

# Nikhil Jadhav

Senior Data Engineer | PySpark, AWS, Iceberg, Databricks, Kafka, GenAI, DBT, MLflow, Docker | 7+ YOE  
+91-9764459911 · [nkhiljadhav77@gmail.com](mailto:nkhiljadhav77@gmail.com) | Pune IN - 411028 | [linkedin.com/in/nikhilofficial](https://linkedin.com/in/nikhilofficial) | [Portfolio](#)

## SUMMARY

Senior Data Engineer with 7+ years of experience building scalable data pipelines and cloud-native analytics solutions. AWS Certified Solutions Architect, skilled in PySpark, Databricks, and Airflow. Contributor to Apache Airflow, Spark, and Databricks open-source projects. Proven track record of reducing costs and driving insights through efficient, high-impact data systems.

## PROFESSIONAL EXPERIENCE

### Brose Group

July 2024 - Present

#### Data engineering Technology Lead

- Designed and implemented streaming data pipelines to monitor and track the product lifecycle and Machine Standstill reporting this improved real-time visibility Also built a **Gen-AI** application based on Bedrock.
- Utilized **PySpark** on **AWS Glue** and **Iceberg on AWS** for efficient data storage and processing.
- Architected and streamlined data infrastructure, reducing costs by **45%** and minimizing operational overhead through enhanced pipeline performance and resource optimization.
- Designed OnPrem based **Spark-Pytest** Testing environment using Docker to test ETL before deployment.

### Deloitte USI

MAR 2021 - July 2024

#### Senior AWS Data Engineer

- Leading a team of 3 cloud data engineers in designing and implementing a robust data pipeline aimed at enhancing the processing efficiency and accuracy of actuarial data.
- Built scalable ETL pipelines in AWS Glue, improving data refresh rate by 80% and saving 12 engineering hours weekly.
- Achieved a 90% reduction in processing time compared to the previous Data engineering approach

**Project Recognition:** Received **Applause Award** and **Spot Award** for technical expertise in delivering this project.

### Knowledge Lens: A Rockwell Automation Company

MAR 2019 - FEB 2021

#### Big Data Engineer

- Designed and implemented a scalable data pipeline using **Databricks(Pyspark)** and **ML** to process 20 TB of user feedback (text and audio) on drug adverse effects monthly.
- Project Contributed ~\$55 million in value to the firm by facilitating improved drug safety analysis through the data pipeline for client.
- Utilized **AWS Aurora as the data publication layer** to ensure efficient data access for analytics and visualization.

**Project Recognition:** This project was recognized as the most valuable within the firm.

### ESDS Software Solutions

Aug 2017 - Aug 2018

#### Junior Data engineer

- Supported as ETL Developer to an Indian government project and created a workflow to automate daily calculations of water and air quality indexes for various Indian cities. This solution eliminated manual calculations. this resulting in a **75% reduction in processing time** and improved data accuracy.
- Developed a user-friendly government bid submission portal, increasing average bids per project by 20%.

## SKILL

- Programming Languages: Python (Proficient), SQL (Expert) , Pyspark (Proficient) , YAML, SCALA
- Cloud Platforms: AWS (**Certified Solutions Architect - Associate**) , Databricks (**certified developer**)
  - S3, EC2, EMR, Lambda, Glue, EKS, DynamoDB, Bedrock, Athena, AppFlow
- Data Engineering & Pipelines :
  - Apache Spark , ETL/ELT Processes, Airflow, Git ,Kafka, Docker ,CloudFormation, Apache Arrow, Data Fusion, DBT, TRINO
- Data Warehousing: Amazon Redshift , Amazon S3 , Apache Iceberg , Google Big Query

## EDUCATION

### Centre for Development of Advanced Computing (2019)

Post Graduate Diploma -Big data analytics

Relevant Coursework: Machine Learning, Data Science Statistics, Big Data Analytics

### Savitribai Phule Pune University (2017)

Bachelor of Engineering in Computer Science