Jan-Benedikt Jagusch

Data Engineer, Data Scientist and Open Source Fan

Heidelberg, Germany

📞 0049176/71201520

✉ jan.jagusch@gmail.com

I am a data engineer that enjoys building great processes, infrastructure and tooling to make his data science colleagues' work more efficient and ready for production.

EXPERIENCE

QuantCo
(2021/01 -)

Data Engineer

- Improving prediction model containerization, deployment and orchestration
- Supporting engineering team strategy and culture

Instaffo
(2019/01 -)

Team Lead Data Science

- Managing and mentoring the 5 person data science team
- Designing and developing machine learning models for matching jobs to candidates
- Planning company data warehouse and implementing ETL processes
- Analyzing and visualizing data for internal and external stakeholders using interactive dashboards
- Introducing best practices regarding code quality, documentation, automation, testing and deployment

PyData
(2019/01 -)

Community Manager and Speaker

- Organizing data science and open source events in the Rhine-Neckar area
- Growing the local community from 300 to more than 1,100 members in one year
- Engaging community members in international charity competition (PyGiving) and finishing in first place
- Speaking at meetups and conferences

EDUCATION

New University of Lisbon
(2016/09 - 2018/08)

Masters in Information Management

- Teaching masters course in natural language processing
- Publishing research about applying genetic programming to deep learning in A rated conference (GECCO)

Baden-Wuerttemberg Cooperative State University
(2013/09 - 2016/08)

Bachelors in International Management for Business and IT

- Completing dual study program together with Evonik Industries
- Studying one semester at University of Tennessee, USA and working in Sao Paulo, Brazil

SKILLS

Python   Data Sciecne   Machine Learning

Data Engineering   Software Engineering

Deep Learning   Genetic Algorithms   Teaching

Mentoring   Public Speaking

TOOLS, PACKAGES AND FRAMEWORKS

Argo   Docker   Gitlab CI   Google BigQuery

Google Cloud Functions   Google Data Studio

Google PubSub   Jupyter   Kubernetes   MLFlow

Numpy   Pandas   Plotly   Poetry   Pytest

RabbitMQ   Redis   Scikit-Learn   Sentry   SQL

Travis CI   Conda