Ankur Varma

United States | ankur.varma@colorado.edu | +1 (551) 359 1942 | Linkedin

EDUCATION

Masters of Science, Computer Science - University of Colorado, Boulder | GPA: 4.00/4.00 **Aug 2024 – May 2026**
Relevant Coursework: Design and Analysis of Algorithms, Deep Learning, Machine Learning, Foundations of Software Engineering, Geospatial ML
Bachelor of Technology (Gold Medalist), Indian Institute of Technology, Dhanbad | GPA: 3.96/4.00

SKILLS

Programming: Java, Python, Rust, Linux | **ML & AI**: PyTorch, sci-kit, Pandas, GPT, CNN, Transformer Models, GenAI, Reinforcement Learning | **Cloud/DevOps:** GCP, Docker | **Frameworks:** SpringBoot, Flask | **Databases**: MySQL, MongoDB | **Applications and Tools:** GitHub, IntelliJ

WORK EXPERIENCE

Teaching Assistant | Data Structures - University of Colorado, Boulder **Jan 2025 - May 2025**

- Teaching data abstractions (e.g., stacks, queues, lists, trees, graphs, heaps, hash tables, priority queues) and their representation techniques (e.g., linking, arrays). Introducing concepts used in algorithm design and analysis including criteria for selecting data structures to fit their applications
- Hold regular office hours to provide academic support and guidance to students

Course Manager | Linear Algebra with CS Apps- University of Colorado, Boulder **Aug 2024 - Dec 2024**

- Manage course logistics by maintaining and updating course materials on Canvas and Gradescope
- Collaborate with the professor and grading team to share grading responsibilities Assist in modifying and refining existing Python-based assignments to enhance their effectiveness

Executive Engineer – Coal India Limited, CCL, Ranchi **Nov 2017 – Jul 2024**

- Led the development of a **Spring Boot**–based job scheduling platform, increasing manufacturing throughput and boosting production output by **$20K/month** across distributed maintenance sites
- Achieved **90%+** test coverage through a robust **JUnit5** and **Mockito-based** testing suite, identifying 50+ defects pre-deployment, significantly reducing QA cycles and post-release bug reports

PROJECT

AI-Powered Resume Analyzer | Flask, PostgreSQL, GCP, Docker, GitHub Actions, Sentence Transformers

- Designed a scalable resume-job matching platform with **Flask** and **GCP Cloud Run**, supporting 100+ concurrent users, and integrated **PostgreSQL** for low-latency transactional queries
- Designed and deployed a pipeline using JobSpy to scrape job listings from multiple platforms and store structured job data via a Collector module. Enabled semantic search through Sentence Transformer-based embedding comparison, delivering **85%** top-5 match accuracy with **<200ms** average latency
- Built modular services for traffic analysis, email notifications (via Cloud Scheduler + Email Trigger + Email Sender), and compatibility scoring; automated end-to-end **CI/CD** workflows using **GitHub Actions**
- Handled resume uploads through **GCS** and built a job alert system that checks user opt-in, processes alerts per user, and sends personalized matches via email with a false positive rate below **10%**

Multimodal Geospatial ML for Climbing Area Prediction | PyTorch, ResNet50, SciBERT, SVM, Random Forest

- Constructed a multimodal ML pipeline by generating embeddings from Sentinel-2 RGB imagery using MoCo **ResNet50**, and MOSAIK (Empirical and Gaussian modes); DEM data processed with MOSAIK and flattening strategies; lithology text descriptions encoded using **SciBERT**
- Trained and benchmarked multiple classifiers (**2-layer MLP, SVM, Random Forest**) on all possible combinations of embeddings; applied ensemble techniques to boost robustness, achieving Precision: **0.92,** False Negative Rate: **0.09** across 5-fold CV
- Validated model predictions against Mountain Project database; **30%** of top-20 predictions corresponded to newly developed climbing sites post-2019, supporting model effectiveness in real-world discovery scenarios

Change Detection in Hyperspectral Imagery | PyTorch, U-Net, 3D CNN

- Designed and evaluated a deep learning pipeline for binary change detection using **2D FCNN U-Net** and **3D FCNN** architectures on hyperspectral imagery from the Onera Satellite Change Detection dataset
- Achieved a **53%** improvement in recall and a **0.12** gain in F1-score by introducing spectral-temporal awareness via **3D convolutions** over stacked multitemporal image cubes
- Built a preprocessing pipeline for class balancing and patch extraction, enabling more efficient **GPU** usage and improved performance on imbalanced geospatial datasets

Interview Outcomes Predictor | PyTorch, BERT, SHAP, LIME

- Developed multimodal models using language (sentiment, BERT) and prosodic audio features (pitch, energy) to predict interview ratings with **0.74** Pearson correlation and <18% relative error across 5-fold CV
- Trained **XGBoost Decision Tree** and **MLP** models on curated features, optimized with multiprocessing for **20×** faster data preprocessing and inference
- Applied **SHAP** for model interpretability, identifying key features in high/low performance predictions and benchmarking against quantized **LLaMA** models