

David Gaikwad

7558228607 | gaikwaddavid2022@gmail.com | Nanded, Maharashtra

PROFILE

Data Engineer with expertise in cloud-based and having 6 months of experience in ETL processes, and scalable solutions. Successful track record in building strong data pipelines for informed decision-making. Looking for a role where I can apply my technical abilities and innovative mindset to contribute to a progressive organization's success and growth.

SKILLS

Technical: SQL, Python, Statistics, ETL, Pyspark, SQL Server, Data Transformations.

Tools: Azure Data Factory, ADLS Gen2, Azure Key vault, Databricks, Git CI/CD, YoloV5, labelimg.

EXPERIENCE

Cynapto Technologies

Mumbai, India

Intern 12/2022 – 6/2023

- Gathered, cleaned, and labeled diverse image datasets, enabling the development of AI models for industrial applications.
- Conducted YOLOv5 model testing and parameter fine-tuning, achieving optimal performance and detection accuracy.
- Ensured the delivery of high-quality models aligned with project objectives, contributing to the team's success.

PROJECTS

On-Premises to Azure Data Lake Storage Gen2 Data Migration

12/2024

- As an Azure Data Engineer, I lead a dynamic project workflow leveraging the advanced capabilities of Azure. Our data journey begins at the source —MySQL Database.
- With the seamless integration of Azure Data Factory (ADF), we establish a robust connection to extract crucial data, which is then meticulously transferred into the ADLS Gen2 Raw Container, marking the inception of our comprehensive data processing journey.
- In our Azure-focused framework, we seamlessly orchestrate a series of sophisticated operations, leveraging Azure Databricks processes, lookup iterations, foreach activities, and data copy processes to ensure the extraction of pristine and transformed datasets.
- The transformative phase, powered by Azure Databricks, encompasses advanced operations on data frames, culminating in the creation of a meticulously curated dataset. This refined data undergoes a pivotal transition into the ADLS Gen2 silver Container, representing a critical juncture in our workflow.
- As the curated data traverses through a gold Container, it aligns seamlessly with our holistic and consumable data processing model. The apex of our project lies in loading the silver (curated) data into the gold(consumption) Container, priming it for comprehensive analysis and insights.
- This strategic integration underscores our commitment to innovation, excellence, and strategic alignment with Azure's dynamic standards.
- Our project not only embodies a data journey but also serves as a testament to our prowess in Azure Data Engineering, encapsulating key skills such as PySpark, SQL, Azure Data Factory, Azure Data Lake Storage Gen2, and Python.

EDUCATION

JSPM's RSCOE Tathwade

Pune, India

Bachelor of Engineering in Mechanical Engineering

2018 – 2022

LANGUAGE

English | Hindi | Marathi