

# Aadish Newrekar

Dallas, TX | 7079556519 | [arnewrekar26@gmail.com](mailto:arnewrekar26@gmail.com) | [LinkedIn](#)

## SUMMARY

---

AI/Data Engineer with **6+** years building scalable ELT/ETL and lakehouse platforms on **Azure & AWS**. Deliver forecasting systems (**MLForecast**, **Optuna**), production **RAG** stacks (**LangChain** + Azure Cognitive Search + **Neo4j/Qdrant**), and analytics enablement with robust data models, governance, and observability.

## PROFESSIONAL EXPERIENCE

---

### AI/Data Engineer

Dec 2022 – Present

*S-Infotech*

*Dallas, TX*

- Owned demand forecasting on **Azure ML** using **MLForecast** tuned by **Optuna**; re-engineered ADF/Databricks pipelines to cut MAE by **#15%** and improve throughput under peak loads.
- Designed a multi-cloud lakehouse on **Redshift Serverless** + **S3** + Spectrum with partitioning & result caching, lowering query cost by **20%**.
- Shipped a production **RAG** platform with **LangChain**, Azure Cognitive Search, **Neo4j** knowledge graph, and **Qdrant** vectors; implemented evaluations and guardrails to raise answer quality by **~10%**.
- Built observability for data quality & SLAs (Great Expectations, custom Python checks, Azure Monitor) to reduce incident MTTR by **35%**.
- Established containerized CI/CD with **Azure DevOps**, **Docker**, and Kubernetes blue/green releases enabling zero-downtime deploys.
- Tech:** Python, SQL, **Azure Data Factory**, **Databricks**, Airflow, Azure ML, Redshift Serverless, Snowflake, Power BI, Tableau, Azure DevOps, Docker, Kubernetes

### Data Engineer

Oct 2020 – Jul 2022

*Byju's*

*Mumbai, India*

- Orchestrated ELT via **ADF/Synapse** ingesting MongoDB logs & LMS CSVs into ADLS; optimized Spark to reduce load latency by **25%**.
- Migrated **PostgreSQL** → **Redshift** using **AWS Glue**; tuned dist/sort keys and batch sizes to cut job times by **30%**.
- Automated OCR document processing with **Azure Document Intelligence** + Logic Apps; persisted outputs to **SQL Server** powering downstream analytics.
- Delivered KPI dashboards in **Power BI**/Tableau, shrinking decision latency and enabling self-service adoption across marketing & retention.
- Tech:** ADF, Synapse, Redshift, AWS Glue, Snowflake, SQL Server, PostgreSQL, Power BI, Tableau, Python

### Data Analyst

Jul 2018 – Oct 2020

*Nuclei Technologies*

*Mumbai, India*

- Automated ETL with **SSIS & Airflow**; centralized ERP/Salesforce data and cut reporting latency by **35%**.
- Implemented CI/CD for SSIS/ADF/Glue jobs (Git, Azure DevOps), removing **30%** manual steps and reducing deployment defects.
- Modeled star/snowflake schemas with **dbt** in **Synapse/Redshift**; cluster tuning yielded **20%** faster queries.
- Built secure **Power BI** dashboards with RLS for sales, finance, and operations stakeholders.
- Tech:** SQL, Python (pandas), SSIS, Airflow, dbt, Synapse, Redshift, Power BI, Tableau, QuickSight

## SELECTED PROJECTS

---

### Sales Dashboard Automation

S-Infotech

*Snowflake, Airflow, Tableau*

- Built nightly ETL to **Snowflake** with optimized star schema; delivered interactive Tableau KPIs with auto-refresh, driving pricing insights and accelerating exec reviews.

### Customer Churn Analysis

Byju's

*Python, PostgreSQL, Power BI*

- Engineered features and trained logistic regression in **Python**; surfaced probability scores in **Power BI**, informing targeted retention campaigns.

### A/B Test Reporting

Nuclei

*Python, statsmodels, SQL, Power BI*

- Designed experiments and analyzed lift/significance in **statsmodels**; executive dashboard guided roadmap & improved conversion outcomes.

TECHNICAL SKILLS

---

- **AI/ML:** Azure ML, **MLForecast**, **Optuna**, **LangChain**, **Neo4j**, **Qdrant**
- **Data/ETL:** Python, SQL, PySpark, **Azure Data Factory**, **Databricks**, **Airflow**, dbt, SSIS
- **Platforms:** **Azure**, **AWS**, Redshift Serverless, Snowflake, Synapse, SQL Server, PostgreSQL, S3
- **Analytics:** Power BI, Tableau, Amazon QuickSight     **DevOps:** Docker, Kubernetes, Azure DevOps, Git

AWARDS, ACHIEVEMENTS & CERTIFICATIONS

---

**Data Excellence Award**, S-Infotech (2024) — Boosted forecasting accuracy and data pipeline reliability.  
**Business Insights Award**, Byju’s (2021) — Self-service dashboards that accelerated decision cycles.  
**Analyst Champion Award**, Nuclei (2019) — Automated reporting and CI/CD for analytics pipelines.  
**Certifications:** Google Data Analytics, Python (Udemy)

EDUCATION

---

<b>Master of Science in Information Systems Management</b> <i>Arizona State University (GPA: 4.0/4.0)</i>	Aug 2022 – Jul 2023 <i>Tempe, AZ</i>
<b>Bachelor of Engineering, Electronics &amp; Telecommunication</b> <i>Vidyalankar Institute of Technology, University of Mumbai</i>	Aug 2016 – Jul 2020 <i>Mumbai, India</i>