

Pacific Symposium on Biocomputing 2020

Fairmont Orchid, Hawaii, USA
3 - 7 January 2020

ISBN: 978-1-7138-2796-2

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

The full collection of papers in these proceedings is licensed under a Creative Commons Attribution 4.0 International License.

License details: <http://creativecommons.org/licenses/by/4.0/>.

The individual papers are © 2020 The Authors. Distributed under the terms of the Creative Commons Attribution Non-Commercial (CC BY-NC) 4.0 License

License details: <https://creativecommons.org/licenses/by-nc/4.0/>.

Printed with permission by Curran Associates, Inc. (2021)

World Scientific Publishing Co., Inc.
27 Warren Street, Suite 401-402
Hackensack, NJ 07601

Phone: 1-201-487-9655

Fax: 1-201-487-9656

www.worldscientific.com

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

ARTIFICIAL INTELLIGENCE FOR ENHANCING CLINICAL MEDICINE	1
<i>R. Daneshjou, L. Kidzinski, O. Afanasiev, J.H. Chen</i>	
GENOME GERRYMANDERING: OPTIMAL DIVISION OF THE GENOME INTO REGIONS WITH CANCER TYPE SPECIFIC DIFFERENCES IN MUTATION RATES	7
<i>A. Young, J. Chmura, Y. Park, Q. Morris, G. Atwal</i>	
PREDICTING LONGITUDINAL OUTCOMES OF ALZHEIMER'S DISEASE VIA A TENSOR- BASED JOINT CLASSIFICATION AND REGRESSION MODEL	19
<i>L. Brand, K. Nichols, H. Wang, H. Huang, L. Shen</i>	
ROBUSTLY EXTRACTING MEDICAL KNOWLEDGE FROM EHRs: A CASE STUDY OF LEARNING A HEALTH KNOWLEDGE GRAPH	31
<i>I.Y. Chen, M. Agrawal, S. Horng, D. Sontag</i>	
INCREASING CLINICAL TRIAL ACCRUAL VIA AUTOMATED MATCHING OF BIOMARKER CRITERIA	43
<i>J.W. Chen, C.A. Kunder, N. Bui, J.L. Zehnder, H.A. Costa, H. Stehr</i>	
ADDRESSING THE CREDIT ASSIGNMENT PROBLEM IN TREATMENT OUTCOME PREDICTION USING TEMPORAL DIFFERENCE LEARNING	55
<i>S. Harati, A. Crowell, H. Mayberg, S. Nemati</i>	
MULTICLASS DISEASE CLASSIFICATION FROM MICROBIAL WHOLE-COMMUNITY METAGENOMES	67
<i>S. Khan, L. Kelly</i>	
LITGEN: GENETIC LITERATURE RECOMMENDATION GUIDED BY HUMAN EXPLANATIONS	79
<i>A. Nie, A.L. Pineda, M.W. Wright, H. Wand, B. Wulf, H.A. Costa, R.Y. Patel, C.D. Bustamante, J. Zou</i>	
FROM GENOME TO PHENOME: PREDICTING MULTIPLE CANCER PHENOTYPES BASED ON SOMATIC GENOMIC ALTERATIONS VIA THE GENOMIC IMPACT TRANSFORMER	91
<i>Y. Tao, C. Cai, W.W. Cohen, X. Lu</i>	
AUTOMATED PHENOTYPING OF PATIENTS WITH NON-ALCOHOLIC FATTY LIVER DISEASE REVEALS CLINICALLY RELEVANT DISEASE SUBTYPES	103
<i>M. Vandromme, T. Jun, P. Perumalswami, J.T. Dudley, A. Branch, L. Li</i>	
MONITORING ICU MORTALITY RISK WITH A LONG SHORT-TERM MEMORY RECURRENT NEURAL NETWORK	115
<i>K. Yu, M. Zhang, T. Cui, M. Hauskrecht</i>	
MULTILEVEL SELF-ATTENTION MODEL AND ITS USE ON MEDICAL RISK PREDICTION	127
<i>X. Zeng, Y. Feng, S. Moosavinasab, D. Lin, S. Lin, C. Liu</i>	
IDENTIFYING TRANSITIONAL HIGH COST USERS FROM UNSTRUCTURED PATIENT PRO WRITTEN BY PRIMARY CARE PHYSICIANS	139
<i>H. Zhang, E. Candido, A.S. Wilton, R. Duchon, L. Jaakkimainen, W. Wodchis, Q. Morris</i>	

OBTAINING DUAL-ENERGY COMPUTED TOMOGRAPHY (CT) INFORMATION FROM A SINGLE-ENERGY CT IMAGE FOR QUANTITATIVE IMAGING ANALYSIS OF LIVING SUBJECTS BY USING DEEP LEARNING	151
<i>W. Zhao, T. Lv, R. Lee, Y. Chen, L. Xing</i>	
ON THE IMPORTANCE OF COMPUTATIONAL BIOLOGY AND BIOINFORMATICS TO THE ORIGINS AND RAPID PROGRESSION OF THE INTRINSICALLY DISORDERED PROTEINS FIELD.....	161
<i>L. Kurgan, P. Radivojac, J.L. Sussman, A.K. Dunker</i>	
DISORDERED FUNCTION CONJUNCTION: ON THE IN-SILICO FUNCTION ANNOTATION OF INTRINSICALLY DISORDERED REGIONS	171
<i>S. Ghadermarzi, A. Katuwawala, C.J. Oldfield, A. Barik, L. Kurgan</i>	
DE NOVO ENSEMBLE MODELING SUGGESTS THAT AP2-BINDING TO DISORDERED REGIONS CAN INCREASE STERIC VOLUME OF EPSIN BUT NOT EPS15	183
<i>N.S. Jagannathan, C.W.V. Hogue, L. Tucker-Kellogg</i>	
MODULATION OF P53 TRANSACTIVATION DOMAIN CONFORMATIONS BY LIGAND BINDING AND CANCER-ASSOCIATED MUTATIONS	195
<i>X. Liu, J. Chen</i>	
EXPLORING RELATIONSHIPS BETWEEN THE DENSITY OF CHARGED TRACTS WITHIN DISORDERED REGIONS AND PHASE SEPARATION	207
<i>R. Somjee, D.M. Mitrea, R.W. Kriwacki</i>	
MUTATIONAL SIGNATURES: ETIOLOGY, PROPERTIES, AND ROLE IN CANCER	219
<i>M.D.M. Leiserson, T.M. Przytycka, R. Sharan</i>	
PHYSIGS: PHYLOGENETIC INFERENCE OF MUTATIONAL SIGNATURE DYNAMICS	226
<i>S. Christensen, M.D.M. Leiserson, M. El-Kebir</i>	
TRACKSIGFREQ: SUBCLONAL RECONSTRUCTIONS BASED ON MUTATION SIGNATURES AND ALLELE FREQUENCIES.....	238
<i>C.F. Harrigan, Y. Rubanova, Q. Morris, A. Selega</i>	
IMPACT OF MUTATIONAL SIGNATURES ON MICRORNA AND THEIR RESPONSE ELEMENTS	250
<i>E. Stamoulakatou, P. Pinoli, S. Ceri, R.M. Piro</i>	
DNA REPAIR FOOTPRINT UNCOVERS CONTRIBUTION OF DNA REPAIR MECHANISM TO MUTATIONAL SIGNATURES	262
<i>D. Wojtowicz, M.D.M. Leiserson, R. Sharan, T.M. Przytycka</i>	
ONGOING CHALLENGES AND INNOVATIVE APPROACHES FOR RECOGNIZING PATTERNS ACROSS LARGE-SCALE, INTEGRATIVE BIOMEDICAL DATASETS.....	275
<i>S.N. Kobren, B. Beaulieu-Jones, C. Darabos, D. Kim, A. Verma</i>	
CLINICAL CONCEPT EMBEDDINGS LEARNED FROM MASSIVE SOURCES OF MULTIMODAL MEDICAL DATA.....	284
<i>A.L. Beam, B. Kompa, A. Schmaltz, I. Fried, G. Weber, N. Palmer, X. Shi, T. Cai, I.S. Kohane</i>	
ASSESSMENT OF IMPUTATION METHODS FOR MISSING GENE EXPRESSION DATA IN META-ANALYSIS OF DISTINCT COHORTS OF TUBERCULOSIS PATIENTS.....	296
<i>C.A. Bobak, L. McDonnell, M.D. Nemesure, J. Lin, J.E. Hill</i>	

TOWARDS IDENTIFYING DRUG SIDE EFFECTS FROM SOCIAL MEDIA USING ACTIVE LEARNING AND CROWD SOURCING.....	308
<i>S. Burkhardt, J. Siekiera, J. Glodde, M.A. Andrade-Navarro, S. Kramer</i>	
MICROVASCULAR DYNAMICS FROM 4D MICROSCOPY USING TEMPORAL SEGMENTATION.....	320
<i>S. Gur, L. Wolf, L. Golgher, P. Blinder</i>	
USING TRANSCRIPTIONAL SIGNATURES TO FIND CANCER DRIVERS WITH LURE.....	332
<i>D. Haan, R. Tao, V. Friedl, I.N. Anastopoulos, C.K. Wong, A.S. Weinstein, J.M. Stuart</i>	
PAGE-NET: INTERPRETABLE AND INTEGRATIVE DEEP LEARNING FOR SURVIVAL ANALYSIS USING HISTOPATHOLOGICAL IMAGES AND GENOMIC DATA.....	344
<i>J. Hao, S.C. Kosaraju, N.Z. Tsaku, D.H. Song, M. Kang</i>	
MACHINE LEARNING ALGORITHMS FOR SIMULTANEOUS SUPERVISED DETECTION OF PEAKS IN MULTIPLE SAMPLES AND CELL TYPES.....	356
<i>T.D. Hocking, G. Bourque</i>	
LEARNING A LATENT SPACE OF HIGHLY MULTIDIMENSIONAL CANCER DATA.....	368
<i>B. Kompa, B. Coker</i>	
SCALING STRUCTURAL LEARNING WITH NO-BEARS TO INFER CAUSAL TRANSCRIPTOME NETWORKS.....	380
<i>H.C. Lee, M. Danieleto, R. Miotto, S.T. Cherng, J.T. Dudley</i>	
PATHFLOWAI: A HIGH-THROUGHPUT WORKFLOW FOR PREPROCESSING, DEEP LEARNING AND INTERPRETATION IN DIGITAL PATHOLOGY.....	392
<i>J.J. Levy, L.A. Salas, B.C. Christensen, A. Sriharan, L.J. Vaickus</i>	
IMPROVING SURVIVAL PREDICTION USING A NOVEL FEATURE SELECTION AND FEATURE REDUCTION FRAMEWORK BASED ON THE INTEGRATION OF CLINICAL AND MOLECULAR DATA.....	404
<i>L. Neums, R. Meier, D.C. Koestler, J.A. Thompson</i>	
BAYESIAN SEMI-NONNEGATIVE MATRIX TRI-FACTORIZATION TO IDENTIFY PATHWAYS ASSOCIATED WITH CANCER PHENOTYPES.....	416
<i>S. Park, N. Kar, J.H. Cheong, T.H. Hwang</i>	
GRAPH-BASED INFORMATION DIFFUSION METHOD FOR PRIORITIZING FUNCTIONALLY RELATED GENES IN PROTEIN-PROTEIN INTERACTION NETWORKS.....	428
<i>M. Pham, O. Lichtarge</i>	
TREE-WEIGHTING FOR MULTI-STUDY ENSEMBLE LEARNERS.....	440
<i>M. Ramchandran, P. Patil, G. Parmigiani</i>	
A LITERATURE-BASED KNOWLEDGE GRAPH EMBEDDING METHOD FOR IDENTIFYING DRUG REPURPOSING OPPORTUNITIES IN RARE DISEASES.....	452
<i>D.N. Sosa, A. Derry, M. Guo, E. Wei, C. Brinton, R.B. Altman</i>	
PTR EXPLORER: AN APPROACH TO IDENTIFY AND EXPLORE POST TRANSCRIPTIONAL REGULATORY MECHANISMS USING PROTEOGENOMICS.....	464
<i>A. Srivastava, M. Sharpnack, K. Huang, P. Mallick, R. Machiraju</i>	
TWO-STAGE ML CLASSIFIER FOR IDENTIFYING HOST PROTEIN TARGETS OF THE DENGUE PROTEASE.....	476
<i>J.T. Stanley, A.R. Gilchrist, A.C. Stabell, M.A. Allen, S.L. Sawyer, R.D. Dowell</i>	

NETWORK REPRESENTATION OF LARGE-SCALE HETEROGENEOUS RNA SEQUENCES WITH INTEGRATION OF DIVERSE MULTI-OMICS, INTERACTIONS, AND ANNOTATIONS DATA	488
<i>N. Tran, J. Gao</i>	
ENHANCING MODEL INTERPRETABILITY AND ACCURACY FOR DISEASE PROGRESSION PREDICTION VIA PHENOTYPE-BASED PATIENT SIMILARITY LEARNING.....	500
<i>Y. Wang, T. Wu, G. Wang</i>	
HADOOP AND PYSPARK FOR REPRODUCIBILITY AND SCALABILITY OF GENOMIC SEQUENCING STUDIES	512
<i>N.R. Wheeler, P. Benchek, B.W. Kunkle, K.L. Hamilton-Nelson, M. Warfe, J.R. Fondran, J.L. Haines, W.S. Bush</i>	
CERENKOV3: CLUSTERING AND MOLECULAR NETWORK-DERIVED FEATURES IMPROVE COMPUTATIONAL PREDICTION OF FUNCTIONAL NONCODING SNPS	524
<i>Y. Yao, S.A. Ramsey</i>	
PRECISION MEDICINE: ADDRESSING THE CHALLENGES OF SHARING, ANALYSIS, AND PRIVACY AT SCALE	536
<i>S.E. Brenner, M.L. Bulyk, D.C. Crawford, A.A. Morgan, P. Radivojac, N.P. Tatonetti</i>	
INTEGRATED CANCER SUBTYPING USING HETEROGENEOUS GENOME-SCALE MOLECULAR DATASETS	540
<i>S. Arslanturk, S. Draghici, T. Nguyen</i>	
ANOMIGAN: GENERATIVE ADVERSARIAL NETWORKS FOR ANONYMIZING PRIVATE MEDICAL DATA.....	552
<i>H. Bae, D. Jung, H.S. Choi, S. Yoon</i>	
FREQUENCY OF CLINVAR PATHOGENIC VARIANTS IN CHRONIC KIDNEY DISEASE PATIENTS SURVEYED FOR RETURN OF RESEARCH RESULTS AT A CLEVELAND PUBLIC HOSPITAL.....	564
<i>D.C. Crawford, J. Lin, J.N.C. Bailey, T. Kinzy, J.R. Sedor, J.F. O'Toole, W.S. Bush</i>	
ASSESSMENT OF COVERAGE FOR ENDOGENOUS METABOLITES AND EXOGENOUS CHEMICAL COMPOUNDS USING AN UNTARGETED METABOLOMICS PLATFORM.....	576
<i>S.W. Kong, C. Hernandez-Ferrer</i>	
COVERAGE PROFILE CORRECTION OF SHALLOW-DEPTH CIRCULATING CELL-FREE DNA SEQUENCING VIA MULTI-DISTANCE LEARNING	588
<i>N.B. Larson, M.C. Larson, J. Na, C.P. Sosa, R. Rowsey</i>	
PGXMINE: TEXT MINING FOR CURATION OF PHARMGKB	600
<i>J. Lever, J.M. Barbarino, L. Gong, R. Huddart, K. Sangkuhl, R. Whaley, M. Whirl-Carrillo, M. Woon, T.E. Klein, R.B. Altman</i>	
NETWORK-BASED MATCHING OF PATIENTS AND TARGETED THERAPIES FOR PRECISION ONCOLOGY	612
<i>Q. Liu, M.J. Ha, R. Bhattacharyya, L. Garmire, V. Baladandayuthapani</i>	
THE POWER OF DYNAMIC SOCIAL NETWORKS TO PREDICT INDIVIDUALS' MENTAL HEALTH.....	624
<i>S. Liu, D. Hachen, O. Lizardo, C. Poellabauer, A. Striegel, T. Milenkovic</i>	

IMPLEMENTING A CLOUD BASED METHOD FOR PROTECTED CLINICAL TRIAL DATA SHARING	636
<i>G. Luthria, Q. Wang</i>	
PHENOME-WIDE ASSOCIATION STUDIES ON CARDIOVASCULAR HEALTH AND FATTY ACIDS CONSIDERING PHENOTYPE QUALITY CONTROL PRACTICES FOR EPIDEMIOLOGICAL DATA	648
<i>K. Passero, X. He, J. Zhou, B. Mueller-Myhsok, M.E. Kleber, W. Maerz, M.A. Hall</i>	
PATHWAY AND NETWORK EMBEDDING METHODS FOR PRIORITIZING PSYCHIATRIC DRUGS	660
<i>Y. Pershad, M. Guo, R.B. Altman</i>	
ATEMPO: PATHWAY-SPECIFIC TEMPORAL ANOMALIES FOR PRECISION THERAPEUTICS	672
<i>C.M. Pietras, L. Power, D.K. Slonim</i>	
ROBUST-ODAL: LEARNING FROM HETEROGENEOUS HEALTH SYSTEMS WITHOUT SHARING PATIENT-LEVEL DATA.....	684
<i>J. Tong, R. Duan, R. Li, M.J. Scheuemie, J.H. Moore, Y. Chen</i>	
FEATURE SELECTION AND DIMENSION REDUCTION OF SOCIAL AUTISM DATA.....	696
<i>P. Washington, K.M. Paskov, H. Kalantarian, N. Stockham, C. Voss, A. Kline, R. Patnaik, B. Chrisman, M. Varma, Q. Tariq, K. Dunlap, J. Schwartz, N. Haber, D.P. Wall</i>	
COMPUTATIONALLY EFFICIENT, EXACT, COVARIATE-ADJUSTED GENETIC PRINCIPAL COMPONENT ANALYSIS BY LEVERAGING INDIVIDUAL MARKER SUMMARY STATISTICS FROM LARGE BIOBANKS	708
<i>J.M. Wolf, M. Barnard, X. Xia, N. Ryder, J. Westra, N. Tintle</i>	
AI ETHICS AND VALUES IN BIOMEDICINE – TECHNICAL CHALLENGES AND SOLUTIONS.....	720
<i>D. Petkovic, L. Kobzik, R. Ghanadan</i>	
NAVIGATING ETHICAL QUANDARIES WITH THE PRIVACY DILEMMA OF BIOMEDICAL DATASETS.....	725
<i>G. Gursoy, M. Doerr, J. Wilbanks, J.K. Wagner, H. Tang, S.E. Brenner</i>	
PACKAGING BIOCOMPUTING SOFTWARE TO MAXIMIZE DISTRIBUTION AND REUSE	728
<i>W.S. Bush, N. Wheeler, B. Beaulieu-Jones, C. Darabos</i>	
TRANSLATIONAL BIOINFORMATICS: BIOBANKS IN THE PRECISION MEDICINE ERA.....	733
<i>M.D. Ritchie, J.H. Moore, J.H. Kim</i>	
MANY-TO-ONE BINDING BY INTRINSICALLY DISORDERED PROTEIN REGIONS	738
<i>W.L. Alterovitz, E. Faraggi, C.J. Oldfield, J. Meng, B. Xue, F. Huang, P. Romero, A. Kloczkowski, V.N. Uversky, A.K. Dunker</i>	

Author Index