

HASWANTH RAJESH

haswanthrajeshn@gmail.com | +1 9452411532 | Farmington Hills, MI | LinkedIn | GitHub

PROFESSIONAL SUMMARY

Data Engineer with 4+ years of experience specializing in **Azure data engineering and Databricks development**. Proven expertise in designing, building, and optimizing **scalable, high-performance data pipelines** using **PySpark, Delta Lake, SQL, and Azure Data Factory**. Experienced in implementing **Unity Catalog, CI/CD pipelines with Azure DevOps**, and **enterprise data governance frameworks** to ensure secure, reliable, and compliant data platforms. Collaborative and agile professional with strong troubleshooting skills, consistently delivering **production-ready data solutions** for enterprise applications. Actively **designing and implementing Generative AI solutions**, leveraging **foundation LLMs, prompt engineering, embeddings, and multimodal tools such as Midjourney** to enhance data processing, automation, and analytical capabilities.

TECHNICAL SKILLS

Cloud Platforms	Azure, Azure Data Factory, Azure Data Lake Storage, Synapse
Big Data & Processing	Delta Lake, Databricks Workflows, Auto Loader, Structured Streaming, Unity Catalog, PySpark, Kafka, Partitioning, Caching, Delta Optimization
Databases	SQL, Data Modeling, Schema Evolution, Data Governance, Snowflake
Programming Languages	Python, Java, JSON
DevOps & CI/CD	Azure DevOps, Git, GitHub Actions, Jenkins, Job Scheduling
Analytics & BI	Power BI
Other Tools & Frameworks	APIs (REST/SOAP), ETL/ELT, Orchestration Pipelines, JIRA, Airflow, SSMS, SSIS, Azure Data Studio, Microsoft Fabric

GENAI & MACHINE LEARNING SKILLS

Foundation LLMs, Midjourney, Vector Pinecone, Chroma, Databases FAISS, Azure AI Search matching, RAG, Methods Chunking, Context Window retrieval, Prompt CoT reasoning, System prompt adaptation, Output control, GenAI Transformers, Attention Custom AI architecture, Embedding models, Token processing, LLM, LangChain, Azure OpenAI, AWS Bedrock, MLflow, Responsible AI Hallucination detection, Output safety, LLMs, fine-tuning, Pinecone, chunking, FAISS, Azure AI Search, vector store

PROFESSIONAL EXPERIENCE

Azure Data Engineer

DatafactZ, Northville, MI | Jul 2023 – Present

- Led the design of end-to-end **Snowflake** data pipelines using **Databricks** and **PySpark**, accelerating processing by 40%.
- Spearheaded the creation of parameterized ETL workflows in **Azure Data Factory** (ADF) and orchestrated them with **Databricks** Workflows.
- Implemented Unity Catalog and **Snowflake** data governance frameworks for access control and metadata management.
- Collaborated in agile sprints to deliver new data features and support production pipelines with **Azure** technologies.
- Developed CI/CD workflows in **Azure DevOps** for automated pipeline deployment and version control, enhancing efficiency.
- Conducted root cause analysis and optimized cluster configurations for improved performance in **Snowflake** environments.
- Utilized machine learning concepts to enhance data processing capabilities at scale with **Databricks** and **Python**.
- Built scaled data science enablement tools to support engineering processes and empower data scientists in **Azure** environments.
- Implemented and optimized **Snowflake** data warehousing solutions to support business intelligence and analytics.

Data Engineer (Databricks)

DatafactZ, Hyderabad, India | Jul 2021 – Aug 2022

- Architected the **Snowflake**-powered Medallion Lakehouse using **DataBricks** Auto Loader, reducing processing time by 30%.
- Implemented scalable ingestion and transformation pipelines with **Kafka**, **ADF**, and Delta Lake, boosting data processing efficiency by 40%.
- Led the development of scalable and reusable data pipelines supporting BI and analytics, streamlining workflows by 30%.
- Enforced data quality, encryption, and compliance policies, ensuring 100% data integrity and regulatory compliance.
- Mentored junior engineers on **PySpark** best practices, resulting in a 25% improvement in code optimization.
- Implemented and optimized data warehousing solutions in **Snowflake**, enhancing data accessibility and reliability by 40%.
- Collaborated with cross-functional teams to deliver data-driven solutions aligned with business goals.
- Applied machine learning concepts to enhance data processing capabilities at scale, improving insights by 25%.
- Developed data science enablement tools to support engineering processes, increasing efficiency and collaboration among teams.

EDUCATION

Master of Science in Computer Science - University of Texas at Arlington

Graduated: 05/2024