

Hao Hu

✉ haohu001108@gmail.com | ☎ +46 763459916 | 📧 himmelhao | 🌐 haohu1108

Certified Data Engineer Associate with experience in Cloud Infrastructure, Data Pipelines, and Machine Learning. Currently working with Azure, Databricks, Spark, and Terraform.

Professional Experience

**Data Engineer**
Aurobay

2024.8-Present
Gothenburg, Sweden

- **Automated warranty data pipelines in Databricks, enabling Power BI integration for analytics:**
 - Processed warranty data using Python and Spark for cleaning, transformation and optimization.
 - Configured Unity Catalog for data governance and access control.
 - Set up automated triggers via Azure Function App upon new file arrivals.
- **Designed and deployed automated Azure and Databricks platform template:**
 - Implemented infrastructure as code using Terraform and Copier with CI/CD pipelines via GitHub Actions.
 - Established role-based access control and security policies across environments.
 - Onboarded 4 engineering teams, standardizing setup and reducing manual effort by 95%.
- **Developed reusable asset bundle Databricks project template:**
 - Created reusable notebook templates with Delta Lake integration.
 - Implemented testing framework and quality checks for Databricks workflows.
 - Configured automated job and pipeline scheduling using Databricks workflows.
- Worked through Hubbau’s young talent program.

Databricks Azure Terraform Python Spark SQL Delta Lake CI/CD GitHub Actions DevOps Agile

**Master Thesis Student**
Smart Eye

2024.1-2024.6
Gothenburg, Sweden

- Developed efficient algorithm for detecting abandoned objects in vehicle cabins using background model and MobileNetV2. Optimized for embedded execution on Jetson Nano and Linux systems.

C++ Machine Learning Computer Vision OpenCV ROS Git Embedded Systems

**Data Engineer Intern**
Volvo Autonomous Solutions

2023.6-2023.8
Gothenburg, Sweden

- Successfully converted a variety of simulation data into PostgreSQL tables with a 100% success rate.
- Created a parallel method based on functional programming, speeding up the conversion of data frames to PostgreSQL tables by a factor of 40.

Python PostgreSQL Git JIRA Parallel processing Docker Documentation

**Data Analyst Intern**
Hikvision

2021.1 - 2021.2
Nanchang, China

- Used Python’s requests for websites, Excel files, and MySQL to gather Data.
- Employed Python and scikit-learn to build a verification model for accuracy assessment.

Python MySQL scikit-learn Beautiful Soup


Education

**Master of Computer Science, Algorithms, Languages and Logic**
Chalmers University of Technology

2022.8 - 2024.8
Gothenburg, Sweden

- Completed courses in functional programming, machine learning and cyber security. Teaching Assistant for “Programming Language Technology” course.

Python Haskell Linux Java C++ Erlang Deep Learning Cyber Security

**Bachelors of Information System**
Dalarna University

2019.8 - 2022.6
Borlänge, Sweden

- Built IT solutions using Java, web technologies, and databases. Applied AI/ML in business strategy and system analysis. Bachelor thesis focused on digitalization strategies for port operations.

Java Python SQL Machine Learning Data Architecture Business Model Web Development

Certificates

Feb, 2025 **Databricks Certified Data Engineer Associate**, Databricks
Nov, 2024 **Microsoft Certified Azure Fundamentals**, Microsoft