- Glossary -- Index -- EULA.

Sommario/riassunto

This book offers a thorough introduction to the framework and dynamics of text analytics and the underlying principles at work and provides an in-depth examination of the interplay between qualitative-linguistic and quantitative, data-driven aspects of data analysis. -- Edited summary from book
