**Ajit Thakur**

**516-426-6543**

**heyajit@yahoo.com**

**Summary**

Trusted partner and results-driven senior technology leader with expertise in enterprise data management, business intelligence, advanced analytics, big data, and cloud solutions. Extensive experience in master data, data quality, data governance, and team building to drive business value and enable data-driven decision-making. Additionally, I have extensive knowledge in ETL, data warehousing, dimensional modeling, and data quality techniques to optimize session load time.

**Data Engineer Lead @ *JP Morgan & Chase, Jersey City, NJ* Feb 2018 – Till Date**

* Led the **migration** of the data processing platform from the **On-Prem Horton** to **AWS** S3 bucket using Data Pipeline and Amazon Athena, resulting in annual **cost savings** of more than $600,000 and an increase in **performance** of 14%.
* Led the migration of **data** **consumption** platforms from **Teradata** to **Snowflake**, enhancing scalability and performance in the cloud. Additionally, spearheaded a POC to enable trusted(silver) zone data access in **Databricks**, allowing data scientists to run their models efficiently.
* **Developed and optimized large-scale ETL pipelines** using **PySpark and SQL** to process terabytes of structured and semi-structured data generated by Bank’s Teller line (***BTL***).
* **Designed and implemented data models and warehouses**, ensuring efficient storage, retrieval, and performance tuning for multiple warehouses and mart like Claims, Disputes and Fraud /Customer and Relationship Data Mart. Implemented **Medallion** architecture in big data platform
* **Built scalable data ingestion frameworks** using JSON, Python, PySpark, and HiveQL, improving data pipeline efficiency and reliability. Designed the data pipeline architecture to support **Schema evolution**. Designed and implemented a **secure data access framework** for handling **PII data**, ensuring compliance with data governance and security policies.
* **Maintained** data pipeline up-time of 99.8% while ingesting *OCR & ICR* streaming and transactional data across 8 different primary data sources using Spark, S3, and Python-based data ingestion framework.
* **Ingested** data from different data sources confirming completeness, consistency, timeliness and accuracy using a combination of SQL, DPL, and real-time Kafka APIs.

**Environment**: Oracle19c, Greenplum/Postgres, AWS tech stack, Snowflake, *Teradata 16.4, AB Initio, Hadoop, Hive, Oracle 11g /19c, Spark, SQL, Kafka, Shell, Squirrel, FTP, Linux, Tidal, WSDL, Web Services, JMS, JSON/JSON-LD*

**Data/ETL/MDM Architect @ *Sirius XM Radio, New York City, NY* Mar 2015 – Feb 2018**

* **Designed and optimized OLTP and OLAP data models**, ensuring high-performance transaction processing for operational systems and efficient analytical querying for enterprise reporting and decision-making
* Designed effective Data engineering and **MDM solutions** to complex technical integrations, factoring non-functional requirements, performing proof-of-concepts to ensure the solution meets predefined objectives.
* Designed & developed loading of semi-structured data (**weblogs**), streaming data from various Kafka topics to **Vertica database** to newly built data lake in **Cloudera** platform utilizing Sqoop / Hive
* Designed MDM data flow for different **metadata**, resolved complex technical issues related to **performance & quality**, and worked with the data engineering team to align the technical architecture framework with the requirement.
* Data Analysis of consumption weblogs for various web audio players. Designed a **dimensional model** to load the weblog data for different management and financial reporting.

**Environment**: *Teradata 14.0, Vertica 7.0, Informatica MDM hub Console 10.1, Hadoop, Hive, Oracle 11g /12c, Informatica Power Center 9.5, Spark, SQL, Kafka, Shell, Squirrel, FTP, Linux, Tidal, WSDL, Web Services, JMS, JSON*

**ETL Architect @ *JP Morgan & Chase, New York City, NY*  Oct 2009 – Mar 2015**

* Working with Investment management’s Back-office operations’ Business Analysts & Business users to understand the feasibility of the requirement. providing Product owners with the effort & estimates
* Ownership of Accounting Data Mart & TLM application, onboarding new recs for cash & holdings. Working with **SOX**, Risk & Compliance & internal auditors for TLM applications
* Managed & mentored resources, both **onsite & offshore**. Co-ordinate with offshore managers to ensure that offshore resources are adequately assigned to backlogs to make sure deliverables are well within budget without compromising quality.
* **Designed and optimized ETL architectures**, ensuring scalable, efficient, and reliable data processing across diverse platforms. Assisting Qlikview team to access Data warehouse objects.

**Environment**: *Oracle 10g/11g, Informatica Power Center 8.6/9.5, TOAD, SQL, SQL Server 16.0, Shell Scripting, DB Artisan, FTP, bcp, CA, Control M, Linux, Sun Solaris, SQL\* Loader, TLM*

**Education**

BS, Electrical Engineering - NIT (REC) Bhopal, India

AWS Cloud Practitioner Certified -2022