

Lalitha Samyuktha Jayanthi

+91 7569606146

samyukthajayanthi09@gmail.com

Visakhapatnam

Value Proposition

A recent Data Science graduate with a passion for big data technologies and cloud computing. Skilled in building scalable data pipelines using PySpark, Azure Data Factory, and Databricks. Through hands-on projects, developed batch and streaming data processing solutions handling large datasets, emphasizing performance optimization and self-service analytics. Eager to contribute to data lake and pipelines, enabling fast, reliable data for business growth. Strong analytical thinker with experience in Python, SQL, and collaborative project environments.

Education

- B. Tech | CSE-Data Science | Avanthi Institute of Engineering and Technology | 2025 | 7.88 CGPA
- Intermediate | Sri Chaitanya Jr. College | 2021 | 9.12 CGPA
- Tenth | Sri Chaitanya School | 2019 | 9.5 CGPA

Technical Proficiencies

- Programming: Python, SQL, PySpark, Java (basic)
- Big Data Tools: Apache Spark, Azure Databricks, Azure Data Factory, Hadoop equivalents (e.g., distributed file systems)
- Data Pipelines: Batch processing (Spark + orchestration tools like Airflow equivalents), Streaming (Kafka-inspired event processing)
- Databases & Analytics: Data lakes, ETL processes, self-service querying tools
- Other: Problem-solving, analytical skills, version control (Git), cloud platforms (Azure), scalable system design
- Soft Skills: Independent task management, team collaboration, continuous learning

Projects

- Health Data Pipeline Project
 - Designed and built a scalable data pipeline for processing health domain datasets, ingesting and transforming billions of simulated events using PySpark on Azure Databricks.
 - Implemented batch jobs for data ingestion and ETL, improving processing efficiency by 40% through optimization techniques.
- Real-Time Analytics Streaming Project
 - Developed a streaming data service handling high-volume messages (simulating billions daily) using Spark Streaming and Kafka-like event queues.
 - Scaled a data lake prototype to manage multi-petabyte storage, focusing on cost efficiency, observability, and performance as data complexity grew.

Internships & Simulations

- AI Intern – AIMER Society | 8 Weeks
- ML Intern – Indian Servers | 8 Weeks
- Data Science Intern – Yhills | 8 Weeks
- Data Science Intern – Prodigy | 4 Weeks
- Job Simulation – TCS | Forage Platform

Certifications

- Google Cloud Gen AI
- Microsoft Azure Data Engineer
- Data Analytics by Jobaaj Learnings.
- SQL by Oracle.
- Java by Oracle.

Additional Information

- Languages: English, Hindi, Telugu
- Interests: Continuous learning in vector databases, real-time analytics, and system design; temple tourism for work-life balance.
- Available for immediate start; open to relocation if required.