Jianlin Cheng

Assistant Professor
School of Electrical Engineering and Computer Science
University of Central Florida
Orlando, FL 32816
Phone: (407) 968-9746
Email: jianlin.cheng@gmail.com
Web: http://www.eecs.ucf.edu/~jcheng

Education

Ph.D. in Information and Computer Science (Area: Bioinformatics), University of California, Irvine, August, 2006. (Advisor: Dr. Pierre Baldi)

M.S. in Computer Science, Utah State University, 2001.

B.S. in Computer Science, Huazhong University of Sci. and Tech., China, 1994.

Experiences

2006 – present   **Assistant Professor**, School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL

2002 – 2006   **Research Assistant**, Institute for Genomics and Bioinformatics, School of Information and Computer Sciences, University of California, Irvine, CA

2001 – 2002   **Software Engineer**, Sony Corporation, Culver City, CA

1999 – 2001   **Research and Teaching Assistant**, Computer Science Department, Utah State University, Logan, UT

1998   **Research Assistant**, Computer Science Department, Huazhong Univ. of Sci. and Tech., China

1994 – 1997   **Software Engineer**, China Construction Bank, Wuhan, China

Research Interests

Bioinformatics and Machine Learning

Research Projects

Protein structure prediction, protein-protein interaction, protein function prediction, biological networks, biological sequence alignment, and machine learning ranking.

Publications

Refereed Papers:

12   M. Tress, **J. Cheng**, P. Baldi, K. Joo, J. Lee, J. Seo, J. Lee, D. Baker, D. Chivian, A. Valencia, I. Ezkurdia. Assessment of Predictions Submitted for CASP7 Domain Prediction Category. *Proteins: Structure, Function, and Bioinformatics*, in press, 2007. (**CASP7 invited contribution**)

11   L. Larson, M. Zhang, K. Kosaka, V. Kovacevic, A. Lamsa, **J. Cheng**, P. Baldi, G.W. Hatfield, and S. Sandmeyer. Ty3 Capsid Scanning Mutations Reveal Early and Late Functions of The Amino-Terminal Domain. *Journal of Virology*. Accepted, 2007.

10   **J. Cheng** and P. Baldi. A Machine Learning Information Retrieval Approach to Protein Fold Recognition. *Bioinformatics*, vol. 22, pp. 1456-1463, 2006.

9   **J. Cheng**, M. Sweredoski, and P. Baldi. DOMpro: Protein Domain Prediction Using Profiles, Secondary Structure, Relative Solvent Accessibility, and Recursive Neural Networks. *Data Mining and Knowledge Discovery*, vol. 13, pp. 1-10, 2006.

8   S. A. Danziger, S. J. Swamidass, J. Zeng, L. R. Dearth, Q. Lu, J. H. Chen, **J. Cheng**, V. P. Hoang, H. Saigo, R. Luo, P. Baldi, R. K. Brachmann, and R. H. Lathrop. Functional Census of Mutation Sequence Spaces: The Example of p53 Cancer Rescue Mutants. *IEEE Transactions on Computational Biology and Bioinformatics*, vol. 3, no. 2, pp. 114-1125, 2006.

7   **J. Cheng**, A. Randall, and P. Baldi. Prediction of Protein Stability Changes for Single Site Mutations Using Support Vector Machines. *Proteins: Structure, Function, Bioinformatics*, vol. 62, no. 4, pp. 1125-1132, 2006.

6   **J. Cheng**, H. Saigo, and P. Baldi. Large-Scale Prediction of Disulphide Bridges Using Kernel Methods, Two-Dimensional Recursive Neural Networks, and Weighted Graph Matching. *Proteins: Structure, Function, Bioinformatics*, vol. 62, no. 3, pp. 617-629, 2006.

5   **J. Cheng**, M. Sweredoski, and P. Baldi. Accurate Prediction of Protein Disordered Regions by Mining Protein Structure Data. *Data Mining and Knowledge Discovery*, vol. 11, no. 3, pp 213-222, 2005.

4   **J. Cheng**, L. Scharenbroich, P. Baldi, and E. Mjolsness. Sigmoid: Towards a Generative, Scalable, Software Infrastructure for Pathway Bioinformatics and Systems Biology. *IEEE Intelligent Systems*, vol. 20, no. 3, pp. 68-75, 2005.

3   **J. Cheng**, A. Randall, M. Sweredoski, and P. Baldi. SCRATCH: a Protein Structure and Structural Feature Prediction Server. *Nucleic Acids Research*, vol. 33 (web server issue), w72-76, 2005.

2   **J. Cheng** and P. Baldi. Three-Stage Prediction of Protein Beta-Sheets by Neural Networks, Alignments, and Graph Algorithms. Proceedings of the 2005 Conference on Intelligent Systems for Molecular Biology (ISMB 2005), *Bioinformatics*, vol. 21(suppl 1), pp i75-84, 2005.

1   P. Baldi, **J. Cheng**, and A. Vullo. Large-Scale Prediction of Disulphide Bond Connectivity. *Advances in Neural Information Processing Systems 17 (NIPS 2004)*, L. Saul,Y. Weiss, and L. Bottou editors, MIT press, pp.97-104, Cambridge, MA, 2004.

Theses:

2   **J. Cheng**. Machine Learning Algorithms for Protein Structure Prediction. *PhD Dissertation*. University of California Irvine, Irvine, CA, 2006.

1   **J. Cheng**. A Comparative Study of the Similarity Measures of Text Categorization. *Master Thesis*, Utah State University, Logan, UT, 2001.

Papers Submitted or to Submit:

1.   **J. Cheng**. A Multi-Template Combination Algorithm for Protein Comparative Modeling. (submitted)

2.   **J. Cheng** and P. Baldi. Improved Residue Contact Prediction Using Support Vector Machine and a Large Feature Set. (submitted)

3.   **J. Cheng**. A Neural Network Approach to Ordinal Regression. (submitted)

4.   **J. Cheng**. DOMAC: A Hybrid, Accurate Protein Domain Prediction Server. (submitted)

5.   A. Randall, **J. Cheng**, M. Sweredoski, and P. Baldi. TMBpro: Secondary Structure, β-contact and Tertiary Structure Prediction of Transmembrane β-Barrel Proteins. (submitted)

6.   **J. Cheng**. ModelEvaluator: Evaluating Protein Model Quality Using Support Vector Machine and Structural Features. (to submit)

Invited Journal (non-refereed)

J. Hecker and **J. Cheng**. Protein Disorder Prediction at Multiple Levels of Sensitivity and Specificity. *Journal of Computational Intelligence in Bioinformatics*, in press.

Grant Proposals

1.   **J. Cheng**. Machine Learning Approaches to Protein Fold Recognition. (to submit to NIH Bioinformatics program).

2.   **J. Cheng**. Protein Domain Prediction Using Homology, Evolutionary Signals, and Neural Networks. (to submit to NIH bioinformatics program).

Major Talks

Presentation at the 13th International Conference on Intelligent Systems for Molecular Biology (**ISMB 2005**), Detroit, Michigan, 2005.

Presentation at the 7th Critical Assessment of Techniques for Protein Structure Prediction (**CASP7**), Asilomar, Monterey, California, 2006.

Bioinformatics Software

Designed and developed the following software tools and web servers free for scientific/academic use:

FOLDpro and *3Dpro*: protein homology modeling and fold recognition (**2nd and 3rd in automated high-accuracy protein structure prediction in CASP7**)

DOMpro: protein domain prediction (**DOMpro-FOLDpro is ranked 1st in automated protein domain prediction in CASP7**)

BETApro: protein beta-sheet topology prediction (**2nd in automated residue contact prediction in CASP7**)

SSpro 4: protein secondary structure prediction

ACCpro 4: protein solvent accessibility prediction

DIpro: protein disulfide bond prediction

CMAPpro 2, BMAPpro, CMAPCon: protein contact map prediction

DISpro: protein disorder region prediction (**2nd in automated protein disorder prediction in CASP7**)

MUpro: protein mutation stability prediction

SVMCon: contact map prediction using support vector machines

SIGMOID 1.0: a biological network database and simulation system (chief designer/developer)

NNRank 1.0: neural network software for ordinal regression and classification

Honors and Awards

2nd in Automated High-Accuracy 3D Protein Structure Prediction in the Seventh Edition of Critical Assessment of Techniques for Protein Structure Prediction (**CASP7**), 2006.

1st in Automated Protein Domain Prediction in the Seventh Edition of Critical Assessment of Techniques for Protein Structure Prediction (**CASP7**), 2006.

2nd in Automated Residue Contact Prediction in the Seventh Edition of Critical Assessment of Techniques for Protein Structure Prediction (**CASP7**), 2006.

2nd in Automated Protein Disorder Prediction in the Seventh Edition of Critical Assessment of Techniques for Protein Structure Prediction (**CASP7**), 2006.

PhD Dissertation Fellowship, University of California Irvine, 2006.

Award for Second Place in the Time Series Competition during the 2005 UC Data Mining Contest, 2005. (Out of over 90 teams representing eight University of California campuses.)

Award for Third Place in the Classification Competition during the 2005 UC Data Mining Contest, 2005. (Out of over 90 teams representing eight University of California campuses.)

Highest score at the Annual Ph.D. Qualifying Exam in Bioinformatics concentration, School of Information and Computer Sciences, University of California Irvine, 2003.

Teaching

Introduction to Bioinformatics. http://www.eecs.ucf.edu/~jcheng/CAP5937.htm. (Fall, 2006)

Statistical Machine Learning Methods for Bioinformatics.
http://www.eecs.ucf.edu/~jcheng/mlbioinfo/mlbioinfo.htm. (Spring, 2007)

Professional Activities

Editorial Board Member: Journal of Computational Intelligence in Bioinformatics

Program Committee Member: IEEE Symposium on Bioinformatics & Bioengineering (BIBE)

Member: International Society for Computational Biology

Reviewing: Nucleic Acids Research, Bioinformatics, Journal of Bioinformatics and Computational Biology, Data Mining and Knowledge Discovery, FEBS letter, Advances in Neural Information Processing Systems, International Conference on Intelligent Systems for Molecular Biology

References

Dr. Pierre Baldi
Professor and Director
Institute for Genomics and Bioinformatics
Donald Bren School of Information and Computer Sciences
University of California Irvine
Irvine, CA 92697-3435
Phone: (949) 824-5809
Email: pfbaldi@ics.uci.edu
Web: http://www.ics.uci.edu/~pfbaldi

Dr. Hal Stern
Professor and Chair
Department of Statistics
Donald Bren School of Information and Computer Sciences
University of California Irvine
Irvine, CA 92697-3435
Phone: (949) 824-1568

Email: sternh@uci.edu
Web: http://www.ics.uci.edu/~sternh

Dr. G. Wesley Hatfield
Professor, Department of Microbiology & Molecular Genetics
Associate Director, Institute for Genomics and Bioinformatics
School of Medicine
University of California Irvine
Irvine, CA 92697-4025
Phone (office): (949) 824-5344
Phone (lab): (949) 824-5858
Email: gwhatfie@uci.edu

Dr. Yang Zhang
Assistant Professor
Center for Bioinformatics
University of Kansas
2030 Becker Dr
Lawrence, KS 66047
Phone: 785-864-1057
Email: yzhang@ku.edu

Dr. Suzanne Sandmeyer
Chair and Professor
Department of Biochemistry
University of California Irvine
Irvine, CA 92697
Phone: 949-824-7571
Email: sbsandme@uci.edu