

ABHISHEK SINGH

SENIOR DATA ENGINEER

hara.xlsx@gmail.com

+91 7007493576

[LinkedIn](#)

abhisheksingh.io

[GitHub](#)

PROFILE

Senior Data Engineer, 4 years' experience, specializing in Cloud Data Engineering, Distributed Data Processing, and Data Platforms.

Currently building the data and analytics infrastructure for enlytical.ai, an AI-powered e-commerce analytics and advertising automation platform.

Expertise in Azure, Google Cloud Platform, Python, SQL, and Apache Spark, designing scalable data pipelines, high-volume data analytics systems, cloud-native architectures, and analytics solutions for production workloads.

WORK EXPERIENCE

Senior Data Engineer, Nuvoretail Enlytical Technologies Pvt. Ltd.

June, 2025 – Present

Designed and built the end-to-end data platform powering enlytical.ai, an AI-enabled ads & analytics platform

- Built a distributed data collection platform that orchestrates data gathering across 50+ heterogeneous APIs and web platforms, landing raw data into the Delta lake for downstream processing.
- Built a centralized job orchestration service that schedules, dispatches and monitors crawling and data-processing workloads across 100+ Azure VMs.
- Built a distributed event-driven data processing/ingestion platform using Azure storage queue events with asynchronous, fault-tolerant execution, that process and catalogues landed data, processing terabyte-scale data through multiple medallion layers.
- Implemented Delta Lake medallion architecture on ADLS, with file-level partitioning and scheduled optimization services to minimize query scans and storage costs, improving data discovery and accessibility.
- Built a config-driven ingestion framework that automatically provisions schemas, transformations, validation rules, and storage layouts for new datasets without code changes and a metadata-driven data reliability suite guarding every layer.
- Built a cross-cloud serving layer via Synapse/Fabric and BigQuery, exposing curated gold datasets to the product's dashboards and FastAPI services.
- Engineered a low-latency analytics layer using pre-aggregations, query-optimized storage and in-memory caching reducing API response time to sub-second for interactive dashboards.
- Designed high-performance analytical SQL models using partition pruning, window functions, and incremental transformations to support low-latency reporting workloads.

Data Engineer, Nuvoretail Enlytical Technologies Pvt. Ltd.

Aug, 2022 – June, 2025

Responsible for Data Architecture, SQL Administration, Management and Development

- Initiated the migration off legacy SQL Server, standing up the first lakehouse foundation - Parquet on ADLS with a Synapse Serverless SQL serving layer which later grew into the production platform.
- Built and maintained complex SQL- and Spark-based ETL pipelines, including stored procedures and functions, to serve business-critical analytics dashboards
- Orchestrated data workflows and refresh schedules while managing SQL Server performance, uptime, and processing capacity.

SKILLS

- **Languages Skills:** Python, SQL (T-SQL, Spark SQL, PL/SQL), PySpark, Apache Spark, Spark Streaming, NoSQL Databases
- **Data Engineering:** Distributed Data Processing, Event-Driven Architecture, ETL/ELT, Delta Lake, Lakehouse (Bronze–Silver–Gold), Medallion Architecture, Schema Design, Metadata-Driven Pipelines, Data Modelling, Data Quality, Incremental Loading Processing, API Architecture, event-driven-architecture, stream processing, data

lineage, CI/CD, Data Factory, Azure Functions, , Logic Apps, OneLake, Playwright/Selenium, Data Warehousing, Parquet, Avro, Apache Airflow,

- **Cloud & Data and Other Platforms:** Azure - ADLS Gen2, Azure Synapse Analytics, Azure Functions, Azure Event Grid, Azure Storage Queues, Azure Blob Storage), Microsoft Azure Fabric, Google Cloud Platform(GCP), Google BigQuery, MongoDB, PostgreSQL, MySQL cloud build, Databricks, Vertex AI, Claude Code, Generative AI
- **Analytics & Performance:** Query Optimization, Analytical SQL, Window Functions, Partition Pruning, Pre-Aggregations, In-Memory Caching, Performance Tuning, Database Management
- **Backend & DevOps:** FastAPI, Github, Github Actions, CI/CD , Docker, Linux, Cloud Build

EDUCATION

Kurukshetra University, Bachelor of Technology
Computer Science and Engineering
2017-21

CERTIFICATES AND LICENSES

Microsoft DP 900 : Azure Data Fundamentals

Databricks for Data Engineering

Level Up : Advanced SQL

Advanced SQL for Query Tuning and Performance Optimization

Big Data Analytics with Hadoop and Apache Spark

HackerRank : Advanced SQL