Title:

Deciphering molecular mechanisms of disease consequent to mutation via semi-supervised learning

Abstract:

A major goal in computational biology is the development of algorithms, analysis techniques, and tools towards deep mechanistic understanding of life at a molecular level. In the process, computational biology must take advantage of the new developments in artificial intelligence and machine learning, and then extend beyond pattern analysis to provide testable hypotheses for experimental scientists. This talk will focus on our contributions to this process and relevant related work. We will first discuss the development of machine learning techniques for partially observable domains such as molecular biology; in particular, methods for accurate estimation of frequency of occurrence of hard-to-measure and rare events. We will show some identifiability results in parametric and nonparametric situations as well as how such frequencies can be used to correct estimated model accuracies. We will then show how these methods play key roles in inferring protein cellular roles and phenotypic effects of genomic mutations, with an emphasis on understanding the molecular mechanisms of human genetic disease. We further assessed the value of these methods in the wet lab where we tested the molecular mechanisms behind selected de novo mutations in a cohort of individuals with neurodevelopmental disorders. Finally, we will discuss implications on future research in machine learning, genome interpretation, and precision health.

=====

Biosketch:

Predrag Radivojac is a Professor of Computer Science at Northeastern University, where he recently moved from Indiana University. Prof. Radivojac received his Bachelor's and Master's degrees in Electrical Engineering from the University of Novi Sad and University of Belgrade, Serbia. His Ph.D. degree is in Computer Science from Temple University (2003) under the direction of Prof. Zoran Obradovic and co-direction of Prof. Keith Dunker. In 2004 he held a post-doctoral position in Keith Dunker's lab at Indiana University School of Medicine, after which he joined Indiana University Bloomington. Prof. Radivojac's research is in the areas of computational biology and machine learning with specific interests in protein function, MS/MS proteomics, genome interpretation, and precision health. He received the National Science Foundation (NSF) CAREER Award in 2007 and is an August-Wilhelm Scheer Visiting Professor at Technical University of Munich (TUM) as well as an honorary member of the Institute for Advanced Study at TUM. At Indiana University, he was Associate Chair of the Department of Computer Science and a co-Director of all of Informatics and Data Science for the multi-campus Prediction Health Initiative. Prof. Radivojac's projects have been regularly supported by NSF and National Institutes of Health (NIH). He is currently an Editorial Board member for the journal Bioinformatics, Associate Editor for PLoS Computational Biology, and serves his third term (elected) on the Board of Directors of the International Society for Computational Biology (ISCB).

=====