

# Shivam Gupta

Pune, India | [shivamai.dev](mailto:shivamai.dev) | [shivam@shivamai.dev](mailto:shivam@shivamai.dev)  
[linkedin.com/in/shivamgupta1812](https://linkedin.com/in/shivamgupta1812) | [github.com/shivam-gupta18](https://github.com/shivam-gupta18)

## PROFESSIONAL SUMMARY

Senior Software Engineer with 7+ years of experience building secure APIs, distributed compute infrastructure, and large-scale data platforms. Reduced cloud compute costs by up to 80% and scaled data pipelines processing 3+ TBs of daily logs. Specializes in Python, Kubernetes, and query optimization (Spark, Trino) to build resilient backend systems and orchestrate data architectures that support analytical and AI/LLM workloads.

## TECHNICAL SKILLS

- **Backend & APIs:** Python, FastAPI, Flask, RESTful APIs, System Design, Microservices, SQLAlchemy, Pydantic
- **Data & AI Infrastructure:** Apache Spark (PySpark), Trino, Apache Airflow, PostgreSQL, LLM Integrations (Gemini API, Claude), Agentic AI Workflows (RAG, Model Context Protocol)
- **Cloud & Infrastructure:** Kubernetes (EKS), AWS (EC2, S3, IAM, SecretsManager, EventBridge, Glue, Lambda, Sagemaker, SQS), Docker, ArgoCD, GitLab CI
- **Observability & Monitoring:** CloudWatch, OpenSearch, Grafana, Prometheus

## PROFESSIONAL EXPERIENCE

### FPL Technologies Pvt Ltd (Onecard)

Pune, India

Senior Software Engineer

Feb 2023 – Present

- Cut infrastructure costs by 80% and improved pipeline reliability by migrating the core data transformation workflows from AWS Glue to Apache Spark on Kubernetes (EKS), processing 5+ TB of transaction logs and events daily via Airflow.
- Eliminated chronic Airflow scheduler crashes and achieved 6 months of zero-incident uptime by resolving S3 sync race conditions, caching variables, and refactoring DAGs for lazy dependency loading to drastically reduce total DAG parse times.
- Reduced compute costs by 40% under peak load by pushing heavy analytical predicates down to Trino and rewriting core Python/Pandas logic into Polars, scaling workloads down from 8xlarge to xlarge AWS instances.
- Dropped managed-service costs by 30% by designing and deploying a self-managed, Docker-isolated JupyterHub environment on EC2, establishing a secure infrastructure for the data science team's machine learning and LLM experimentation.
- Eliminated SQL injection risks in legacy code by architecting a custom SQLAlchemy translation layer that intercepts and validates user-defined parameters before database execution.
- Secured multi-tenant access across thousands of automated Git jobs by building a Python authentication library that integrates OAuth2 with AWS Secrets Manager for unified RBAC.
- Enforced organization-wide LLM cost controls by configuring and testing a LiteLLM gateway, establishing multi-tenant budget thresholds that centralized token tracking across all internal GenAI workloads.
- Accelerated initial PR feedback cycles by 25% by integrating an AI-powered reviewer into the Git CI pipeline, engineering custom Claude skills (MCP) to inject repository-wide context from Airflow, dbt, and core data pipelines.

### Bigfoot Retail Solutions Pvt Ltd (Shiprocket)

Delhi, India

Software Engineer

Jun 2021 – Feb 2023

- Eliminated 95% of manual accounting intervention by architecting a rule-based decision engine that fully automated the fulfillment billing reconciliation system.
- Maintained 100% platform uptime during high-traffic peak seasons by utilizing latency profiling to identify and patch critical database bottlenecks and optimize slow queries.

### Triyama Software Pvt Ltd

Pune, India

Full Stack Developer

Mar 2020 – Jun 2021

- Cut manual developer intervention by 90% by building a dynamic internal web-page builder that automated the delivery of UI components for client projects.
- Accelerated client reporting by 100% by building automated data processing pipelines that ingested raw business metrics and generated recurring visual dashboards.

### GAK E Solution Pvt Ltd

Hyderabad, India

Software Engineer

Jun 2019 – Mar 2020

- Raised the application SEO score from 64% to 97% by optimizing MySQL schema designs to improve read response times and deploying targeted UI structure enhancements.
- Secured user authentication and transactions by integrating third-party REST APIs for social logins and payment gateways.

## EDUCATION

### IFHE Hyderabad

Hyderabad, India

Bachelor of Technology

Aug 2015 – Jul 2019