1. Record Nr. UNINA9911020081403321

Titolo	Data mining techniques in grid computing environments / / editor, Werner Dubitzky
Pubbl/distr/stampa	Hoboken, NJ, : J. Wiley, 2008
ISBN	9786611939458 9781281939456 1281939455 9780470699904 0470699906 9780470699898 0470699892
Edizione	[1st edition]
Descrizione fisica	1 online resource (290 p.)
Altri autori (Persone)	DubitzkyWerner <1958->
Disciplina	004/.36
Soggetti	Data mining Computational grids (Computer systems)
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	Data Mining Techniques in Grid Computing Environments; Contents; Preface; List of Contributors; 1 Data mining meets grid computing: Time to dance?; 1.1 Introduction; 1.2 Data mining; 1.2.1 Complex data mining problems; 1.2.2 Data mining challenges; 1.3 Grid computing; 1.3.1 Grid computing challenges; 1.4 Data mining grid - mining grid data; 1.4.1 Data mining grid: a grid facilitating large-scale data mining; 1.4.2 Mining grid data: analyzing grid systems with data mining techniques; 1.5 Conclusions; 1.6 Summary of Chapters in this Volume; 2 Data analysis services in the knowledge grid 2.1 Introduction2.2 Approach; 2.3 Knowledge Grid services; 2.3.1 The Knowledge Grid architecture; 2.3.2 Implementation; 2.4 Data analysis services; 2.5 Design of Knowledge Grid applications; 2.5.1 The VEGA visual language; 2.5.2 UML application modelling; 2.5.3 Applications and experiments; 2.6 Conclusions; 3 GridMiner: An advanced support for e-science analytics; 3.1 Introduction; 3.2 Rationale behind the design and development of GridMiner; 3.3 Use Case; 3.4 Knowledge

discovery process and its support by the GridMiner; 3.4.1 Phases of knowledge discovery; 3.4.2 Workflow management
3.4.3 Data management3.4.4 Data mining services and OLAP; 3.4.5 Security; 3.5 Graphical user interface; 3.6 Future developments; 3.6.1 High-level data mining model; 3.6.2 Data mining query language; 3.6.3 Distributed mining of data streams; 3.7 Conclusions; 4 ADaM services: Scientific data mining in the service-oriented architecture paradigm; 4.1 Introduction; 4.2 ADaM system overview; 4.3 ADaM toolkit overview; 4.4 Mining in a service-oriented architecture; 4.5 Mining web services; 4.5.1 Implementation architecture; 4.5.2 Workflow example; 4.5.3 Implementation issues
4.6 Mining grid services4.6.1 Architecture components; 4.6.2 Workflow example; 4.7 Summary; 5 Mining for misconfigured machines in grid systems; 5.1 Introduction; 5.2 Preliminaries and related work; 5.2.1 System misconfiguration detection; 5.2.2 Outlier detection; 5.3 Acquiring, pre-processing and storing data; 5.3.1 Data sources and acquisition; 5.3.2 Pre-processing; 5.3.3 Data organization; 5.4 Data analysis; 5.4.1 General approach; 5.4.2 Notation; 5.4.3 Algorithm; 5.4.4 Correctness and termination; 5.5 The GMS; 5.6 Evaluation; 5.6.1 Qualitative results; 5.6.2 Quantitative results
5.6.3 Interoperability5.7 Conclusions and future work; 6 FAEHIM: Federated Analysis Environment for Heterogeneous Intelligent Mining; 6.1 Introduction; 6.2 Requirements of a distributed knowledge discovery framework; 6.2.1 Category 1: knowledge discovery specific requirements; 6.2.2 Category 2: distributed framework specific requirements; 6.3 Workflow-based knowledge discovery; 6.4 Data mining toolkit; 6.5 Data mining service framework; 6.6 Distributed data mining services; 6.7 Data manipulation tools; 6.8 Availability; 6.9 Empirical experiments; 6.9.1 Evaluating the framework accuracy
6.9.2 Evaluating the running time of the framework

Sommario/riassunto

Based around eleven international real life case studies and including contributions from leading experts in the field this groundbreaking book explores the need for the grid-enabling of data mining applications and provides a comprehensive study of the technology, techniques and management skills necessary to create them. This book provides a simultaneous design blueprint, user guide, and research agenda for current and future developments and will appeal to a broad audience; from developers and users of data mining and grid technology, to advanced undergraduate and postgraduate students inte