

SUMMARY

Databricks Certified Associate Data Engineer with overall 3+ years of experience in architecting and optimizing data pipelines, Delta Lake workflows, and distributed processing using Apache Spark. Experienced in delivering end-to-end data engineering solutions on Microsoft Fabric, including lakehouse design, medallion architecture, data ingestion and transformation. Proficient in Python, SQL, PySpark, notebooks-based development, and cloud-native orchestration. Additionally skilled in AI/ML development—fine-tuning LLMs, building NLP pipelines, and implementing embeddings/vector search—to extend analytical capabilities. Previously held Azure Data Engineer certification (expired).

TECHNICAL SKILLS

Programming & Data Engineering: Python, SQL, PySpark, Delta Lake, Apache Spark, Databricks, Microsoft Fabric

Libraries & Frameworks: TensorFlow, PyTorch, Hugging Face Transformers, LangChain, Scikit-learn

GenAI & NLP: BERT, T5, GPT-based LLMs, Prompt Engineering, Fine-tuning, Embeddings, RAG Pipelines

Vector Databases: FAISS, Pinecone, Azure Cognitive Search

Cloud Platforms: Azure ML Studio, Azure Databricks, Azure OpenAI Service, Azure Functions, Azure DevOps

MLOps & Deployment: Docker, Flask/FastAPI (REST APIs), CI/CD Pipelines, Git, Model Deployment & Monitoring

Tools & Platforms: Jupyter, VS Code, PyCharm, GitHub, Postman

INTERNSHIPS

Intern	Mar '16 - Jun '15
ECIL	Hyderabad, India
Participated for 1 month	
I collaborated on an industrial project for Hindustan Petroleum involving the Ramanmandi-Bahadurgarh pipelines (RBPL)	

PROFESSIONAL EXPERIENCE

Data scientist	Sep '24 - Present
YASH Technologies	Hyderabad, India

Working as a Data Engineer at YASH, supporting a UK-based client (Binks) in developing and maintaining scalable data transformation pipelines on Microsoft Fabric. I design and optimize ETL workflows and analytical models using Python, SQL, and PySpark

Binks

- Developed **end-to-end data pipelines for 5 ERP systems** on **Microsoft Fabric**, enabling unified data ingestion, transformation, and analytics at scale
- Implemented **Medallion Architecture (Bronze → Silver → Gold)** for structured data processing and governance
- Ingested data from **MS SQL source systems** into the **Bronze layer**, storing raw datasets in **Parquet format** for optimized storage and incremental processing
- Built **Silver layer pipelines** to transform Bronze data by applying **column renaming, duplicate removal, validation rules, and data quality checks**, ensuring clean and standardized datasets
- Converted client-provided **CTEs (Common Table Expressions)** into optimized **Spark SQL transformations**, leveraging Silver tables to generate curated **Gold layer** datasets
- Delivered **Gold layer models** for downstream analytics and reporting, ensuring high performance and scalability in the Fabric Lakehouse environment
- Designed and implemented a **metadata-driven solution**, reducing manual configuration, improving pipeline reusability, and enabling dynamic onboarding of new data sources
- Leveraged **Python, SQL, and PySpark within Microsoft Fabric notebooks** to implement scalable data processing and transformation workflows

○ **Worked closely with cross-functional teams** to maintain data quality, enforce best practices, and support CI/CD for Fabric-based data pipelines

Data scientist Jan '23 - Sep '24

Dartin Technologies Hyderabad, India

Working as a Data scientist for Lenia Healthcare. I've been working for healthcare system in building various transformation models. I leverage, Python, SQL, Flask

I.A.E Bot Platform Paris, France

- **EPM Chatbot**
 - **Developed a chatbot using Google services for EPM systems**, enabling end users to retrieve real-time weather reports on demand, **improving user query response efficiency by 40%**
 - **Empowered users to create, modify, execute, and delete MDX scripts**, providing full control over analytical operations and **reducing manual report generation effort by 35%**
 - **Ensured seamless integration with BI dashboards**, delivering executed responses that **enhanced data visualization comprehension for 90% of end users**
 - **Integrated a Gen-AI model into the chatbot**, allowing end users to interact via general conversation and AI-driven insights, **increasing engagement rates by 25%**
 - **Implemented the complete solution using a Python Flask application**, managing backend logic, API endpoints, and chatbot functionality, ensuring 99% uptime and stable performance
- **OVH cloud platform**
 - **Developed a cloud-native chatbot on OVH Cloud** using **Python and Azure services**, enabling automated **virtual machine provisioning and billing operations**, improving operational efficiency by **30%**
 - **Integrated backup and invoice generation features**, streamlining end-user workflows and **enhancing automation across key tasks**
 - **Designed a RAG-like (Retrieval-Augmented Generation) interaction pipeline** to handle both structured and unstructured data, **ensuring accurate and context-aware responses**

Lenia Healthcare Australia

- **Integrated GPT-based Gen-AI model into a healthcare chatbot** using **Azure and Flask**, enabling natural dialogue and efficient medical data retrieval
- **Developed transformation models to convert HL7 messages into FHIR bundles**, ensuring standardized healthcare interoperability
- **Created optimized transformation models** to map fields from HL7 message segments of various versions into FHIR bundles
- **Built RESTful APIs for model inference and data processing**, deployed using **Docker on Azure VMs**
- **Implemented prompt engineering strategies** to enhance user interaction with embedded LLM in chatbot systems

Jr. Software Engineer Sep '21 - Sep '22

Cognizant Technology Solutions Bangalore

Perform development projects and enhancements on Financial Regulatory reporting systems (FIREG).Execute on the development projects and enhancements of Fire Loss Corporate (LF) CM Bureau Loss (BL)/ CM Bureau Premium (BP) and COBRA (LO) systems. Learn and pick up work in our AM and OCAS BI Finance systems.

- **Translate business requirements into technical specifications** and develop reliable, scalable programs for machine learning applications
- **Designed methods to efficiently transform and clean data**, optimizing memory usage and calculating key metrics for downstream analytics
- **Reviewed and interpreted functional and technical documents**, preparing clear and actionable documentation for development teams
- **Analyzed and enhanced existing programs**, improving performance, reliability, and maintainability
- **Performed coding and unit testing**, ensuring robust and error-free implementations
- **Prepared Unit Test Plans (UTP) and documented Unit Test Results (UTR)** to validate code functionality and completeness
- **Resolved defects promptly** and communicated updates or changes to application modules for seamless release management
- **Reviewed peer code and work**, documenting feedback and improvements to maintain high-quality standards
- **Validated and debugged data pipelines**, ensuring accurate processing and publishing of data into live environments
- **Actively participated in knowledge-sharing sessions**, contributing insights and best practices within the team and project

Trainee Oct '19 - Dec '20

Flipchip Electronics Pvt. Ltd Bangalore

I served as a Trainee at FlipChip Electronics Pvt. Ltd., an Electronic Division of YASH Technologies.

- **Identified and debugged complex electronic circuits**, ensuring accurate functionality and performance
- **Contributed to the Quality Control Department**, upholding strict quality standards across all production stages
- **Maintained manufacturing quality at 98% accuracy** by verifying precise component placement on circuit boards
- **Tested and debugged advanced electronic circuit boards**, resolving issues to ensure reliable and defect-free operation

EDUCATION

Master of Science (Artificial Intelligence and Machine Learning)	Oct '24 - May '25
Liverpool John Moores University (LJMU)	Liverpool, United Kingdom
B.Tech (Electronics and Communication Engineering)	Aug '15 - Jun '19
AVN Institute of Engineering and Technology	Hyderabad
Intermediate (MPC)	Jun '13 - Mar '15
Sri Chaitanya jr College	Hyderabad
10th Standard (SSC)	Jul '11 - Mar '12
Bhashyam public School	Hyderabad

CERTIFICATIONS/TRAINING

Microsoft Certified: Azure Data Engineer Associate

Google Certified: Data to AI with Gemini & Gemma - Google Developer Groups

IBM Certified: Python for Data Science – [dad361320f7841f185cf131c7beff636](#)