AMBICA DASARI

✉ ambdasa@iu.edu 📞 4088022340 📍 Bloomington, Indiana in linkedin.com/in/ambica-dasari

EDUCATION

Master of Science in Data Science   Aug 2023 – May 2025 | Bloomington, Indiana
Indiana University, Bloomington

Bachelor of Technology in Computer Science   Jul 2017 – May 2021 | Puducherry, India
National Institute of Technology, Puducherry

PROFESSIONAL EXPERIENCE

AWS Cloud Developer   Jan 2025 – present | Bloomington, Indiana
Indiana University

- **Cloud Infrastructure:** Designed and deployed a real-time data processing pipeline on AWS using Apache Kafka, efficiently handling 10,000+ seismic data points per second for ambient noise tomography under the FADS program.
- **AWS Solutions:** Engineered an auto-scaling, fault-tolerant cloud architecture, optimizing computational efficiency and ensuring high availability for large-scale seismological data processing and analysis.
- **Performance Optimization:** Implemented AWS Lambda and S3-based event-driven workflows, reducing data ingestion latency by 40% and improving the scalability of geophysical data analysis pipelines.

Data Engineer   Jul 2021 – Jul 2023 | Hyderabad, India
Larsen and Toubro Infotech

- **Scalable Data Processing:** Built and deployed an ingestion framework handling 1,000+ data sources with parallel processing, reducing execution time by 50% for 10TB+ daily data, and automating workflows using Python and Apache Airflow.
- **Security & Compliance:** Implemented GDPR-compliant security protocols and an ML-based anomaly detection system, reducing fraud by 30% across 500K+ transactions and securing 1M+ records.
- **Advanced Analytics:** Developed monitoring solutions leveraging machine learning for real-time anomaly detection, improving data integrity and fraud prevention strategies.
- **Cloud Infrastructure Optimization:** Enhanced system reliability and scalability, improving processing efficiency for enterprise data workflows by 40%, ensuring seamless performance under high data loads.

Data Analyst Intern   Jun 2019 – Jul 2019 | Hyderabad, India
Lancius IT Solutions

- **Interactive Dashboards** – Assisted in designing and generating Power BI dashboards, supporting data-driven decision-making and improving business operations efficiency by 35% through real-time visualizations and user-friendly reports.
- **Data Cleaning & Reporting** – Contributed to data cleaning and reporting efforts, reducing inconsistencies by 40% and increasing actionable insights by 20% by applying standardized data transformation and validation techniques.
- **Cross-Functional Collaboration** – Collaborated with stakeholders and analysts to translate data insights into business objectives, helping improve project delivery timelines by 20% through effective communication and requirement gathering.

PROJECTS

Stock Market Real-Time Data Analysis 🔗   Aug 2024 – Nov 2024

Architected and deployed a fault-tolerant real-time stock analysis pipeline using Apache Kafka and AWS MSK, implementing dead-letter queues and retry mechanisms for zero data loss. Integrated comprehensive monitoring using CloudWatch and Grafana dashboards, enabling real-time system health tracking and achieving 99.9% uptime.
Skills: Apache Kafka, Real-Time Data Processing, Stream Processing, Spark, SQL

YouTube Data Pipeline Optimization 🔗   Apr 2024 – Jul 2024

Designed and implemented a serverless ETL pipeline using AWS Step Functions and Lambda functions, reducing processing costs by 60% through optimized S3 storage patterns and Athena partition pruning. Integrated automated data quality checks using Great Expectations framework, achieving 99% anomaly detection rate in data processing workflows.
Skills: AWS (S3, Glue, Athena, Lambda), Data Lake Creation, ETL, SQL, Data Preprocessing, Cloud Security

SKILLS

Programming Languages: Python, C++, Java, R, C, HTML, CSS, PHP, JavaScript

Development and Data Tools: Apache Kafka, Apache Spark, Git, Docker, Real-Time Data Processing, ETL & Data Pipelines

Cloud & Tools: AWS(ECS, Lambda,S3, Glue, Redshift CloudWatch), GCP

Frameworks & Libraries: Pandas, NumPy, PySpark, Matplotlib, Express, Flask, Power BI, Tableau

Databases: PostgreSQL, MySQL, SQL, Supabase, Snowflake, Firebase

Others:: Data Warehousing & Governance, Big Data Processing, Machine Learning & Analytics, System Design, Agile & Scrum

CERTIFICATES

AWS Certified Data Engineer- Associate | IBM Data Engineering Professional