Lukas **Pfannschmidt**

Applied Scientist · Software Engineer

lukas@lpfann.me | lpfann.me | lpfann | lpfannschmidt | Google Scholar

Summary

Finishing PhD candidate with 5 years of experience dedicated to efficient and maintainable solutions to interesting problems. Studies of bioinformatics and machine learning with focus on feature selection. Very familiar with handling and representing diverse data.

Skills

Machine Learning	Feature Selection and Representation, Data Science, Model Selection and Design
Development	Software and algorithm design, parallel and efficient computing, Python, Databases, JAVA, Julia, C, LaTeX
Research	Scientific Writing and Presentation
DevOps	Docker, Kubernetes, Travis, Git, GitHub Actions, Linux
Languages	German (native), English (fluent)

Education

PhD in Machine Learning

Bielefeld, Germany
Vancouver, Canada

Bielefeld University, CITEC, SFU Vancouver

2016 - planned 2020

- Thesis: Relevance Learning for Redundant Features
- Member of *Prof. Hammer's* machine learning group
- Research stay at SFU Vancouver in *Prof. Ester's* datamining group

B. Sc. & M. Sc. in Bioinformatics and Genome Research

Bielefeld, Germany

Bielefeld University

2011 - 2016

- Master thesis: Interactive feature selection for biomedical data analysis
- Bachelor thesis: Survey of the cuckoo-RNA family beyond the Alphaproteobacteria

Projects

Feature Relevance Intervals - Python application

PhD project

- Developed parallelized and fast implementation of theoretical feature selection algorithm
- In comparison achieved the best accuracy and best scaling for big data sets and includes automatic hyperparameter tuning
- Released with *scikit-learn* API compatibility and extendable via modules
- Deployed via *GitHub* and *PyPi* package repository with continuous testing

Price Prediction in Dynamic Online Game Economy using Deep Learning

Side project

- Achieved prediction of prices of unseen in-game items based on historical data
- Made possible by efficient scraping of global marketplace streaming data and compressed database storage
- Developed novel set representation of in-game item features used in deep learning model with on average 30% better accuracy than alternatives

Endoscope Management Terminal - Professional Health App

Team Competition - Technical Lead - Winning Team

- We created Android app for medical professionals handling endoscopes in a clinical setting with high quality constraints
- I designed modular technical architecture and delegated appropriate tasks to a team of 10 co-workers in agile fashion
- Achieved first place in competition by integrating all requirements from endoscope manufacturer *Miele Professional*

Parallel K-Means Clustering - High Throughput Library

Study project

- Developed highly parallel and efficient implementation of clustering running on CPUs and GPUs – NVIDIA and AMD hardware
- Achieved linear speedup – performance scaling near perfectly with number of compute units

Adverse Drug Reactions Checker - User Health App

Study project

- Created user facing Android app warning against possible harmful interactions between medications in collaboration with local hospital
- Designed accessible, appealing but instructional user interface by integrating feedback of user studies
- Enabled up-to-date information by utilizing database backed infrastructure

Research

Center for Cognitive Interaction Technology — *Bielefeld, Germany*

PhD Candidate — *Oct. 2018 – 2020*

- Research in Prof. Hammer's machine learning group
- Research of feature relevance and potential applications

Simon Fraser University — *Vancouver, Canada*

Guest Researcher — *May. 2018 – Oct. 2018*

- Research stay at Prof. Martin Esters Datamining group
- Focus on feature representation and use of non-linear models

German-Canadian DFG International Research Training Group (1906/1) — *Bielefeld, Germany*

PhD Candidate & Research Fellow — *Oct. 2015 – April. 2018*

- Bioinformatics focused application research and development

Presentation

IEEE International Conference on Computational Intelligence in Bioinformatics and Computational Biology — *Sienna, Italy*

Speaker - Representing Co-Authors — *Jul. 2019*

- Introduced novel view on feature selection to bioinformatics researchers
- Announced software tool FRI to enable this feature selection for biomedical data

DFG Exchange Workshop for Research Training Groups — *Dagstuhl, Germany*

Speaker - Representing German-Canadian RTG DiDy — *Jun. 2017*

- Gave broad overview about DiDy's research topics and presented my project in detail

15th Bioinformatics Research and Education Workshop — *Bergen, Norway*

Speaker — *May. 2017*

- Presentation of theoretical feature relevances for applied feature selection

14th Bioinformatics Research and Education Workshop — *Helsinki, Finland*

Speaker — *May. 2016*

- Presentation of redundancy preserving feature selection paradigm

Honors & Awards

2019	**Fellow**, Bielefeld Young Researchers' Fund	*Bielefeld, Germany*
2019	**Student Speaker**, GRK 1906/1	*Bielefeld, Germany*
2017	**Student Speaker**, GRK 1906/1	*Bielefeld, Germany*
2016	**Fellow**, DFG Fast Track Fellowship	*Bielefeld, Germany*
2013	**Winning Team**, Miele Endoscope App	*Gütersloh, Germany*

Publications

Feature Relevance Determination for Ordinal Regression in the Context of Feature Redundancies and Privileged Information
Lukas Pfannschmidt, Jonathan Jakob, Fabian Hinder, Michael Biehl, Peter Tino, Barbara Hammer
Neurocomputing (Apr. 9, 2020). 2020

FRI – Feature Relevance Intervals for Interpretable and Interactive Data Exploration
Lukas Pfannschmidt, Christina Göpfert, Ursula Neumann, Dominik Heider, Barbara Hammer
2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2019

Feature Relevance Bounds for Ordinal Regression
Lukas Pfannschmidt, Jonathan Jakob, Michael Biehl, Peter Tino, Barbara Hammer
ESANN 2019, 2019, Bruges

Interpretation of Linear Classifiers by Means of Feature Relevance Bounds
Christina Göpfert, Lukas Pfannschmidt, Jan Philip Göpfert, Barbara Hammer
Neurocomputing 298 (July 12, 2018) pp. 69–79. Elsevier, 2018

Feature Relevance Bounds for Linear Classification
Christina Göpfert, Lukas Pfannschmidt, Barbara Hammer
Proceedings of the ESANN, 24th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2017, Bruges