

Shivam Sharma

10sharmashivam@gmail.com | +91-8285635909 | Linked-In | GitHub

Education

Amity University, Noida, B.Sc. Physics (Hons.)	Jul 2018 – Jul 2021
Financial Risk Manager, FRM (GARP)	Candidate

Experience

AI Engineer Intern, Vitto (Fintech, part of **Bank of America's Breakthrough Lab** – 2024 APAC) – Gurugram, Haryana India Nov 2024 – Present

- **LLM-Based Fraud Detection:** Built and deployed fraud detection microservices using **Docker, AWS Lambda, and REST APIs**; fine-tuned LLMs to run in low-latency environments in production systems.
- **Cloud & Big Data Infrastructure:** Spearheaded migration of **12+ microservices** from **AWS (EC2, Lambda, S3)** to **GCP (Dataproc, Cloud Functions)**, cutting cloud costs by **80%** and ensuring **100%** uptime. Streamlined ETL pipelines using **Hadoop, Hive, PySpark, and PostgreSQL**, boosting data processing efficiency by **1.6x**.
- **ETL & Data Warehousing:** Designed and optimized scalable ETL pipelines using **Apache Spark, Scala, Hive, and HDFS on GCP Dataproc**, enabling near real-time data ingestion and warehousing for high-throughput fintech analytics.
- **DevOps & Observability:** Utilized **Linux CLI** for server management, configured **Docker Compose** and **YAML**-based deployments, and used **Git CLI** for version control; deployed observability stack using tools like **Grafana** and **ELK** for system diagnostics.
- **Enterprise System Integrations:** Integrated **Account Aggregator (FinVU, CAMS)**, **CRIF reports**, and **KYC APIs (Decentro, Digio)** to automate credit risk assessment, accelerating approval processes by **40%**. Developed REST APIs to connect **EASEBUZZ** payment gateway and **Care Health Insurance (CHIL)** systems, ensuring real-time compliance for regulated workflows.
- **Distributed Data Engineering & Workflow Orchestration:** Worked on prototypes involving on-prem **Hadoop distributions (Hortonworks, MapR)** and **EMR/HDInsight** clusters. Used **Pig, Hive, Sqoop, and Spark SQL** for batch ETL pipelines; experimented with **Spark Streaming and Kafka** for ingestion and **HBase** for real-time access. Delivered work in Agile/Scrum teams with CI/CD best practices.

Projects

Cloud Native Resource Monitoring Python App on KubernetesGitHub

- Built a Python-based system resource monitoring tool using **Flask + Plotly**, containerized with **Docker**, and deployed on **Kubernetes** for scalable performance tracking.
- Implemented real-time **CPU/memory visualization** dashboards with Bootstrap UI, YAML-based deployments, and exposed services via **K8s**.

Open Source Contributions

Flyte (Linux Foundation AI and Data graduate Project) GitHub

- Integrated PyTorch Profiler Renderer into FlyteDeck for seamless visualization of training bottlenecks.
- Enhanced MPI Operator error handling, caching documentation mechanism.
- Revamped test data hosting with 500GB of data by reducing duplicates, & implementing version control.

SKILLS and TOOLKIT

Cloud & Big Data: Spark, Hadoop, Hive, Scala, PostgreSQL, MongoDB, AWS (EC2, Lambda, S3, Beanstalk), GCP(Dataproc, App Engine, Cloud Build), automation (Selenium, cron, workflow orchestration), Docker, Kubernetes

Machine Learning/AI and Infra: Scikit-Learn, PyTorch, TensorFlow, HuggingFace, spaCy, NLTK, anomaly detection, GenAI, Docker, Large Language Models (LLMs), multimodal AI

Data and Analytics: Python, SQL, Pandas, Excel, Tableau, KPI design, Data Visualization (Tableau), GraphQL, Django & Flask, Istio, Prometheus, OpenTelemetry, Glue, ADF, Dataflow