

1. Record Nr.	UNINA9910484457903321
Titolo	Big Data Analytics : 4th International Conference, BDA 2015, Hyderabad, India, December 15-18, 2015, Proceedings // edited by Naveen Kumar, Vasudha Bhatnagar
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2015
ISBN	3-319-27057-5
Edizione	[1st ed. 2015.]
Descrizione fisica	1 online resource (XII, 267 p. 104 illus. in color.)
Collana	Information Systems and Applications, incl. Internet/Web, and HCI ; ; 9498
Disciplina	006.312
Soggetti	Data mining Health informatics Database management Information storage and retrieval Algorithms Data Mining and Knowledge Discovery Health Informatics Database Management Information Storage and Retrieval Algorithm Analysis and Problem Complexity
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Bibliographic Level Mode of Issuance: Monograph
Nota di contenuto	Intro -- Preface -- Organization -- Contents -- Big Data: Security and Privacy -- Privacy Protection or Data Value: Can We Have Both? -- 1 Introduction -- 2 Big Data -- 3 Privacy -- 3.1 Individuals Own Data About Themselves -- 3.2 Corporations Own Data Collected -- 3.3 Shared Ownership of Data -- 3.4 Summary -- Ownership -- 4 Privacy Principles -- 5 A Call for a New Approach -- 5.1 User and Corporate Obligations -- 5.2 System Obligations -- 5.3 Data Acquisition Process -- 5.4 Managing Private Queries -- 6 Summary -- References -- Open Source Social Media Analytics for Intelligence and Security Informatics Applications -- 1 Introduction -- 1.1 Online Social Media -- 1.2 Intelligence and Security Informatics -- 2 Technical and Computational

Challenges -- 3 Machine Learning and Data Mining Techniques -- 4 Case Studies -- 4.1 Identification of Extremist Content, Users and Communities -- 4.2 Event Forecasting for Civil Unrest Related Events -- 5 Characterization and Classification of Related Work -- 6 ISI Leading Conferences and Journals -- 7 Conclusions -- References -- Big Data in Commerce -- Information Exploration in E-Commerce Databases -- 1 Introduction -- 2 Search Interface in E-Commerce -- 3 Obstacles in Exploratory Search -- 3.1 Limited Feedback in the Query Panel -- 3.2 Low Adaptivity of the Result Panel -- 3.3 Information Loss from Hidden Attributes -- 4 Extensions to Facilitate Exploratory Search -- 4.1 Exploratory Search in Query Panel -- 4.2 Exploratory Search in Result Panel -- 5 Evaluation -- 6 Conclusion -- References -- A Framework to Harvest Page Views of Web for Banner Advertising -- 1 Introduction -- 2 Related Work -- 3 Proposed Approach -- 3.1 Existing Framework -- 3.2 Proposed Framework -- 4 Design Issues -- 5 Summary and Conclusions -- References -- Utility-Based Control Flow Discovery from Business Process Event Logs. 1 Research Motivation and Aim -- 2 Related Work and Research Contributions -- 3 Research Framework and Solution Approach -- 4 Experimental Analysis and Results -- 4.1 Experimental Dataset -- 4.2 Experimental Results -- 4.3 Airport Dataset Results -- 4.4 BPI 2014 Dataset Results -- 5 Conclusion -- References -- An Efficient Algorithm for Mining High-Utility Itemsets with Discount Notion -- 1 Introduction -- 2 Related Work -- 3 Background -- 3.1 Preliminary -- 3.2 UP-Hist Tree and UP-Hist Algorithm -- 3.3 Utility-List Data Structure and FHM Algorithm -- 3.4 Three-Phase Algorithm for Mining High-Utility Itemsets with Discount Strategies -- 4 Integration of Discount Notion in State-of-the-art Algorithms -- 4.1 UP-Hist Discount Algorithm -- 4.2 FHM Discount Algorithm -- 5 Mining High-Utility Itemsets -- 6 Experiments and Results -- 7 Conclusion and Future Work -- References -- Big Data: Models and Algorithms -- Design of Algorithms for Big Data Analytics -- 1 Introduction -- 2 Building FCA Lattice -- 2.1 Inefficient Map-Reduce Based Design -- 2.2 Efficient Map-Reduce Based Design -- 3 High Velocity Datasets -- 4 Conclusions -- References -- Mobility Big Data Analysis and Visualization (Invited Talk) -- 1 Introduction -- 2 Public Transportation Data Analysis -- 2.1 Framework for Analyzing Smart Card Data -- 2.2 Visual Fusion Analysis Environment -- 3 Vehicle Recorder Data Analysis -- 3.1 Relationships Between Drivers' Behaviors and Their Past Driving Histories -- 3.2 Visual Exploration of Caution Spots -- 4 Conclusions -- References -- Finding Periodic Patterns in Big Data -- 1 Introduction -- 2 Periodic Pattern Mining in Time Series -- 2.1 The Basic Model of Periodic Patterns -- 2.2 The Limitations of Basic Model -- 2.3 Research Efforts to Address the Limitations -- 3 Periodic Pattern Mining in Transactional Databases. 3.1 The Basic Model of Periodic-Frequent Patterns -- 3.2 The Performance Issues of the Periodic-Frequent Pattern Model -- 3.3 Research Efforts to Address the Performance Issues -- 4 Experimental Results -- 4.1 Experimental Setup -- 4.2 Generation of Periodic-Frequent Patterns -- 4.3 Interesting Patterns Discovered in Accidents Database -- 4.4 Interesting Patterns Discovered in Shop-4 Database -- 5 Conclusions -- References -- VDMR-DBSCAN: Varied Density MapReduce DBSCAN -- 1 Introduction -- 2 Related Work -- 3 Proposed Work -- 3.1 Partitioning Phase -- 3.2 Map Phase -- 3.3 Reduce Phase -- 3.4 Merge Phase -- 3.5 Relabeling of Data Points -- 4 Results and Discussions -- 4.1 Experimental Settings -- 4.2 Experimental Results -- 4.3 Clustering Results on Zahn\_compound Dataset (DS1) -- 4.4 Clustering Results on Spiral Dataset (DS2) -- 4.5 Large Synthetic

Dataset (DS3) -- 5 Conclusions and Future Work -- References --  
Concept Discovery from Un-Constrained Distributed Context --  
Abstract -- 1 Introduction -- 2 Definitions and Properties in FCA -- 3  
Related Work -- 4 Methodology -- 4.1 Data Format and Input -- 4.2  
Data Aggregation -- 4.3 Concept Finding Approach -- 4.4 Additional  
Important Insights -- 5 Implementation Using Spark -- 6 Evaluation  
and Results -- 7 Discussion -- 7.1 Scalability with Respect to Context  
-- 7.2 Comparison with Distributed Frequent Item-Set Mining -- 7.3  
Concept Explosion -- 7.4 Mining Links Between Concepts -- 8  
Conclusion -- Acknowledgments -- References -- Khanan:  
Performance Comparison and Programming -Miner Algorithm in  
Column-Oriented and Relational Database Query Languages -- 1  
Research Motivation and Aim -- 2 Related Work and Novel  
Contributions -- 2.1 Implementation of Mining Algorithms in Row-  
Oriented Databases -- 2.2 Implementation of Mining Algorithms in  
Column-Oriented Databases.  
2.3 Performance Comparison of Mining Algorithms in Column-Oriented  
and Graph Databases -- 3 -Miner Algorithm -- 4 Implementation of -  
Miner Algorithm in SQL on Row-Oriented Database (MySQL) -- 5  
Implementation of -Miner Algorithm on NoSQL Column-Oriented  
Database Cassandra -- 6 Experimental Dataset -- 7 Benchmarking and  
Performance Comparison -- 8 Conclusion -- References -- A New  
Proposed Feature Subset Selection Algorithm Based on Maximization of  
Gain Ratio -- Abstract -- 1 Introduction -- 2 Related Work -- 3  
Theoretical Background -- 4 Algorithm and Analysis -- 5 Empirical  
Study -- 5.1 Datasets -- 5.2 Experimental Results and Comparisons --  
5.3 Dimension Wise Comparison of Feature Selection Algorithms -- 5.4  
Domain Wise Comparison of Feature Selection Algorithms -- 6  
Conclusion -- References -- Big Data in Medicine -- Genomics 3.0:  
Big-Data in Precision Medicine -- 1 Introduction -- 2 Genomics 3.0 --  
3 Dimensions of Big-data -- 4 Genomic Big-data -- 4.1 Patient Data --  
4.2 Background Databases -- 5 Exploratory Data Analysis -- 6 Big-  
data to Information -- 7 Information to Knowledge -- 8 Integrative  
Systems Biology -- 9 Patient Stratification -- 10 Characteristics and  
Life of Genomic Big-data -- 11 Mark-up Languages to Represent  
Biological Data -- 12 The Big-data Genomics Platform iOMICS -- 13  
Live Genome Research Case -- 14 Live Genomics and Precision  
Medicine Case -- References -- CuraEx - Clinical Expert System Using  
Big-Data for Precision Medicine -- 1 Introduction -- 2 CuraEx Overview  
-- 3 Patient Specific Information -- 4 Analytical Engine -- 4.1 Cancer  
Staging -- 4.2 Prognosis -- 4.3 Therapeutics -- 5 Knowledge Database  
-- 6 Software Stack and Data Management -- 7 Conclusion --  
References -- Multi-omics Multi-scale Big Data Analytics for Cancer  
Genomics -- 1 Introduction -- 2 Available Data -- 2.1 DNA Level Data  
-- 2.2 RNA Level Data.  
2.3 Clinical Data -- 2.4 Background Databases -- 3 Key Aims -- 4  
Exploratory Data Analysis -- 4.1 Mutation Association with Cancer  
State -- 4.2 Differential Gene Expression -- 4.3 Patient Stratification --  
5 Multi-scale Integrative Analysis -- 6 Data Integration and Network  
Analysis -- 6.1 Functional Characterization Databases -- 6.2 Metabolic  
Network Reconstruction and Protein Interactions -- 7 Conclusions --  
References -- Class Aware Exemplar Discovery from Microarray Gene  
Expression Data -- Abstract -- 1 Introduction -- 2 Overview of Our  
Approach -- 2.1 Gene Data -- 2.2 Gene-Gene Similarity -- 2.3 Class  
Aware Preference -- 2.4 Class Aware Message Passing -- 3  
Experimental Evaluation -- 3.1 Description of Experimental Datasets --  
3.2 Comparison of Class Aware Exemplar Discovery with Affinity  
Propagation -- 3.3 Comparison of Class Aware Exemplar Discovery

with Standard Feature Subset Selection Techniques -- 3.4 Comparison of Class Aware Exemplar Discovery with All Features -- 4 Conclusions -- References -- Multistage Classification for Cardiovascular Disease Risk Prediction -- Abstract -- 1 Introduction -- 2 Related Work -- 3 Proposed Work -- 3.1 Data Set -- 4 Experiments and Results -- 4.1 Pre-processing -- 4.2 Feature Extraction -- 4.3 Classification -- 5 Conclusion -- Acknowledgement -- References -- Author Index.

---

Sommario/riassunto

This book constitutes the refereed conference proceedings of the Fourth International Conference on Big Data Analytics, BDA 2015, held in Hyderabad, India, in December 2015. The 9 revised full papers and 9 invited papers were carefully reviewed and selected from 61 submissions and cover topics on big data: security and privacy; big data in commerce; big data: models and algorithms; and big data in medicine.

---