Shubham Sarkar
Machine Learning Engineer
email: shubhamsarkarthe1@gmail.com | Github: http://github.com/SSarkar-1 | linkedin: https://www.linkedin.com/in/shubham-sarkar-57438b27a | Leetcode: https://leetcode.com/u/ShubhamSarkar243

---

## SUMMARY

Machine Learning Engineer with hands-on experience designing and deploying NLP-driven systems for document analysis, semantic search, and intelligent insights. Strong background in text preprocessing, embedding-based retrieval, multimodal learning, and scalable ML pipelines. Proven ability to collaborate across data, product, and infrastructure teams to deliver low-latency, production-ready AI systems using Python and cloud platforms. Passionate about experimenting with LLMs and building end-to-end AI applications.

---

## SKILLS

- **Machine Learning & AI:** Supervised & Unsupervised Learning, NLP, Computer Vision, Multimodal Models, Feature Engineering, Model Evaluation
- **Frameworks & Libraries:** TensorFlow, PyTorch, scikit-learn, Pandas, NumPy
- **MLOps & Deployment:** Docker, MLflow, FastAPI, Flask, CI/CD
- **Data & Systems:** SQL, ETL Pipelines, Linux, AWS, API Integration
- **Programming:** Python, C, JavaScript

---

## WORK EXPERIENCE

### Machine Learning Engineer, Freelance

*Oct 2025 - Nov 2025*

- Built an end-to-end sentiment analysis pipeline to extract and analyze user comments from YouTube videos using Python and popular ML libraries, improving analysis speed by 30%.
- Preprocessed and cleaned large-scale text data, applying NLP techniques such as tokenization and lemmatization, resulting in a 25% increase in model performance.
- Used DVC and Git for version control and MLFlow for experiment tracking, enhancing project collaboration efficiency by 40%.
- Containerized the application for reproducible deployment with Docker, enabling seamless updates and maintenance.
- Developed Flask RESTful APIs for backend integration, facilitating a user-friendly interface.

### Machine Learning Engineer, Freelance

*July 2025 - Aug 2025*

- Constructed a production-ready ML pipeline from ingestion to deployment, demonstrating strong ML system design principles.
- Engineered structured features and optimized models via hyperparameter tuning, resulting in a 20% improvement in prediction accuracy.
- Deployed a real-time inference application with monitoring, achieving 99% uptime.

---

## PROJECTS

### Golden Crossover Backtester

- Collected and cleaned historical stock data for NIFTY-listed equities using Python, Pandas, and NumPy.
- Developed and automated the Golden Crossover (20/50 SMA) trading strategy to generate buy/sell signals, leading to a 15% increase in profitability.

- Built a backtesting framework to evaluate strategy performance, including profitability and win/loss ratios, with results summarized for reporting.

**Smart Product Pricing Model, Amazon ML Challenge 2025**

- Built a multimodal document intelligence system combining product descriptions and images, enhancing semantic understanding.
- Engineered transformer-based text embeddings and CNN/CLIP image embeddings for semantic representation, achieving a 142 rank in the challenge.
- Designed efficient feature storage and retrieval pipelines for large-scale inference, improving retrieval speed by 35%.

---

## EDUCATION

**Bachelor of Technology, Jadavpur University**

*2023 - 2027*

**High School, Hariyana Vidya Mandir**

*2021 - 2023*

---

## EXTRACURRICULARS

**Core Member, Jadavpur University Entrepreneurship Cell**

- Organized national-level flagship events such as E-Summit 2025 and Hult Prize 2025, achieving 5,000+ registrations and 1,000+ on-campus attendees.
- Contributed to establishing an Incubation Center at Jadavpur University under the Institution's Innovation Council (IIC).

**Coordinator, Jadavpur University Finance Club**

- Led planning and execution of Finspire 2025, a national-level finance event with 1,000+ registrations and 500+ on-ground attendees.
- Delivered high-impact trading and investment courses to 100+ students, enhancing engagement in financial markets.