

# Sai Kiran Adlluru

850 Greenside Dr, Richardson, TX 75080

[saikiranadlluru2408@yahoo.com](mailto:saikiranadlluru2408@yahoo.com) | [LinkedIn](#) | [GitHub](#) | [My portfolio](#) | [Leetcode](#) | (469)-370-9743

## Education

### University of Texas at Dallas

MS, Information Technology & Management (3.3/4.0)

Aug 2023 – May 2025

Richardson, Texas

### Vellore Institute of Technology

Bachelor of Mechanical Engineering (8.21/10.0)

Aug 2018 – Jun 2022

Vellore, India

## Relevant Coursework

- Data Structures & Algorithms
- Statistics & Probability
- Distributed Systems & Big Data
- Database Design
- Cloud computing Platform
- Data Science
- Data Pipelines & ETL tools
- Data Visualization

## Work Experience

### Dpi Ind services

Aug 2022 – Apr 2023

Data Analytics Engineer

Bangalore, India

- **Built AWS workflows** using **EMR** and **Amazon Kinesis workflows**, optimizing data pipelines to handle over **2Gb of real time data**, enhancing fault tolerance and achieving **99.9% uptime**.
- **Implemented Spark Streaming on EMR** to consume data from **Kinesis Data Streams**, applying real-time transformations and aggregations, reducing data processing latency by **20%** and scaling to handle over **2 thousand events per minute**.
- **Amazon Kinesis Firehose** to deliver streaming data to Amazon S3 in **Parquet format**, optimizing data storage and query efficiency. This **reduced storage costs by 25%** and **improved data retrieval speed by 40%**, supporting large-scale analytics workloads in S3.
- Leveraging **Apache Spark** to process real-time user data and provide actionable insights, increasing data processing speed by **60%**.
- Enhanced **real-time analytics** with minimal downtime and **data security** by leveraging AWS services such as **AWS IAM roles** for fine-grained access control, **VPCs** for network isolation, and **KMS encryption** for data at rest in S3 and in transit.

## Projects

### Data Pipeline for Twitter Data([GitHub](#))/ Python

Aug 2024 – Present

- **Built a Twitter data ingestion pipeline** using **Apache Airflow**, automating real-time tweet extraction and storage, increasing data throughput to **10,000 tweets per minute** with a **25%** improvement in processing speed.
- **Developed a scalable ETL pipeline** combining **Lambda** and **Airflow DaGs** to handle event-driven tasks, reducing operational costs by **40%** and improving query performance by **35%** for Twitter data stored in **Amazon S3**, reducing manual intervention by **50%** and improving pipeline efficiency by **30%**, ensuring real-time tweet processing and transformation.
- **Amazon Redshift** enabling storage of over **50 thousand tweets** in Redshift and reducing query times by **30%** using **Pandas** for data transformations. With **Quick Sight** dashboards providing real-time insights, improving data-driven decisions by **40%**.

### Apache Spark processing with AWS EMR ([GitHub](#))/ PySpark

Apr 2024 – Jul 2024

- **Implemented Apache Flume configurations**, enabling real-time streaming of **1000+ events/second** and improving throughput by **20%**.
- **Designed Kafka-Flume workflows**, cutting **data latency by 40%** and processing **200M+ records daily** for real-time analytics, enhancing decision-making and response times.
- **Automated error handling** in ETL pipelines, reducing downtime by 15% and increasing data throughput by 25%. Enabled faster error detection and reliable data flow for analytics.

### CI/CD on AWS and Terraform ([GitHub](#))

Feb 2024– Mar 2024

- **Implemented CI/CD pipelines for data workflows** using AWS Code Pipeline and Terraform, automating deployments and reducing manual errors by **70%**.
- **Containerized infrastructure with Terraform on AWS**, managing data resources like AWS S3, Redshift, and RDS, improving scalability by **30%**.

## Technical Skills & Certifications

**Languages:** Python, SQL, R, Unix/Linux, Data Structures & Algorithms, NumPy, Pandas, Matlab, Skit learn,

**Technologies & Tools:** MongoDB, RESTful APIs, PostgreSQL, AWS Lambda, S3, RDS, Redshift, Sage Maker, VPC, Quick sight, Tableau, Docker, Apache Spark, Power BI, HTML/CSS, Git, GitHub.

**Certifications:** AWS Cloud Practitioner, AWS Data Engineer, Data science with R, Python for Data science.