Jay Shah

Mumbai, India | +91 8879566776 | jayshah36262@gmail.com

linkedin.com/in/jay-shah-de | github.com/jayshah-12 | medium.com/@jayshah36262

Professional Summary

Data Engineer with 1.5+ years of experience building AWS-based data pipelines and platform solutions. Skilled in ETL development, data integration, and serverless architectures with exposure to monitoring and observability systems.

Technical Skills

Programming: Python, SQL, PySpark, Bash

Data Engineering & ML: ETL Pipelines, Data Warehousing, Batch & Real-time Processing, Feature Engineering, Exploratory Data Analysis, Time Series Forecasting, Hyperparameter Tuning

Cloud & Databases: AWS (Glue, Lambda, Step Functions, S3, Athena), PostgreSQL, Snowflake, MySQL, DynamoDB

DevOps & Tools: Docker, Terraform, GitHub Actions, Concourse CI, Git, MkDocs, Selenium, Playwright

Professional Experience

Data Engineer — **Nov 2025 – Present**

Go Digital Technology Consulting LLP — *Mumbai, India*

- Architected cross-account AWS infrastructure connectivity using Terraform, enabling 6 products across isolated AWS accounts to securely publish events and share logs to a central observability account
- Built UI uptime and API performance monitoring using Playwright - tracking metrics like LCP, response times, and payload sizes to detect degradation before it impacts users
- Created an automated compliance monitoring script for a legally regulated government-integrated map, performing hourly component checks and triggering instant Microsoft Teams alerts on failures, mitigating legal risk for the organization
- Built a Documentation Agent POC using GitHub Copilot to auto-generate structured docs from multiple repositories, hosted on GitHub Pages via MkDocs with CI/CD integration for auto-updates on every push

Associate Data Engineer — **Jul 2024 – Nov 2025**

Go Digital Technology Consulting LLP — *Mumbai, India*

- Appointed as migration lead for transitioning 25+ CI/CD workflows across 5 products from Concourse CI to GitHub Actions; completed ahead of schedule, received client appreciation, and accelerated org-wide adoption by authoring documentation and reusable boilerplate templates
- Co-built the Product Health Dashboard, a centralized observability platform monitoring 6 enterprise products — implementing telemetry for AWS Glue, Lambda, and Step Functions with granular execution metadata to reduce mean time to debug failures
- Established CloudWatch log ingestion pipelines into a central S3-backed database and built cost tracking dashboards with automated 30% spike alerts to tech leads, enabling proactive cost governance
- Reduced AWS RDS infrastructure costs by 60% through Terraform optimization and right-sizing; also migrated a critical RDS instance from end-of-life v1 to v2 with zero pipeline disruption

Projects

Product Health Dashboard | *AWS , S3, Glue, Athena, Terraform, Playwright* — **2024 – Present**

- Designed a centralized observability platform aggregating events, logs, and cost data across 6 enterprise products, with cross-account infrastructure provisioned via Terraform
- Engineered event capture for AWS Glue, Lambda, and Step Functions, storing execution metadata in Athena-queryable S3 layers for failure debugging
- Built UI uptime and API performance monitoring using Playwright and integrated automated cost anomaly alerts with 30% spike detection and Teams notifications

Load Forecasting Model | *Python, Pandas, Scikit-learn, Time Series* — **2025 – Present**

- Working on a load forecasting model to predict electricity demand using historical time-series data
- Performed exploratory data analysis and feature engineering including time-based features, lag variables, and rolling statistics
- Applied hyperparameter tuning and model evaluation techniques to improve forecasting accuracy

Education

Vivekanand Education Society's Institute of Technology — **Aug 2020 – May 2024**

Bachelor of Engineering in Information Technology, Honours in AI/ML | *CGPA: 8.75/10* — *Mumbai, India*