



Kaspar Paul Westenthanner

Data Scientist & Engineer

- 09.06.1994
- Keuslinstraße 4, 80798 München
- +49 174 5476560
- <https://westenthanner.dev>
- kaspar@westenthanner.dev
- LinkedIn
- Github

About Me

Freelancer who likes data, machine learning, open source and clean code.

Languages

- Python ●●●●●
- SQL ●●●●●
- Scala ●●●●●
- R ●●●●●
- Bash ●●●●●

Skills

- Apache Spark ●●●●●
- Machine Learning ●●●●●
- Software Engineering ●●●●●
- Analytics ●●●●●
- Apache Kafka ●●●●●
- Git ●●●●●
- Relational Databases ●●●●●
- Docker ●●●●●
- Benchling ●●●●●
- DevOps - CI/CD ●●●●●
- Tableau ●●●●●
- Cloud ●●●●●

Working Experience

- 10/2020 – currently **Data Engineer (Freelance)** Roche, Penzberg
I'm in charge of digitalizing and automating processes in a bio-chemistry lab.
I develop web applications and REST interfaces to automate analytical assays on a high throughput lab robotics system. Furthermore, I design and implement data models within the Benchling ELN system.
Technologies: Python, REST (flask), Plotly-Dash, DevOps, Docker, MongoDB, HTML, CSS, Benchling.
- 10/2022 – 02/2023 **Data Engineer (Freelance)** SICK, Freiburg i. Br.
As a data engineer I worked on ingesting SAP data into S3 and a front-end for enabling end-users to configure load jobs.
Technologies: Scala, NiFi, Spark, Minio S3, CI/CD, SAP.
- 10/2018 – 09/2020 **Data Engineer (Freelance)** Telefonica, München
As the lead backend developer of a time series problem detection framework, I assured high code quality, trained junior developers and supported end users. Furthermore I advised other projects on software architecture, efficient computing and theoretical topics in machine learning.
Technologies: PySpark, Hadoop Stack, Oracle, Apache Kafka, R-Shiny, Tableau.
- 04/2018 – 09/2018 **Data Scientist (Freelance)** Adidas, Herzogenaurach
I worked on a variety of data driven Proof-of-Concepts like improving product recommendations, click-stream analysis and purchase propensity.
Technologies: Python, SQL (Exasol), AWS.
- 12/2017 – 04/2018 **Machine Learning Engineer (Freelance)** OTTO, Hamburg
I developed a fraud detection model based on click-stream data. The model improved the training time by two orders of magnitude and was capable of real-time predictions.
Technologies: Scala, Apache Spark, Apache Hadoop, Apache Kafka.
- 03/2016 – 04/2017 **Data Scientist (Intern)** Vodafone Group, London
I built key parts of a framework for ETL and feature engineering. In this context, I unit-tested the code, created and reviewed pull requests and trained colleagues and external consultants to use the tool.
Technologies: Python, Scala, Hadoop, Spark.

Education

- 2015 – 2018 **Mathematics (Master - Grade 1.7)** Technical University of Munich
Thesis: "Predicting Travel from Smartphone Data using Machine Learning" in collaboration with Vodafone Group London.
Topics in Abstract Algebra, Optimization, Statistics, Machine Learning and Computer Science
- 2012 – 2015 **Mathematics (Bachelor - Grade 1.7)** Technical University of Munich
Thesis: Algorithms in the representation theory of finite groups

Extra-Curricular Activities

- Volunteer** Maintainer of sklearn category encoders with over 1 million downloads per month.
- Interests** Bike repairer at Ackermannbogen Repair Cafe. Chess (playing and chess programming), Cycling, Running, Philosophy, History