Gabriel Vinicius Santos de Oliveira

gvinicius2013.gv@gmail.com | +55 11 989003485 | São Paulo, Brazil | linkedin.com/in/gabrielviniciusoliveira

Experience

Tiger Analytics

Senior Data Engineer
02/2024 - Present
United States, Remote

- Designed and developed a framework to apply unit tests in data transformations using Pytest. Increased the delivery time by 10% for new data assets;
- Transferred knowledge and helped lead a team of 5 junior data engineers;
- Developed complex ETL pipelines and transformation reusable functions using PySpark, Databricks, SQL Server, and Data Factory;
- Built gold layer data assets, used by the data science team and the third-party software coworkers.
- Created automated reports to apply statistics metrics and reconciliation in gold layer tables, enhancing stakeholders' visibility.

Niteo

Senior Data Engineer
03/2023 - 02/2024
São Paulo, Brazil

- Ingested over 30 tables from multiple data sources, collaborating with Data Analysts to improve data quality to create Power BI dashboards;
- Identified opportunities to automate manual data analysis tasks, over 10 tasks;
- Solved technical debts by implementing new approaches to connect to data sources, reducing costs by 50%;
- Provisioned CI/CD pipelines to help the development team speed up by 5% the creation of new features;
- Built a data lineage to identify issues on daily ingestions, reducing issue frequency by 15%.

XP Inc.

Data Engineer
08/2021 - 03/2023
São Paulo, Brazil

- Designed an architecture to ingest mobile and web user events data, approximately 100 million events each day;
- Identified bottlenecks in the user journey registration process created metrics to compare marketing campaigns and defined KPIs to track leads;
- Built Bronze, Silver, and Gold tables on Databricks, using Pyspark, and Spark SQL;
- Applied Data Quality checks using Great Expectations and Soda Python libraries, reducing maintenance workload by more than 30%.

Klabin SA

Jr. Data Engineer
02/2019 - 08/2021
São Paulo, Brazil

- Implemented two machine learning models to predict product quality and specifications.
- Using Apache Kafka and Python to stream data from over 10 IoT sensors to AWS S3;
- Data visualization using Power BI connected to AWS Redshift and user interface using React JS with Node as the backend;
- Predictions made using AWS Lambda and Python, 91% precision for product quality prediction;
- Models trained using XGBOOST. These models helped to reduce time spent developing a new product by 10% and reduce costs with material waste by bad quality test results.

Education

FIAP
MBA in Data Engineering
03/2022 - 06/2023

Faculdade de Tecnologia Termomecânica
B.S of Control and Automation Engineering
08/2016 - 06/2021

Certifications

DEA-C01 AWS Data Engineer Associate
AWS
05/2024

10th place in the Datathon ranking of more than 100 teams
Factored
09/2023

DP-203 Azure Data Engineer Associate
Microsoft
07/2022

DP-900 Azure Data Fundamentals
Microsoft
06/2022

Skills

Hard Skills: Python, SQL, Scala, Apache Spark, ETL, ELT, Hadoop Framework, Apache Kafka, Databricks, Microsoft Azure, AWS Lambda, AWS S3, AWS Redshift, GCP Pub/Sub, GCP Cloud Functions, GCP BigQuery, Shell Scripting, CI/CD pipelines, Azure DevOps, Github, APIs, SFTP, JavaScript, TypeScript, ReactJS, C#, Power Automate, Power Apps, Microsoft SQL Server.

Main Soft Skills: Communication, Problem-solving, Adaptability, Emotional Intelligence, Active Listening, Customer Service, and Mentorship.