

Haresh Chawla

DATA ENGINEER

Phone: 7083828269

Email: hareshchawla11@gmail.com

Location: Mumbai

LinkedIn: <https://linkedin.com/in/haresh-chawla>

Having 3+ years of experience with different industry verticals in building and implementing data analytics solutions, data migration and ingestion process on scalable data platforms using Data bricks, Snowflake, Big Data and Cloud technologies.

WORK EXPERIENCE

EY | Role: Data Engineer Consultant

Apr'22 - Present

Insurance Domain- Built a Data Lakehouse on Azure using Databricks Delta, Spark Pool, Synapse Dedicated SQL Pool, Logic apps

- Designed and developed a data ingestion spark framework using lambda architecture which is generic for all data sources.
- Maintained real time data pipeline for streaming CDC and transactional data across 16 different data sources with SLA of 90.0%.
- Created the Unified layer on Synapse SQL Pool which is used further for data visualization.
- Designed the DQ framework to check the data quality which helps client to create machine learning model with efficiency of 40%.
- Automated ETL processes across thousands of SFTP data files into ADLS which reduced the manual workload by 30% per month.

EY | Role: Data Engineer Consultant

Dec'21 - May'22

Mortgage Loan Domain- Built a Data Warehouse on Snowflake using AWS S3, AWS Lambda, ETL tool.

- Migrated the data from PostgreSQL to Snowflake using Talend dynamic pipelines with S3 and AWS.
- Designed a real-time data pipeline to process semi-structured data with 1 million records/month from various sources.
- Implemented more than 30 plus BIU scripts using snowflake stored procedure which provides the valuable insights for overcome KPI failures of customer acquisitions or loan default.
- Developed 100+ SCD1, SCD2 pipelines on ETL tool which fetch the data from various data source (Salesforces, Google Analytics, Databases) and loading it into snowflake with audit mechanism.

Accenture | Role: Associate Data Engineer

Aug'19 - Nov'21

HealthCare Domain- Built a Data Lakehouse on AWS using Snowflake, Airflow, Spark, Glue, Databricks on AWS.

- Designed a dynamic data ingestion spark framework which is used to process ~500GB/Day. (Architecture - Lambda)
- Modified the process of gathering and transforming the data from 11 different application into refinement metrics.
- Also, consolidated the data from multiple distributed heterogenous data sources into a single point of reference i.e., Data Lake on S3.
- Automated the process of creating dynamic cluster using data bricks job API and airflow as job orchestration tools and maintenance of the metadata was achieved using glue catalog crawler on data lake.
- This architecture helped the client to deliver results with more efficiency in time & cost by 70% and it's being utilized by 18 dashboards.

TECHNICAL SKILLS

- | | | | | |
|-----------------------------------|--------------|------------------|------------------------|-------------|
| • Azure Services - ADF, Functions | • Databricks | • Hadoop / Hive | • On Premise DB | • Dremio |
| • Synapse Analytics | • Spark | • Python | • Airflow / logic apps | • ETL Tools |
| • AWS Services - EMR, DMS, Lambda | • Snowflake | • Bash Scripting | • Delta Lake | • Power BI |

EDUCATION

Bachelor of Engineering - Computer Science

Vivekanand education society institute of technology - VESIT (2015-2019)

7.86 GPA

COURSES AND CERTIFICATIONS

- Microsoft Data Engineer Certified (DP-200, DP-201, DP-203) and Azure Fundamental Certified (DP-900, AZ-900, AI-900).
- AWS Developer Associate Certified.
- Snowflake Snow pro trained certification and Certified Data Engineer from IIT Roorkee.
- Databricks Data Engineering Associate Certified.

VOLUNTEERING ACTIVITIES AND ACHIEVEMENTS

- Received the client manager appreciation for the individual project.
- Received Kudos award for continuous Quarterly in EY for handling individual project.