

# Shubham Sarkar

Kolkata, India

 shubhamsarkarthe1@gmail.com |  LinkedIn: <https://www.linkedin.com/in/shubham-sarkar-57438b27a/>

 GitHub: <https://github.com/SSarkar-1> | LeetCode: <https://leetcode.com/u/ShubhamSarkar243/>

---

## SUMMARY

Junior LLM / Machine Learning Engineer with hands-on experience building NLP and multimodal ML systems using transformer-based models. Skilled in fine-tuning, evaluation, and deployment of language models, with practical exposure to RAG pipelines, prompt engineering, and MLOps workflows. Strong Python background with experience shipping ML features into production.

---

## SKILLS

**LLMs & NLP:** Transformers, Prompt Engineering, RAG, Text Classification, Summarization, Embeddings, Tokenization

**Frameworks & Libraries:** PyTorch, TensorFlow, Hugging Face Transformers, scikit-learn

**MLOps & Deployment:** Docker, MLflow, DVC, FastAPI, Flask, Streamlit

**Data & Systems:** Python, SQL, Pandas, NumPy, AWS, Linux

---

## WORK EXPERIENCE

### Machine Learning Engineer (Freelance) | Oct 2025 – Nov 2025

- Built and deployed an NLP-based sentiment analysis system using transformer models to extract structured insights from large-scale YouTube comment data.
  - Designed text data pipelines including cleaning, tokenization, and prompt-style preprocessing to improve model robustness.
  - Tracked experiments and evaluations with MLflow; versioned datasets and models using DVC and Git.
  - Containerized and deployed inference services using Docker and FastAPI on AWS.
-

## PROJECTS

### Smart Product Pricing Model – Amazon ML Challenge 2025

- Developed a multimodal ML system combining transformer-based text embeddings and CNN-based image representations on a 150K+ product dataset.
- Integrated text and vision features via neural fusion architecture; evaluated using SMAPE, achieving **global rank 142**.

### YouTube Sentiment Analysis & Reporting Engine

- Built an end-to-end NLP pipeline for sentiment classification and automated report generation from raw text data.
- Exposed model outputs via REST APIs, enabling integration with downstream product features.

### Customer Churn Prediction System

- Built and deployed a supervised ML model classifying customers into five churn-risk tiers; delivered real-time predictions via a Streamlit app.

### Golden Crossover Trading Strategy Backtester

- Implemented and evaluated a data-driven trading strategy with automated performance reporting.

## EDUCATION

**B.Tech, Engineering** – Jadavpur University (2023–2027)

**Higher Secondary** – Hariyana Vidya Mandir (2021–2023)

---

## LEADERSHIP & EXTRACURRICULARS

**Coordinator – Jadavpur University Finance Club:** Led Finspire 2025 (1,000+ registrations); delivered trading workshops to 100+ students.

**Core Member – Jadavpur University Entrepreneurship Cell:** Organized E-Summit & Hult Prize 2025 (5,000+ registrations).