Vishwas Vikas Kothari

Boulder, CO | +1 (720) 388-4856 | vishwasvkothari@gmail.com | LinkedIn | GitHub | Portfolio

Graduating May 2027 | Available June 2027

SUMMARY

M.S. Computer Science candidate at CU Boulder building and evaluating ML systems in Python: LLM fine-tuning, retrieval, document AI and credit modeling, each backed by statistical evaluation, tested code and reproducible pipelines.

EDUCATION

University of Colorado Boulder Aug 2025 – May 2027
Professional Master's in Computer Science, ***GPA: 3.74/4.0*** Boulder, CO
Coursework: Algorithms, Database Systems, Neural Networks & Deep Learning, Natural Language Processing.
Database coursework: Spark SQL over 1M records; Redis caching cut repeat-query latency from **176 to 20 ms** on PostgreSQL.

Vellore Institute of Technology Aug 2021 – Jun 2025
B.Tech in Information Technology, GPA: 8.49/10.0 Vellore, India

EXPERIENCE

Space Applications Centre, Indian Space Research Organisation (ISRO) Ahmedabad, India
Research Intern, Biological Oceanography Division Jan 2025 – Jul 2025

- **Retrieved chlorophyll-a** at log-space $\mathbf{R^2}$ **= 0.879** across 105 NASA NOMAD in-situ samples (reflectance fit $\mathbf{R^2}$ **= 0.987**) by building a Levenberg–Marquardt inversion of a semi-analytical bio-optical model.
- **Stabilized convergence** for parameters spanning three orders of magnitude by reparameterizing the optimization in log space, and benchmarked it against my own Particle Swarm Optimization implementation using RMSE, MAPD and bias.
- **Quantified OCM vs. MODIS-Aqua agreement** (pixel-level RMSE 0.29, bias −0.14) by automating SNAP reprojection, resampling, xarray/NetCDF merging and region-based error metrics; documented both studies in a 44-page report.

SELECTED PROJECTS

SemEval-2027 Task 3 | LLM Fine-Tuning for Word-Sense Detection GitHub

- **Reached 0.779 macro-F1** on Swedish, **+0.10 over zero-shot** with no Swedish training data, by LoRA fine-tuning Gemma 4 31B on 8,510 Russian/English sense-matching pairs and scoring the Yes/No logit margin.
- **Removed seed luck** from the final system: single runs ranged 0.724 to 0.774, so I averaged five adapters, gaining +0.012 (95% CI 0.005 to 0.021) under nested leave-one-word-out CV with predeclared success rules.

FinAgent-Eval | Document AI and Vision-Language Model Evaluation GitHub

- **Exposed a 4.8x gap** between Qwen2.5-VL-7B (0.738 ANLS) and the best legacy document model on 397 financial-document questions, yet Qwen solved only **10 of 65** arithmetic questions, by building a five-model benchmark with bootstrap CIs.
- **Re-credited 35 of 151** wrongly failed answers by rebuilding unit-aware grading ("$85,269 million" = "85,269"), then pinned every published number with 96 tests, a 13-label failure taxonomy and CI that fails on metric drift.

FinRAG Auditor | Hybrid Retrieval and Validated Tool Execution GitHub

- **Raised MRR from 51.6 to 61.3** over BM25 on 883 FinQA questions and 8,783 chunks by fusing BM25 and BGE dense retrieval with cross-encoder reranking; the BGE reranker beat TinyBERT by 10.4 points (95% CI 8.3 to 12.6).
- **Abstained on 10 of 10** source-withheld questions via a LangGraph workflow that validates citations and operands and computes with restricted Decimal arithmetic; fixed a wrong-company evidence bug in my evaluator; 114 tests, 93% coverage.

XAI Credit Lens | Credit-Default Modeling and Explainability GitHub

- **Reached 0.782 test ROC AUC** and 0.559 average precision on 30,000 credit-card records by tuning Random Forest, XGBoost and LightGBM with Optuna and stratified five-fold cross-validation.
- **Measured disparate impact** of 0.96 to 1.03 across four demographic attributes, above the 0.80 four-fifths threshold, and built a Streamlit prototype with SHAP explanations and what-if queries for reviewing decisions.

Implicit Bias of Adam vs. SGD | Optimization and Robustness GitHub

- **Confirmed optimizer geometry**: across five seeds, gradient descent converged to the L2 max-margin direction (cosine **0.990**) and Adam to the L-infinity one (**0.988**), checked against exact CVXPY solutions.
- **SGD lost 2.2x more** out-of-distribution accuracy than Adam (**30.2 vs. 13.6 points**) when a color patch's label correlation flipped from 99% to 1% in a five-seed MNIST CNN study, at near-identical in-distribution accuracy.

TECHNICAL SKILLS

Languages: Python, SQL, Java, C++, JavaScript, R, MATLAB
ML and LLMs: PyTorch, Hugging Face Transformers, PEFT/LoRA, scikit-learn, XGBoost, LightGBM, Optuna, LangGraph
Retrieval and Evaluation: RAG, LangChain, BM25, dense embeddings, reranking, nested CV, bootstrap intervals, SHAP, LIME
Data Systems: PostgreSQL, Redis, Apache Spark, PySpark, Spark SQL, Pandas, NumPy, xarray, NetCDF, Linux
Engineering: Git, Docker, GitHub Actions (CI), pytest, Ruff, mypy, Streamlit, CVXPY, SNAP

PUBLICATION, PEER REVIEW AND CERTIFICATION

Publication: Kothari, V. and Krishwanth, B. (2025). *Empowering Survivors: Ethical Artificial Intelligence for Countering Violence Against Women (VAW).* Springer.
Peer Reviewer: *Public Health*; *Social Sciences & Humanities Open* (Elsevier, 2026–Present).
Certification: AWS Certified Cloud Practitioner (2024).