

# 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

Robotics Engineering Intern

Marathon, FL (Remote)

08/2025 - Present

- Developed an advanced pill counter pipeline with **progressive K-Means**, achieving **95%+ segmentation accuracy**
- Introduced **HSV-based chromatic thresholding** with **localized DT** and connected components, boosting robustness by **30%**
- Implemented **SLAM-based navigation** in ROS2 and **YOLO object detection** in Docker for real-time perception

### Mercor

Data Science Expert (Independent Contractor)

San Francisco, CA (Remote)

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

Data Scientist

Boulder, CO (US)

01/2025 - 05/2025

- Engineered SBA loan classifiers, achieving **87% accuracy** on ineligible applications to **streamline** loan approvals
- Reduced **false approvals** by **22%** using class rebalancing, early stopping, and learning rate tuning

### National Institute of Technology, Tiruchirappalli

Undergraduate Research Intern

Tiruchirappalli (India)

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

### Semantic Book Recommender | [GitHub]

05/2025 - 06/2025

- Crafted a **RAG-based** semantic retrieval pipeline using **HuggingFace sentence transformers + ChromaDB** for **5K+ books**
- Designed **context orchestration logic** for emotion tagging and genre clustering, **improving personalization** by **22%**
- Achieved **78% genre** (BART-MNLI) and **66% emotion** (DistilRoBERTa) accuracy, resolving cold-start issues for **1.2K+ books**

### 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%**

### TextSum | [GitHub]

09/2024 - 12/2024

- Built a novel **Bi-LSTM + dynamic attention model** for extractive summarization on **287k+ CNN/DailyMail articles**
- Designed NLP pipeline with **tokenization, embeddings, and positional encoding**, enabling efficient long-text processing
- Achieved **93.8% accuracy**, delivering high-fidelity summaries suitable for **legal document review and contract analysis**

### 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**