Arvind Reddy Narra

Ottawa, Ontario, Canada | +1 (438) 855-7696 | narra.arvind.reddy@gmail.com | LinkedIn | Portfolio

Professional Summary

Strategic and impact-driven Data Engineer with 5+ years of experience designing, building, and optimizing cloud-native data platforms in Azure, AWS, and hybrid environments. Expert in crafting end-to-end data pipelines, delivering real-time streaming and batch processing solutions using ADF, Databricks, Apache Spark, Airflow, and Kafka.

Demonstrated excellence in data modeling, data warehousing, and metadata-driven frameworks across domains like finance, operations, and compliance. Proven ability to drive CI/CD, infrastructure automation, and cloud governance using Terraform, Azure DevOps, Docker, and Key Vault. Passionate about delivering actionable analytics with Power BI, Tableau, and integrating Microsoft Fabric Lakehouse architectures.

Thrives in agile teams, values mentorship, and continuously contributes to innovation through internal projects and technical events such as Hackathons.

Technical Skills

Languages: Python (Pandas, PySpark, NumPy, FastAPI), SQL (T-SQL, PL/pgSQL, PostgreSQL), Bash, Scala
Cloud Platforms: Microsoft Azure (ADF, Databricks, Synapse, Data Lake Gen2, Azure SQL), Microsoft Fabric, AWS (S3, Glue, Redshift, Lambda, IAM)
ETL & Orchestration: Azure Data Factory, Apache Spark, Apache Airflow, Databricks, DBT, SSIS, Event-Driven Architectures (Kafka, Event Hubs)
Data Warehousing & Modeling: Snowflake, Redshift, Azure SQL; Star/Snowflake Schema, SCD Types 1 & 2, Dimensional Modeling, OLAP/OLTP Design
DevOps & Infrastructure: Terraform, Azure DevOps, Docker, GitHub Actions, Azure Monitor, Key Vault, CI/CD Pipelines, Monitoring & Alerting Systems
Analytics & BI: Power BI, Tableau, SSRS, Excel, DAX, Data Visualization, Report Optimization

Professional Experience

Teaching Assistant Sep 2023 – May 2024
Carleton University — Ottawa, ON, Canada

- Instructed and mentored graduate students in Computer Vision and Deep Learning, with real-world applications in streaming data pipelines, feature extraction, and efficient model evaluation.
- Guided 40+ students in developing PyTorch-based pipelines and optimizing memory usage for scalable training.
- Created custom lab experiments to simulate real-time data workflows using OpenCV, Python, and Apache Spark.
- Collaborated with faculty to modernize curriculum focused on cloud-scale data engineering and introduced Databricks workflows for lab evaluations.

Software Engineer Jun 2022 – Dec 2022
Deloitte

- Led migration of legacy batch pipelines to modern ADF and Databricks pipelines using PySpark and SQL.
- Tuned Spark jobs for performance using adaptive execution, broadcast joins, and caching strategies in Databricks.
- Integrated multiple Azure Data Factory pipelines with triggers, error handling, and parameterized datasets.
- Enabled Delta Live Tables (DLT) and Unity Catalog for enterprise data governance and lineage tracking.
- Worked closely with stakeholders to implement Power BI reports with real-time monitoring and historical trend analysis.
- Contributed significantly to ADF, Databricks, Spark, Power BI, and Azure SQL solutions during this project.

Associate Software Engineer Jun 2020 – May 2022
Deloitte

- Designed and deployed ETL/ELT pipelines using ADF, Apache Spark, and Databricks, integrating with Azure SQL, Delta Lake, and Azure Data Lake Gen2.
- Built modular PySpark notebooks in Databricks for incremental and historical load processing with schema validation.
- Developed metadata-driven SQL stored procedures for parameterized pipeline control and implemented SCD Type 1/2 logic.
- Orchestrated complex data workflows using Airflow DAGs and built CI/CD pipelines using Terraform, Azure DevOps, and GitHub Actions.
- Migrated 15+ SSIS jobs to ADF, increasing modularity and improving maintainability.
- Instrumented Azure Monitor and Key Vault for secure pipeline automation and monitoring.
- Developed Power BI dashboards with embedded DAX metrics, real-time SLA compliance tracking, and drill-through capabilities.
- Integrated Kafka streams into transformation logic for semi-structured log processing.
- Extensively utilized Power BI, ADF, Azure SQL, Spark, and Databricks across all phases of the data lifecycle.

Solution Delivery Analyst Jun 2019 – May 2020
Deloitte

- Engineered Python-based ingestion scripts for consuming Excel, XML, and JSON from FTP and REST APIs; used ADF for ingestion.
- Developed PySpark data validation layers and optimized SQL procedures for transformation-heavy workloads.
- Designed dimensional models using Star Schema and implemented SCD Type 2 logic for audit traceability.
- Built Power BI dashboards integrated with Azure Synapse and Azure SQL for operational KPIs.
- Partnered with DevOps team to configure Terraform IaC, Docker containers, and error logging pipelines.
- Actively worked on Azure Data Factory, Azure SQL, Spark, Power BI, and ADF orchestration tasks.

IT Intern Jan 2019 – Apr 2019
Deloitte

- Created SQL views, stored procedures, and triggers supporting reconciliation reports.
- Built Excel automation macros and contributed to migration plans for legacy data workflows.
- Introduced Power BI dashboards and developed new ADF pipelines under supervision.

Education
Carleton University — *Master of Science, Software Engineering*
Ottawa, Ontario • Jun 2024
Relevant Coursework: Distributed Systems, Cloud Infrastructure, Big Data Analytics, Neural Networks, RDBMS
TA Experience: Led labs in Deep Learning and Computer Vision (Python, PyTorch)

Certifications

- Microsoft Certified: Fabric Data Engineer Associate (DP-700)
- Databricks Certified Data Engineer Associate
- Microsoft Certified: Azure Data Fundamentals (DP-900)

Training & Platforms
Pluralsight: Apache Spark 3, Querying with Snowflake, Exploratory Data Analysis in Python, Terraform Infrastructure
Internal (Deloitte): Agile Scrum Essentials, JIRA/Confluence, Unstructured Data Digitization, Cloud Fundamentals

References available upon request