

HARISH ARAVINDH M

DATA ENGINEER

CONTACT

Email harisharavindhm@gmail.com; harisharavindhm@outlook.com
Phone +91 7395842051; +91 7904009417
Location Chennai, India.
LinkedIn <https://www.linkedin.com/in/harish-aravindh-m-331288129>

PROFESSIONAL SUMMARY

Data Engineer with **5 years of experience** building and optimizing **cloud-based data platforms** and ETL pipelines for telecom and banking domains. Strong expertise in **Python, PySpark, SQL, and cloud data processing (AWS/Databricks)**, with proven success in **SAS-to-Python migrations**, performance optimization, and **production pipelines processing millions of records**. Experienced in owning **end-to-end data workflows**, building Python applications for **API-driven ingestion, ETL workflow orchestration, automated monitoring and GenAI proof of concepts**. Hands-on experience with **time-series forecasting, anomaly detection, and data visualization** using Python libraries.

CORE SKILLS

| | |
|------------------------|---|
| Languages & Frameworks | Python, PySpark, Pandas, NumPy, SQL, Spark (Scala), Shell Scripting, Base SAS |
| Data Engineering | Databricks, AWS – Glue, Athena, Redshift, Apache Airflow, ETL Design |
| Cloud Platforms | AWS (S3, Lambda, EC2, ECR), Azure Databricks |
| Databases | Oracle (PL/SQL), MySQL, DB2, MongoDB, Neo4j, DuckDB, Vector Databases |
| Big Data & Analytics | Apache Spark, Databricks Lakehouse |
| ML & GenAI | Scikit-Learn, Prophet, LLMs, MCP, LangChain, LangGraph |
| DevOps & Tools | Git, Linux, Docker, GitHub Actions, Terraform |

PROFESSIONAL EXPERIENCE

Comcast India Engineering Center I LLP

Engineer 3, Software Development and Engineering (September 2025 – Present)

Engineer 2, Software Development and Engineering (November 2023 – August 2025)

- Designed and owned **Spark (PySpark & basic Scala exposure) ETL pipelines on Databricks** to process large-scale telecom datasets for analytics and reporting.
- Developed **AWS Glue jobs** to ingest, transform, and publish data to **Athena and Redshift**, enabling ad-hoc and downstream consumption.
- Built **backward-compatible Redshift views** as part of legacy data migration initiatives, ensuring zero disruption to consuming applications.
- Optimized Spark jobs through **partitioning, caching strategies, and efficient data formats**, improving runtime and reducing compute cost.
- Developed **Pandas-based ETL workflows** for data transformation, validation, and enrichment outside of Spark for lightweight and ad-hoc processing needs.

- Implemented **AWS Lambda functions** for payload validation, orchestration logic, and triggering downstream data processing workflows.
- Built **time-series forecasting (Prophet)** and **anomaly detection (Isolation Forest)** models to analyze trends and identify outliers, and developed **GenAI proof-of-concepts using LLM-based frameworks** to automate internal data workflows and experimentation use cases.
- Collaborated with **US stakeholders and analytics teams** to translate business requirements into scalable, production-ready data solutions.

Impact: Improved reliability and performance of production data pipelines while supporting cloud-native modernization initiatives.

Accenture Solutions Private Limited

Data Engineering Analyst (June 2022 – November 2023)

Data Engineering Associate (December 2020 – May 2022)

- Played a key role in **SAS-to-Python (Pandas) migration** projects for major North American banking clients (**TD Bank, Synchrony Financials**), converting complex ETL, analytics, and reporting workloads.
- Designed and optimized **Python-based data processing frameworks** using multiprocessing, vectorization, and memory-efficient data structures.
- Enhanced runtime performance of migrated pipelines to match SAS Grid execution for **millions of records in production**.
- Modified and standardized **Unix/Linux shell scripts** to orchestrate Python workflows, supporting platform modernization efforts.
- Built Python utilities to perform **SAS inventory analysis**, identifying workload patterns, data lineage, and migration readiness.
- Conducted **unit testing and data validation**, ensuring parity between SAS and Python outputs during client acceptance phases.
- Supported production issues and stabilization activities until successful **SAS grid decommissioning**.

Impact: Enabled large-scale legacy modernization, reducing dependency on proprietary platforms and accelerating cloud adoption.

CERTIFICATIONS

- Databricks Certified – Data Engineer Associate
- AWS Certified – Developer Associate (DVA-C02)
- Microsoft Azure Fundamentals Certified (AZ900)

EDUCATION

Bachelor of Engineering - ELECTRONICS AND COMMUNICATION ENGINEERING (8.04/10)

Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College (Affiliated to Anna University), Chennai.
2016 – 2020.

AWARDS AND RECOGNITION

Recognized with Outstanding Rating, promotion, and ACE award for high-impact client delivery in Accenture. Received multiple HIGH FIVE Champ awards and high ratings for performance and business impact in Comcast.