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)

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

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