

CHRISTINA S. LESLIE

Center for Computational Learning Systems, Columbia University
Center for Computational Biology and Bioinformatics, Columbia University
475 Riverside Dr., Mail Code 7717, New York, NY 10115
Phone: 212-870-1278, Email: cleslie@cs.columbia.edu
Home page: <http://www.cs.columbia.edu/~cleslie>
Group page: <http://www.cs.columbia.edu/compbio>

Visa Status

United States Permanent Resident

Professional Experience

- 2004-present Research Scientist, Center for Computational Learning Systems, Columbia University
- 2002-present Faculty, Center for Computational Biology and Bioinformatics, Columbia University
- 2000-2003 Assistant Professor, Department of Computer Science, Columbia University
- 1999-2000 Postdoctoral Research Fellow, Department of Mathematics, Columbia University

Education

- 1998 University of California, Berkeley, Ph.D. (Mathematics), 1998
Advisor: Joseph A. Wolf
- 1992 University of Waterloo (Canada), B.Math.

Awards

- 2003 PhRMA Foundation Award in Informatics, Research Starter Grant
- 1998 Natural Sciences and Engineering Research Council of Canada (NSERC) Postdoctoral Fellowship in Mathematics
- 1996 Outstanding Graduate Student Instructor Award (UC Berkeley)
- 1992 NSERC 1967 Science and Engineering Scholarship
- 1987 René Descartes Mathematics Competition Fellowship

Refereed Conference and Journal Publications

1. *A classification-based framework for predicting and analyzing gene regulatory response.* A. Kundaje, M. Middendorf, M. Shah, C. Wiggins, Y. Freund and C. Leslie. BMC Bioinformatics, Special issue on “New Problems and Methods in Computational Biology”. In press. **ISI Impact Factor: 5.42**
2. *Protein Ranking by Semi-Supervised Network Propagation.* J. Weston, R. Kuang, C. Leslie and W. S. Noble. BMC Bioinformatics, Special issue on “New Problems and Methods in Computational Biology”. In press. **ISI Impact Factor: 5.42**

3. *Motif-based protein ranking by network propagation.* R. Kuang, J. Weston, W. S. Noble and C. Leslie. *Bioinformatics* 21(19):3711-3718 (2005). **ISI Impact Factor: 5.742**
4. *Multi-Class protein fold detection using adaptive codes.* E. Ie, J. Weston, W. S. Noble, C. Leslie. Proceedings of the International Conference on Machine Learning (ICML 2005). August 7-11, 2005, Bonn, Germany. **Oral presentation, totals: 144 papers accepted (29%), 491 submitted.**
5. *Motif Discovery through Predictive Modeling of Gene Regulation.* M. Middendorf, A. Kundaje, M. Shah, Y. Freund, C. Wiggins, C. Leslie. Proceedings of the Ninth Annual International Conference on Research in Computational Molecular Biology (RECOMB 2005). **Oral presentation, totals: 39 papers accepted (18%), 217 submitted.**
6. *Semi-supervised protein classification using cluster kernels.* J. Weston, C. Leslie, E. Ie, D. Zhou, A. Elisseff and W. S. Noble. *Bioinformatics* 21:3241-3247 (2005). **ISI Impact Factor: 6.701**
7. *Dichotomous Splicing Signals in Exon Flanks.* X. Zhang, C. Leslie, and L. Chasin. *Genome Research* 15:768 - 779 (2005). **ISI Impact Factor: 9.863**
8. *Combining Sequence and Time Series Expression Data to Learn Transcriptional Modules.* A. Kundaje, M. Middendorf, C. Wiggins, C. Leslie. *IEEE Transactions in Computational Biology and Bioinformatics.* In press, 2005.
9. *Remote protein homology detection and motif extraction using profile kernels.* R. Kuang, E. Ie, K. Wang, K. Wang, M. Siddiqi, Y. Freund, C. Leslie. *Journal of Bioinformatics and Computational Biology* 3(3):527-550 (2005).
10. *Predicting genetic regulatory response using classification: Yeast stress response,* M. Middendorf, A. Kundaje, C. Wiggins, Y. Freund, and C. Leslie. In Proceedings of the RECOMB 2004 Satellite Workshop on Regulatory Genomics. *Lecture Notes in Bioinformatics*, 2005.
11. *Predicting Genetic Regulatory Response Using Classification,* M. Middendorf, A. Kundaje, C. Wiggins, Y. Freund, C. Leslie. Proceedings of the Twelfth International Conference on Intelligent Systems for Molecular Biology (ISMB 2004). *Bioinformatics* 20 (Supp 1):I232-I240 (2004). **Long paper presentation, totals: 67 long and short papers accepted (14%), 496 submitted.**
12. *Profile-based String Kernels for Remote Homology Detection and Motif Extraction,* R. Kuang, E. Ie, K. Wang, K. Wang, M. Siddiqi, Y. Freund, C. Leslie. Proceedings of the Computational Systems Biology Conference (CSB 2004). **Oral presentation, totals: 30 oral (15%), 202 submitted.**
13. *Protein ranking: from local to global structure in the protein similarity network,* J. Weston, A. Elisseff, D. Zhou, C. Leslie, W. S. Noble. Proceedings of the National Academy of Sciences USA 101(17):6559-63 (2004). **ISI Impact Factor: 10.700**
14. *Semi-Supervised Protein Classification using Cluster Kernels,* J. Weston, C. Leslie, D. Zhou, A. Elisseff, W. Noble. *Advances in Neural Information Processing Systems* 16 (NIPS 2003), 2004. **Spotlight presentation, totals: 65 oral and spotlight (9%), 198 poster, 717 submitted.**

15. *Protein Backbone Angle Prediction with Machine Learning Approaches*, R. Kuang, C. Leslie, A. Yang. *Bioinformatics* 20(10):1612-1621 (2004). **ISI Impact Factor: 6.701**
16. *Sequence Information for the Splicing of Human pre-mRNA Identified by Support Vector Machine Classification*, X. Zhang, K. Heller, I. Hefter, C. Leslie, L. Chasin. *Genome Research* 13:2637-2650 (2003). **ISI Impact Factor: 9.863**
17. *Fast String Kernels using Inexact Matching for Protein Sequences*. C. Leslie, R. Kuang. *Journal of Machine Learning Research* 5:1435-1455, 2004. **ISI Impact Factor: 4.317**
18. *Mismatch String Kernels for Discriminative Protein Classification*. C. Leslie, E. Eskin, A. Cohen, J. Weston, W. S. Noble. *Bioinformatics* 20(4):467-76 (2004). **ISI Impact Factor: 6.701**
19. *Fast Kernels for Inexact String Matching*. C. Leslie and R. Kuang. *Proceedings of The 16th Annual Conference on Computational Learning Theory (COLT 2003)*. Volume 2777 of *Lecture Notes in Computer Science*, 2003. **Oral presentation, totals: 26 oral (28%), 49 accepted, 92 submitted**
20. *Mismatch String Kernels for SVM Protein Classification*. C. Leslie, E. Eskin, J. Weston, W. S. Noble. *Advances in Neural Information Processing Systems 15 (NIPS 2002)*, 2003. **Oral presentation, totals: 26 oral (4%), 207 poster, 694 submitted.**
21. *Dealing with Large Diagonals in Kernel Matrices*. B. Schölkopf, J. Weston, E. Eskin, C. Leslie, W. S. Noble. *Annals of the Institute of Statistical Mathematics* 55(2), 391-408 (2003).
22. *A Kernel Approach for Learning Almost Orthogonal Patterns*. B. Schölkopf, J. Weston, E. Eskin, C. Leslie, W. S. Noble. *Proceedings of the 13th European Conference on Machine Learning (ECML 2002) and Proceedings of the 6th European Conference on the Principles and Practice of Knowledge Discovery in Databases (PKDD 2002)*, Helsinki. *Lecture Notes in Computer Science*, Vol. 2430/2431, Springer, 2002.
23. *The Spectrum Kernel: A String Kernel for SVM Protein Classification*. C. Leslie, E. Eskin, W. S. Noble. *Proceedings of the Pacific Symposium on Biocomputing (PSB 2002)*, 2002.

Invited Book Chapters and Reviews

- *Semi-Supervised Protein Classification using Cluster Kernels*. J. Weston, C. Leslie, E. Ie, D. Zhou, A. Elisseeff and W. S. Noble. Book Chapter in “Semi-Supervised Learning”. O. Chapelle, B. Scholkopf and A. Zien, ed. MIT Press, 2006.
- *Identifying remote protein homologs by network propagation*. W. S. Noble, R. Kuang, C. Leslie, J. Weston. *FEBS Journal* 272(20) (2005).
- *Inexact Matching String Kernels for Protein Classification*, C. Leslie, R. Kuang, E. Eskin. In *Proceedings of the RECOMB Workshop on Kernel Methods in Computational Biology*. B. Schoelkopf, K. Tsuda and J.-P. Vert, ed. MIT Press, 2004.

Selected Invited Talks

- Princeton University, PICASso Seminar Series, March 2006
- UC San Diego, ITA Inaugural Workshop, February 2006
- New York Academy of Sciences, Systems Biology Group, November 2005
- McMaster University, Hamilton, Canada. Workshop on Mathematical Programming in Data Mining and Machine Learning, June 2005
- NEC Labs, Workshop on Machine Learning, May 2005
- Memorial Sloan-Kettering Computational Biology Center, January 2005
- University of Washington, Seattle WA, September 2004
- IBM Watson Research Center, KDD Group, June 2004
- Columbia University, Computer Science Colloquium, April 2004
- New York Academy of Sciences, Computational Biology Group, October 2003
- IBM Watson Research Center, Computational Biology Group, July 2003
- New York University, Computer Science Colloquium, May 2003
- Workshop on Kernel Methods for Bioinformatics, Berlin, Germany, April 2003
- National Institutes of Health, National Center for Biotechnology Information (NCBI), Bethesda, MD, July 2002

Selected Grants

- National Center for Biomedical Computing: Center for Multi-scale Analysis of Genetic and Cellular Networks (NIH). \$18.5M from 9/2005 to 8 2010. Senior personnel.
- Recognizing protein folds using discriminative learning (NIH). \$1.4M from 5/2005 to 4/2010. Principal Investigator.
- ITR: Machine learning approaches to protein sequence comparison: discriminative, semi-supervised, scalable algorithms (NSF). \$300K from 9/2003 to 8/2006. Principal Investigator.
- PhRMA Foundation Award in Informatics, Research Starter Grant. \$60K from 1/2003 to 12/2004. Principal Investigator.
- Center for Computational Biology and Bioinformatics (NIH). \$2.3M from 9/2001 to 8/2004. Collaborator.
- Combined Research-Curriculum Development: Curriculum for Genomic Engineering, NSF. \$500K from 2/2001 to 1/2005. Collaborator.

Selected Activities

- Head of Computational Biology Group, in the Center for Computational Learning Systems and affiliated with the Computer Science Department at Columbia University
- Co-organizer, DIMACS Workshop on Machine Learning Approaches for Understanding Gene Regulation, DIMACS Center, Rutgers University, August 2005
- Co-editor, Special issue of BMC Bioinformatics, Proceedings of the NIPS 2004 Computational Biology Workshop
- Co-organizer, Computational Biology Workshop, Neural Information Processing Systems (NIPS 2003-2005)
- NIH Panelist, 2004
- NSF Panelist, 2004
- Reviewer (Conferences): International Conference on Intelligent Systems in Molecular Biology (ISMB 2003-2005), European Conference on Computational Biology (ECCB 2005), Conference on Research in Computational Molecular Biology (RECOMB 2004-2006), Pacific Symposium on Biocomputing (PSB 2004-2005), Neural Information Processing Systems (NIPS 2003-2005), International Conference on Machine Learning (ICML 2004), Computational Learning Theory/Kernel Workshop (COLT/KW 2003)
- Reviewer (Journals): Bioinformatics, Journal of Machine Learning Research, Proteins, BMC Bioinformatics, IEEE Transactions on Computational Biology and Bioinformatics, Journal of Bioinformatics and Computational Biology, Neural Processing Letters, Journal of Neural Networks, IEEE Signal Processing Letters
- Reviewer (Grants): National Science Foundation, 2003-2005; Israel-U.S. Binational Science Foundation, 2003

Teaching Experience

- CBMF W4761 Computational Genomics
- COMS W3137 Data Structures and Algorithms

Advising

- Graduate students: Rui Kuang, 2002-present (PhD CS); Anshul Kundaje, 2002-present (PhD CS); Eugene Ie, 2003-2005 (PhD CS); David Quigley, 2005-present (Masters Biomedical Informatics); Steve Lianoglou, 2006-present (Masters CS); Girish Rao, 2006-present (Masters CS); Vishakh Vishakh, 2006-present (Masters CS); Kai Wang, 2003 (PhD Biomedical Informatics); Omar Antar, 2002-2003 (MD/PhD); Katherine Heller, 2002 (PhD CS); Ilana Hefter, 2001-2002 (Masters CS);
- Undergraduate project students: Ben Neuwirth (Fall 2005 - Spring 2006); Joshua Gordon (Spring 2006); Carolina Gabbai (Fall 2004); Dong Lou (Fall 2003 - Spring 2004); Mahira Siddiqi (Fall 2003 - Spring 2004); Ke Xu (Fall 2002 - Spring 2003); Meena George (Summer 2002 - Spring 2003); Richard Mallah (Spring 2002); Luis Tello (Fall 2001); Ilan Wapinski (Summer 2001).