Attila Kertesz-Farkas, Ph.D.

Postdoctoral fellow
Protein Structure and Bioinformatics
International Centre for Genetic
Engineering and Biotechnology
Padriciano 99
34012 Trieste - ITALY

E-mail: kfattila@icgeb.org
Lab tel. +39-0403757341
Central fax +39-040226555

Education:

1999-2004	University of Szeged, Hungary. MSc (2004) in Computer Science and Mathematics (specializing in Fundamentals of Computer Science), (summa cum laude), Thesis title: „Compact Representation of Finite Languages with Nondeterministic Finite Automata", Advisor: Prof. Zoltán Fülöp,

Qualifications:

2009-	Postdoctoral fellow at Protein Structure and Bioinformatics Group, International Centre of Genetic Engineering and Biotechnology, Trieste, Italy.
2008-2009	Postdoctoral fellow at Division of Imaging and Applied Mathematics, Center for Devices and Radiological Health, U.S. Food and Drug Administration, MD, USA joint affiliation to Department of Biology, University of Maryland Baltimore County, MD, USA.
2004-2008	PhD student at the Doctoral School of Computer Science, University of Szeged, Hungary. PhD (2009) in Computer Science (summa cum laude) Thesis: "Protein classification in machine learning framework", Advisors: Prof. János Csirik, Dr. András Kocsor

Teaching Activity:

2011	Lecturer at Theoretical and Practical Course on Advanced Bioinformatics, ICGEB, Trieste, Italy
2010-2012	Teaching assistant at Theoretical and Practical Course on Bioinformatics, ICGEB, Trieste, Italy
2007	Teaching assistant Operational Research, University of Szeged, Hungary
2006	Teaching assistant Algorithms and Data Structures, University of Szeged, Hungary
2004,2005	Teaching assistant Introduction to Pascal, C and Java programming, University of Szeged, Hungary

Scientific Interest:

On account of recent advances in high-throughput techniques vast amount of raw data is produced daily, however the information usually is hidden for naked eyes. My scientific interests include development of mathematical algorithms that can reveal and learn subtle relationships and regularities among data using statistical analysis, machine learning and data mining techniques. I believe that sophisticated data representation and knowledge integration into mathematical models are cornerstones of accurate and practical software applications.

Scholarships and awards:

2006-2007	Erasmus scholarship, one semester at Technische Universitaet Dresden, Germany
2004	1st prize at a Scientific Conference of Students for the article entitled "Kernel-based learning with feature weighting".
2002	1st prize at a Scientific Conference of Students for the article entitled "Compact representation of finite languages with nondeterministic automata".

Membership:

2012-present	Editorial Board of International Journal of Advances in Life Sciences
2011-present	Consultant at ManuSTAT central (a statistical consulting service)
2006-present	John von Neumann Computer Science Society, Hungary
2005-2007	PhD student representative at the Board of Doctoral Schools at Faculty of Science, University of Szeged,
2005-2007	PhD student representative at the Board of the Institute of the Informatics, University of Szeged,
2003-2008	Eötvös Collegium (Honors College) at University of Szeged, Hungary

LIST OF PUBLICATIONS

Theses:

1. PhD: "Protein Classification in a Machine Learning Framework", 2008,
2. MSc: "Compact Representation of Finite Languages with Nondeterministic Finite Automata", 2004

Book chapter

3. Attila Kertész-Farkas1, András Kocsor3 and Sándor Pongor4,5, The application of Data Compression-based Distance to Biological Sequences, In: Frank Emmert-Streib. (Ed.) Information Theory and Statistical Learning, (2008), p. 73-88

Refereed journal papers, first, shared first or senior authorships

4. Attila Kertész-Farkas, Beáta Reiz, Michael P. Myres, Sándor Pongor, PTMSearch: A Greedy Tree Traversal Algorithm for Finding Protein Post-Translational Modifications in Tandem Mass Spectra. In: ECML PKDD 2, Vol. 6912 Springer (2011), p. 162-176.
5. Attila Kertész-Farkas, Beáta Reiz, Michael P. Myers, Sándor Pongor Database searching in mass spectrometry based proteomics. 2011, in press
6. Dmitry V. Nikitin, Attila Kertész-Farkas, Alexander S. Solonin and Marina L Mokrishcheva, Bifunctional restriction-modification enzymes: products of natural genetic engineering, 2012, submitted
7. Bruna Marini, Attila Kertész-Farkas, Hashim Ali, Mauro Giacca, Marina Lusic Nuclear Topology drives HIV-1 integration in the nuclear periphery, 2012, under preparation.
8. Emily Doughty, Attila Kertész-Farkas, Olivier Bodenreider, Gary Thompson, Asa Adadey, Thomas Peterson, and Maricel G. Kann Toward an automatic method for extracting cancer- and other disease-related point mutations from the biomedical literature, Bioinformatics, 27(3), 2011, p. 408-415,
9. József Dombi and Attila Kertész-Farkas, Applying Fuzzy Technologies to Equivalence Learning in Protein Classification, Journal of Computational Biology, 16(4), 2009, p. 611-623
10. Attila Kertész-Farkas, András Kocsor and Sándor Pongor, Equivalence Learning in Protein Classification, In: P. Perner (Ed.) Machine Learning and Data Mining in Pattern Recognition, LNAI (4571), Springer Verlag, Heidelberg, 2007, p. 824-837
11. Attila Kertész-Farkas, Somdutta Dhir, Paolo Sonego, Mircea Pacurar, Sergiu Netoteia, Harm Nijveen, Arnold Kuzniar, Jack A.M. Leunissen, András Kocsor and Sándor Pongor Benchmarking protein classification algorithms via supervised cross-validation, Journal of Biochemical and Biophysical Methods, 70(6), 2008, pp. 1215-1223
12. János Z. Kelemen, Attila Kertész-Farkas, András Kocsor and László G. Puskás, Kalman Filtering for Disease-State Estimation from Microarray Data, Bioinformatics (22), 2006, pp 3047-3053
13. László Kaján, Attila Kertész-Farkas, Dino Franklin, Nelly Ivanova, András Kocsor and Sándor Pongor, Application of a simple log likelihood ratio approximant to protein sequence classification, Bioinformatics (22), 2006, pp 2865-2869
14. András Kocsor, Attila Kertész-Farkas, László Kaján and Sándor Pongor, Application of compression-based distance measures to protein sequence-classification: a methodological study, Bioinformatics (22), 2006, pp 407-412
15. Attila Kertész-Farkas and András Kocsor, Kernel-based Classification of Tissues using Feature Weightings, Applied Ecology and Environmental Research 4(2), 2006, pp. 63-71
16. Attila Kertész-Farkas, Zoltán Fülöp and András Kocsor, Compact Representation of Hungarian Corpora (in Hungarian), Hungarian Journal of Applied Linguistics (1-2), 2005, pp. 63-70

Refereed journal papers,

17. Ádam Kerényi, Dóra Bihary, Zsolt Gelencsér, Sergiu Netotea, Attila Kertész-Farkas, Vittorio Venturi and Sándor Pongor, Simulation of communication and cooperation in multispecies bacterial communities with an agent based model, in press
18. Beáta Reiz, Attila Kertész-Farkas, Sándor Pongor and Michael P. Myers. Data preprocessing and filtering in mass spectrometry based proteomics. 2011, in press
19. Beáta Reiz, Attila Kertész-Farkas, Somdutta Dhir, Sándor Pongor, Michael P. Myers, Chemical rule-based filtering of MS/MS spectra, 2012, submitted

20. Somdutta Dhir, Mircea Pacurar, Dino Franklin, Zoltán Gáspári, Attila Kertész-Farkas, András Kocsor, Frank Eisenhaber, Sándor Pongor(2010) Detecting atypical examples of known domain types by sequence similarity searching: The SBASE domain library approach, Current Protein Peptide Science, 11, 2010, p. 538-49
21. Róbert Busa-Fekete, Attila Kertész-Farkas, András Kocsor and Sándor Pongor, Balanced ROC analysis (BAROC) protocol for the evaluation of protein similarities, Journal of Biochemical and Biophysical Methods, 70(6), 2008, p. 1210-1214
22. Paolo Sonego, Mircea Pacurar, Somdutta Dhir, Attila Kertész-Farkas, András Kocsor, Zoltán Gáspári, Jack A.M. Leunissen and Sándor Pongor, A Protein Classification Benchmark collection for Machine Learning, Nucleic Acids Research (35), 2006, D232-6

Conferences

23. Attila Kertész-Farkas, PTMSearch: A Greedy Tree Traversal Algorithm for Finding Protein Post-Translational Modifications in Tandem Mass Spectra. In: ECML PKDD 2, Vol. 6912 Springer (2011), p. 162-176.
24. Attila Kertész-Farkas, Equivalence Learning in Protein Classification, International Conference on Machine Learning and Data Mining in Pattern Recognition MLDM 2007, 18-20 July 2007, Leipzig (Germany)
25. Attila Kertész-Farkas, Classification of Tissues using Feature Weightings (in Hungarian), VII. Hungarian Conference on Biometrics and Biomathematics, 5th-6th July 2005, Budapest (Hungary).
26. Attila Kertész-Farkas, Compact Representation of Hungarian Vocabulary with Nondeterministic Finite Automata I. Conference on Hungarian Computational Linguistic (in Hungarian, I. Magyar Számítógépes Nyelvészet Konferencia) 10th-11th December 2003, Szeged pp. 231-237, (Hungary)
27. Attila Kertész-Farkas, Kernel-based learning with dimension weighting (in Hungarian), Scientific Conference of Students, spring 2004, Supervisor: András Kocsor1
28. Attila Kertész-Farkas, Compact representation of finite languages with nondeterministic automata (in Hungarian), Scientific Conference of Students, autumn 2002, Supervisors: Zoltán Fülöp9 and András Kocsor1

Talks

29. Protein databases and similarity measures for protein classification Bioinformatics seminar at Rényi Mathematical Institute, Budapest, Hungary, 2007 Oct
30. Equivalence Learning in Protein Classification, Young Researcher Symposium on Intelligent Systems at the John von Neumann Computer Society, Budapest, Hungary, 2007 Nov
31. A new algorithm for identification of post-translational modification in tandem mass spectrometry data, Symposium at International Centre for Genetic Engineering and Biotechnology, Trieste Italy, 2011 Jun

Posters

32. Attila Kertész-Farkas, Olivier Bodenreider, Trevor C. Suznick, Gary Thompson, Yanan Sun and Maricel G. Kann, EMU: A tool for the Extraction of MUtations with disease associations from literature Growth Factor and Signal Transduction Conferences: System Biology, Integrative, Comaparitve and Multi-scale Modeling, 11–14 June, Iowa, USA
33. Bruna Marini, Attila Kertesz-Farkas, Hashim Ali, Mauro Giacca, Marina Lusic, The impact of nuclear topology on HIV-1 integration, 4th Conference on Nuclear Structure and Dynamics, EMBO Converence Series, 2011,