Liyan Jiang

Data Engineer (Current)

Belgradstraße 140   Email   Phone +49 01746292254
80804, Munich   liyann.jiang@gmail.com

Summary

Liyan is a Pythonista who has a good understanding of **natural language processing** and **data mining**. With more than 5 years of **Python** experience, she is proficient with **data analysis** libraries (NumPy, Pandas), **OOP** paradigm, **SQL** and **Spark** - skills that every data scientist has in their toolkit. She enjoys coding and solving engineering problems in her day-to-day work. Currently she's teaching herself **Rust** in her spare time.

Work Experience

10.2020-now **Data Engineer**, Cobrainer GmbH, Munich

Built and maintained a skill knowledge graph using workflow management tools (Airflow, Prefect) that incorporates external data sources on a regular basis.

- Data Wrangling, ETL of skills data from raw Wikipedia dumps
- Implemented data processing pipeline
- Performance improvement using concurrent computing

Developed and maintained a service endpoint written in Python using FastAPI

- Storage: MongoDB, Redis
- Background task scheduler: celery
- Design pattern: event-driven system, OOP

01.2020 - 05.2020 **Working Student**, Siemens AG, Munich

Built a POC of information retrieval system leveraging modern clustering methods and SOTA machine learning models.

- Document content matching based on embeddings generated by a multi-lingual BERT model
- Data visualisation of the document similarity in a PCA-reduced 3 dimensional space
- AWS SageMaker

02.2018 - **Chatbot Developer**, rb2, Amsterdam

01.2019   Intent identification and name entity recognition(NER) using RASA-NLU library

- Trained and tuned a LSTM model using 10,000+ user utterance data to learn to identify user intentions and the name entities
- Designed a rule-based question answering model using RASA-core library

Education

2018-2020   **Technical University of Munich (TUM), Germany**

Master's Degree of Informatics

- Master thesis: Deep Active Learning Methods for Opinion Mining

2013-2017   **Xiamen University, China**

Bachelor's Degree of Software Engineering

Tools

Git, Jupyter Notebook, VSCode, PyCharm, JIRA, GitLab-CI

Skills

Python • Rust • Natural Language Processing (NLP) • Machine Learning • Data Mining • Knowledge Graph • AWS • CI/CD • Docker • RESTful • SQL • TypeScript

Languages

Chinese - Native Speaker   **English** - C1   **German** - B2

Links

GitHub   LinkedIn   Kaggle

Liyan Jiang • Data Engineer   last updated: 09.08.2022