

Machine Learning Methods in Life Sciences

Graduate Course

Lecturer: Prof. dr. hab. Jan Komorowski

Time: Fridays, 13:15-15:00 on the following dates:

October 3rd, lecture hall 3089

October 17th, lecture hall 3089

November 14st, lecture hall 3075

January (to be announced)

Localization: Interdisciplinary Centre for Mathematical and Computational Modelling,
University of Warsaw, Al. Żwirki I Wigury 93

Course description

This course will introduce the student to Machine Learning methods applied to problems in Life Sciences. The syllabus of the course is based on the work done by Professor Komorowski's group and his collaborators. The aims of the course are to introduce graduate students to transdisciplinary research in bioinformatics and to teach how selected machine learning methods can be applied to solve complex problems in Life Sciences.

The course will be taught in a new and innovative way. There will be one introductory lecture to the rough set methodology followed by a set of exercise with the Rosetta computer system. The sequel to the first meeting will consist of presentations of selected research articles. The presenters will be the course participants in teams: one team of Life Science students and one of Computer/Mathematics/Statistics students. The LS partners will present the biological problem studied in the selected article while the C/M/S partners will present the computational part of the work. Both teams are then expected to repeat the computational experiment, if possible, and to present new or related findings based on literature search for the problem.

In order to PASS the course every person has to give an oral presentation and be present at all but one meeting, at least. The order of the presentations will be established at the first meeting. The number of participants is limited. The language of the course will be English.

Syllabus

I. October 3rd, Course Organization and Introduction to Rough Sets

- *Rough sets in bioinformatics*, TR Hvidsten and J Komorowski. **Transactions on Rough Sets VII** edited by E. Orłowska, J. F. Peters and A. Skowron, LNCS 4400, pp. 225-243, Springer-Verlag Berlin Heidelberg New York, 2007.
- *Rough Sets – A Tutorial*, Jan Komorowski, Lech Polkowski and Andrzej Skowron, In: *Rough Fuzzy Hybridization: A New Trend in Decision-Making*, Sankar K. Pal, Andrzej Skowron, Eds, Springer-Verlag New York, Incorporated 1999.

II. Modelling protein properties

- *Rough Set-Based Proteochemometrics Modeling of G-protein-Coupled Receptor-Ligand Interactions*, Helena Strömbergsson, Peteris Prusis, Herman Midelfart, Maris Lapinsh, Jarl E. S. Wikberg, Jan Komorowski, **Proteins: Structure, Function and Bioinformatics** 2006 63:1, 24-34.
- *Generalized modeling of enzyme-ligand interactions using proteochemometrics and local protein substructures*. *Proteins: Structure, Function and Bioinformatics*, H. Strömbergsson, A. Kryshtafovych, P. Prusis, K. Fidelis, J. E. S. Wikberg, J. Komorowski and T. R. Hvidsten. **Proteins: Structure, Function and Bioinformatics** 2006 65:3, 568-579.

- *Proteochemometric analysis of small cyclic peptides' interaction with wild-type and chimeric melanocortin receptors*, Aleksejs Kontijevskis, Ramona Petrovska, Ilze. Mutule, Staffan.Uhlen, Jan Komorowski, Peteris Prusis and Jarl E. Wikberg. **Proteins: Structure, Function and Bioinformatics** 2007 69:1 83-96.

III. Modeling HIV-1 protease and reverse transcriptase

- *Computational Proteomics Analysis of HIV-1 Protease Interactome*, Aleksejs Kontijevskis, Jarl E.S. Wikberg, Jan Komorowski. **Proteins: Structure, Function and Bioinformatics**. 2007 68:305–312.
- *A Look Inside HIV Resistance through Retroviral Protease Interaction Maps*, Aleksejs Kontijevskis, Peteris Prusis, Ramona Petrovska, Sviatlana Yahorava, Felikss Mutulis, Ilze. Mutule, Jan Komorowski and Jarl E. Wikberg. **PloS Computational Biology**, March 9, 2007.

IV. Cancer and Feature Selection

- *Markers of Adenocarcinoma Characteristic of the Site of Origin – Development of a Diagnostic Algorithm* J. L Dennis, T.R. Hvidsten, J. Komorowski, E. C Wit, A. Bell, I. Downie, J. Mooney, C. Verbeke, C. Bellamy, W.N. Keith and K.A. Oien, **Clin Cancer Res** 2005 11:10 3766-72.
- *Monte Carlo feature selection for supervised classification*, Michal Draminski, Alvaro Rada-Iglesias, Stefan Enroth, Claes Wadelius, Jacek Koronacki and Jan Komorowski. **Bioinformatics** 2008 24: 110-117.
- *Feature Synthesis and Extraction for the Construction of Generalized Properties of Amino Acids*, W. R. Rudnicki, J. Komorowski: *Rough Sets and Current Trends in Computing* 2004: Lecture Notes in Computer Science 3066, pp.786-791.

V Transcriptomics

- *Predicting Gene Ontology Biological Process From Temporal Gene Expression Patterns* Astrid Lægreid, Torgeir R. Hvidsten, Herman Midelfart, J. Komorowski and Arne K. Sandvik, **Genome Research**. 2003 13:5 965-79
- *Learning rule-based models of biological processes from gene expression time profiles using Gene Ontology*, T.R. Hvidsten, A. Lægreid, J. Komorowski, **Special Issue of Bioinformatics**, 2003 19:1116-23.
- *Discovering regulatory binding site modules using rule-based learning*, T. R. Hvidsten, B. Wilczyński, A. Kryshtafovych, J. Tiuryn, J. Komorowski and K. Fidelis, **Genome Research**, 2005 15:6 856-866.
- *Using local gene expression similarities to discover regulatory binding site modules*, Bartek Wilczyński, Torgeir R Hvidsten, Andriy Kryshtafovych, Jerzy Tiuryn, Jan Komorowski and Krzysztof Fidelis, **BMC Bioinformatics** 2006, 7:505
- *Using local gene expression similarities to discover regulatory binding site modules*, Bartek Wilczyński, Torgeir R Hvidsten, Andriy Kryshtafovych, Jerzy Tiuryn, Jan Komorowski and Krzysztof Fidelis, **BMC Bioinformatics** 2006, 7:505
- *Revealing cell cycle control by combining model-based detection of periodic expression with novel cis-regulatory descriptors* Claes R Andersson, Torgeir R Hvidsten, Anders Isaksson, Mats G. Gustafsson and Jan Komorowski, **BMC Systems Biology** 2007, 1:45 (16 Oct 2007).