

1. Record Nr.	UNISA996213835303316
Titolo	Computational Intelligence Methods for Bioinformatics and Biostatistics [[electronic resource]] : 10th International Meeting, CIBB 2013, Nice, France, June 20-22, 2013, Revised Selected Papers // edited by Enrico Formenti, Roberto Tagliaferri, Ernst Wit
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2014
ISBN	3-319-09042-9
Edizione	[1st ed. 2014.]
Descrizione fisica	1 online resource (XIII, 275 p. 99 illus.)
Collana	Lecture Notes in Bioinformatics ; ; 8452
Disciplina	006.3
Soggetti	Bioinformatics Pattern recognition Data mining Computers Optical data processing Algorithms Computational Biology/Bioinformatics Pattern Recognition Data Mining and Knowledge Discovery Computation by Abstract Devices Image Processing and Computer Vision 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 -- Keynote Speaker -- Dynamic Gaussian Graphical Models for Modelling Genomic Networks -- 1 Introduction -- 2 Graphical Models -- 3 Dynamic Gaussian Graphical Model for Networks -- 3.1 Sparsity Restrictions of the Precision Matrix -- 3.2 Model Restrictions of the Precision Matrix -- 3.3 Maximum Likelihood -- 4 Max Determinant Optimization Problem -- 5 Application to T-Cell Data -- 6 Conclusions -- References -- Bioinformatics Regular Session -- Molecular Docking for Drug

Discovery: Machine-Learning Approaches for Native Pose Prediction of Protein-Ligand Complexes -- 1 Introduction -- 1.1 Background -- 1.2 Related Work -- 1.3 Key Contributions -- 2 Materials and Methods -- 2.1 Compound Database -- 2.2 Compound Characterization -- 2.3 Decoy Generation and Formation of Training and Test Sets -- 2.4 Conventional Scoring Functions -- 2.5 Machine Learning Methods -- 3 Results and Discussion -- 3.1 Evaluation of Scoring Functions -- 3.2 ML vs. Conventional Approaches on a Diverse Test Set -- 3.3 ML vs. Conventional Approaches on Homogeneous Test Sets -- 3.4 Impact of Training Set Size -- 4 Conclusion -- References -- BioCloud Search EnGene: Surfing Biological Data on the Cloud -- Abstract -- 1 Introduction -- 2 Background and Motivations -- 3 Architectural Aspects -- 3.1 Query Contextualization -- 3.2 Technical Details -- 4 BSE Functionalities -- 5 Conclusions -- Acknowledgments -- References -- Genomic Sequence Classification Using Probabilistic Topic Modeling -- 1 Introduction -- 2 Related Work -- 3 Methods -- 3.1 Probabilistic Topic Models -- 3.2 Document Model and DNA Sequences -- 4 Experimental Tests -- 4.1 Bacteria Dataset -- 4.2 Training and Testing Pipelines -- 4.3 Classification Pipeline -- 4.4 Results and Discussion -- 5 Conclusion and Future Work -- References. Community Detection in Protein-Protein Interaction Networks Using Spectral and Graph Approaches -- 1 Introduction -- 2 Communities Discovering in Networks Using Graph Analysis -- 3 Communities Discovering in Networks Using Clustering Approaches -- 4 The FSM Community Detection Method -- 5 Saccharomyces cerevisiae PPIs Discovery -- 5.1 Dataset -- 5.2 Experimental Results and Discussion -- 6 Conclusions -- References -- Weighting Scheme Methods for Enhanced Genomic Annotation Prediction -- 1 Introduction -- 2 Annotation Data Considered -- 2.1 Genomic and Proteomic Data Warehouse -- 2.2 Annotation Unfolding -- 3 Computational Methods -- 3.1 Weighting Schemes for Annotation Matrices -- 3.2 Latent Semantic Indexing by Singular Value Decomposition -- 3.3 Semantic Improvement (SIM) -- 3.4 Anomaly Correction -- 3.5 Computational Complexity -- 4 Validation of Annotation Predictions -- 5 Validation Results -- 6 Conclusions -- References -- French Flag Tracking by Morphogenetic Simulation Under Developmental Constraints -- 1 Introduction -- 1.1 Morphogenesis: Emerging of Interests -- 1.2 Below Genetic Expression -- 2 Morphological Dynamic of Cells -- 2.1 Mathematical Model -- 2.2 Shapes Emergence -- 3 Simulation Tool -- 3.1 Architecture -- 3.2 Features -- 4 Algorithms -- 4.1 Software Architecture -- 4.2 Genomes Base Construct -- 4.3 Initialization -- 4.4 Description of the Algorithm -- 5 Evaluation -- 5.1 Test Conditions -- 5.2 Output Results -- 6 Conclusion -- 6.1 Relevance to Biological Issues -- 6.2 Future Works -- References -- Biostatistics Regular Session -- High-Dimensional Sparse Matched Case-Control and Case-Crossover Data: A Review of Recent Works, Description of an R Tool and an Illustration of the Use in Epidemiological Studies -- 1 Introduction -- 2 Conditional Logistic Regression -- 2.1 The Model -- 2.2 Conditional Likelihood. 2.3 L1 Penalized Conditional Likelihood -- 3 Complexity Tuning -- 4 Uncertainty Measures -- 5 Standardization, Bias Correction, Unpenalized Predictors -- 6 Capabilities of clogitLasso -- 6.1 Example 1 -- 6.2 Example 2 -- 7 Conclusion -- References -- Piecewise Exponential Artificial Neural Networks (PEANN) for Modeling Hazard Function with Right Censored Data -- 1 Introduction -- 2 Methods -- 3 Breast Cancer Survival Study -- 4 Conclusions -- References -- Writing Generation Model for Health Care Neuromuscular System Investigation -- Abstract -- 1 Introduction -- 2 Neuromuscular System Function

Transfer minus the Delta-Log Model -- 3 Neuromuscular System Function Transfer -- the Sigma-Log Model -- 4 Experimental Setup for Writing Generation Model -- 5 Conclusion -- Acknowledgment -- References -- Clusters Identification in Binary Genomic Data: The Alternative Offered by Scan Statistics Approach -- 1 Introduction -- 2 Methods -- 2.1 DBSCAN -- 2.2 Kulldorff Spatial Scan Statistics for Bernoulli Model -- 2.3 Experimental Data -- 3 Results -- 4 Conclusions -- References -- Special Session: Knowledge Based Medicine -- Reverse Engineering Methodology for Bioinformatics Based on Genetic Programming, Differential Expression Analysis and Other Statistical Methods -- Abstract -- 1 Introduction -- 2 Enhanced GP RODES Methodology -- 2.1 Data Fitting with Smoothing Spline and Time Derivative Computing -- 2.2 Temporal Differentially Expressed Genes -- 3 Results -- 3.1 Data Fitting and Time Derivative Computing -- 3.2 Results of Time Course Study -- 3.3 Structure and Parameters Discovery from the GSE35074 miRNA Time Series Data with GP RODES -- 4 Conclusions -- Acknowledgments -- References -- Integration of Clinico-Pathological and microRNA Data for Intelligent Breast Cancer Relapse Prediction Systems -- 1 Introduction -- 2 Methods -- 2.1 Data Set. 2.2 Experimental Framework -- 3 Results -- 3.1 Patient Differentiation Based on ER Status -- 3.2 Exclusion of HS Sequences -- 4 Biological Significance -- 5 Conclusions -- References -- Superresolution MUSIC Based on Marcenko-Pastur Limit Distribution Reduces Uncertainty and Improves DNA Gene Expression-Based Microarray Classification -- 1 Introduction -- 2 Methods -- 2.1 Datasets Used -- 2.2 Gene Selection and Usage -- 2.3 Bootstrap Root MUSIC Classifier -- 2.4 Classification Runs Using Bootstrap Root MUSIC (BRM) -- 3 Results -- 4 Discussion -- References -- Special Session: Data Integration and Analysis in Omic-Science -- Prediction of Single-Nucleotide Polymorphisms Causative of Rare Diseases -- 1 Introduction -- 2 Materials -- 2.1 Formats and Tools for Managing NGS Data -- 2.2 Pooling -- 3 Methods -- 3.1 Decision Rules -- 3.2 Features -- 3.3 Generalized Eigenvalue Classification -- 4 Results and Discussion -- 4.1 Benchmark Dataset -- 4.2 Other Tools Results -- 4.3 Decision Rules Results -- 4.4 Standard Classifiers Performance -- 4.5 Prediction -- 4.6 SIFT -- 5 Concluding Remarks -- References -- A Framework for Mining Life Sciences Data on the Semantic Web in an Interactive, Graph-Based Environment -- 1 Introduction -- 2 Material and Methods -- 2.1 Semantic Reconciliation of Ondex and RDF Data Models -- 2.2 SPARQL Query Console -- 2.3 Interactive Browsing -- 2.4 SPARQL Commands Configuration File -- 3 Results and Discussion -- 3.1 Example Use Case Identifying Interacting Proteins with IPR002048 and IPR003527 Domains -- 3.2 Discussion -- 4 Conclusion -- References -- Combining Not-Proper ROC Curves and Hierarchical Clustering to Detect Differentially Expressed Genes in Microarray Experiments -- 1 Introduction -- 2 ROC Curve and the TNRC Statistics -- 2.1 Properties of TNRC -- 2.2 Interpreting TNRC Using Information from Hierarchical Clustering. 3 Results -- 4 Conclusions -- References -- Fast and Parallel Algorithm for Population-Based Segmentation of Copy-Number Profiles -- 1 Introduction -- 2 Linearization of Dynamic Programming for Segmentation and Segmentation/Clustering -- 2.1 Original Dynamic Programming Algorithm for Segmentation -- 2.2 A Linear Dynamic Programming Algorithm for the Classification Cost Function -- 2.3 A Bound on the Quality of the Approximation -- 3 Joint Segmentation and Parallelization of the Algorithm -- 4 Correctness, Computational Footprint and Scalability -- References -- Identification of Pathway Signatures in Parkinson's Disease with Gene Ontology and Sparse

Regularization -- 1 Introduction -- 2 Materials and Methods -- 2.1 Feature Selection Framework -- 2.2 The KDVS Pipeline -- 2.3 The Standard Pipeline -- 2.4 Benchmark Lists -- 2.5 Precision, Recall and F-measure -- 3 Results and Discussion -- 3.1 The KDVS Pipeline -- 3.2 Comparison Between Lists: KDVS and Benchmark -- 3.3 The Standard Pipeline -- 3.4 Data Analysis -- 3.5 Functional Analysis -- 3.6 Comparison Between the Lists: Standard and Benchmark -- 3.7 Comparison of KDVS and Standard Pipelines -- 4 Conclusions -- References -- Author Index.

Sommario/riassunto

This book constitutes the thoroughly refereed post-conference proceedings of the 10th International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics, CIBB 2013, held in Nice, France in June 2013. The 19 revised full papers presented were carefully reviewed and selected from 35 submissions. The papers are organized in topical sections on bioinformatics, biostatistics, knowledge based medicine, and data integration and analysis in omic-science.
