

1. Record Nr.	UNINA9910349464803321
Autore	Altman Russ B
Titolo	Biocomputing 2019 - Proceedings Of The Pacific Symposium
Pubbl/distr/stampa	World Scientific Publishing Co, 2018 Singapore : , : World Scientific Publishing Company, , 2018 ©2019
ISBN	981-327-982-6 981-327-981-8
Descrizione fisica	1 online resource (471 pages)
Altri autori (Persone)	DunkerA. Keith <1943-> (Alan Keith) HunterLawrence <1961-> RitchieMarylyn D MurrayTiffany A KleinTeri E
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- PATTERN RECOGNITION IN BIOMEDICAL DATA: CHALLENGES IN PUTTING BIG DATA TO WORK -- Session introduction -- Introduction -- References -- Learning Contextual Hierarchical Structure of Medical Concepts with Poincaré Embeddings to Clarify Phenotypes -- 1. Introduction -- 2. Methods -- 2.1. Source Code -- 2.2. Data Source -- 2.3. Data Selection and Preprocessing -- 2.3.1. Reference ICD9 Example -- 2.3.2. Real Member Analyses -- 2.4. Poincaré Embeddings -- 2.5. Processing and Evaluating Embeddings -- 3. Results -- 3.1. ICD9 Hierarchy Evaluation -- 3.2. Poincaré Embeddings on 10 Million Members -- 3.3. Comparison with Euclidean Embeddings -- 3.4. Cohort Specific Embeddings -- 4. Discussion and Conclusion -- 5. Acknowledgments -- References -- The Effectiveness of Multitask Learning for Phenotyping with Electronic Health Records Data -- 1. Introduction -- 2. Background -- 2.1. Multitask nets -- 3. Methods -- 3.1. Dataset Construction and Design -- 3.2. Experimental Design -- 4. Experiments and Results -- 4.1. When Does Multitask Learning Improve Performance? -- 4.2. Relationship Between

Performance and Number of Tasks -- 4.3. Comparison with Logistic Regression Baseline -- 4.4. Interaction between Phenotype Prevalence and Complexity -- 5. Limitations -- 6. Conclusion --

Acknowledgments -- References -- ODAL: A one-shot distributed algorithm to perform logistic regressions on electronic health records data from multiple clinical sites -- 1. Introduction -- 1.1. Integrate evidence from multiple clinical sites -- 1.2. Distributed Computing -- 2. Material and Method -- 2.1. Clinical Cohort and Motivating Problem -- 2.2. Algorithm -- 2.3. Simulation Design -- 3. Results -- 3.1. Simulation Results -- 3.2. Fetal Loss Prediction via ODAL -- 4. Discussion -- References.

PVC Detection Using a Convolutional Autoencoder and Random Forest Classifier -- 1. Introduction -- 2. Methods -- 2.1. Data Set and Implementation -- 2.2. Proposed PVC Detection Method -- 2.2.1. Feature Extraction -- 2.2.2. Classification -- 3. Results -- 3.1. Full Database Evaluation -- 3.2. Timing Disturbance Evaluation -- 3.3. Cross-Patient Training Evaluation -- 3.4. Estimated Parameters and Convergence -- 4. Discussion -- References -- Removing Confounding Factors Associated Weights in Deep Neural Networks Improves the Prediction Accuracy for Healthcare Applications -- 1. Introduction -- 2. Related Work -- 3. Confounder Filtering (CF) Method -- 3.1. Overview -- 3.2. Method -- 3.3. Availability -- 4. Experiments -- 4.1. lung adenocarcinoma prediction -- 4.1.1. Data -- 4.1.2. Results -- 4.2. Segmentation on right ventricle(RV) of Heart -- 4.2.1. Data -- 4.2.2. Results -- 4.3. Students' confusion status prediction -- 4.3.1. Data -- 4.3.2. Results -- 4.4. Brain tumor prediction -- 4.4.1. Data -- 4.4.2. Results -- 4.5. Analyses of the method behaviors -- 5. Conclusion -- 6. Acknowledgement -- References --

DeepDom: Predicting protein domain boundary from sequence alone using stacked bidirectional LSTM -- 1. Introduction -- 2. METHODS -- 2.1 Data Set Preparation -- 2.2 Input Encoding -- 2.3 Model Architecture -- 2.4 Evaluation criteria -- 3. RESULTS AND DISCUSSION -- 3.1 Parameter configuration experiments on test data -- 3.2 Comparison with Other Domain Boundary Predictors -- 3.2.1 Free modeling targets from CASP 9 -- 3.2.2 Multi-domain targets from CASP 9 -- 3.2.3 Discontinuous domain target from CASP 8 -- 4. CONCLUSION -- 5.

ACKNOWLEDGEMENTS -- REFERENCES -- Res2s2aM: Deep residual network-based model for identifying functional noncoding SNPs in trait-associated regions -- 1. Introduction -- 2. Background theory. 3. Dataset for training and testing -- 3.1. Source databases -- 3.2. Dataset generation -- 4. Methods -- 4.1. ResNet architecture in our model -- 4.2. Tandem inputs of forward- and reverse-strand sequences -- 4.3. Biallelic high-level network structure -- 4.4. Incorporating HaploReg SNP annotation features -- 4.5. Training of models -- 5. Results -- 6. Conclusions and discussion --

Acknowledgements -- References -- DNA Steganalysis Using Deep Recurrent Neural Networks -- 1. Introduction -- 2. Background -- 2.1. Notations -- 2.2. Hiding Messages -- 2.3. Determination of Message-Hiding Regions -- 3. Methods -- 3.1. Proposed DNA Steganalysis Principle -- 3.2. Proposed Steganalysis RNN Model -- 4. Results -- 4.1. Dataset -- 4.2. Input Representation -- 4.3. Model Training -- 4.4. Evaluation Procedure -- 4.5. Performance Comparison -- 5. Discussion -- Acknowledgments -- References -- Bi-directional Recurrent Neural Network Models for Geographic Location Extraction in Biomedical Literature -- 1. Introduction -- 2. Related Work -- 3. Methods -- 3.1. Toponym Detection -- 3.1.1. Recurrent Neural Networks -- 3.1.2. LSTM -- 3.1.3. Other Gated RNN Architectures -- 3.1.4. Hyperparameter search and optimization -- 3.2. Toponym

Disambiguation -- 3.2.1. Building Geonames Index -- 3.2.2. Searching Geonames Index -- 4. Results and Discussion -- 4.1. Toponym Disambiguation -- 4.2. Toponym Resolution -- 5. Limitations and Future Work -- 6. Conclusion -- Acknowledgments -- Funding -- References -- Automatic Human-like Mining and Constructing Reliable Genetic Association Database with Deep Reinforcement Learning -- 1. Introduction -- 2. Related Work -- 3. Method -- 3.1. Model Framework -- 3.2. Deep Reinforcement Learning for Organizing Actions -- 3.3. Preprocessing and Name Entity Recognition with UMLS -- 3.4. Bidirectional LSTM for Relation Classification. 3.5. Algorithm -- 3.6. Implementation Specification -- 4. Experiments -- 4.1. Data -- 4.2. Evaluation -- 4.3. Results -- 4.3.1. Improved Reliability -- 4.3.2. Robustness in Real-world Situations -- 4.3.3. Number of Articles Read -- 5. Conclusions and Future Work -- 6. Acknowledgement -- References -- Estimating classification accuracy in positive-unlabeled learning: characterization and correction strategies -- 1. Introduction -- 2. Methods -- 2.1. Performance measures: definitions and estimation -- 2.2. Positive-unlabeled setting -- 2.3. Performance measure correction -- 3. Experiments and Results -- 3.1. A case study -- 3.2. Data sets -- 3.3. Experimental protocols -- 3.4. Results -- 4. Conclusions -- Acknowledgements -- References -- PLATYPUS: A Multiple-View Learning Predictive Framework for Cancer Drug Sensitivity Prediction -- 1. Introduction -- 2. System and methods -- 2.1. Data -- 2.2. Single views and co-training -- 2.3. Maximizing agreement across views through label assignment -- 3. Results -- 3.1. Preliminary experiments to optimize PLATYPUS performance -- 3.2. Predicting drug sensitivity in cell lines -- 3.3. Key features from PLATYPUS models -- 4. Conclusions -- Acknowledgments -- References -- Computational KIR copy number discovery reveals interaction between inhibitory receptor burden and survival -- 1. Introduction -- 2. Materials and Methods -- 2.1 Data collection -- 2.2 K-mer selection -- 2.3 NGS pipeline and k-mer extraction -- 2.4 Data cleaning -- 2.5 Normalization of k-mer frequencies -- 2.6 Copy number segregation and cutoff selection -- 2.7 Validation of copy number -- 2.8 Survival analysis -- 2.9 Additional immune analysis -- 3. Results and Discussions -- 3.1 Establishing unique k-mers -- 3.2 Varying coverage of KIR region by exome capture kit -- 3.3 Inference of KIR copy number -- 3.4 Population variation of the KIR region. 3.5 KIR inhibitory gene burden correlates with survival in cervical and uterine cancer -- 5. Conclusions -- 6. Acknowledgements -- 7. Supplementary Material -- References -- Exploring microRNA Regulation of Cancer with Context-Aware Deep Cancer Classifier -- 1. Introduction -- 2. Data -- 2.1. Preprocessing -- 3. Deep Cancer Classifier -- 3.1. Training & testing -- 3.2. Parameter tuning -- 3.3. Feature importance -- 4. Results and Discussion -- 4.1. Model selection -- 4.2. Classifier performance -- 4.3. Comparison with other methods -- 4.4. Feature importance -- 5. Conclusion -- References -- Implementing and Evaluating A Gaussian Mixture Framework for Identifying Gene Function from TnSeq Data -- 1. Introduction -- 1.1. TnSeq Motivation and Background -- 1.2. Motivation and New Methods -- 2. Methods -- 2.1. TnSeq Experimental Data -- 2.2. Mixture framework -- 2.3. Classification methods -- 2.3.1. Novel method - EM -- 2.3.2. Current method - t-statistic -- 2.3.3. Bayesian hierarchical model -- 2.3.4. Data partitioning for the Bayesian model -- 2.4. Simulation -- 2.5. Real data -- 3. Results -- 3.1.1. Classification rate -- 3.1.2. False positive rate -- 3.1.3. Positive classification rate -- 3.1.4. Cross entropy -- 3.2. Simulation Results -- 3.3. Comparisons on

real data -- 3.4. Software -- 4. Discussion -- References --  
SNPs2ChIP: Latent Factors of ChIP-seq to infer functions of non-coding  
SNPs -- 1. Introduction -- 2. Results -- 2.1. SNPs2ChIP analysis  
framework overview -- 2.2. Batch normalization of heterogeneous  
epigenetic features -- 2.3. Latent factor discovery and their biological  
characterization -- 2.4. SNPs2ChIP identifies relevant functions of the  
non-coding genome -- 2.4.1. Genome-wide SNPs coverage of the  
reference datasets -- 2.4.2. Non-coding GWAS SNPs of systemic lupus  
erythematosus -- 2.4.3. ChIP-seq peaks for vitamin D receptors.  
2.5. Robustness Analysis in the latent factor identification.

---

## Sommario/riassunto

The Pacific Symposium on Biocomputing (PSB) 2019 is an international, multidisciplinary conference for the presentation and discussion of current research in the theory and application of computational methods in problems of biological significance. Presentations are rigorously peer reviewed and are published in an archival proceedings volume. PSB 2019 will be held on January 3 – 7, 2019 in Kohala Coast, Hawaii. Tutorials and workshops will be offered prior to the start of the conference. PSB 2019 will bring together top researchers from the US, the Asian Pacific nations, and around the world to exchange research results and address open issues in all aspects of computational biology. It is a forum for the presentation of work in databases, algorithms, interfaces, visualization, modeling, and other computational methods, as applied to biological problems, with emphasis on applications in data-rich areas of molecular biology. The PSB has been designed to be responsive to the need for critical mass in sub-disciplines within biocomputing. For that reason, it is the only meeting whose sessions are defined dynamically each year in response to specific proposals. PSB sessions are organized by leaders of research in biocomputing's "hot topics." In this way, the meeting provides an early forum for serious examination of emerging methods and approaches in this rapidly changing field.

---