

SUPRITH M

Data Engineer

8197723170
Hoskote

@ suprithvasista829@gmail.com

https://www.linkedin.com/in/suprithm1/

EXPERIENCE

Data Engineer

MBB Labs Private Ltd

- 03/2023 Bengaluru, India
- Implemented Slowly Changing Dimensions (SCD)/SCD Recovery frameworks to manage historical data changes effectively, ensuring data accuracy over time
- Developed and optimized incremental/incremental recovery data processing frameworks, enabling efficient updates to large datasets with minimal disruption
- Enhanced the alt measure engine for reporting purposes, ensuring accurate and timely insights
- Led enhancements to the Spark DQ framework, optimizing performance and enabling real-time processing for large datasets reducing job execution time by 60%
- Led the production deployment of multiple UDL versions, ensuring robust implementation of critical business changes with zero downtime
- Designed and implemented scalable data extraction frameworks, improving operational efficiencies and SLA adherence across multiple teams
- Developed enterprise libraries to facilitate cross-platform data connectivity, promoting code reusability and reducing development time
- Engineered solutions in oracle for high-volume data loading for 100M records in under 30 seconds, enhancing system performance and responsiveness

Data Engineer Intern

MBB Labs Private Ltd

- 06/2022 - 02/2023 Bengaluru, India
- Developed and implemented data quality checks and automation scripts in Python, enhancing data integrity and operational efficiency
- Created proof of concept for Oracle Data Integration (ODI), streamlining metadata extraction and improving data handling processes
- Automated complex installer preparation and reconciliation workflows for REDW and UDL, significantly reducing manual effort and error rates
- Analyzed industry data transformations, designing automated scripts for data mapping and reporting, facilitating seamless transitions from development to production

Projects

Recursive Job Matching

- 06/2024 - 08/2024
- Developed and deployed a web application on Streamlit Cloud that delivers customised resume feedback by leveraging data from job portals and the Google Gemini Pro API for enhanced job alignment

Pandas-Consolidated-Loader

- 01/2024
- Developed and deployed a data consolidation utility in Python to the PyPI community

Streamlit-Filepath-Loader

- 02/2024
- Created a custom Streamlit file uploader widget that delivers upload paths for metadata-focused use cases, contributing it to the Streamlit community for enhanced functionality

SUMMARY

With over two years of experience as a Data Engineer, I have a strong background in building scalable data processing frameworks and optimizing data pipelines. Proficient in Python, Scala, and big data technologies, specialize in enhancing data integrity, process efficiency, and process automation. Looking forward to apply my technical expertise and problem-solving skills to deliver impactful, data-driven solutions in a dynamic and challenging environment

TRAINING / COURSES

BE in Computer Science and Engineering AUG

Visvesvaraya Technological University

SKILLS

API	Artificial Intelligence	Azure
Azure Cloud	Big Data	Dart
Data Extraction	Data Integration	
Data Integrity	Data Lake	
Data Mapping	Data Quality	Flutter
Google Cloud	Hadoop	Hive
Machine Learning	Mongo	
Mongo DB	Oracle	Oracle Database
Plsql	Process Automation	Pyspark
Python	Scala	Shell Scripting
Spark	SQL	Streamlit
		Teradata
Unix		

FIND ME ONLINE

- Linkedin
https://www.linkedin.com/in/suprithm1/
- Portfolio
https://suprithm.vasistatech.in/