

Gopi Naga Sai Thattukolla

 Data Engineer

 (205) 607-2131  gopinagasai40@gmail.com  Birmingham, AL  gopinagasai



SUMMARY

Data Engineer with more than four years of experience focused on developing, implementing, and optimizing data solutions for finance-related objectives. Adept at building strong ETL/ELT data pipelines and improving the quality and reliability of extensive financial datasets. Proficient with cloud technologies, including AWS and Snowflake, and well-versed in a diverse set of tools and programming languages such as Python (NumPy, Pandas), SQL, Hive, as well as BI tools (Tableau, PowerBI), Apache Spark, Kafka, and Scala. Known for excellent teamwork and consistently delivering significant solutions that support data-driven decision-making in the finance sector.

EXPERIENCE

BlackRock, Alabama / August 2023 - Current / Data Engineer

- Led the integration of data from various enterprise systems (Dynamics 365 CRM, AX Finance and Operations, NetSuite, Salesforce) into AWS Data Lake using AWS Glue, facilitating advanced analytics in a major investment firm.
- Utilized AWS Redshift for performing complex analytics, accelerating decision-making and improving investment and client portfolio management.
- Executed complex SQL queries and optimized ETL processes to ensure data integrity and reliability for analytics within a large-scale data infrastructure.
- Developed PySpark applications to manage and process data from multiple sources, with final storage solutions implemented in Snowflake on AWS.
- Worked in collaboration with interdisciplinary teams to design and enhance data models, optimizing database performance to meet specific business needs.
- Automated various data workflows using Python, increasing operational efficiency and reducing data processing errors across the data ecosystem.
- Specialized in manipulating JSON data formats and crafted custom Python scripts for efficient data parsing and integration into existing data systems.
- Designed and deployed Power BI dashboards that visually represented crucial financial indicators, aiding strategic decisions and providing stakeholders with actionable insights.
- Implemented statistical analyses for detecting patterns indicative of risk and fraud within financial systems, bolstering the firm's capabilities in risk management.
- Merged technical solutions with business insights through the creation of customized data reports, enhancing organizational decision-making.
- Maintained rigorous data governance and compliance protocols, protecting sensitive financial data and ensuring compliance with regulatory standards.

PayU, Bengaluru, India / Jan 2019 - July 2022 / Data Engineer

- Led the development of a comprehensive data integration framework within a Hadoop ecosystem, employing Cassandra for NoSQL data storage, enhancing the infrastructure for advanced payment processing with PayU.
- Enabled real-time analytics capabilities, improving fraud detection, dynamic pricing, and customer insights, thereby boosting e-commerce transaction security and user engagement.
- Created AWS Lambda functions to perform real-time data validation, filtering, and sorting upon updates in DynamoDB, facilitating efficient data transformation and migration to secondary storage solutions.
- Boosted operational efficiency of payment processing systems, ensuring robust data integrity and agile transaction management for smoother, more reliable operations.
- Achieved a 99.1% uptime for databases through strategic replication, enhancing database stability and accessibility in critical usage scenarios.
- Developed multiple MapReduce jobs in Java for data cleaning and preprocessing.
- Involved in crafting Spark SQL queries and developing DataFrames, importing data, executing transformations, and performing read/write operations, saving results to HDFS.
- Optimized Hive fact tables, improving integration of financial data from various retailers through strategic partitioning by time, retailer, and data supplier, resulting in a 30% improvement in query performance and a 20% increase in data processing speed.

- Deployed AWS cloud services (EC2, S3, RDS, etc.) to architect and maintain a robust cloud infrastructure, enhancing system performance and reliability by 50%.
- Led ETL projects that integrated diverse datasets into SQL Server and Oracle databases using SCD Types 1 and 2, speeding up transaction processing by 20% and supporting refined analytics for dynamic pricing and market expansion.
- Enhanced the integrity and performance of AWS databases, optimizing relational systems across the infrastructure.
- Utilized Python with Spark for scalable data processing, efficiently managing diverse data formats and improving storage on HDFS.
- Streamlined database maintenance through optimized PL/SQL scripts, automating routine tasks to enhance operational efficiency.
- Improved database capacity planning, reducing lags and meeting increased demands 43% faster.
- Increased productivity by 34% through task automation with AWS Lambda, eliminating repetitive workflows.
- Developed a real-time data processing pipeline with Kafka and PySpark, integrating over 1 million records daily from diverse sources, significantly increasing analytical throughput.
- Pioneered a cloud-first data ingestion strategy using Python, SQL, and Spark, accelerating data processing speeds by 74% and significantly improving data readiness for analytics.

TECHNICAL SKILLS & INTERESTS

Programming Languages: JavaScript, Python, java ,C

BigData Eco Systems : Hadoop, HDFS, MapReduce, Hive,Apache Spark, Pig

Database Management: MongoDB, DynamoDB, MySQL, PostgreSQL

Cloud Services: AWS (S3, Lambda, SES, EC2), Docker, AWS Data Lake, AWS Glue, AWS Redshift

Operating Systems: Windows, MAC, Linux

Version Control: Gitlab, Git, GitHub, Bitbucket

Tools: Tableau, Power BI, SSRS, Jira

Methodology: SDLC, Agile, SaFe

Packages: NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow

EDUCATION

University of Alabama at Birmingham
Master of Science in Computer Science

Aug 2022 – Dec 2023
GPA:3.8/4

CERTIFICATIONS

[AWS Certified Solutions Architect – Associate](#)

[Certified Developer on Legacy JavaScript Algorithms and Data Structures from FreeCodeCamp](#)