1.

Record Nr.	UNINA9910138969003321
Autore	Elloumi Mourad
Titolo	Biological knowledge discovery handbook : preprocessing, mining and postprocessing of biological data / / edited by Mourad Elloumi, Albert Y. Zomaya ; cover Design, Michael Rutkowski ; Jad Abbass [and one hundred twenty eight others], contributors
Pubbl/distr/stampa	Hoboken, New Jersey : , : Wiley, , 2014 ©2014
ISBN	1-118-85372-5 1-118-61715-0 1-118-61711-8
Descrizione fisica	1 online resource (1192 p.)
Collana	Wiley Series in Bioinformatics : Computational Techniques and Engineering
Classificazione	COM021040
Altri autori (Persone)	ElloumiMourad RutkowskiMichael AbbassJad
Disciplina	572.80285
Soggetti	Bioinformatics Computational biology Data mining
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 at the end of each chapters and index.
Nota di contenuto	BIOLOGICAL KNOWLEDGE DISCOVERY HANDBOOK: Preprocessing, Mining, and Postprocessing of Biological Data; CONTENTS; PREFACE; CONTRIBUTORS; SECTION I: BIOLOGICAL DATA PREPROCESSING; PART A: BIOLOGICAL DATA MANAGEMENT; 1 GENOME AND TRANSCRIPTOME SEQUENCE DATABASES FOR DISCOVERY, STORAGE, AND REPRESENTATION OF ALTERNATIVE SPLICING EVENTS; 1.1 INTRODUCTION; 1.2 SPLICING; 1.2.1 Mechanism of Splicing; 1.2.2 Regulation of Splicing; 1.3 ALTERNATIVE SPLICING; 1.3.1 Introduction to Alternative Splicing; 1.3.2 Mechanism of Alternative Splicing; 1.3.3 Regulation of Alternative Splicing 1.3.4 Evolution and Conservation of Splicing and Alternative Splicing 1.4 ALTERNATIVE SPLICING DATABASES; 1.4.1 Genomic and

Transcriptomic Sequence Analyses; 1.4.2 Literature Overview of Various Alternative Splicing Databases; 1.4.3 SDBs; 1.5 DATA MINING FROM ALTERNATIVE SPLICING DATABASES; 1.5.1 Implementation of dbASQ and Utility of SDBs; 1.5.2 Identification of Transcript-Initial and Transcript-Terminal Variation; ACKNOWLEDGMENTS; WEB RESOURCES; REFERENCES; 2 CLEANING, INTEGRATING, AND WAREHOUSING GENOMIC DATA FROM BIOMEDICAL RESOURCES; 2.1 INTRODUCTION; 2.2 RELATED WORK

2.3 TYPOLOGY OF DATA QUALITY PROBLEMS IN BIOMEDICAL RESOURCES 2.4 CLEANING, INTEGRATING, AND WAREHOUSING BIOMEDICAL DATA; 2.4.1 Lessons Learned from Integrating and Warehousing Biomedical Data on Liver Genes and Diseases; 2.4.2 Data Quality-Aware Solutions; 2.4.3 Biological Entity Resolution and Record Linkage; 2.4.4 Ontology-Based Approaches; 2.5 CONCLUSIONS AND PERSPECTIVES; WEB RESOURCES; REFERENCES; 3 CLEANSING OF MASS SPECTROMETRY DATA FOR PROTEIN IDENTIFICATION AND QUANTIFICATION; 3.1 INTRODUCTION; 3.2 PREPROCESSING APPROACH FOR IMPROVING PROTEIN IDENTIFICATION; 3.2.1 Existing Approaches

3.2.2 New Dynamic Wavelet-Based Spectra Preprocessing Method 3.3 IDENTIFICATION FILTERING APPROACH FOR IMPROVING PROTEIN IDENTIFICATION; 3.3.1 Existing Approaches; 3.3.2 New Target-Decoy Approach for Improving Protein Identification; 3.4 EVALUATION RESULTS; 3.4.1 Evaluation of New Proprocessing Method; 3.4.2 Evaluation of New Identification Filtering Method; 3.5 CONCLUSION; REFERENCES; 4 FILTERING PROTEIN-PROTEIN INTERACTIONS BY INTEGRATION OF ONTOLOGY DATA; 4.1 INTRODUCTION; 4.2 EVALUATION OF SEMANTIC SIMILARITY; 4.2.1 Gene Ontology; 4.2.2 Survey of Semantic Similarity Measures

4.2.3 Correlation with Functional Categorizations 4.3 IDENTIFICATION OF FALSE PROTEIN-PROTEIN INTERACTION DATA; 4.3.1 Classification Method; 4.3.2 Accuracy of PPI Classification; 4.3.3 Reliability of PPI Data; 4.4 CONCLUSION; REFERENCES; PART B: BIOLOGICAL DATA MODELING; 5 COMPLEXITY AND SYMMETRIES IN DNA SEQUENCES; 5.1 INTRODUCTION; 5.2 ARCHAEA; 5.3 PATTERNS ON INDICATOR MATRIX; 5.3.1 Indicator Matrix; 5.3.2 Test Sequences; 5.4 MEASURE OF COMPLEXITY AND INFORMATION; 5.4.1 Complexity; 5.4.2 Fractal Dimension; 5.4.3 Entropy; 5.5 COMPLEX ROOT REPRESENTATION OF DNA WORDS

5.5.1 Pseudorandom Sequence on Unit Circle

Sommario/riassunto

The first comprehensive overview of preprocessing, mining, and postprocessing of biological data Molecular biology is undergoing exponential growth in both the volume and complexity of biological data-and knowledge discovery offers the capacity to automate complex search and data analysis tasks. This book presents a vast overview of the most recent developments on techniques and approaches in the field of biological knowledge discovery and data mining (KDD)- providing in-depth fundamental and technical field information on the most important topics encountered. Written by top