Harish Devathraj

Data Engineer

Harish Devathraj

Bangalore India,

Phone: 7760679910

Email: devathrajharish@gmail.com

Github: https://github.com/devathrajharish

LinkedIn: https://www.linkedin.com/in/devathrajharish/

Skills

Programming Languages: Python, SQL.

Cloud & Applications: Azure, Databricks, AWS, Tableau, PowerBi, Git, DevOps, Grafana.

Databases: Snowflake, MySQL, Postgres.

Big Data Technologies: Spark.

Experience

Tredence (Pepsico) / Associate Manager Data Engineering

Jan 2024 - Present, Bangalore, IN

- Designed and implemented ETL processes to manage large-scale datasets, ensuring data accuracy and consistency across multiple sources.

- Vastly improved the performance of existing code using optimization techniques, which saved a huge amount of company's time on daily operations.

- Led migration of legacy systems to cloud-based data platforms, leveraging tools like Azure ADF and Databricks to enhance scalability and performance.

- Developed automated data validation processes using ICEDQ to ensure the quality and integrity of production data.

- Built and owned Grafana dashboards that capture metrics and operational KPIs for data pipeline quality and performance.

- Led a team of 3-4 Data engineers, mentoring them in designing, building and maintaining pipelines and ensuring timely delivery of high-quality data solutions.

Capgemini / Data Engineer

Jun 2022 - Jan 2024, Houston, TX

- Designed and implemented an Azure data pipeline to process structured and semi structured data from various sources using the Python framework, Databricks.

- Developed and designed Api's using Azure function app to integrate with Power plan, Datahub and Copperleaf, which resulted in 75% cost reduction.

- Worked with clients to understand business needs and translate those business needs into meaningful data that is useful for actionable reports on PowerBI using Azure databricks and Sql.

Mindtree Ltd / Data Engineer

Nov 2020 - Jun 2022, Seattle, WA

- Evangelize engineering best practices and proactively identify and resolve technical debt on Azure services.

- Collaborate with team members on data design and reporting strategy across business functions to ensure cohesive alignment utilizing Python and SQL.

- Facilitate ingestion of raw data, build ETL pipeline, and perform joining, and

aggregation to facilitate consumption downstream using ADF, Databricks and Spark.

Brahm Works / Software Developer

July 2017 - July 2018, Bangalore, IN

- Designed and developed secure, maintainable and scalable application- ensuring that the services and the components are: well-defined and modularized, secure, reliable and reusable.

- Experienced in designing RESTful services and APIs for the application.

- Containerized apps using Docker for deployment.

- Ensured test coverage for unit testing and support integration and performance testing using CI/CD.

Education

University of Houston / Master of Science in Computer Science

AUG 2018 - MAY 2020, Houston, TX

Visvesvaraya Technological University / Bachelor of Engineering in Computer Science

AUG 2013 - MAY 2017, Bangalore, IN

Projects

Big Data Analytics (Python, HDFS, Apache Spark, Hadoop, MapReduce)

- Performed MapReduce on the large dataset consisting of 200 thousand files and to extract the useful information for further analysis.

- Implemented various operations on the dataset and stored the results onto Avro, Parquet and Text files and performed Snappy compression which helped in the whole understanding the data and formulating various hypotheses over it.

- Utilized Python's API PySpark to perform the task on the Whale Cluster.

Data Analysis of cross-disciplinary evolution of genomics revolution

- Research based project where the social and cultural capital generated because of Human Genome Project (HGP) was estimated by bridging biology and computing departments in various universities.

- Tech Stack: R, Gephy, Plotly, ggplot.

Activities

Social media officer of Graduate Indian Student Organization at University of Houston

Volunteered at the Oriental English school NGO to teach students basic Math and English.

Completed Data Engineering - Nanodegree program from Udacity.
https://github.com/devathrajharish/udacity-dataengineer-nanodegree/tree/master/Projects