Vamshi G

Tampa, FL
-Email me on Indeed: http://www.indeed.com/r/Vamshi-G/121a9d47981db001

Work Experience

GCP Data Engineer

Freddie Mac-McLean, VA
November 2023 to Present

• Developed scalable data pipelines using Apache Beam/Dataflow for efficient data processing and integration of financial datasets, ensuring seamless data flow across services like loans and transactions.

• Implemented Cloud Composer for automating workflows and orchestrating complex ETL processes, ensuring smooth integration and optimized data processing in the GCP environment.

• Leveraged BigQuery for data warehousing and analytics, providing comprehensive insights into customer interactions, which enabled more personalized financial services and improved profitability.

• Migrated on-prem Hadoop systems to GCP, ensuring scalable and secure data storage while optimizing data transfer using tools like DISTCP for large dataset handling.

• Designed Spark jobs using Scala for batch processing and developed near-real-time data pipelines using Kafka and Spark Streaming to handle high-volume data streams efficiently.

AWS Data Engineer

L.L. Bean, Inc.-New Jersey
March 2021 to October 2023

• Designed and implemented scalable ETL workflows using AWS Glue and orchestrated processes with AWS Step Functions, ensuring efficient data integration into Amazon S3 and Redshift for centralized storage of mortgage-related data.

• Developed real-time data streaming solutions using Amazon Kinesis and AWS Lambda for event-driven processing, allowing the system to respond to real-time market changes and optimize decision-making.

• Optimized cloud infrastructure by leveraging AWS services like EC2, S3, Redshift, and DynamoDB, ensuring high availability, fault tolerance, and auto-scaling for seamless data pipeline operations.

• Ensured data governance and security by implementing IAM policies, encryption, and CloudTrail, maintaining strict access control and protecting sensitive mortgage data.

• Monitored and optimized performance of data pipelines and Spark Databricks clusters, adjusting cluster size for cost-effective scaling and ensuring the reliability and efficiency of data processing workflows.

Data Engineer

Dignity Health-Phoenix, AZ
December 2019 to February 2021

• Designed and implemented large-scale data pipelines using Spark and Python for efficient ingestion and transformation of healthcare data, ensuring timely and accurate analysis to improve patient outcomes.

• Developed real-time data processing systems with Spark Streaming and Kafka, enabling the rapid processing of streaming healthcare data for real-time analytics and operational efficiency.

• Optimized data infrastructure by managing big data tools like Snowflake and Hadoop, ensuring scalable and reliable data pipelines to handle large healthcare datasets.

• Enhanced data governance and security by implementing encryption, access control, and auditing through AWS Lambda and DynamoDB, ensuring the confidentiality and integrity of sensitive healthcare data.

ETL Developer

Optum-Tampa, FL

December 2017 to November 2019

• Developed a region-specific SMS delivery system using AWS SNS, Google Maps API, and Flask, enhancing communication efficiency and ensuring seamless delivery across regions.

• Designed and implemented scalable ETL processes using Informatica PowerCenter to manage and transform financial data, ensuring compliance with industry regulations and integrating it into centralized data warehouses.

• Optimized database structures with SQL Server and implemented security measures to ensure data integrity and protection, deploying solutions on Linux virtual servers.

• Automated ETL job execution using Tidal Enterprise Scheduler, streamlining data processing, reducing manual intervention, and ensuring timely data loads.

Data Analyst

Cybage Software Private Limited-Hyderabad, India

January 2016 to September 2017

• Developed predictive models using machine learning algorithms like Linear Regression, Naive Bayes, and Random Forests, leveraging large datasets to forecast trends and generate actionable insights for business decisions.

• Created interactive dashboards using Tableau and Power BI, integrating data from multiple sources and utilizing Power Query Editor to communicate insights and drive informed decision-making.

• Optimized ETL processes using Informatica PowerCenter, ensuring seamless data extraction, transformation, and loading, while maintaining data integrity across platforms.

• Collaborated with stakeholders to design dimensional models and assess ER diagrams, improving reporting efficiency and aligning data models with business requirements.

SQL Developer

1 Wipro Technologies-Bangalore, India

June 2014 to December 2015

• Developed and optimized ETL processes using SQL Server Integration Services (SSIS) for seamless data extraction, transformation, and loading, ensuring data integrity and accuracy across platforms.

• Designed and developed complex reports using SQL Server Reporting Services (SSRS), providing actionable insights through drill-down, tabular, and crosstab reporting.

• Created and maintained SQL and PL/SQL procedures for database optimization, improving performance and scalability in Oracle and SQL Server environments.

• Led database migration and upgrade projects, ensuring smooth transitions with minimal downtime and implementing error-handling mechanisms to maintain data integrity.

Education

Bachelor's degree in Computer Science
JNTUH

Skills

- AWS
- SQL
- Python
- Tableau
- Linux
- Scripting

- APIs
- MySQL
- Business Requirements
- Data Warehouse
- S3
- Hadoop
- XML
- Jenkins
- Git
- User Interface (UI)
- Apache
- Java
- Jira
- Agile
- Analytics
- Shell Scripting
- Microsoft SQL Server
- Data Modeling
- Apache Hive
- JavaScript
- Google Cloud Platform
- Redshift
- Snowflake
- MongoDB
- Cassandra