

Wyatt T. Clark

Post Doctoral Associate

266 Whitney Ave.

New Haven, CT 06511

☎ (812)-320-3951

✉ wyatt.clark@yale.edu

🌐 tinyurl.com/WyattClark

Education and Work Experience

- 2013–present **Post Doctoral Associate**, *Laboratory of Mark Gerstein*.
Department of Molecular Biophysics & Biochemistry, Yale University
- 2013 **Ph.D. Informatics**, *Indiana University*, Bloomington, IN.
Track: Bioinformatics
- 2006 **B.S. Informatics**, *Indiana University*, Bloomington, IN.
Minors: Biology, Philosophy

Technical Skills

- Programming **MATLAB***, **Perl***, **CGI**, **R**, **HTML**, **CSS**, (** fluent*).
- Machine Learning **Multi-output Neural Networks**, **Random Forests**, **Hidden Markov Models**, **Support Vector Machines**, **Bayesian Networks**, **Natural Language Processing**.

Selected Research Experience

- 2013–present **Post Doctoral Associate**, *Laboratory of Mark Gerstein*.
Department of Molecular Biophysics & Biochemistry, Yale University
- *Interactions between heterogeneous vertices*: Many methods that predict vertex labels and edges in graphs are based on the concept of guilt by association. Under this framework similar vertices are more likely to be connected by an edge and connected vertices are more likely to share the same labels. I am currently investigating the strength of this assumption in protein interaction networks and developing tractable methodology for utilizing statistically significant interactions between heterogeneous vertices to predict missing edges and vertex labels.
- 2007–2013 **Research Assistant**, *Laboratory of Predrag Radivojac*.
School of Informatics and Computing, Indiana University
- *Protein function prediction*: Utilized multi-output neural networks and kernel methods to predict Gene Ontology functions from sequence derived properties.
 - *Testing the ortholog conjecture*: Performed evolutionary analysis of how speciation and gene duplication influence the rate of functional divergence between homologous sequences.
 - *Information theoretic evaluation of multi-label classification*: Utilized a bayesian network in order to account for dependencies between terms in the Gene Ontology and provide a probabilistic framework for evaluating multi-label classification.
 - *Detection of inauthentic text*: Originated a novel method using compression profiles to detect classes of inauthentic computer generated text.

Professional Membership

2009–present **International Society for Computational Biology (ISCB).**

Academic Recognition and Awards

- 2013 **Ian Lawson Van Toch Memorial Award for Outstanding Student Paper**, *ISMB 2013*, Berlin, Germany.
- 2013 **Travel Fellowship**, *ISMB 2013*, Berlin, Germany, National Science Foundation.
- 2012 **Research Fellowship**, Center for Bioinformatics Research, Indiana University.
- 2011 **Travel Fellowship**, *ISMB 2011*, Vienna, Austria, National Institutes of Health.
- 2006 **Travel Fellowship**, *SDM 2006*, Bethesda, Maryland, Lawrence Livermore National Laboratory.

Extracurricular activities

IU Bioinformatics Club, *Co-Founder*.

IU Caving Club, *Member*.

Publications

P. Radivojac, **W.T. Clark**, et al. A large-scale evaluation of computational protein function prediction. *Nature Methods*, 10(3):221–227, 2013.

W.T. Clark and P. Radivojac. Information-theoretic evaluation of predicted ontological annotations. *Bioinformatics*, 29(13):i53–i61, 2013. (**Ian Lawson Van Toch Memorial Award for Outstanding Student Paper**).

N.L. Nehrt, **W.T. Clark***, P. Radivojac, and M.W. Hahn. Testing the ortholog conjecture with comparative functional genomic data from mammals. *PLoS Computational Biology*, 7(6):e1002073, 2011. (*** Co-first author**).

W.T. Clark and P. Radivojac. Analysis of protein function and its prediction from amino acid sequence. *Proteins: Structure, Function, and Bioinformatics*, 79(7):2086–2096, 2011.

Y. Zhao, **W.T. Clark**, M. Mort, D.N. Cooper, P. Radivojac, and S.D. Mooney. Prediction of functional regulatory snps in monogenic and complex disease. *Human Mutation*, 32(10):1183–1190, 2011.

M.M. Dalkilic, J.C. Costello, **W.T. Clark**, and P. Radivojac. From protein-disease associations to disease informatics. *Frontiers in Bioscience*, 13:3391–3407, 2008.

P. Radivojac, K. Peng, **W.T. Clark**, B.J. Peters, A. Mohan, S.M. Boyle, and S.D. Mooney. An integrated approach to inferring gene-disease associations in humans. *Proteins: Structure, Function, and Bioinformatics*, 72(3):1030–1037, 2008.

M.M. Dalkilic, **W.T. Clark**, J.C. Costello, and P. Radivojac. Using compression to identify classes of inauthentic texts. In *Proceedings of the 2006 SIAM Conference on Data Mining*, pages 604–608, April 2006.