

SACHIN LODDIYA KARTHIK

☎ 517-243-9938 ✉ sachin.loddiyakarthik@wmich.edu in www.linkedin.com/in/sachin-lk/ 🏠 Kalamazoo, MI

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), 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, Accenture Jul 2021 - Jul 2023

- Constructed comprehensive ETL pipelines using Azure Data Factory to migrate and transform critical supply chain data into optimized Parquet format, boosting processing efficiency by 40% with minimal operational disruption.
- Utilized Azure Databricks to process raw supply chain data from Data Lake Gen2, enhancing accuracy by 25% and halving transformation times.
- Created external tables and views within Azure SQL Database and designed interactive Power BI dashboards, enabling seamless data access and yielding insights that improved post-upgrade operational efficiency by 25%.
- Optimized data workflows by combining Azure Data Factory triggers with Databricks and establishing CI/CD pipelines via Azure DevOps, while bolstering data governance through GDPR compliance measures, resulting in 15% higher team efficiency and 30% faster SCPO upgrade project delivery.

Data Engineer, Claritrics India Pvt Limited Nov 2020 - May 2021

- Developed 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 implemented 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 processes using Azure Data Factory triggers and Databricks job scheduling, boosting team efficiency by 20% and accelerating project delivery by 30%.

EDUCATION

Master of Science in Data Science Aug 2023 - Apr 2025

Western Michigan University, Kalamazoo, Michigan

GPA - 3.84/4

Relevant Coursework: Machine Learning, RDBMS, Azure Databricks and Spark For Data Engineers (PySpark / SQL), Applied Linear Models, Big Data Analysis, R programming for Data Science.

Bachelor of Engineering in ECE Aug 2016 - May 2020

Anna University, Chennai, India

GPA - 7.3/10

PROJECTS

COVID-19 Analysis (Azure Data Factory, Azure Storage, Databricks, Kafka, DevOps) Jun 2024 - Jul 2024

- Architected end-to-end data pipelines in Azure Data Factory (ADF) to process COVID-19 data from various sources, including real-time Kafka streams, establishing a comprehensive analysis foundation.
- Crafted complex data transformations using ADF features, cutting pipeline development time by 25% and enhancing data analysis capabilities through efficient processing of Kafka and other data sources.
- Established a robust CI/CD process with Azure DevOps, employing Kafka for real-time data streaming and ADF for orchestration, reducing release cycle time by 35% and system downtime by 20%, enabling near real-time COVID-19 insights.

F1 Data Analytics Project (Azure Synapse, Databricks, Pyspark, SparkSQL) May 2024 - Jun 2024

- Built end-to-end ETL pipelines in Azure Synapse to integrate Ergast API data, automating execution with triggers and refining workflows using PySpark and SparkSQL, while setting up Azure Monitor for pipeline metrics tracking.
- Harnessed Azure Synapse for advanced analytics and data warehousing, incorporating schema enforcement, versioning, and monitoring to ensure data quality and reliability, facilitating actionable insights into F1 data trends.

CERTIFICATIONS

- Foundations: Data, Data, Everywhere - Google
- Data Analytics Essentials, Introduction to Data Science - Cisco
- Databases and SQL for Data Science with Python, ETL and Data Pipelines with Shell, Airflow and Kafka - IBM