NIRMAL KUMAR P

Data Engineer | Microsoft Fabric & Databricks | SQL, PySpark, Power BI

Chennai, India | [Phone] | [Personal Email] | linkedin.com/in/nirmalkumar6 | github.com/nirmalkcse | IEEE: ieeexplore.ieee.org/document/10544873

PROFESSIONAL SUMMARY

Data Engineer specializing in Microsoft Fabric and Databricks lakehouse delivery across ingestion, CDC/SCD2, validation, BI migration, and data-product workflows.

Impact highlights: built Fabric migration assets for 3 telecom source systems and 18-24 daily incremental tables; validated 15+ BI reports with 90% fewer discrepancies; migrated 6 Tableau reports to Power BI; currently building a Databricks-hosted MDM/entity-resolution product.

Recognized with Outstanding performance ratings, improving from 4.4/5 to 4.8/5.

TECHNICAL SKILLS

Platforms & Storage: Microsoft Fabric (Lakehouse, Warehouse, Data Factory pipelines, OneLake), Databricks (Apps, SQL Warehouse, Jobs API, Secrets API, Unity Catalog), Delta Lake, Power BI

Data Engineering: SQL, Spark SQL, PySpark, metadata-driven ETL/ELT, Medallion Architecture, incremental/CDC pipelines, SCD Type 2, schema enforcement, data quality

Analytics & BI: semantic modeling, DAX, Power Query, Power BI Report Builder, KPI modeling, Tableau-to-Power BI migration, MicroStrategy validation

Application & Workflow: TypeScript, Next.js, React, Zod, TanStack Query, REST API routes, Git/GitHub, GitHub Copilot, Copilot Studio, SSMS

PROFESSIONAL EXPERIENCE

iLink Systems - Software Engineer Trainee, Data Engineering & Product Development | Sep 2024 - Present

iCORE Enterprise MDM / Entity Resolution Platform | Jan 2026 - Present

- Own end-to-end development of iCORE, a Databricks-hosted MDM/entity-resolution product using Next.js 15, TypeScript, Databricks Apps, SQL Warehouse, Jobs API, Secrets API, Unity Catalog, and Delta control tables.
- Re-architected the accelerator from connector-centric workflows into a Master Entity model with reusable connectors, source mappings, canonical attributes, golden records, golden-record history, and resolution templates.
- Built server-side Databricks integration APIs for parameterized SQL execution, notebook/job submission, run status polling, catalog access, secrets-backed credentials, proxy-header auth, RBAC, and audit logging.
- Developed entity-scoped pipeline launch flows passing entity IDs, source mappings, entity names, and column mappings into Databricks orchestration for future multi-source entity resolution.
- Implemented stewardship and governance workflows for match review, review queues, approval chains, SLA policies, delegation, escalation, overrides, compliance scorecards, evidence packs, and decision lineage.
- Added explainability features including match reasons, score breakdowns, golden-record suggestions, triage guidance, feature-level model configuration, and audit trails for model-backed decisions.
- Designed a feature-flagged rollout and migration plan for Master Entity, entity pipelines, adaptive dashboard, and template-library phases, enabling controlled release and rollback.

Tier-1 European Telecom Provider - Microsoft Fabric Data Migration PoC | Oct 2025 - Dec 2025

- Built PoC artifacts for a Teradata-to-Microsoft Fabric modernization, moving legacy data flows toward a scalable Lakehouse/Warehouse architecture.
- Engineered metadata-driven ingestion for 3 source systems - GRASP SFTP, AMDOCS Oracle, and AIDA SFTP - supporting daily incremental loads across roughly 18-24 tables.
- Implemented Bronze/Silver/Gold processing using Fabric Data Factory pipelines, PySpark, SQL notebooks, Delta tables, and a dual-Lakehouse/single-Warehouse design.
- Developed SCD Type 2 historization with MD5 hash-based CDC, rec_start/rec_end validity windows, active/deleted status tracking, and optimized unchanged-record handling.
- Processed Gzip text/CSV, pipe-delimited CSV, and Binary JSON sources using regex parsing, JSON flattening, schema correction, UTF-8 standardization, and table normalization.
- Refactored Teradata/BTEQ procedural logic into Fabric SQL/PySpark patterns using CTEs, temp tables, MERGE statements, and orchestration-level error handling.
- Created Subject Area and Solution Mart transformations for Contract, Subscription, Product Instance, IMSI/ICCID, Gross Adds, and Churn reporting logic.
- Designed validation-ready stages for Landing, Transformation, Validation, and Gold loading with idempotent and retry-friendly execution.

BI Migration, Validation & Analytics Delivery | Sep 2024 - Sep 2025

- Validated 15+ MicroStrategy-to-Fabric reports for a Middle East manufacturing client using SQL regression checks, Gold-to-BI comparisons, and Power BI Report Builder layout/data validation; reduced discrepancies by 90%.
- Migrated 6 Tableau reports to Power BI for beverage and life sciences clients, rebuilding semantic models, relationships, hierarchies, DAX measures, calculations, report layouts, and accessibility/usability checks.
- Built an insurance analytics PoC with synthetic datasets, semantic models, Power Query/DAX transformations, in-report Copilot, verified response patterns, prompt libraries, and a parameterized SQL data agent.
- Built the FabCon25 Europe 2025 Talk to Your Data demo integrating Lakehouse data, semantic models, Power BI Chat with Data, in-report Copilot, Data Agent RCA, Copilot Studio orchestration, and JIRA MCP incident context.
- Contributed to a team solution recognized by senior company and Microsoft leadership for applied Microsoft Fabric, Power BI, Copilot, and data-agent capabilities.

SELECTED PROJECTS

IEEE Research - Semiconductor Chip Analysis using Medallion Architecture

- Published research on optimizing semiconductor chip data processing using Bronze, Silver, and Gold Medallion layers with Microsoft Fabric and Power BI.
- Designed analytical flows for multi-source ingestion, transformation, and Power BI reporting over Gold-layer chip performance metrics.

VITAC - Skincare Market Analysis

- Transformed unstructured skincare-market data in Microsoft Fabric/Azure Data Factory and built Power BI dashboards to identify market opportunities.
- Performed TAM, SAM, and SOM analysis and structured insights using SCQA and Pyramid Principle methods for executive-style decision support.

CERTIFICATIONS

- Microsoft Certified: Fabric Data Engineer Associate (DP-700)
- Microsoft Certified: Fabric Analytics Engineer Associate (DP-600)
- Databricks Certified Data Engineer Associate
- Databricks UC Upgrade Delivery Specialization
- Databricks Accredited Databricks Fundamentals
- Microsoft GitHub Copilot
- Google Data Analytics Professional Certificate
- Oracle Cloud Infrastructure Foundations Associate

EDUCATION

Bachelor of Engineering in Computer Science and Engineering, Sathyabama Institute of Science and Technology, 2024 | CGPA: 9.11/10