Akshata Madavi

✉ akshatamadavi@gmail.com | 🌐 linkedin.com/in/akshata-madavi | ⌂ github.com/Akshata4

Summary

Graduate student in Software Engineering (Data Science) at San Jose State University with industry experience in **Python, data engineering, and cloud technologies** (GCP). Skilled in building **scalable data pipelines, optimizing performance**, and collaborating across teams. Actively seeking a Data engineering internship to apply technical and analytical expertise to real-world problems.

Education

San Jose State University, MS in Software Engineering   Aug 2025 – Present
Relevant Coursework - Data Analytics, Data Mining, Recommender Systems, Enterprise Sofware Platforms

College of Engineering Pune (COEP), BS in Computer Engineering   Aug 2014 – May 2018
Relevant Coursework - Programming, Databases, Algorithms, Software Design

Skills

- **Languages-** Python, SQL
- **Machine Learning & Analytics**: PySpark ML, scikit-learn, XGBoost, Tableau
- **Data warehouse/Databases**- BigQuery, MySQL, ClickHouse, Presto
- **Data pipeline/Data architecture**- Dataflow, Cloudrun, Cloud Storage, Compute engine, Spark, Kafka, PubSub
- **Orchestration**- Workflows, Cloud Composer, Airflow

Work Experience

Senior Data Engineer, LTIMindtree – Pune, India   May 2023 – Mar 2025

- Developed and **optimized** near real-time data pipelines to seamlessly migrate on-premises traditional databases to a centralized BigQuery platform, facilitating **efficient analytical workloads** and insights. Processed approximately **10 million records.**
- Designed and implemented an **event-driven pipelin**e to process and load complex, inconsistent data from multiple Excel sheets into GCP, reducing processing time from a week to under 10 minutes with **parallel processing of 30 GB**
- Established **orchestration** using **Google Workflow** to ensure smooth execution and management of data pipelines.

Data Engineer, AI Adventures – Pune, India   Aug 2021 – Apr 2023

- Led migration of course data from Moodle to MySQL, enabling business intelligence views that **drove a 20% increase in revenue**.
- Expanded insights into **sales, student retention, and curriculum performance**, improving data-informed decisions across the organization.

Data Engineer, STL – Pune, India   Jun 2018 – May 2021

- Built **near real-time data pipelines** to migrate on-prem databases (Oracle, SAP and Salesforce) to BigQuery, **enabling efficient analytics** and processing approximately **1 TB** of data.
- Provided Tableau reports and data-driven insights to **support business functions**, including scrap detection and DOE analysis, **driving process improvements** and informed decision making.
- Collaborated with **cross-functional teams** to ensure effective data usage, driving successful operational outcomes, and achieve business objectives.

Projects

- Philosophy-based Recommender System for Mental Wellness (in progress) : Developing a **philosophy-driven mental wellness recommender** using **LLMs** and **RAG architecture** to deliver personalized philosophical insights based on emotional context.
- Built a distributed **PySpark ML** pipeline on NYC TLC trip data (30M rows), engineered geospatial & time-based features, trained Gradient-Boosted Trees for fare prediction, and tracked experiments with MLflow.

Certifications

Introduction to Data Science in Python, Coursera. | Exploring and Preparing your Data with BigQuery, Coursera. | Creating New BigQuery Datasets and Visualizing Insights by Google Cloud, Coursera. | Achieving Advanced Insights with BigQuery, Coursera. | Applying Machine Learning to your Data with GCP, Coursera.