

# 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

<b>Void Robotics</b> Robotics Engineering Intern	<i>Marathon, FL (Remote)</i> 08/2025 - Present
<ul style="list-style-type: none"><li>Developed an advanced pill counter pipeline with <b>progressive K-Means</b>, achieving <b>95%+ segmentation accuracy</b></li><li>Introduced <b>HSV-based chromatic thresholding</b> with <b>localized DT</b> and connected components, boosting robustness by <b>30%</b></li><li>Implemented <b>SLAM-based navigation</b> in ROS2 and <b>YOLO object detection</b> in Docker for real-time perception</li></ul>	
<b>Mercor</b> Data Science Expert (Independent Contractor)	<i>San Francisco, CA (Remote)</i> 08/2025 - 09/2025
<ul style="list-style-type: none"><li>Executed <b>20+ ML</b> projects covering hypothesis testing, time series forecasting, and predictive modeling</li><li>Crafted prompt engineering workflows for <b>Gemini 2.5 Pro</b>, benchmarking reasoning with <b>custom rubrics</b></li><li>Developed evaluation frameworks to score the LLM on crafted rubrics, achieving <b>27% across all tasks</b></li></ul>	
<b>Parlay Finance</b> Data Scientist	<i>Boulder, CO (US)</i> 01/2025 - 05/2025
<ul style="list-style-type: none"><li>Engineered SBA loan classifiers, achieving <b>87% accuracy</b> on ineligible applications to <b>streamline</b> loan approvals</li><li>Reduced <b>false approvals</b> by <b>22%</b> using class rebalancing, early stopping, and learning rate tuning</li></ul>	
<b>National Institute of Technology, Tiruchirappalli</b> Undergraduate Research Intern	<i>Tiruchirappalli (India)</i> 06/2022 - 08/2022
<ul style="list-style-type: none"><li>Devised <b>ResCoWNet</b> CNN to denoise OCT scans, boosting <b>SSIM</b> by <b>12%</b> over baseline models for <b>clearer retinal imaging</b></li><li>Contributed to peer-reviewed research <b>cited 8+ times</b>, showcasing applied expertise in <b>medical AI and computer vision</b></li></ul>	

## TECHNICAL PROJECTS

<b>Semantic Book Recommender</b>   [GitHub]	05/2025 - 06/2025
<ul style="list-style-type: none"><li>Crafted a <b>RAG-based</b> semantic retrieval pipeline using <b>HuggingFace sentence transformers + ChromaDB</b> for <b>5K+ books</b></li><li>Designed <b>context orchestration logic</b> for emotion tagging and genre clustering, <b>improving personalization</b> by <b>22%</b></li><li>Achieved <b>78% genre</b> (BART-MNLI) and <b>66% emotion</b> (DistilRoBERTa) accuracy, resolving cold-start issues for <b>1.2K+ books</b></li></ul>	
<b>Privacy Enhancing ML for Job Interview Analysis</b>   [GitHub]	11/2024 - 12/2024
<ul style="list-style-type: none"><li>Classified job interview responses into four explanation levels for <b>38 interviewees</b>, analyzing linguistic patterns.</li><li>Delivered <b>79% balanced accuracy</b> for <b>Random Forest</b> (Over-explained vs. Comprehensive) and <b>66.67%</b> for <b>Conv1D</b> (Under-explained vs. Succinct).</li><li>Engineered <b>1500+ linguistic features</b> (TF-IDF, POS tagging, embeddings), reducing speaker bias by <b>30%</b></li></ul>	
<b>TextSum</b>   [GitHub]	09/2024 - 12/2024
<ul style="list-style-type: none"><li>Built a novel <b>Bi-LSTM + dynamic attention model</b> for extractive summarization on <b>287k+ CNN/DailyMail</b> articles</li><li>Designed NLP pipeline with <b>tokenization, embeddings, and positional encoding</b>, enabling efficient long-text processing</li><li>Achieved <b>93.8% accuracy</b>, delivering high-fidelity summaries suitable for <b>legal document review and contract analysis</b></li></ul>	
<b>Wine Quality Prediction with MLOps</b>   [GitHub]	10/2023 - 11/2023
<ul style="list-style-type: none"><li>Logged <b>30+ experiments</b> via <b>MLflow</b> UI for <b>ElasticNet</b> tuning using <b>AWS Sagemaker</b>, reducing training time by <b>27%</b></li><li>Deployed via <b>Docker</b> on <b>AWS EC2</b> with GitHub actions, delivering predictions in <b>under 3s</b> per input</li><li>Monitored <b>latency, drift, and availability</b> through <b>CloudWatch</b> dashboards and alerts</li></ul>	

## EDUCATION

<b>University of Colorado Boulder</b> Master of Science in Data Science	08/2023 - 05/2025 GPA: 3.73
<ul style="list-style-type: none"><li><b>Key Coursework:</b> Neural Networks and Deep Learning, Machine Learning, Big Data Analytics, Statistical Methods</li></ul>	
<b>National Institute of Technology, Tiruchirappalli</b> Bachelor of Technology in Electronics and Communication Engineering	07/2019 - 05/2023 GPA: 3.22
<ul style="list-style-type: none"><li><b>Key Coursework:</b> Linear Algebra and Calculus, Statistical Theory of Communication, Deep Learning</li></ul>	

## CERTIFICATION

- SnowPro Core Certification, Snowflake** - Aug 2025