


# SATYAM GOVILA

☎ (+91)844-584-1444 ✉ govila.satyam@gmail.com  linkedin.com/in/satyam-govila  github.com/satyamgovila  
satyamgovila.github.io/portfolio/

## Work Experience

---

### Zupee (*India's fastest Real Money gaming platform*)

August 2022 – Present

Senior Data Engineer

Full-time

- Revamped Data Migration Architecture with MSK
  - \* Solved S3 rate limits from MongoDB/MySQL/Third-Party Marketing tools by integrating MSK with Apache Hudi, ensuring smooth data migration, incremental data processing and 30% reduced Athena scans via ZSTD compression.
  - \* Created automated API-driven framework for DMS-Hudi Migration by storing metadata in PostgreSQL to decouple manual efforts.
- Automated Reconciliation Data Platform
  - \* Built an automated reconciliation platform with AWS (Lambda, Glue, Step Functions), SQL, and Tableau, unifying PG and internal transaction records into a single source of truth for FinTech operations
- Data Infra Optimisations
  - \* Developed highly available real-time compaction system reducing S3 API costs by 47% (\$7,000 monthly) by optimizing high-frequency Kafka events.
  - \* Achieved 55% data scan reduction and 30% faster query processing through sharding, automated partition management, saving 14% monthly AWS costs by optimising ListAPI and Network calls on S3.
- Realtime Feature Store Integration
  - \* Architected real-time feature store with Kafka, Flink, and Aerospike for 10M+ daily events, serving 100+ features across 20+ ML models with <50ms p99 latency.
- Fraud Detection Platform using GraphDB
  - \* Developed graph-based fraud detection with Flink, AWS Neptune, and Gremlin, processing 10M+ events/day with sub-second queries, preventing \$2M+ fraud at 95%+ accuracy.
- Integration of Trino into Data Architecture
  - \* Deployed Trino as scalable Athena fallback, seamless integration, and workload migration ensuring compatibility with Glue catalogs and Hudi tables.
- Game mechanics data logging & Integration of DBT and Data Modelling
  - \* Led data logging across multiple games and drove end-to-end infra setup for multiple games, ensuring timely cross-functional alignment with Product/Tech Team.
  - \* Designed scalable dbt data models with modular transformations, dependency management, and CI/CD validation for finance and marketing analytics.

### Concirus (*Marine insurance based AI platform*)

July 2019 - August 2022

Data Engineer(SDE-2)

Full-time

- Built ETL APIs for real-time and batch data pipelines using Python, Spark and AWS. Designed optimal data pipeline architectures and developed automated vessel history tracking system.

## Education

---

### University of Hyderabad

February 2021 - April 2022

PGD - Machine Learning and AI

### ZHCET , Aligarh Muslim University

July 2015 - July 2019

Bachelor of Technology - Electronics Engineering

## Projects

---

### Real Time Log Data Analytics | *Lambda Architecture, Spark, Kafka*

- Built AWS real-time website monitoring system with streaming log analysis and Lambda triggers for events using Kinesis, Aurora, DynamoDB.

### Thesis Project - ML Based Virtual Gym Assistant | *Python, AWS EC2, Pose Estimation*

- Developed AI system for home workouts providing real-time posture feedback and virtual training using pose estimation techniques.

Blogs <https://satyamgovila.medium.com/>

## Technical Skills

---

**Languages and Frameworks:** Python, SQL, Scala, Gremlin, DBT

**Streaming & Big Data:** Kafka, Flink, Spark, Hudi, Trino

**Databases & Storage:** PostgreSQL, MongoDB, MySQL, Aerospike, AWS Neptune, Amazon S3