

Shouvik Sharma

Chicago, Illinois

+1-312-459-2008 | shouvik19@gmail.com | <http://www.linkedin.com/in/shouvik-sharma>

WORK EXPERIENCE

AVANT, LLC

Chicago, Illinois

Senior Associate, Data

August 2021 - Present

- Engineered efficient and scalable data pipelines using Snowflake, dbt and Apache Spark, to process structured and unstructured data, improving data processing efficiency by 20% over a six-month period.
- Implemented Apache Airflow for workflow orchestration, resulting in a 30% increase in pipeline reliability and reducing system downtime by 50% within the first quarter of implementation.
- Utilized Airflow, Apache Spark, EMR and Amazon SQS to build and optimize data pipelines for large-scale data processing, enabling the implementation of machine learning models in PySpark that improved predictive accuracy by 15%.
- Developed Looker dashboards for marketing attribution, contributing to a 4% increase in application rates over a three-month period.
- Created source attribution funnel dashboards in Looker, helping business stakeholders improve customer application experience and increase conversion rates by 7%.
- Collaborated with cross-functional teams to implement data quality checks using SODA, further utilizing AWS services including Redshift and S3 for large-scale data processing and storage, resulting in a 35% improvement in data accuracy and a 20% reduction in processing time.

CNH INDUSTRIAL INC.

Racine, Wisconsin

Data Engineer Intern

August 2019 - March 2022

- Designed and implemented ETL processes using Microsoft Access and SQL that reduced the data processing time by 30%.
- Collaborated with data scientists to define data requirements, implemented ETL processes using Databricks, optimized SQL queries, and developed predictive models using Python and Databricks, enhancing data-driven insights.

DATEN SOLUTIONS INC.

Chicago, Illinois

Data Engineer

May 2020 - March 2021

- Developed and automated data migration pipeline from SQL Server to Snowflake using SnowSQL and SnowPipe, and further enhanced data quality by performing dimensional modeling on the migrated data.
- Automated ETL processes using Prefect (Python), enhancing data wrangling capabilities and achieving a 40% reduction in time through large-scale data conversions. Facilitated the seamless transfer of BAAN data into standardized formats for integration into Snowflake.

CARTESIAN CONSULTING

Mumbai, India

Big Data Developer

April 2018 - July 2019

- Designed and implemented ETL processes using DynamoDB and SQL, reducing data processing time by 50%.
- Developed dimensional data models and a data warehouse powered by MariaDB, supporting the creation of a campaign datamart and a comprehensive customer one-view, which improved marketing campaign efficiency by 20%.
- Implemented data transformation and aggregation processes using R and MariaDB, improving reporting and analysis efficiency by 25%.

EDUCATION

ILLINOIS INSTITUTE OF TECHNOLOGY

Chicago, Illinois

Master of Computer Science and Mathematics

August 2017 - December 2018

Selected Courses: Big Data Technologies, Applied Statistics, Database Management, Data Preparation and Analysis.

NARSEE MONJEE INSTITUTE OF MANAGEMENT STUDIES

Mumbai, India

Master of Statistics

August 2016 - May 2018

SKILLS

Programming Languages: Python, SQL, Scala, Java, HTML, Excel VBA (Macros), Spark, Hadoop, MapReduce, Hive, Pig, Kafka.

Big Data Ecosystem: Spark, Hadoop, MapReduce, Hive, Pig, Kafka, Flume, Hbase, Microsoft Azure, Big Query.

Cloud Technologies: GCP (BigQuery, Dataflow, DataProc, Cloud Composer, Cloud Storage, GKE), AWS (Redshift, S3, EMR), Snowflake, Airflow, Prefect.

ACHIEVEMENTS AND CERTIFICATIONS

Databricks Data Engineer Certification

2024