CURRICULUM VITAE

RYAN J. URBANOWICZ, Ph.D.

PERSONAL

Home Mailing Address:	732 N Orlando Ave, Apt 201, Los Angeles CA 90069
	Mobile: 802-299-9461
Work Mailing Address:	700 N. San Vicente Blvd. Pacific Design Center Suite G541E West Hollywood, CA 90069
	Office: 310-432-3202 Primary Email: Ryan.Urbanowicz@cshs.org Secondary Email: ryanurb@upenn.edu Permanent Email: ryanurbanowicz@gmail.com Personal Website: www.ryanurbanowicz.com Twitter: www.twitter.com/DocUrbs YouTube: www.youtube.com/c/ryanurbanowicz Lab GitHub: https://github.com/UrbsLab Personal GitHub: https://github.com/ryanurbs
Date and place of birth:	January 28, 1982 New Milford, Connecticut, USA

EDUCATION

2005-2012	Dartmouth College Hanover, New Hampshire Mentor: Dr. Jason Moore, Department of Genetics	Ph.D. in Genetics/Computational Biology
2004-2005	Cornell University Ithaca, New York Mentor: Dr. Richard Durst, Department of Biological and Environmental Eng.	M.Eng. in Biological and Environmental Eng.
2000-2004	Cornell University Ithaca, New York Mentor: Dr. Ruth Collins, Department of Molecular Medicine	B.S. in Biological and Environmental Engineering

POSTDOCTORAL TRAINING

2012-2015	Dartmouth College	Postdoctoral Research Associate

Hanover, New Hampshire
Mentors: Dr. Jason H. Moore, Department of Genetics & Dr. Eugene Demidenko, Department of Biostatistics and Epidemiology

LICENSURE AND CERTIFICATION

none

ACADEMIC APPOINTMENTS

2021-present	Research Scientist II (Assistant professor tile pending internal approval), Department of Computational Biomedicine. Cedars-Sinai Medical Center, Los Angeles, California
2018-2021	Assistant Professor, Clinical Educator (CE), Division of Informatics of the Department of Biostatistics, Epidemiology, and Informatics. University of Pennsylvania, Philadelphia, Pennsylvania
2015-2018	Research Associate with Dr. Jason Moore, Division of Informatics of the Department of Biostatistics, Epidemiology, and Informatics. University of Pennsylvania, Philadelphia, Pennsylvania
2012-2015	Postdoctoral Research Associate with Dr. Jason Moore, Department of Genetics and Dr. Eugene Demidenko, Department of Biostatistics and Epidemiology, Dartmouth College, Hanover, New Hampshire
2013	Faculty of the Citizen Science Program at Bard College, Annandale-on-Hudson, New York
2006-2012	Graduate Research Assistant with Dr. Jason Moore, Department of Genetics, Computational Biology Laboratory, Geisel School of Medicine, Hanover, New Hampshire
2004-2005	Graduate Research Assistant with Dr. Richard A. Durst and Dr. Antje J. Baeumner, Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York
2004	Undergraduate Research Assistant with Dr. Ruth N. Collins, Department of Molecular Biology, Cornell University, College of Veterinary Medicine, Ithaca, New York
2003	Undergraduate Environmental Engineering Intern with Larry Marsicano and Joe Polulech (Engineer), Candlewood Lake Authority, Sherman, Connecticut
2000	High School Research Intern at Boehringer Ingleheim Pharmaceuticals, Ridgefield, Connecticut
1997-1999	High School Research Intern with Dr. Jacquline Burns, Cytogenetics Laboratory, Danbury Hospital, Danbury, Connecticut

OTHER PROFESSIONAL POSITIONS

none

MAJOR COMMITTEE ASSIGNMENTS AND CONSULTATIONS

International, National and Regional

2020-present	Consultant for AEgis – Rule-based ML for RIF Battle Manager
2021	NIH Review Panel for IMST-14: Small Business - Computational Modeling and Biodata Management (SBIR/STTR). July.
2019	Grant Reviewer, American Heart Association (AHA)
2015-2019	Cornell Alumni Admissions Ambassadors Network (CAAAN), Alumni Ambassador of the greater Philadelphia, Pennsylvania region.
2016	Lead-Organizer, International Workshop on Evolutionary Rule-Based Machine Learning at the Genetic and Evolutionary Computing Conference (GECCO'17), Berlin, Germany
2016	Co-Organizer, International Workshop on Evolutionary Rule-Based Machine Learning at the Genetic and Evolutionary Computing Conference (GECCO'16), Denver, Colorado
2015	Chair, Hot off the Press Session at the Genetic and Evolutionary Computing Conference (GECCO'15), Madrid, Spain
2014-2015	Co-Chair, Biological and Biomedical Applications Track at the Genetic and Evolutionary Computing Conference (GECCO'15), Madrid, Spain
2014-2015	Algorithm/Prediction Consultant, MedicFP, LLC
2010-2014	Co-Organizer, International Workshop on Learning Classifier Systems (IWLCS)
2013	Chair, Session on Genetics Based Machine Learning at the Genetic and Evolutionary Computing Conference (GECCO'12), Amsterdam, The Netherlands
2013	Co-Chair, International Workshop on Learning Classifier Systems (IWLCS) at the Genetic and Evolutionary Computing Conference (GECCO'13), Amsterdam, The Netherlands
2012	Chair, Session on Genetics Based Machine Learning at the Genetic and Evolutionary Computing Conference (GECCO'12), Philadelphia, Pennsylvania
2012	Co-Chair, International Workshop on Learning Classifier Systems (IWLCS) at the Genetic and Evolutionary Computing Conference (GECCO'12), Philadelphia, Pennsylvania
2011	Co-Chair, International Workshop on Learning Classifier Systems (IWLCS) at the Genetic and Evolutionary Computing Conference (GECCO'11), Dublin, Ireland
2009	Chair, Session on Session on Bioinformatics and Computational Biology at the Genetic and Evolutionary Computing Conference (GECCO'09), Montreal, Canada

Institutional

2019-present	Member of Graduate Group in Genetics and Computational Biology (GCB)
2020-present	Member of Graduate Group in Epidemiology and Biostatistics (GGEB)
2020-2021	GCB Retreat Organizing Committee

2019-2021	Actively participated in GCB qualifying exams, candidate interviews, and panel sessions.
2007-2008	Graduate Student Representative for the Molecular and Cellular Biology (MCB) Graduate Committee, Dartmouth College, Hanover, New Hampshire

MEMBERSHIPS IN PROFESSIONAL SOCIETIES

2019-present	Institute of Electrical and Electronics Engineers (IEEE)
2012-present	International Genetic Epidemiology Society (IGES)
2007-present	Association for Computing Machinery (ACM)
2007-present	American Society of Human Genetics (ASHG)
2006-present	International Society of Computational Biology (ISCB)

EDITORIAL BOARDS

2019-present	PLOS ONE Academic Editor
2019-2020	Guest Associate Editor for Frontiers in Genetics, Special Issue on Bioinformatics and Computational Biology
2012-2015	Guest Associate Editor for Evolutionary Intelligence, Special Issue on Advancements in Learning Classifier Systems

AWARDS AND HONORS

2016	Best Paper Award, (Co-Author) Genetic and Evolutionary Computation Conference (GECCO'16), Denver, Colorado
2016	Best Paper Award, (Co-Author) EvoBIO, Porto, Portugal
2012-2015	NIH R25 Post-Doctoral Training Grant CA134286
2012	Best Paper Award, (First Author) 2012 Translational Bioinformatics Conference, Jeju, South Korea
2010	Best Paper Award, (First Author) Genetic and Evolutionary Computation Conference (GECCO'10), Portland, Oregon
2009-2010	Graduate Research Fellowship from the William H. Neukom 1964 Institute for Computational Sciences at Dartmouth College
2008	Best Paper Nomination, (First Author) Genetic and Evolutionary Computation Conference (GECCO'18), Atlanta, Georga
2007	PSB Department of Energy Travel Award, Pacific Symposium on Biocomputing, Maui, Hawaii
2004	New Horizons Summer Research Grant, Cornell University, Ithaca, New York
2004	Teaching Research Specialist Grant, Cornell University, Ithaca, New York
2000	Eagle Scout, Boy Scouts of America, Sherman, Connecticut

REFEREE ACTIVITIES

2011-present referee for Genetic and Evolutionary Computation Conference (GECCO), International Workshop on Learning Classifier Systems (IWLCS) , and Pacific Symposium on Biocomputing (PSB), Evolutionary Intelligence (EI), Evolutionary Computation Journal (ECJ), International Journal of Molecular Sciences, American Journal of Human Genetics (AJHG), PeerJ Computer Science, Artificial Intelligence in Medicine (AIME), Big Data, Translational Bioinformatics Conference (TBC), Evostar, Bioinformatics, PLOS ONE, Journal of Biomedical Informatics (JBI), Journal of the American Medical Informatics Association (JAMIA), IEEE Transactions on Evolutionary Computing, Evolutionary Computation Journal (ECJ), Journal of Investigative Dermatology (JID), BioData Mining, Methods of Information in Medicine, International Joint Conference on Rules and Reasoning (RuleML+RR), Journal of Testing and Evaluation, Congress on Evolutionary Computing (CEC), Annals of Family Medicine, Machine Learning with Applications, Briefings in Bioinformatics, BMC Bioinformatics

MAJOR RESEARCH INTERESTS

As director of the Unbounded Research in Biomedical Systems (URBS) lab, my research focuses on the development of informatics, machine learning, artificial intelligence, data mining, and visualization methods, and their application to various biomedical data analyses, including genomics, other 'omics', and clinical data. This research is motivated by myriad challenges native to biomedical data including large-scale data, integrated data types, noisy and incomplete data, data imbalance, considering covariates, rare-variant analysis, and the need to detect and characterize complex pattens of associations such as epistasis and genetic heterogeneity. My research places particular emphasis on interpretability, reproducibility, flexibility, and artificial intelligence automation in machine learning methods and applications.

More specifically, my research has focused on (1) the development of rule-based machine learning approaches, i.e. learning classifier systems (LCS), for the interpretable modeling of complex, heterogeneous biomedical associations This work epitomizes my interest in developing methods which limit the number of assumptions made about the data, and instead allows the data to speak for itself for detecting complex or heterogeneous patterns. This work led to the development of 'ExSTraCS', an LCS specifically designed to address the challenges of biomedical data mining, and 'LCS-DIVE' an automated pipeline to identify candidate patient subgroups driven by heterogeneous feature subsets; (2) developing methods and software for the simulation of complex genetic models and associated datasets for testing, evaluating, and comparing machine learning strategies. This work led to the development of 'GAMETES' a software suite for the rapid customized generation of simulated single nucleotide polymorphism (SNP) datasets with complex patterns of association; (3) developing automated machine learning (AutoML) algorithms and pipelines to facilitate rigorous application of machine learning methodologies in a reproducible and user friendly manner. This work led to the development of three cutting edge AutoML approaches, TPOT, PennAI, and AutoMLPipe-BC; (4) the development of feature importance estimation algorithms to address the bottleneck of feature

selection in large scale data with complex patterns of association. This has led to the development of ReBATE a scikit-learn compatible software suite of Relief-based feature importance/selection algorithms that are sensitive to epistatic feature interactions.

In summary, my research interests lie at the intersection of genomics, biostatistics, epidemiology, machine learning, and computer science. I have adopted a quantitative biomedical research direction that embraces, rather than ignores, the complexity of the relationship between predictive factors and disease endpoints and emphasizes the importance of model interpretability.

RESEARCH FUNDING
(dollar amounts are annual direct costs)

U01-AI152690 (Keating, B) 09/30/20 – 06/30/25 1.2 Calendar
NIH $2,749,140.00
MHC and KIR Sequencing and Association Analyses in the iGeneTRAiN Studies
Our overarching aim of this application is to accelerate the discovery of MHC and KIR genetic variants underpinning transplant-related phenotypes/outcomes. To achieve this aim, we will use complementary genomic approaches (short-read and amplicon-based HLA Class I/II second generation sequencing, and long-read sequencing approaches). We will also employ a number of innovative analyses pipelines to maximize the associative potential of very large genetic (genotyping and sequencing) and phenotypic datasets that are derived from well-curated solid organ samples
Role: Co-Investigator

R01-LM010098 (Moore, J) 03/05/19 – 02/28/24 1.20 Calendar
NIH $1,882,624
Bioinformatics Strategies for Genome Wide Association Studies
Genome-wide association studies (GWAS) and next-generation sequencing are now commonplace despite a lack of comprehensive bioinformatics approaches for relating genotype to phenotype. The common method of analysis is to employ parametric statistics and then adjust for the large number of tests performed to limit false-positives. This agnostic approach is preferred by some because no assumptions are made about which genes or genomic regions might be important. The goal of our proposed research program continuation is to develop and evaluate a bioinformatics approach that analyzes genetic associations in the context of expert knowledge about biochemical pathways, gene function and experimental results using gene set enrichment (GSE) methods.
Role: Co-Investigator

N/A (Yoon, S) 09/01/18 – 08/31/22 0.12 Calendar
NSF $598,266
Professional Development Supports for Teaching Bioinformatics through Mobile Computing
The goal of this grant is to co-design and field test multidisciplinary high school courses organized around concepts in bioinformatics, computational literacies, and mobile learning activities.
Role: Co-Investigator

R01-HL143790 (Gehrman, P) 07/01/20 – 06/30/22 1.44 Calendar
NIH $2,865,429
Elucidation of Genetic Effects on Sleep and Circadian Traits
Sleep and circadian rhythms play critical roles in the maintenance of physical and mental health, with disturbances in these domains associated with major public health consequences. Leveraging a combination of genetic, functional and animal model approaches, we aim to characterize the genomic underpinnings of this trait. This study will fill a critical need to uncover their biologic mechanisms underlying these traits in order to improve diagnostic, preventive and treatment approaches.
Role: Co-Investigator

R01-HL134015 (Pack, A/Robishaw, J) 08/15/16 – 04/30/21 2.40 Calendar
NIH $3,151,287
Approaches to Genetic Heterogeneity of Obstructive Sleep Apnea
Obstructive sleep apnea (OSA) is now recognized as an extremely common disorder that results in excessive daytime sleepiness and an increased risk of crashes. It is also an independent risk factor for insulin resistance, hypertension, cardiovascular events, increased cancer rates and cancer mortality, and accelerates the progression of neurodegenerative disorders. OSA is a systemic disorder and the cyclical intermittent hypoxia that occurs during sleep affects every organ system. Obstructive sleep apnea is heritable, with first degree relatives of individuals with OSA having a two-fold increased risk of the disorder.
Role: Co-Investigator

R03-OD030600-01 (Taylor, D) 09/18/20 – 08/31/21 1.20 Calendar
NIH $363,306
Leveraging Common Fund Data for Feature Selection in Kids First Studies
This project will pilot a process for identifying multi-variant interactions contributing structural birth defect and childhood cancer disorders.
Role: Co-Investigator

P01-HD-070454 (Goldmuntz,E) 07/01/19 – 06/30/21 2.40 Calendar
NIH/Sub to CHOP/Einstein $374,455
Genes for Non-Syndromic Congenital Heart Disease
Congenital heart defects (CHDs) are the most common serious birth defect and a leading cause of infant mortality. Affected individuals who survive infancy require substantial medical care and experience lifelong morbidity and early mortality. Despite their prevalence and impact upon public health, the etiology of CHDs is poorly understood. This gap in understanding limits our ability to prevent CHDs, assess risk and predict outcome. Our long-term goal is to address this gap by defining the genetic basis of CTRDs and using this information to develop preventive measures and provide precise medical management based on genotype.
Role: Co-Investigator

R01-AI116794 (Moore,J) 03/01/16 – 02/28/21 1.20 Calendar
NIH $2,520,126
Biomedical Computing and Informatics Strategies for Infectious Disease Research

The overall goal of this proposal is to develop innovative new computational methods for the genetic analysis of infectious diseases. We will focus on the development of methods that are able to detect synergistic effects of multiple genetic variants regardless of whether they are rare of common in human populations. We will apply these methods to the study of HIV/AIDS vaccination response.
Role: Co-Investigator

W81XWH-16-PRCRP-CDA (Lynch,S/Moore,J) 08/01/17 – 07/31/20 0.60 Calendar
DOD/Sub to Fox Chase Ctr $13,888
Towards Precision Prevention: Testing a Novel Risk Prediction Algorithm in Pancreatic Cancer
In this career development award, Dr. Shannon Lynch, under the advice and mentorship of Dr. Rachael Stolzenberg-Solomon at the National Cancer Institute and Dr. Jason Moore and his study team at the University of Pennsylvania, will adapt a learning classifier system approach(LCS) from machine learning to evaluate the individual and joint effects of 100s of risk factors for pancreatic cancer in order to: 1) develop a more comprehensive pancreatic cancer prediction model; 2) empirically identify high risk subgroups of etiologic factors in 350 cases and 1400 controls from the Prostate, Lung, Colorectal, and Ovarian Cancer Trial (PLCO). This machine learning approach is unique because it can identify patient subgroups indicative of heterogeneous disease pathways. Thus, findings could provide the groundwork for improvements in prevention, screening and treatment for individuals with high risk factor profiles for pancreatic cancer.
Role: Co-Investigator

2015-2019 N/A (P.I. – Polsky)
PA CURE $3,006,8552.4 cal (beginning in 2018)
PA CURE: Smarter Big Data for a Healthy Pennsylvania

N/A (Polsky) 06/1/15 – 05/31/19 2.40 Calendar
PA CURE $3,006,855
PA CURE: Smarter Big Data for a Healthy Pennsylvania
The goal of this project is to improve the health of Pennsylvanians at an individual, community, and population level by changing the paradigm of medical care and health care delivery to predicting and preventing onset, exacerbation, and advancement of disease rather than principally reacting to clinical events with expensive treatments. Using medical record data with linkages to administrative claims, wearable monitor data, and social media data, we will develop algorithms to better predict clinical events in the hospital, at home, and in the community. The proposed project will expand an established and highly successful minority health services research training program to provide opportunities for training in big-data research to support the career development of under-represented minorities within the Commonwealth of Pennsylvania.
Role: Co-Investigator

TEACHING EXPERIENCE/CURRENT TEACHING RESPONSIBILITIES

<u>The Perelman School of Medicine at the University of Pennsylvania (UPenn)</u>

Courses and Lectures

2020-present	Course director and primary lecturer for 'Foundations of Artificial Intelligence' (BMIN 520)
2019-present	Lecturer for K12 Grant Program teaching Bioinformatics content to Philadelphia area high school teachers. (PI: Susan Yoon)
2020-present	Lecturer for Graduate Data Science Course (BMIN503/EPID600), UPenn (Evolutionary Algorithms) (Host: Blanca Himes, PhD)
2019-present	Lecturer for Introduction to Biomedical and Health Informatics (BMIN 505), UPenn (Host: Mary Regina Boland)
2021	Lecturer for Statistics for Genomics and Biomedical Informatics (GCB 533), UPenn (Host: Pablo Gonzalez-Camara)
2021	Lecturer for Seminar in Genomics (GCB/CAMB 752), UPenn (Host: Kara Maxwell, MD, PhD and Sharon Diskin, PhD)
2020-2021	Revised curriculum, lectures, assignments, and recorded/edited class lectures for 'Foundations of Artificial Intelligence' (BMIN 520)
2020	Lecturer for Tutorials in Cell and Molecular Biology (CAMB 698), UPenn (Host: Jason Moore, PhD)
2019-2020	Developed curriculum, course lectures, assignments, and answer keys for new course 'Foundations of Artificial Intelligence' (BMIN 520)
2018-2020	Module Director for Nature-Inspired Computing in Special Topics in Biomedical and Health Informatics (BMIN 504), UPenn (Host: Jason Moore, PhD)
2018-2019	Lecturer for Graduate Data Science Course (EPID600), UPenn (2 lectures: Evolutionary Algorithms and Enrichment Analysis) (Host: Blanca Himes, PhD)
2019	Lecturer for AI 2: Machine Learning (BMIN 522) (Host: Li Shen, PhD)
2018	Lecturer for 'Exploring Data Science Methods with Health Care" (NURS 849), School of Nursing, UPenn (1 lecture: Decision Trees and Random Forests) (Hosts: Kathryn Bowles and Michael Milo)
2015-2017	Teaching Assistant and Lecturer for Blanca Himes, PhD, Graduate Data Science Course (EPID600), UPenn
2017	Recorded Online Lecture modules on topic of 'Data Science' for Nursing 165: Nursing Informatics, UPenn
2017	Lecturer for Nursing Informatics (165), UPenn, (1 lecture: Data Science) (Host: Kathryn Bowles)

High-School Students Mentored

2021-present	Satvik Desariraju - Lawrenceville School – Student Researcher (Published)
2021-present	Advaith Hari - Peddie School (Contact: Shani Peretz) Advanced Science Research Program (EXP)
2021-present	Richard Zhang - Conestoga High School – Student Researcher
2021-present	Wilson Zhang - Conestoga High School – Student Researcher
2020	Michael (Suyu) Ye – Peddie School (Contact: Shani Peretz) Advanced Science Research Program (EXP) (Publication in preparation)

Undergraduate Students Mentored

2021-present	Arjun Shah, UPenn – (PURM Summer Intern)
2021-present	Riju Datta, UPenn – (PURM Summer Intern)
2021-present	Keshav Ramji, UPenn – (PURM Summer Intern)
2019-present	Robert Zhang, UPenn – (PURM Summer Intern) Engineering & Applied Sciences,Wharton/2022 (Published)
2020	Yijie Lu, UPenn – (PURM Summer Intern) Engineering & Applied Sciences, /2023
2020	Cindy Xinyu, UPenn – (PURM Summer Intern) Engineering & Applied Sciences,Wharton/2023
2017-2020	Siddharth Verma, University of Delhi - Computer Engineering, Delhi (published)
2018	Alex Xu, UPenn - Computer Science, Philadelphia, PA
2019	Pranshu Suri, UPenn – (PURM Summer Intern) Engineering & Applied Sciences,Wharton/2022 (Publication submitted)
2019	Seungmin (Kevin) Han – (VIPER Summer Intern) Engineering & Computer Science/2023 (Publication in prep)
2019	Gabrielle Hemlick, UPenn - School of Arts and Sciences/2021
2019	Sherjeel Arif, Upenn – Electrical Engineering/2020
2018-2019	Maria Turner, UPenn – Engineering and Computer Science, Philadelphia, PA
2016-2019	Christopher Lo, UPenn – Bioengineering, Philadelphia, PA (Published)
2017-2018	Riley Wong, UPenn – Computer Science, Philadelphia, PA
2017	James Xue, UPenn (Wharton) – Economics/Statistics, Philadelphia, PA
2016	Tuan A. Nguyen, Swarthmore College – Mathematics, Swarthmore, PA
2016-2017	Ben Yang, UPenn – Computational Biology, Philadelphia, PA (Published)
2017	Melissa Meeker, Ursinus College - Mathematics, Collegeville, PA (Published 2)

Masters Students Mentored

2019-2020	Ziquing Ye, UPenn – Biotechnology, Philadelphia PA (independent study)
2019-2020	Yili Du, UPenn – Biotechnology, Phildelphia PA (independent study)
2019-2020	Kai (Jeffery) Jin, UPenn – Biotechnology, Philadelphia, PA (independent study)
2018-2020	Piyush Borole, UPenn – Computer Science, Philadelphia, PA (post-grad RA)
2018-2019	Yuhan Cui, UPenn – Biotechnology, Philadelphia, PA (independent study)
2018	Saurav Bose, UPenn – Scientific Computing and Mechanical Engineering, Philadelphia, PA (highlight: programmer job at CHOP)
2016	Ted Fujimoto, UPenn – Computer Science, Philadelphia, PA
2016	Vishal Murali, UPenn - Mechanical Engineering, Philadelphia, PA

Graduate Students Mentored (Rotation or Other)

2019-present	Vivek Shiram – GCB Rotation Student (First in my new lab), Philadelphia PA

2019-present Alexa Woodward – (Head of Thesis Committee), GGEB Rotation Student, Philadelphia PA
2018-present Zihe (Emma) Zheng – (Thesis Committee Advisor) Epidemiology PhD Student, Philadelphia, PA
2018 Lin Xi, UPenn – MD student, Philadelphia, PA (research intern)
2018 John Gregg, UPenn – GCB Rotation Student, Philadelphia, PA

Ph.D. Committees Served On

2019-2021 Zihe (Emma) Zheng – Center for Clinical Epidemiology and Biostatistics (CCEB), UPenn
2020-present Alexa Woodward – Center for Clinical Epidemiology and Biostatistics (CCEB), UPenn (Committee Chair)
2020-present Yun Hao – Genetics and Computational Biology (GCB), UPenn
2021-present Francesca Mandel - Center for Clinical Epidemiology and Biostatistics (CCEB), UPenn
2020-2021 Pankhuri Singhal – Cell and Molecular Biology Graduate Group (CAMB) G&E, Left committee in 2021 due to student's decision to change research direction.

<u>The Geisel School of Medicine at Dartmouth</u>

Courses and Lectures

2013 Bioinformatics E-Learning Module Development: Hypothesis Testing
2008 Teaching Assistant, Undergraduate Biology (BIOL 11), Dartmouth College
2007 Center for the Advancement of Learning – Teaching Series, Dartmouth College

High School Students Mentored

2010-2015 Ambrose Granizo-MacKenzie, Hanover High School, Hanover, NH (Published 2)
2008-2009 Delaney Granizo-MacKenzie, Hanover High School, Hanover, NH (highlight: attended Princeton University). (Published)

Undergraduate Students Mentored

2014-2015 Niranjan Ramanand, class of 2017, Dartmouth College (highlight: student awarded Dartmouth Sophomore Science Scholarship) (Published)
2013 Dzung Pham, class of 2014 Saint Anselm College, NH, and class of 2015 University of Notre Dame, IN
2013 Ryan Amos, class of 2016, Dartmouth College
2013-2014 Gediminas Bertasius, class of 2014, Dartmouth College (Published)
2009-2013 Delaney Granizo-MacKenzie, Princeton University, NJ (Published)
2009-2012 Nicholas Sinnott-Armstrong, Brown University, RI (Published)
2012 Christine Cuddemi, SURF student, Emmanuelle College, MA

2010 Tyler Perry, class of 2012, Dartmouth College

Graduate Rotation Students Mentored

2014 Elizabeth Piette, Quantitative Biomedical Sciences (QBS), Dartmouth College
2013 Jie Tan, Molecular and Cellular Biology (MCB), Dartmouth College (Published)
2012-2013 James Rudd, Quantitative Biomedical Sciences (QBS), Dartmouth College (Published)

Masters Students Mentored

2009 Tamra Heberling, Mathematics, Montana State University (Published)

Bard College

Courses and Lectures

2013 Course Developer and Lecturer for "Citizen Science" (Required Freshman Course Teaching a Universal Scientific Skill Set - 4.5 Lecture Hours/Day for Three Weeks), Bard College, Annandale-on-Hudson, New York

Cornell University

Courses and Lectures

2005 Teaching Assistant, Personal Finance (HADM 3200), Cornell University
2004 Teaching Assistant, Physiological Engineering (BEE 4540), Cornell University

Other Teaching Experience

2019-2021 Facilitator for Responsible Conduct in Research (RCR) workshops
2021 Recorded and shared lectures for "Fundamentals of Artificial Intelligence" Course (BMIN 520) as YouTube Video Playlist: https://www.youtube.com/playlist?list=PLafPhSv1OSDew4qhV9YeNFwf-1eU0NHng
2000-2016 Adult Leader and Merit Badge Counselor for Troop 48 of the Boy Scouts of America (Engineering, Fire Safety, Emergency Preparedness, Reading, Scholarship, and Astronomy), Sherman, Connecticut
2016 Created Educational YouTube Video: 'Learning Classifier Systems in a Nutshell' https://www.youtube.com/watch?v=CRge_cZ2cJc
2008-2015 Science Educator for the "Camp-in at the Museum" Program, Astronomy/Star-Lab and Chemistry Stations, Montshire Children's Museum of Science, Norwich, Vermont

2008-2009 Science Mentor for the Montshire-Rivendell-Dartmouth Howard Hughes Medical Institute Science Camp, Norwich, Vermont

2002-2005 Substitute Teacher and Teaching Assistant for the Sherman Elementary School (K-8), Sherman, Connecticut

INVITED PRESENTATIONS

International, National and Regional

2021 "A Rigorous Automated Machine Learning Pipeline for Biomedical Classification" (Invited Speaker) Cedars-Sinai Health System (Host: David Ouyang). Los Angeles, CA. (Remote). November 10.

2021 "Automated Machine Learning Analysis Pipeline for Biomedical Data Mining" (Invited Speaker) Mid-Atlantic Bioinformatics Conference. (Host: Deanne Taylor). Philadelphia, PA. (Remote). November 8.

2021 "A Rigorous Automated Machine Learning Pipeline for Biomedical Classification" Research Informatics Seminar Series (Host: Eric Ross) Fox Chase Cancer Center. Philadelphia PA. (Remote). November 3.

2021 "Machine Learning Approach for Binning Donor-Recipient HLA Amino Acid Position Mismatches to Detect Association with Kidney Graft Failure" American Society for Histocompatibility and Immunogenetics Annual Meeting. (Abstract Invited Talk). Orlando FL. (Remote). September 30.

2021 "Learning Classifier Systems in an Automated Machine Learning Framework", RuleML Webinar Series. (Host: Tomas Kliegr) University of Economics Prague. (Remote). September 29.

2021 "AutoMLPipe-BC for Biomedical Data Mining" University of Kansas Medical Center (UKMC), KS (Host: Diego Mazzotti), (Remote), September 15.

2021 "Interpretable Machine Learning with Rule-Based Modeling" (Keynote for Declarative AI Conference 2021) Leuven, Belgium. (Remote). (Host: Sotiris Moschoyiannis), September 13.

2021 "RARE: Evolutionary Feature Engineering for Rare-variant Bin Discovery" GECCO Workshop on Evolutionary Algorithms for Problems with Uncertainty. Lille, France. (Remote). July 11.

2020 "Machine learning strategies for detecting epistatic and heterogeneous associations in precision histocompatibility". Topics in Histocompatibility and Transplantation Audio Seminar Series. Georgetown University, Washington DC. (Remote). Sept. 15

2020 "Interpretability challenges and opportunities in rule-based machine learning". Workshop Keynote. GECCO, IWLCS. Cancun, Mexico. (Remote). July 9.

2020 "Evolutionary Algorithms in Biomedical Data Mining: Challenges, Solutions, and Frontiers". Tutorial. GECCO, IWLCS. Cancun, Mexico. (Remote) July 8.

2020 "Coevolving Artistic Images using OMNIREP" Evo* - EuroMUSART, (remote pre-recorded lecture for digital conference) Seville, Spain. (Remote) April 16.

2020 "Epistasis and machine learning" Invited talk for EDGE meeting. Key West, FL, Feb. 3.

2019 "Machine learning strategies for detecting epistatic and heterogeneous associations in precision histocompatibility" American Society for Histocompatibility and Immunogenetics, Pittsburgh, PA. Sept 25.

2019 "The ExSTraCS Machine Learning Algortihm" (Remote presentation). GECCO, IWLCS. Prague, Czech Republic. July 13.

2019 "Adverse Events and Medical History: Data Harmonization for CMREF" (Remote presentation) Pulmonary Vascular Research Institute (PVRI) 6th Annual Drug Discovery and Development Symposium. Paris, France. July 1.

2019 "Learning Classifier Systems in Bio/Medical/Clinical – Informatics" Victoria University of Wellington, New Zealand. June 14.

2019 "Solution and fitness evolution (SAFE): A study of multiobjective problems" IEEE Computational Evolutionary Conference (CEC), Wellington, New Zealand. June 11.

2019 "Solution and fitness evolution (SAFE): Coevolving solutions and their objective functions" Evo* - EuroGP, Leipzig, Germany. April 25.

2018 "Attribute Tracking: Strategies towards improved detection and characterization of complex associations" Genetic and Evolutionary Computing Conference (GECCO'18), Kyoto, Japan, July 17.

2018 Tutorial: "Introducing Learning Classifier Systems: Rules that Capture Complexity", Genetic and Evolutionary Computing Conference (GECCO'18), Kyoto, Japan, July 15.

2018 "A Machine Learning Aperitif", invited speaker, NIMBioS Bioacoustic Workshop, Knoxville, TN, June 25.

2017 "Avoiding Common Machine Learning Mistakes", DataPhilly Meetup, Philadelphia, PA, November 7. (https://www.meetup.com/DataPhilly/events/244140696/)

2017 Tutorial: "Introducing Rule-Based Machine Learning: Capturing Complexity", Genetic and Evolutionary Computing Conference (GECCO'17). Berlin, Germany, July 15.

2017 "Co-evolving Genetic Programming Trees and Rules in a Learning Classifier System for Problem Driven Machine Learning" Genetic Programming Theory and Practice (GPTP). Ann Arbor, Michigan, May 18.

2017 Webinar - "An Introduction to Bioinformatics" for multi-site T32 on genetics of sleep and sleep disorders (John Hopkins, University of Michigan, and Stanford), Philadelphia, PA, April 24.

2017 "An Introduction to Data Science", invited lecturer, Nursing Informatics 651, UPenn, Philadelphia, PA, April 3.

2017 "Feature Selection with REBATE", invited speaker, Epistasis Discovery in Genetics and Epidemiology (EDGE) workshop, Key West, Florida, February 9. (Host: Marylyn Ritchie).

2016 "Pareto-Inspired Multi-Objective Rule Fitness for Adaptive Rule-Based Machine Learning", International Workshop on Evolutionary Machine Learning at the Genetic and Evolutionary Computing Conference (GECCO'16). Denver, Colorado, July 20.

2016 Tutorial: "Introducing Rule-Based Machine Learning: Capturing Complexity", Genetic and Evolutionary Computing Conference (GECCO'16). Denver, Colorado, July 20.

2016 "A New 'Front' in Rule-Based Machine Learning", invited speaker, Epistasis Discovery in Genetics and Epidemiology (EDGE) workshop, Key West, Florida, February 4. (Host: Marylyn Ritchie).

2015 “ExSTraCS 2.0: Description and Evaluation of a Scalable Learning Classifier System” Hot Off The Press Selection at the Genetic and Evolutionary Computing Conference (GECCO’15). Madrid, Spain, July 13.

2015 “Continuous Endpoint Data Mining with ExSTraCS: A Supervised Learning Classifier System", International Workshop on Evolutionary Machine Learning at the Genetic and Evolutionary Computing Conference (GECCO’15). Madrid, Spain, July 12.

2015 Tutorial: “Introducing Rule-Based Machine Learning: A Practical Guide” Co-presented with Will Browne PhD, Genetic and Evolutionary Computing Conference (GECCO’15). Madrid, Spain, July 11.

2015 “Modeling Complex Disease Associations with ExSTraCS: A Study of Scalability”, invited speaker, Epistasis Discovery in Genetics and Epidemiology (EDGE) workshop, Key West, Florida, February 5. (Host: Marylyn Ritchie).

2014 “An Introduction to Learning Classifier Systems", invited speaker for Evolutionary Computation Course, University of Vermont, October 23. (Host: Margaret J. Eppstein).

2014 “The Role of Genetic Heterogeneity and Epistasis in Bladder Cancer Susceptibility and Outcome: A Learning Classifier System Approach” Hot Off The Press Selection at the Genetic and Evolutionary Computing Conference (GECCO’14). Vancouver BC, Canada, July 15.

2014 “ExSTraCS: An Extended Michigan-Style Learning Classifier System for Flexible Supervised Learning Classification and Data Mining", International Workshop on Learning Classifier Systems at the Genetic and Evolutionary Computing Conference (GECCO’14). Vancouver BC, Canada, July 12.

2013 “An Introduction to Learning Classifier Systems", invited speaker for Evolutionary Computation Course, University of Vermont, October 24. (Host: Margaret J. Eppstein).

2013 “A Simple Multi-Core Parallelization Strategy for Learning Classifier System Evaluations", International Workshop on Learning Classifier Systems at the Genetic and Evolutionary Computing Conference (GECCO’13). Amsterdam, The Netherlands, July 6.

2013 “ExSTraCS: A Genetics-Based Supervised Machine Learning Tool for Epidemiological Data Mining", International Workshop on Learning Classifier Systems at the Genetic and Evolutionary Computing Conference (GECCO’13). Amsterdam, The Netherlands, July 6.

2013 Tutorial: “Introduction to Learning Classifier Systems” Co-presented with Will Browne PhD, Genetic and Evolutionary Computing Conference (GECCO’13). Amsterdam, The Netherlands, July 6.

2013 “Classification and Data Mining in Genetic Epidemiology: A Learning Classifier System Approach”, invited speaker for Bioinformatics course, University of Southern Maine, April 11. (Host: Clare B. Congdon).

2013 “GAMETES: Complex Genetic Model Simulation”, invited speaker, Epistasis Discovery in Genetics and Epidemiology (EDGE) workshop, Key West, Florida, January 31. (Host: Jason Moore).

2013 “Tackling Complexity in Common Human Disease Research: Machine Learning Disease Risk”, Invited Faculty Lecture, Bard College, Annandale-on-Hudson, New York, January 22. (Host: Amy Savage)

2012 “The Role of Genetic Heterogeneity and Epistasis in Bladder Cancer Susceptibility and Outcome: A Learning Classifier System Approach”, Translational Bioinformatics Conference, Jeju, South Korea, Oct 13-16.

2012 “Instance-Linked Attribute Tracking and Feedback for Michigan-Style Supervised Learning Classifier Systems”, Genetic and Evolutionary Computing Conference (GECCO’12), Genetics Based Machine Learning Track, Philadelphia, Pennsylvania, July 7-11.

2012 “An Analysis Pipeline with Visualization-Guided Knowledge Discovery for Michigan Style Learning Classifier Systems: Interpreting the Black Box”, International Workshop on Learning Classifier Systems at the Genetic and Evolutionary Computing Conference (GECCO’12), Philadelphia, Pennsylvania, July 7-11.

2011 “Random Artificial Incorporation or Noise in a Learning Classifier System Environment", International Workshop on Learning Classifier Systems: Genetic and Evolutionary Computing Conference (GECCO’11), Dublin, Ireland, July 12-16.

2010 “The Application of Michigan-Style Learning Classifier Systems to Address Genetic Heterogeneity and Epistasis in Association Studies", Genetic and Evolutionary Computing Conference (GECCO’10), Bioinformatics Track. Portland, Oregon, July 7-11.

2008 “Mask Functions for the Symbolic Modeling of Epistasis Using Genetic Programming", Genetic and Evolutionary Computing Conference (GECCO’08), Bioinformatics Track. Atlanta, GA, July 12-16.

Institutional

2021 “Introduction to Boostrap and Randomization” For Statistics for Genomics and Biomedical Informatics (GCB 533), UPenn (Host: Pablo Gonzalez-Camara), Dec. 2.

2021 “Machine Learning for Big Data in Biomedical Research” For Seminar in Genomics (GCB 752), UPenn (Host: Sharon J. Diskin & Kara Maxwell), April 29.

2021 “Should Journals Pay Peer Reviewers?”, Lunch and Learn Series, UPenn, April 5.

2021 “Feature Selection and ReBATE”, Lunch and Learn Series, UPenn, March 21.

2021 “Machine Learning Strategies for Mining Complex Patterns from Biomedical Data” – Departmental Flash Talk. March 16.

2020 “URBS Lab Research”, GCB Orientation Talk. August 28.

2020 “Frontiers of Machine Learning and AI in Medicine” UPenn School of Dentistry (Host: Hyun Koo) – Informational talk/meeting to foster collaboration. Feb 24.

2020 “Expert systems and PyKE” Led by my student (Ziquing Ye). Lunch and Learn Series, UPenn. Feb 24.

2020 “Machine Learning: Building an Analysis Pipeline” Health Policy Analytics Seminar Series (MEHP) (Host: Jingsan Zhu) Jan. 8.

2019 “History of Artificial Intelligence” Lunch and Learn Series, UPenn, October 28.

2019 “Evolutionary Machine Learning” For AI III: Advanced Methods and Health Applications in Machine Learning (BMIN 522) UPenn (Host: Li Shen) October 17.

2019 “Foundations to Frontiers of Biomedical Data Mining” GCB Orientation talk. Sept. 20.

2019 “Data Harmonization of Adverse Events and Medical History Terms in Clinical Trials Targeting Arterial Hypertension” Natural Language Processing (NLP) Seminar Series (Host: Graciela Gonzalez) July 11.
2019 “Gene by Environment Associations” K12 Bioinformatics Teaching. (PI: Susan Yoon) July 10.
2019 “Indoor Air Pollution” K12 Bioinformatics Teaching. (PI: Susan Yoon) July 10.
2019 “Outdoor Air Pollution” K12 Bioinformatics Teaching. (PI: Susan Yoon) July 10.
2019 “Exposome” K12 Bioinformatics Teaching. (PI: Susan Yoon) July 10.
2019 “Introduction to Machine Learning” Workshop speaker. (Host: Helen Loeb) Children’s Hospital of Philadelphia (CHOP). June 19.
2019 “Advanced machine learning for biomedical/clinical informatics” (Host: Elizabeth Goldmuntz) P01 grant retreat. Children’s Hospital of Philadelphia (CHOP). June 3.
2019 “Machine learning strategies targeting personalized medicine and complexity” Informatics day invited research talk. (Host: John Holmes) May 23.
2019 “Introduction to undergraduate research” Upenn, May 17.
2019 “Machine Learning: An Analysis Pipeline” Workshop Lecture organized jointly by IBI and LDI. (Host: PA CURE Steering Committee), May 1.
2019 “Solution and fitness evolution (SAFE): Coevolving solutions and their objective functions” Lunch and Learn Series, UPenn, April 8.
2019 “Informatics Methods: Machine Learning and AI” Guest Lecture for Introduction to Biomedical and Health Informatics (BMIN 505) UPenn (Host: Mary Boland PhD), Feb 5.
2019 “Machine Learning for Biomedical Informatics” Recruitment Lecture for GCB graduate group (Host: Ben Voight PhD), January 11.
2018 “An Introduction to Machine Learning” Workshop Lecture organized jointly by IBI and LDI. (Host: PA CURE Steering Committee), December 17.
2018 “Introducing Decision Trees and Random Forests” Guest Lecture for Exploring Data Science Methods with Health Care Data (NURS 849) at UPenn (Hosts: Kathryn Bowles and Michael Milo), November 28.
2018 “Functional Annotation and Enrichment Analysis” Guest Lecture for Data Science (BMIN 503/EPID600) at UPenn (Host: Blanca Himes PhD), November 15.
2018 “Evolutionary Algorithms and Rule-Based Machine Learning” Guest Lecture for Data Science (BMIN 503/EPID600) at UPenn (Host: Blanca Himes PhD), October 16.
2018 “Feature Selection and ReBATE” Lunch and Learn Series, UPenn, August 13.
2018 “An Introduction to Machine Learning: Practical Guidelines and Algorithm Selection” Lunch and Learn Series, UPenn, May 21.
2017 “Evolutionary Algorithms and Rule-Based Machine Learning” Guest Lecture for Data Science (BMIN 503/EPID600) at UPenn (Host: Blanca Himes PhD), November 16
2017 “Functional Annotation and Enrichment Analysis” Guest Lecture for Data Science (BMIN 503/EPID600) at UPenn (Host: Blanca Himes PhD), November 14 .
2017 “A New Paradigm for Data Mining in Bioinformatics: Embracing Genetic Heterogeneity” Invited candidate faculty seminar, UPenn (Host: John H. Holmes PhD), June 1.
2016 “Functional Annotation and Enrichment Analysis” Guest Lecture for Data Science (EPID600) at UPenn (Host: Blanca Himes PhD), November 17.

2016 “An Introduction to Rule-Based Machine Learning”, Lunch and Learn Series, UPenn, October 31.
2016 “Evolutionary Algorithms and Rule-Based Machine Learning” Guest Lecture for Data Science (EPID600) at UPenn (Host: Blanca Himes PhD), October 27.
2016 “Simulation Studies: Model/Dataset Generation and Study Design” Lunch and Learn Series, UPenn, June 13
2016 “Relief-Based Algorithms”, Lunch and Learn Series, UPenn, April 25.
2016 “An Introduction to Machine Learning: Practical Guidelines and Algorithm Selection”, Lunch and Learn Series, UPenn, January 11.
2015 “Functional Annotation and Enrichment Analysis” Guest Lecture for Data Science (EPID600) at UPenn (Host: Blanca Himes PhD), November 12.
2014 “Modeling Complex Disease Associations with ExSTraCS: A Study of Scalability”, Pizza Talk Series, Dartmouth College, December 12.
2014 “The Rise of Machine Learning: Advancing Epidemiological Data Mining, Prediction, and Knowledge Discovery”, Invited Faculty Interview Talk, Dartmouth College, April 8.
2013 “GAMETES: Software for Simulating Complex Epistatic Models and Heterogeneous Datasets for the Advancement of Algorithm Development", Pizza Talk Series, Dartmouth College, December 13.
2013 “ExSTraCS: The Development of a Genetics Based Supervised Machine Learning Tool for Epidemiological Data Mining", Pizza Talk Series, Dartmouth College, February 22.
2012 “The Detection and Characterization of Epistasis and Heterogeneity: A Learning Classifier System Approach", Thesis Defense, Dartmouth College. Feb. 20.
2011 “The Detection, Characterization, and Modeling of Genetic Heterogeneity and Epistasis: A Learning Classifier System Approach", MCB Research in Progress, Dartmouth College. Sept. 28.
2010 “The Detection, Characterization, and Modeling of Genetic Heterogeneity and Epistasis: A Learning Classifier System Approach", Pizza Talk Series, Dartmouth College, February 11.
2010 “The Detection, Characterization, and Modeling of Genetic Heterogeneity and Epistasis: A Learning Classifier System Approach", Institute for Quantitative Biomedical Sciences Applicant, Dartmouth College, December 20.
2010 “The Detection, Characterization, and Modeling of Genetic Heterogeneity and Epistasis: A Learning Classifier System Approach", MCB Research in Progress, Dartmouth College, November, 10.
2010 “The Detection, Characterization, and Modeling of Genetic Heterogeneity and Epistasis: A Learning Classifier System Approach", MCB Research in Progress, Dartmouth College. February, 17.
2009 “The Detection, Characterization, and Modeling of Genetic Heterogeneity and Epistasis", MCB Research in Progress, Dartmouth College. February, 11.
2008 “The Detection, Characterization, and Modeling of Genetic Heterogeneity and Epistasis", MCB Research in Progress, Dartmouth College. April, 9.
2007 “Mask Functions for Symbolic Modeling of Epistasis", MCB Research in Progress, Dartmouth College. May 23.

BIBLIOGRAPHY

Refereed Publications

1. **Urbanowicz, R.J.**, White, B.C., Barney, N., Moore, J.H. Mask functions for the symbolic modeling of epistasis using genetic programming. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, pp. 339-346 (2008). (Highlight: nominated for Best Paper Award, Bioinformatics and Computational Biology Track).
2. **Urbanowicz, R.J.**, Moore, J.H. Learning classifier systems: A complete introduction, review and roadmap. Journal of Artificial Evolution and Applications 2009, 1-25 (2009). (Highlight: cited 70 times as of November 2013)
3. **Urbanowicz, R.J.**, Moore, J.H. The application of Michigan-style learning classifier systems to address genetic heterogeneity and epistasis in association studies. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, pp. 195-202 (2010). (Highlight: won best paper award at GECCO'10).
4. **Urbanowicz, R.J.**, Moore, J.H. The application of Pittsburgh-style learning classifier systems to address genetic heterogeneity and epistasis in association studies. Springer Lecture Notes in Computer Science pp. 195-202 (2010).
5. **Urbanowicz, R.J.**, Sinnott-Armstrong, N., Moore, J.H. Random artificial incorporation of noise in a learning classifier system environment. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, pp. 369-374 (2011).
6. **Urbanowicz, R.J.**, Granizo-Mackenzie, A., Moore, J.H. An analysis pipeline with statistical and visualization-guided knowledge discovery for Michigan-style learning classifier systems. Computational Intelligence Magazine 7, 35-45 (2012).
7. **Urbanowicz, R.J.**, Granizo-Mackenzie, A., Moore, J.H. Instance-linked attribute tracking and feedback for Michigan-style supervised learning classifier systems. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, pp. 927-934 (2012).
8. **Urbanowicz, R.J.**, Granizo-Mackenzie, A., Moore, J.H. Using expert knowledge to guide covering and mutation in a Michigan-style learning classifier system to detect epistasis and heterogeneity. Springer Lecture Notes in Computer Science 7491, 266-275 (2012).
9. **Urbanowicz, R.J.**, Kiralis, J., Sinnott-Armstrong, N.A., Heberling, T., Fisher, J.M., Moore, J.H. GAMETES: A fast, direct algorithm for generating pure, strict, epistatic models with random architectures. BioData Mining 5, 16 (2012). (Highlight: Top cited paper published in BioData Mining since 2012)
10. **Urbanowicz, R.J.**, Kiralis, J., Fisher, J.M., Moore, J.H. Predicting the difficulty of pure, strict, epistatic models: Metrics for simulated model selection. BioData Mining 5, 15 (2012).
11. **Urbanowicz, R.J.**, Andrew, A.S., Karagas, M.R., Moore, J.H. The role of genetic heterogeneity and epistasis in bladder cancer susceptibility and outcome: A learning classifier system approach. Journal of the American Medical Informatics Association, 20(4), 603-612 (2013). (Highlight: Won a best paper award at the 2013 Translational Bioinformatics Conference, South Korea)
12. Rudd, J., Moore, J.H., **Urbanowicz, R.J.** A simple multi-core parallelization strategy for learning classifier system evaluations. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, 1259-1266 (2013).

13. Tan, J., Moore, J.H., **Urbanowicz, R.J.** Rapid rule compaction for knowledge discovery in a supervised learning classifier system. Advances in Artificial Life, 12. 110-117 (2013).
14. Rudd, J., Moore, J.H., **Urbanowicz, R.J.** A multi-core parallelization strategy for statistical significance testing in learning classifier systems. Evolutionary Intelligence. 6(2), 127-134 (2013).
15. **Urbanowicz, R.J.**, Granizo-Mackenzie, A., Kiralis, J., Moore, J.H. A classification and characterization of two-locus pure, strict epistatic models for simulation and detection. BioData Mining. 7(1), 8 (2014).
16. **Urbanowicz, R.J.**, Bertasius, G., Moore, J.H. An extended michigan-style learning classifier system for flexible supervised learning, classification, and data mining. Springer Lecture Notes in Computer Science 8672, 211-221 (2014).
17. **Urbanowicz, R.J.**, Moore, J.H. ExSTraCS 2.0: Description and evaluation of a scalable learning classifier system. Evolutionary Intelligence. 8(2-3), 89-116 (2015). (Highlight: Solved the extremely complex 135-bit benchmark multiplexer problem directly for the first time reported in literature).
18. **Urbanowicz, R.J.,** Ramanand, N., Moore, J.H., Continuous endpoint data mining with ExSTraCS. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, 1029-1036 (2015).
19. **Urbanowicz, R.J.,** Moore, J.H., Retooling fitness for noisy problems in a supervised Michigan-style learning classifier system. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, 591-598 (2015).
20. Olson, R.S, **Urbanowicz, R.J.,** Moore, J.H., Automating biomedical data science through tree-based pipeline optimization. Springer Lecture Notes in Computer Science 9597, 123-137 (2016). (Highlight: Won a best paper award in the EvoBIO track)
21. **Urbanowicz, R.J.,** Olson, R.S, Moore, J.H., Pareto inspired multi-objective rule fitness for noise-adaptive rule-based machine learning. Springer Lecture Notes in Computer Science 9921, 514-524 (2016).
22. Olson, R.S, **Urbanowicz, R.J.,** Moore, J.H., Evaluation of a tree-based pipeline optimization tool for automating data science. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, 485-492 (2016). (Highlight: Won a best paper award in the Evolutionary Machine Learning Track at GECCO'16)
23. Olson, R.S., LaCava W., Orzechowski, P., **Urbanowicz, R.J.,** Moore, J.H. PMLB: A large benchmark suite for machine learning evaluation and comparison. BioData Mining, 10(1): 36 (2017)
24. Olson, R.S., Sipper, M., LaCava, W., Tartarone, S., Vitale S., Fu, W., Orzechowski, P., **Urbanowicz, R.J.**, Holmes, J.H., Moore, J.H. A system for accessible artificial intelligence. Genetic Programming Theory and Practice XV. Springer, Cham, 121-134. (2018)
25. **Urbanowicz, R.J.**, Yang, B., Moore, J.H. Problem driven machine learning by co-evolving genetic programming trees and rules in a learning classifier system. Genetic Programming Theory and Practice XV. Springer, Cham, 55-71 (2018)
26. Verma, S.S., Lucas, A., Zhang, X., Veturi, Y., Dudek, S., Li, B., Li, R., **Urbanowicz, R.J**, Moore, J.H., Kim, D., Ritchie, M.D. Collective feature selection to identify crucial epistatic variants. BioData mining, 11(1): 5 (2018)
27. **Urbanowicz, R.J.**, Meeker, M., LaCava, W., Olson, R.S., Moore, J.H. Relief-based feature selection: Introduction and review. Journal of Biomedical Informatics. 85, 189-203 (2018) (PMID:30031057)

28. **Urbanowicz, R.J.**, Olson, R.S., Schmitt, P., Meeker, M., Moore, J.H. Benchmarking Relief-based feature selection methods. Journal of Biomedical Informatics. 85, 168-188 (2018) (PMID:30030120)
29. Le, T. T., **Urbanowicz, R. J.**, Moore, J. H., & McKinney, B. A.. Statistical Inference Relief (STIR) feature selection. Bioinformatics. (2018) (PMID:30239600)
30. **Urbanowicz, R.J.**, Lo, C., Holmes J.H, Moore, J.H. Attribute tracking: strategies towards improved detection and characterization of complex associations. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, 553-560 (2018)
31. Sipper, M., **Urbanowicz, R.J.,** Moore J.H. Editorial: To know the objective is not (necessarily) to know the objective function. BioData Mining. 21 (2018)
32. Moore, J.H., Boland M.R., Camara, P.G., Chervitz, H., Gonzalez, G., Himes, B.B., Kim, D., Mowery D.L., Ritchie, M.D., Shen, L., **Urbanowicz, R.J.**, Holmes, J.H. Preparing next-generation scientists for biomedical big data: Artificial intelligence approaches. Personalized Medicine (0) (2019)
33. Sipper, M., Moore J.H., **Urbanowicz, R.J.** Solution and Fitness Evolution (SAFE): Coevolving Solutions and Their Objective Functions. European Conference on Genetic Programming. Springer. 146-161 (2019)
34. Lo, Y., Lynch, S.F., **Urbanowicz, R.J.,** Olson, R.S., Ritter, A.Z., Whitehouse, C.R., O'Connor, M., Keim, S.K., McDonald, M., Moore, J.H. and Bowles, K.H., Using Machine Learning on Home Health Care Assessments to Predict Fall Risk. Studies in health technology and informatics, 264, pp.684-688. (2019)
35. Sipper, M., Moore J.H., **Urbanowicz, R.J.** Solution and Fitness Evolution (SAFE): A Study of Multiobjective Problems. IEEE Congress on Evolutionary Computation (CEC) 1868-1874, (2019)
36. Sipper, M., Moore J.H., **Urbanowicz, R.J.** New Pathways in Coevolutionary Computation. Genetic Programming Theory and Practice XVII. Springer pp. 295-305 (2020)
37. Liu, Y., Huang, J., **Urbanowicz, R.J.,** Chen, K., Manduchi, E., Greene, C.S., Moore, J.H., Scheet, P. and Chen, Y., Embracing study heterogeneity for finding genetic interactions in large-scale research consortia. Genetic epidemiology, 44(1), pp.52-66. (2020)
38. Verma, S., Borole, P., **Urbanowicz, R.J**. Evolving genetic programming trees in a rule-based learning framework. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, pp. 233-234. (2020)
39. Zhang, R.F., **Urbanowicz, R.J**. A Scikit-learn Compatible Learning Classifier System. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, pp. 1816-1823. (2020)
40. Moore JH, Barnett I, Boland MR, Chen Y, Demiris G, Gonzalez-Hernandez G, Herman DS, Himes BE, Hubbard RA, Kim D, Morris JS, Mowery DL, Ritchie MD, Shen L, **Urbanowicz R.J.**, Holmes JH. Ideas for how informaticians can get involved with COVID-19 research. BioData Min. (2020) May 12; 13:3. doi: 10.1186/s13040-020-00213-y. PMID: 32419848; PMCID: PMC7216865.
41. Diwadkar, A.R., Yoon, S., Shim, J., Gonzalez, M., **Urbanowicz, R.J.**, Himes, B.E. Integrating Biomedical Informatics Training into Existing High School Curricula. AIMIA Annual Symposium Proceedings, Vol 2021, pp. 190-199. American Medical Informatics Association. (2021) May 17. PMCID: PMC8378629; PMID: 34457133

42. Dasariraju, R. **Urbanowicz, R.J**., RARE: Evolutionary Feature Engineering for Rare-variant Bin Discovery. Proceedings of the Genetic and Evolutionary Computing Conference. ACM Press, pp. 1816-1823. (2021) Jul 8. doi: 10.1145/3449726.3463174
43. Holmes J.H., Beinlich J., Boland M.R., Bowles K., Chen Y., Cook T., Demiris G., Draugelis M., Fluharty L,. Gabriel P., Grundmeier R., Hanson W., Herman D.S., Himes B.E., Hubbard R.A., Kahn C., Kim D., Koppel R., Long Q., Mirkovic N., Morris J., Mowery D., Ritchie M., **Urbanowicz R.J**., Moore J.H. Why is the EHR So Challenging for Clinical Care and Research. Methods of Information in Medicine. 60(01/02), pp.032-048. (2021) DOI: 10.1055/s-0041-1731784
44. **Urbanowicz R.J**., Holmes J.H., Appleby D., Narasimhan V., Durborow S., Al-Naamani N., Fernando M., Kawut S.M. A Semi-Automated Term Harmonization Pipeline Applied to Pulmonary Arterial Hypertension Clinical Trials. Methods of Information in Medicine. (2021)
45. Min J, Appleby DH, McClelland RL, Minhas J, Holmes JH, **Urbanowicz RJ**, Pugliese SC, Mazurek JA, Smith KA, Fritz JS, Palevsky HI, Moutchia J, Al-Naamani N, Kawut SM. Secular and regional trends among pulmonary arterial hypertension clinical trial participants. Ann Am Thorac Soc. (2021) Dec 10;

Non-Refereed Letters and Invited Publications

1. Loiacono, D., Orriols-Puig, A., **Urbanowicz, R.J.** Special issue on advances in learning classifier systems. Evolutionary Intelligence, 1-2 (2012).
2. Shafi, K., **Urbanowicz, R.J.,** Iqbal, M. Special issue on advances in learning classifier systems. Evolutionary Intelligence, 1-2 (2013).
3. **Urbanowicz, R.J.** ExSTraCS: Rule Based Machine Learning, Classification and Knowledge Discovery for Complex Problems. SIGEVOlution Newsletter of the ACM Special Interest Group on Genetic and Evolutionary Computation. 8(1) http://www.sigevolution.org/ (2014)
4. Hu, T., Darabos, C., **Urbanowicz R.** Machine learning in genome-wide association studies. Frontiers in Genetics. (2020);11.

Book Chapters

1. **Urbanowicz, R.J.**, Moore, J.H. Learning classifier systems: The rise of genetics-based machine learning in biomedical data mining. In. Sarkar, N., (Eds.) Methods in Biomedical Informatics, 1st Edition, Elsevier. (2014).
2. Cole, B.S., Hall, M.S., **Urbanowicz, R.J.**, Gilbert-Diamond, D., Moore, J.H. Analysis of Gene-Gene Interactions. Current Protocols of Human Genetics (2018)
3. Yoon, S.A., Shim, J., Miller, K., Cottone, A., Noushad, N., Yoo, J., Gonzalez, M.V., **Urbanowicz, R.J**., Himes B. Professional Development for STEM Integration: Analyzing Bioinformatics Teaching by Examining Teachers' Qualities of Adaptive Expertise. Book Chapter in Bioinformatics PD (In Press) (2021)

Books and Edited Proceedings

1. **Urbanowicz, R.J.** Reassessment of a ganglioside-liposome biosensor for the detection of biological toxins. Durst, R.A, Baeumner, A.J. (Eds.) Masters Dissertation, Cornell University. (2005).
2. **Urbanowicz, R.J.** The detection and characterization of epistasis and heterogeneity: a learning classifier system approach. Moore, J.H., Whitfield, M.L., Eppstein M.J., Gross, R.H., Thornton-Wells, T.A. (Eds.) Genetics PhD Thesis, Dartmouth College. (2012).
3. **Urbanowicz, R.J.,** Browne, W. Book: Introduction to learning classifier systems. Springer, New York, NY (2017). (Available on amazon.com)

Other Non-Refereed Publications

1. **Urbanowicz, R.J.**, Kiralis, J., Fisher, J.M., Moore, J.H. GAMETES User's Guide. BioData Mining Supplemental Materials (2012).
2. **Urbanowicz, R.J.**, Bertasius, G., Moore, J.H. ExSTraCS 1.0 User's Guide. http://sourceforge.net/projects/exstracs/ (2014).
3. **Urbanowicz, R.J.**, Moore, J.H. ExSTraCS 2.0 User's Guide. http://sourceforge.net/projects/exstracs/ (2014).
4. **Urbanowicz RJ**, Olson RS, Schmitt P, Meeker M, Moore JH. Benchmarking relief-based feature selection methods. arXiv preprint arXiv:1711.08477. (2017) Nov.
5. **Urbanowicz RJ**, Meeker M, Cava WL, Olson RS, Moore JH. Relief-based feature selection: Introduction and review, CoRR abs/1711.08421. arXiv preprint arXiv:1711.08421. (2017) Nov.
6. Olson RS, La Cava W, Orzechowski P, **Urbanowicz RJ**, Moore JH. PMLB: A large benchmark suite for machine learning evaluation and comparison. arXiv preprint arXiv:1703.00512. (2017) March.
7. **Urbanowicz RJ**, Suri P, Cui Y, Moore JH, Ruth K, Stolzenberg-Solomon R, Lynch SM. A Rigorous Machine Learning Analysis Pipeline for Biomedical Binary Classification: Application in Pancreatic Cancer Nested Case-control Studies with Implications for Bias Assessments. arXiv preprint arXiv:2008.12829. (2020) Aug 28.
8. **Urbanowicz, R.J.** Documentation for I2C2 Computing Cluster at UPenn. https://github.com/UrbsLab/I2C2-Documentation (2021) March.
9. Zhang R, Stolzenberg-Solomon R, Lynch SM, **Urbanowicz RJ**. LCS-DIVE: An Automated Rule-based Machine Learning Visualization Pipeline for Characterizing Complex Associations in Classification. arXiv preprint arXiv:2104.12844. (2021) Apr 26.
10. **Urbanowicz, R.J.** Interpretable Machine Learning with Rule-Based Modeling (Keynote Abstract) LNCS proceedings of Declarative AI Conference (In Press) (2021)

Abstracts

1. Wojcieszynski, A., La Cava, W., **Urbanowicz R.J.,** Ying, X., Metz, J., Lin, A., Lukens, J., Fotouhi Ghiam, A.,Swisher-McClure, S., Moore, J.M., Baumann, B. Machine learning to predict toxicity in head and neck cancer patients treated with definitive chemoradiation**.** American Society for Radiation Oncology (ASTRO) (2019)
2. Lo, Y., Lynch, S., **Urbanowicz, R.J.**, Olson, R.S., Ritter, A.Z., Whitehouse, C.R., Connor, M.O., Keim, S.K., McDonald, M., Moore, J.H., Bowles, K.H. Using Machine

Learning on Home Health Care Assessments to Predict Fall Risk Studies in health technology and informatics, 264, 684-688, (2019)

3. Mazzotti, D.R., Keenan, B.T., **Urbanowicz, R.** and Pack, A.I., 0832 Evaluating Supervised Machine Learning Models for Cardiovascular Disease Prediction Using Conventional Risk Factors, Apnea-Hypopnea Index and Epworth Sleepiness Scale. Sleep, 42(Supplement_1), pp.A334-A334. (2019)
4. J. Minhas, D. Appleby, R. McClelland, J. Holmes, **Urbanowicz R.J.**, C. Burwell, C. L. Archer-Chicko, S. Pugliese, J. A. Mazurek, A.Smith, J. S. Fritz, H. Palevsky, S. Kawut , N. Al-Naamani. Obesity and its relationship with exercise capacity and hemodynamics in pulmonary arterial hypertension. A56. LEFT HEART DISEASE, METABOLISM AND OBESITY IN PULMONARY HYPERTENSION: CLINICAL STUDIES. American Thoracic Society, 2020. A2080-A2080. (2020)
5. J. Minhas, D. Appleby, R. McClelland, J. Holmes, **Urbanowicz R.J.**, C. Burwell, C. L. Archer-Chicko, S. Pugliese, J. A. Mazurek, A.Smith, J. S. Fritz, H. Palevsky, S. Kawut , N. Al-Naamani. Secular trends in subject characteristics of pulmonary arterial hypertensions clinical trials. B27. UP-TO-DATE PAH ASSESSMENT AND MANAGEMENT (pp. A2915-A2915). American Thoracic Society (2020)
6. Matuk, C., Yoon, S., Polman, J., Amato, A., Barton, J., Bulalacao, N. M., Cafaro, F., Haldar, L. C., Cottone, A., Cortes, K., DesPortes, K., Erickson, T., Finzer, W., Taylor, K. H., Herbel-Eisenmann, B., Graville, C., Gutiérrez, K., Higgins, T., Himes, B., Lanouette, K., Lee, H., Lim, V., Lopez, M. L., Lyons, L., Milz, D., Olivares, M. C., Osche, E., Parikh, T. S., Philip, T., Rubel, L., Shelley, J., Rivero, E., Roberts, J., Roberto, C., Petrosino, T., Rubin, A., Shim, J., Silander, M., Sommer, S., Stokes, D., Tes, M., Trajkova, M., **Urbanowicz, R.,** Vacca, R., Van Wart, S., Vasudevan, V., Wilkerson, M., & Woods, P. J. Data Literacy for Social Justice. In Gresalfi, M. and Horn, I. S. (Eds.), The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences (ICLS) 2020, Volume 1 (pp. 343-349). Nashville, Tennessee: International Society of the Learning Sciences. (2020)
7. Desariraju, S., Kamoun, M., Gragert, L., Wager, G., **Urbanowicz, R.J.** Machine Learning Approach for Binning Donor-Recipient HLA Amino Acid Position Mismatches to Detect Association with Kidney Graft Failure. American Society of Histocompatibility and Immunogenetics. (Podium Presentation) (In Press) (2021)
8. Wager, G., Kamoun, M., **Urbanowicz, R.J**., Biagini, D,. Dribus, M., Gragert, L. Visualization of Linkage Disequilibrium Patterns of Frequent Mismatched Amino Acid Positions Across HLA Class II Loci: Implications for Kidney Transplant Outcomes Studies. (In Press) (2021)
9. Z Zheng, H Li, AH Anderson, **R Urbanowicz**, D Xie, CM Rebholz, E Rhee, W Yang, S Waikar, S Navaneethan, J Hsu, R Mehta, S Rosas, A Ricardo, R Nelson, V Shah, C Hsu, J Sondheimer, J He, J Chen, V Ramachandran, C Parikh, P Kimmel, HI. Feldman. The Adiposity and Obesity Related Metabolomics Analysis in the CRIC Study (Podium Presentation) (In Press) (2021)

Submitted Abstracts

None

Submitted Publications

1. Zhang RF, Stolzenberg-Solomon R, Lynch SM, and **Urbanowicz R.J.** LCS-DIVE: An Automated Rule-based Machine Learning Visualization Pipeline for Characterizing Complex Associations in Classification. Journal of Machine Learning Research (Submitted) (2021)
2. Woodward A.A., **Urbanowicz R.J.**, Moore J.H. Biomedical Heterogeneity: A Categorized Review of Terminology, Challenges, and Methods. BioEssays (Submitted) (2021)
3. Woodward A.A. Taylor D., Goldmuntz E., Mitchell L.E., Agopian A.J., Moore J.H., **Urbanowicz R.J.** Gene-Interaction-Sensitive Enrichment Analysis in Congenital Heart Disease. BioData Mining (Submitted) (2021)
4. Ventetuolo, C.E., Moutchia, J., Baird, G.L., Appleby, D.H., Min, J., Holmes, J.H., **Urbanowicz, R.J.**, Al-Naamani N., Kawut, S.M. Baseline Sex Differences in Pulmonary Arterial Hypertension Randomized Clinical Trials. Eur Resp Journal. (Submitted) (2021)