Eric Bach

Curriculum Vitae

Muusantori 5A 18, 00350 Helsinki
+358 40 5893344
eric.bach@aalto.fi

About me

I am a Computer Scientist and Machine Learning expert with strong knowledge in Computational Metabolomics.

I work quality centered and solution oriented with the attitude to provide the right computational tools for biological questions.

I am a team player with effective communication skills and ready tell my opinion.

Keywords Friendly, Knowledgable, Curios, Self critical, Foresightful

Education

8/2016 – present **Doctoral Candidate**, *Aalto University · School of Science*, Espoo, Finnland.
In KEPACO bioinformatics research group under supervision of Prof. Juho Rousu
Topic: Machine Learning for Computational Metabolomics

4/2013 – 8/2016 **M.Sc. Computer Science**, *FSU*, Jena.
Major: Digital Image Processing, Minor: Machine Learning & Data-Mining

10/2015 – 7/2016 **Master's thesis**, *Aalto University · School of Science*, Espoo, Finnland.
Metabolite Identification using Magnitude-preserving Input Output Kernel Regression

10/2009 – 4/2013 **B.Sc. Applied Computer Science**, *FSU*, Jena.
Minor (40 ECTS): Computational Neuroscience, Studies encompassed 210 ECTS
Thesis: Adaption of a Semantic Segmentation Algorithm for Remote Sensing Data

Doctoral studies

Abstract I am developing machine learning based metabolomics data analysis pipelines to identify small molecules in biological samples using data arising in mass spectrometry analyses. My main focus lays in the integration of orthogonal compound information, e.g. liquid chromatography retention times, to improve the de-novo discovery of molecules.

Methods I am using kernel methods, e.g. Support Vector Machines, as machine learning framework. This covers the theoretical development of new algorithms until the practical implementation. I talk to mass spectrometry domain experts, evaluate the relevant literature and learn about the technical aspects of mass spectrometry devices, to ensure the relevance of my research. I integrate publicly available data to allow joint use as databases for my experiments. I manage my data using relational databases. I effectively communicate my research with others using presentations and reports.

Responsibilities

I am defining the "What?" and "How?" in my research. I structure my tasks and plan in foresighted manner, how to conduct my research. I do quality assurance for my experiments and results, e.g. automated code-testing and review processes. I am teaching university students in machine learning and bioinformatics. I design educational tasks in kernel methods, evaluate students and give constructive feedback. I organize contact teaching events for students, e.g. Question & Answer sessions.

Relevant work-experience

4/2011 – 7/2013
10/2014 – 8/2015

Research assistant, *Computer Vision Group*, FSU, Jena.

I researched in the fields of semantic image segmentation (image classification), object detection and deep-neural-networks. This included designing, implementing and extending algorithms in C++, R and Matlab in a Linux environment. I practiced understanding others' source code and tailoring implementations to specific tasks. I was responsible for the scheduling, the efficiency and the documentation of my work. I could apply knowledge from my university courses straight in practice and gained experience in presentations and problem orientated discussions. I pre-processed data-sets, conducted experiments (e.g. parameter estimation and performance measurements) and reported the evaluation results (e.g. to external industry collaborators). I read and utilized state-of-the-art scientific publications.

Technical expertise

Programming Languages

Matlab (>4 years)	Machine Learning (kernel methods, graphical models), Prototyping
R (>4 years)	Data visualization & organization (`ggplot, data.frame`), Processing molecular structures (`rcdk`), Statistical analyses
bash (>4 years)	Automatization of experiments, Operating Linux environments
Python (>2 years)	Machine learning (`sklearn`), Database administration (`sqlite3`), Molecule feature processing (`rdkit`), Jupyter Notebooks, Python education for students
C++ (>2 years)	algorithm implementation & software development
Java (>1 year)	Molecular structure processing cheminformatics (`CDK`)

General

Distributed computing	SLURM	*experience using workload manager to run large scale machine learning tasks on cluster systems*
Version control	git	*experience in using version control for various projects*
Typesetting	LaTeX	*reports, documentations & presentations*
Relational Databases	SQLite	*Organizing experimental data, Fast query of molecular structures, Unified data representation from different sources*

Publications

Eric Bach, Simon Rogers, John Williamson, and Juho Rousu. "Probabilistic framework for integration of mass spectrum and retention time information in small molecule identification". In: *Bioinformatics* (Nov. 2020). ISSN: 1367-4803

Sandor Szedmak and **Eric Bach**. "On the generalization of Tanimoto-type kernels to real valued functions". In: *arXiv preprint arXiv:2007.05943* (2020)

Eric Bach, Sandor Szedmak, Céline Brouard, Sebastian Böcker, and Juho Rousu. "Liquid-chromatography retention order prediction for metabolite identification". In: *Bioinformatics* 34.17 (2018), pp. i875–i883

Céline Brouard, **Eric Bach**, Sebastian Böcker, and Juho Rousu. "Magnitude-Preserving Ranking for Structured Outputs". In: *Asian Conference on Machine Learning*. 2017, pp. 407–422

Björn Fröhlich, **Eric Bach**, Irene Walde, Sören Hese, Christiane Schmullius, and Joachim Denzler. "Land cover classification of satellite images using contextual information". In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 3 (2013)