

SACHIN LODDIYA KARTHIK

☎ 517-243-9938 ✉ sachinlkece@gmail.com in www.linkedin.com/in/sachin-lk/ 🏠 Kalamazoo, MI

SUMMARY

Data Engineer with 3+ years of experience building scalable data pipelines, optimizing ETL workflows, and enabling real-time analytics. Proficient in PySpark, SQL, Azure Databricks, and Synapse, with hands-on experience in data modeling and lakehouse architecture. Skilled at transforming raw data into actionable insights to support business intelligence and strategic decisions.

TECHNICAL SKILLS

Data visualization: Power BI, Tableau **Machine learning modules:** Supervised Learning (Decision Tree, SVM, Random Forest, Linear Regression, Naive Bayes, KNN) **Statistics:** ANOVA, Hypothesis Testing, Regression Analysis

Data Engineering and Cloud Technologies : Apache Spark, Apache Kafka, Airflow, Databricks, Azure (Data Factory, Synapse Analytics, DevOps), AWS (EC2, S3, AWS Glue, Athena, Redshift), GCP (BigQuery, Pub/Sub, Dataflow), Docker

Database: Azure SQL, Oracle SQL, MySQL, MS SQL, PostgreSQL

Libraries: Pandas, NumPy, Matplotlib, PyTorch, TensorFlow, Scikit-Learn, Seaborn, OpenCV

Languages, tools and frameworks: Python, R, T-SQL, MS Office, Java, C, C++, HTML, CSS, JavaScript, Pycharm, VS Code, Winscp, R Studio, Putty, Jupyter Notebook

EXPERIENCE

Data Engineer, WMU - USA

Jan 2025 - Apr 2025

- Designed an automated class scheduling solution using Google OR-Tools and constraint programming, optimizing faculty-course assignments by factoring in availability, preferences, and time-slot constraints.
- Created and deployed a Streamlit-based UI for versatile CSV input processing, real-time validation, and interactive visualizations of optimized schedules.
- Minimized scheduling conflicts and manual effort through the integration of advanced optimization algorithms and a user-friendly frontend.

Data Engineer, Green Expectations LLC - USA

Jan 2024 - Apr 2024

- Constructed the backend infrastructure for an AI-powered web application that evaluates home sustainability, transforming user data into structured pipelines through a rule-based House Sustainability Calculator.
- Spearheaded a secure user authentication framework with persistent storage, facilitating individualized access and enabling real-time sustainability tracking.
- Orchestrated data flows for a conversational AI chatbot to enhance NLP-driven recommendations and improve user interaction.

Data Engineer, Accenture - India

Jul 2021 - Jul 2023

- Engineered ETL pipelines with Azure Data Factory to migrate and reformat critical supply chain data into Parquet format, boosting processing efficiency by 40% while minimizing disruption.
- Utilized Azure Databricks to process raw supply chain data from Data Lake Gen2, enhancing accuracy by 25% and halving transformation times.
- Developed external tables and views within Azure SQL and implemented interactive Power BI dashboards, enhancing data access and delivering insights that resulted in a 25% increase in operational efficiency.
- Improved data workflows by integrating ADF triggers, Databricks notebooks, and CI/CD pipelines via Azure DevOps, boosting team efficiency by 15% and 30% faster SCPO upgrade project delivery.

Data Engineer, Claritrics India Pvt Limited - India

Nov 2020 - May 2021

- Pioneered an advanced ETL pipeline using Azure Data Factory and Azure Databricks to create an OCR solution for extracting text from images, enhancing extraction accuracy by 30%.
- Crafted and scheduled an end-to-end automated workflow using Azure Data Factory for orchestration and Azure Databricks for scalable image processing, achieving a 40% reduction in text extraction processing time.
- Enhanced development and deployment cycles with Azure Data Factory triggers and Databricks job scheduling, resulting in a 20% gain in team efficiency and a 30% acceleration of project timelines.

EDUCATION

Master of Science in Data Science

Aug 2023 - Apr 2025

Western Michigan University, Kalamazoo, Michigan

GPA - 3.86/4

Relevant Coursework: Machine Learning, RDBMS, Azure Databricks and Spark For Data Engineers (PySpark / SQL), Applied Linear Models, Big Data Analysis, Google Cloud Big Data and Machine Learning Fundamentals.