Ashwath Guru

☎ +91-7200545236

✉ ashwathguru.s@gmail.com

in LinkedIn

Portfolio

Summary

Results-driven **Data & AI Engineer** with over 6 years of hands-on experience in **data engineering** and **data science**, delivering robust data solutions across **on-premise** and **cloud environments**.

- Proven expertise in designing and implementing scalable **data pipelines**, **ETL workflows**, and **data architectures** using tools such as **Apache Spark**, **Python**, **SQL**, **Databricks**, **AWS Glue**, **GCP** and **Azure Data Factory**.
- Skilled in processing high-volume, complex datasets with advanced knowledge of **PySpark**, **SQL optimization**, **NoSQL, LLMs** and cloud-native services, ensuring optimal performance and reliability.
- Adept at developing and managing end-to-end **data workflows**, enabling actionable insights through comprehensive **data analysis** and **visualization**.
- Collaborative and communicative team player, experienced in working with cross-functional, globally distributed agile teams to meet business objectives and stakeholder expectations.
- Passionate about leveraging **Generative AI** and **machine learning** technologies to create cutting-edge, AI-enabled data solutions for innovation and efficiency.

Experience

Senior Data Engineer
ACL Digital | March 2022 – Present

- **Client 1: Barnes & Noble |** Domain**: Supply Chain**
  - Developed a new reporting tool from scratch, enabling users to generate tailored reports from a legacy Order Management System.
  - Designed and developed a **delta ingestion pipeline** leveraging **AWS Glue**, **S3**, and **Postgres** to ingest and transform data from **SFTP sources**, supporting robust **ETL workflows** for legacy Order Management Systems.
  - Built **APIs using AWS Lambda** to enable report generation and scheduling in **xlsx format**, significantly improving reporting efficiency and user experience.
  - Engineered the first custom reporting solution, resolving critical technical challenges to establish a **scalable reporting framework** for future use cases.
  - Enhanced the supply chain process with efficient data orchestration and automation using AWS native services.

- **Client 2**: Sure Scripts LLC | Domain: **Healthcare IT**
  - Developed robust solutions tailored for **doctors**, **pharmacies**, and **insurance providers**, leveraging **Scala** and **GCP-based applications** to streamline healthcare data workflows.
  - Emphasized integration with **EMR** and **EHR systems**, ensuring compliance with healthcare standards and delivering high-quality data insights.
  - Built secure **authentication systems** and developed **APIs** for generating personalized recommendations within the healthcare ecosystem.

- **Gen AI Labs**: Building Gen AI solutions for existing clients
  - **Project 1: Image-Based Rack Awareness and Restocking**
    - Developed an advanced **computer vision model** using **PyTorch**, **YOLOv5**, and **RoboFlow** to analyse rack images, identify inventory levels, and generate restocking alerts.
    - Automated inventory reporting and notifications, reducing manual intervention and improving operational efficiency for retail clients.
  - **Project 2: Custom Chatbots for Document-Based Query Resolution**
    - Engineered both **paid and open-source chatbot solutions** using **OpenAI**, **LangChain**, **Streamlit**, **LLMs**, and **Hugging Face**, with a **FAISS DB** for efficient vector storage.
    - Integrated **Guardrails** and the **Olama Framework** to enhance response accuracy, security, and user interaction reliability for enterprise applications.

Member Technical Staff - Python
OpsVeda Asia | October 2020 – March 2022

- Developed and optimized **supply chain data solutions** for a SaaS platform, including **real-time monitoring systems**, automated **Python-based reporting workflows**, and robust **data pipelines** using **PostgreSQL**, **Liquibase**, and **Jenkins**.
- Contributed to **product engineering** with bi-weekly **feature upgrades** and **bug fixes**, ensuring continuous improvement and operational excellence.

Data Engineer (Intern)
Reliance Industries | June 2019 – May 2020

- Enhanced ETL performance by 75% through optimization of **Sqoop** and **Scala** code.
- Improved query response times by 40% by optimizing **Hive** and **Phoenix** configurations.

Software Engineer – Trainee: SQL & PL/SQL Solutions for Oil & Gas Domain.
Value Global | April 2017 – May 2018

Certifications

Certification	Issued on
AWS Certified Cloud Practitioner (CLF-C02)	Oct 2024
Data Virtualization Champion – Denodo	Mar 2024
Databricks Lakehouse Fundamentals	Feb 2024
Langchain & Vector Databases in production	Oct 2023
Dataiku ETL Advanced Designer	Jan 2023
Oracle Certified Professional (PL/SQL Developer)	Jan 2017

Education

M.E. in Big Data & Data Analytics | *Manipal School of Information Sciences, Manipal University*
June 2018 – July 2020
CGPA: 9.05/10

B.E. in Electronics and Communication Engineering | *Sri Venkateswara College of Engineering (SVCE), Anna University*
August 2012 – May 2016