

# Lukas Pfannschmidt

APPLIED SCIENTIST · SOFTWARE ENGINEER

✉ [lukas@lpfann.me](mailto:lukas@lpfann.me) | 🏠 [lpfann.me](https://lpfann.me) | 📄 [lpfann](https://lpfann.com) | 📺 [lpfannschmidt](https://www.youtube.com/channel/UCpfnnschmidt) | 🎓 [Google Scholar](https://scholar.google.com/citations?user=...)

## 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

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

## Education

### PhD in Machine Learning

BIELEFELD UNIVERSITY, CITEC, SFU VANCOUVER

- 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

*Bielefeld, Germany*

*Vancouver, Canada*

*2016 - planned 2020*

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

BIELEFELD UNIVERSITY

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

*Bielefeld, Germany*

*2011 - 2016*

## Projects

### Feature Relevance Intervals - Python application

PHD PROJECT

- Developed **parallelized** and **fast** implementation of theoretical feature selection algorithm
- In comparison achieved the **highest accuracy** and **best scaling** for big data sets and includes **automatic hyperparameter tuning**
- Released with **scikit-learn API compatibility** and 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
- Enabled by **efficient scraping** of extensive global marketplace streaming data and **compressed database storage**
- Developed novel set representation of **complex item features**; used in deep learning model with **30% better accuracy** than others

### 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**
- As Technical Lead 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 friendly Android app** warning against harmful interactions between medications in **collaboration** with local hospital
- Designed accessible, appealing and instructional user interface by integrating feedback of **user studies**
- Provides **up-to-date information** by utilizing database backend

## Research

---

### Center for Cognitive Interaction Technology

PHD CANDIDATE

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

*Bielefeld, Germany*

*Oct. 2018 – 2020*

### Simon Fraser University

GUEST RESEARCHER

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

*Vancouver, Canada*

*May. 2018 – Oct. 2018*

### German-Canadian DFG International Research Training Group (1906/1)

PHD CANDIDATE & RESEARCH FELLOW

- Bioinformatics focused application research and development

*Bielefeld, Germany*

*Oct. 2015 – April. 2018*

## Presentation

---

### IEEE International Conference on Computational Intelligence in Bioinformatics and Computational Biology

SPEAKER - REPRESENTING CO-AUTHORS

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

*Sienna, Italy*

*Jul. 2019*

### DFG Exchange Workshop for Research Training Groups

SPEAKER - REPRESENTING GERMAN-CANADIAN RTG DiDY

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

*Dagstuhl, Germany*

*Jun. 2017*

### 15th Bioinformatics Research and Education Workshop

SPEAKER

- Presentation of theoretical feature relevances for applied feature selection

*Bergen, Norway*

*May. 2017*

### 14th Bioinformatics Research and Education Workshop

SPEAKER

- Presentation of redundancy preserving feature selection paradigm

*Helsinki, Finland*

*May. 2016*

## 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