Shreyash Sahare

Boulder, CO | work.shreyash1001@gmail.com | 720-255-4636 | LinkedIn | GitHub

SKILLS

Languages and Frameworks: Python, C++, SQL, PyTorch, TensorFlow, OpenCV, YOLO, Scikit-learn, Hugging Face
Deep Learning Concepts: CNNs, Transfer Learning, Autoencoders, Vision Transformers, Mixed Precision Training

EXPERIENCE

Void Robotics *Marathon, FL (Remote)*
Robotics Engineering Intern *08/2025 - Present*

- Engineered **ROS2-based** perception modules containerized in **Docker**, to support **robotic automation**
- Introduced **HSV-based chromatic thresholding** with **localized DT** and connected components, boosting robustness by **30%**

Mercor *San Francisco, CA (Remote)*
Data Science Expert (Independent Contractor) *08/2025 - 09/2025*

- Executed **20+ ML** projects covering hypothesis testing, time series forecasting, and predictive modeling
- Crafted prompt engineering workflows for **Gemini 2.5 Pro**, benchmarking reasoning with **custom rubrics**
- Developed evaluation frameworks to score the LLM on crafted rubrics, achieving **27% across all tasks**

Parlay Finance *Boulder, CO (US)*
Data Scientist *01/2025 - 05/2025*

- Built loan classifiers using **XGBoost**, **Random Forests** and **TabTransformers** with **87% accuracy** on **200k+ financial records**
- Reduced false approvals by **22%**, directly improving user trust and conversion strategies
- Conducted **exploratory data analysis** and identified variables that increased model interpretability by **35%** through **SHAP**

National Institute of Technology, Tiruchirappalli *Tiruchirappalli (India)*
Undergraduate Research Intern *06/2022 - 08/2022*

- Devised **ResCoWNet** CNN to denoise OCT scans, boosting **SSIM** by **12%** over baseline models for **clearer retinal imaging**
- Contributed to peer-reviewed research **cited 8+ times**, showcasing applied expertise in **medical AI and computer vision**

TECHNICAL PROJECTS

Palantir Foundry Assistant | *[GitHub]* *06/2025 - 07/2025*

- Fine-tuned **LLaMA-3.2-1B** using **QLoRA** and **PEFT** to scale training via **RunPod** with **distributed multi-GPU setup**
- Generated **1,000+ synthetic QA pairs** with **Qwen-2.5-14B** and refined to **700+**, **boosting answer clarity** and **relevance**
- Improved assistant accuracy by **28%**, demonstrating advances in **context window efficiency & agentic workflows**

Scaleflow - Supply Chain Analytics | *[GitHub]* *02/2025 - 05/2025*

- Architected multi-threaded ETL pipelines to process **3M+ trade records**, improving data ingestion speed by **58%**
- Connected **Pinecone vector DB + Supabase DB** for contextual reasoning, achieving an API response success rate of **95%**
- Automated **Airflow-based orchestration** and **Docker** deployments, enabling scalable and reproducible model execution

Privacy Enhancing ML for Job Interview Analysis | *[GitHub]* *11/2024 - 12/2024*

- Classified job interview responses into four explanation levels for **38 interviewees**, analyzing linguistic patterns.
- Delivered **79% balanced accuracy** for **Random Forest** (Over-explained vs. Comprehensive) and **66.67%** for **Conv1D** (Under-explained vs. Succinct).
- Engineered **1500+ linguistic features** (TF-IDF, POS tagging, embeddings), reducing speaker bias by **30%**

Wine Quality Prediction with MLOps | *[GitHub]* *10/2023 - 11/2023*

- Logged **30+** experiments via **MLflow** UI for **ElasticNet** tuning using **AWS Sagemaker**, reducing training time by **27%**
- Deployed via **Docker** on **AWS EC2** with GitHub actions, delivering predictions in **under 3s** per input
- Monitored **latency**, **drift**, and **availability** through **CloudWatch** dashboards and alerts

EDUCATION

University of Colorado Boulder *08/2023 - 05/2025*
Master of Science in Data Science *GPA: 3.73*

- **Key Coursework:** Neural Networks and Deep Learning, Machine Learning, Big Data Analytics, Statistical Methods

National Institute of Technology, Tiruchirappalli *07/2019 - 05/2023*
Bachelor of Technology in Electronics and Communication Engineering *GPA: 3.22*

- **Key Coursework:** Linear Algebra and Calculus, Statistical Theory of Communication, Deep Learning

CERTIFICATION

- **SnowPro Core Certification**, *Snowflake - Aug 2025*