Shivam

+91-7307602433 — shivamgravity@gmail.com
linkedin.com/in/shivam-gravity — github.com/shivamgravity — kaggle.com/shivamgravity

TECHNICAL SKILLS

Programming:	Python, SQL
AI/ML:	TensorFlow, PyTorch, Scikit-learn, Pandas, LangChain, OpenCV
Data Engineering:	Apache Spark, Airflow, Databricks, Snowflake, ETL/ELT, n8n
Cloud & Infrastructure:	AWS (Lambda, S3, EventBridge), Docker, Git
Databases:	PostgreSQL, MySQL, Neo4j

EXPERIENCE

Blackcoffer Pvt. Ltd. — Data Scientist Associate (Probationary) *Dec 2025 – Mar 2026*

- Built scalable retail data solutions for a US-based client, processing pricing and product availability data across chains with 1,000+ locations in North America.
- Designed automated ETL pipelines using n8n and Python for data extraction, transformation, integration, and workflow optimization.
- Migrated and organized product data in PostgreSQL, enabling efficient SQL querying and exploratory analysis across large retail datasets.
- Partnered with product and engineering teams on data-driven solutions, system design, and workflow optimization.

OPEN SOURCE CONTRIBUTIONS

Kaggle CLI — Open Source Contributor *Aug 2026*

Python, pytest, Black, Git, GitHub PR #1183

- Fixed misleading code-competition submission validation by adding API-level input checks and regression tests, preventing invalid requests from reaching the backend.
- Incorporated maintainer feedback, refined validation architecture, and contributed the fix to the official Kaggle CLI repository; full test suite passed 1,292 tests.

PROJECTS

Enterprise Document AI Assistant — RAG Application *May 2026*

github.com/shivamgravity/langchain-rag-chatbot — Live Demo — Demo Video

- Developed and deployed a multi-PDF RAG assistant using LangChain, HuggingFace embeddings, ChromaDB, and Groq/Ollama LLMs for contextual document question answering.
- Built persistent authentication, SQLite/SQLAlchemy data management, structured logging, and a modular retrieval pipeline supporting both cloud and local inference.

Serverless Data ETL Pipeline & Dashboard *Dec 2025*

AWS Lambda, S3, EventBridge, Python, Streamlit, Pandas, ETL/ELT
github.com/shivamgravity/aws-serverless-data-pipeline

- Architected a serverless ETL pipeline using AWS Lambda, S3, and EventBridge to generate, ingest, transform, and store large-scale transaction data.
- Developed a Streamlit analytics dashboard and automated event-driven ingestion workflow to demonstrate production-style data processing and visualization.

PhyCosmos — Production Science Publication Platform (phycosmos.com) *Apr 2020 – Present*

Astro, JavaScript, Markdown, Git, GitHub, Cloudflare Pages, SEO

- Built and maintain a production science publication platform with Astro, GitHub, and Cloudflare Pages, supporting articles, categories, tags, pagination, search, related content, and responsive navigation.
- Implemented technical SEO infrastructure including sitemap, RSS, OpenGraph metadata, structured data, canonical URLs, and performance optimizations.

EDUCATION

VIT Bhopal University, Bhopal *Oct 2022 – May 2026*

Bachelor of Technology — Computer Science Engineering *CGPA: 8.69*

CERTIFICATIONS

Data Management in Databricks — Databricks, Datacamp *Aug 2026*
Snowflake Architecture — Snowflake, Datacamp *Aug 2026*
Machine Learning with Python — iamneo *Feb 2025 – Apr 2025*