Paul Westenthanner

Work Experience:

Freelance
Data Scientist / Machine Learning Engineer
Munich, Germany
Apr. 2017 – present

As freelancer I focus on delivering big data pipelines in Scala, Python and R.
Projects include:
- *Recommendation system for Adidas*:
    I am in charge of improving the recommendation engine used on adidas.com
    in a data-driven way using Python and a variety of data sources including
    order and product related data. Furthermore I have an advisory role in a
    project to predict purchases from clickstream data.
- *Fraud detection model for OTTO*:
    Development and productionizing of a fraud-detection-model in e-commerce
    based on clickstream data. The model was implemented and tested in Scala
    and parallelized using Apache Spark on a hadoop cluster. The model was also
    deployed for real time predictions in an Apache Kafka framework.
- *Analytical engine for Moubi Analytics*:
    Automate data-profiling and data pre-processing of new data-sources. This
    includes generating insights by visualizing the data, applying dimensionality
    reduction and clustering techniques.
- *Payment matching for a DAX30 company*:
    Development of an AI algorithm for matching open positions to incoming
    payments. Implementation was done in R. Crucial parts include sophisticated
    data pre-processing (including extracting information from PDF via OCR),
    feature generation and applying machine learning models for predicting
    matches.

Vodafone Group
Data Scientist
London, UK
Mar. 2016 – Apr. 2017 (full time)

As data scientist I was a main contributor to an end-to-end data science machine.
My responsibilities included ensuring data quality, understanding the raw data
and crafting of suitable features for predictive-analytic machine learning models.
Use-cases were for example travel prediction or device recommendation.
Furthermore I made key contributions to a smart Python and Scala framework to
implement data flows in production on an hadoop cluster. In this context, I unit-
tested the code, created and reviewed pull requests and trained colleagues and
external consultants to use the tool.

Vodafone Kabel Deutschland
Working Student in Data Science and Business Intelligence
Munich, Germany
Aug. 2014 – Mar. 2016 (part time, 16h/week)

As working student I supported the data science team with analysis of various technical data. My main focus was on finding multivariate correlations that could be exploited to optimize the assurance process. Other responsibilities included generation of KPIs or statistical support for other department's campaigns. As software tools I used Python and SQL.

Education:

Oct. 2015 – Present, Technical University of Munich
- Candidate for Master of Science in Mathematics
- Thesis: Predicting Travel from Smartphone Data (grade 1.3)
- Topics in Abstract Algebra, Optimization, Statistics, Machine Learning and Computer Science

Oct. 2012 – Sept. 2015, Technical University of Munich
- Bachelor of Science in Mathematics (grade 1.7, top 15%)
- Thesis: Algorithms in the representation theory of finite groups
- Topics in Analysis, Optimization Theory, Statistics and Algebra
- Minor subject: Economics

Sept. 2004 – July 2012, Maximilian-von-Montgelas Gymnasium Vilsbiburg
- Abitur (A level) (grade 1.4)

Skills:

Programming:
- *Expert*: Python (incl. pandas, sklearn, pyspark), Scala
- *Intermediate:* R, SQL
- *Basic*: MATLAB and C++

Tech Stack: Apache Spark, Hadoop, Apache Kafka, Impala, Jenkins, Keras, Tensorflow, CARTO, TwitterAPI, Microsoft Office, Linux, Git, OracleSQL, AWS, Jupyter Notebook, Intellij IDEA

Languages: German (native), English (fluent), Latin (basic), Spanish (basic)

Other:

Speaker at @databeers London, code at github.com/PaulWestenthanner
Interests: Chess (club player), reading, cycling