Avigya Paudel

avigyapaudel045@gmail.com | linkedin.com/in/avigya-paudel | github.com/Avi161 | avigya.com

Education

Union College *Sept 2024 – June 2028 (expected)*
Bachelor of Science in Computer Science & Mathematics; Minor in Statistics Schenectady, NY

- **GPA:** Cumulative - 3.98/4.0 | Computer Science - 4.0/4.0 | Mathematics - 3.95/4.0
- **Honors:** Generation Google Scholar, Union Scholar, Dean's List, Uber Career Prep Fellow
- **Relevant Coursework:**
  - **Computer Science:** Object-Oriented Programming, Data Structures, Algorithm Design and Analysis
  - **Mathematics:** Probability Theory, Linear Algebra, Discrete Mathematics, Enriched Differential Vector Calculus (Honors), Integral Vector Calculus, Differential Equations, Logic and Set Theory, Geometry
- **Academic Activities:** ICPC Team Member, ACM Member, Mathematics Club Member, Robotics Club Member

Research Experience

Machine Learning Researcher *May 2025 – Sept 2025*
Algoverse Remote

- Implemented **Multi-Layer Activation Steering (MLAS)** on Gemma-2-2b-it using TransformerLens hooks across 26 residual layers, computing anti-sycophancy directions via PyTorch tensor operations. Reduced sycophancy from (78% → 0%) while restoring accuracy from (45% → 68%), with a team of 4 researchers.
- Co-Engineered a **Sparse Activation Fusion** pipeline by integrating HuggingFace Transformers with custom Sparse Autoencoders, applying prompt-conditioned bias subtraction in latent feature space. Reduced sycophancy (63% → 39%) and doubled accuracy when users held incorrect opinions.
- Developed an **inference-time intervention framework** with custom PyTorch hooks, batch processing optimizations, and evaluation on StrategyQA, GSM8K, MMLU, and SVAMP. Achieved training-free bias mitigation with only a 3.75% average accuracy trade-off.

Undergraduate Research Assistant *Jan 2025 – Present*
Union College Department of Computer Science Schenectady, NY

- Implemented **Strassen's algorithm** in Python using recursive divide-and-conquer, achieving an 8% empirical runtime improvement and reducing time complexity from $O(n^3)$ to $O(n^{2.81})$.
- Modeled 3×3 matrix multiplication as a **SAT problem** by translating ∼729 Brent equations into ∼130,000 CNF clauses, successfully re-deriving Laderman's 23-product algorithm with a custom NumPy-based tensor framework.
- Built a testing harness using Python's **multiprocessing** module to analyze solver performance through randomized exclusion experiments, logging unique solutions to JSON for validation.

Publications & Presentations

Publications

- **Adityo, N.*, Min, P. P.*, Paudel, A.*, Rufail, A.*, Zhu, A.*, Blondin, C., Zhu, K., Dev, S., & O'Brien, S.** (2025). Mitigating Sycophancy in Language Models via Sparse Activation Fusion and Multi-Layer Activation Steering. ***NeurIPS 2025 Workshop on Mechanistic Interpretability***. *Equal contribution. [PDF]

Presentations & Posters

- **Poster:** *Finding Tensor Rank*. Presented at the Union College Undergraduate Summer Research Poster Session, Lally Reading Room (Schaffer Library), Schenectady, NY (July 24, 2024). Co-authors: Shravani Kulkarni, Brodie Flaherty. Adviser: Dr. Matthew Anderson.

Technical Skills

Languages: Python, C/C++, Java, JavaScript/TypeScript, Swift, SQL
Web Development: HTML/CSS, Tailwind CSS, React, Node.js, Express.js, RESTful APIs, Flask
Tools & Technologies: Git/GitHub, Arduino, Docker, MongoDB, MySQL, PostgreSQL
Machine Learning & Data Science: PyTorch, NumPy, TensorFlow, Pandas, Matplotlib, Scikit-learn, Einops

Projects

FloodNet: Multi-Modal Disaster Prediction | GitHub | *PyTorch, CNN, Geospatial Data* Fall 2025

- Developed a multi-modal deep learning model using PyTorch to detect flood events by fusing Sentinel-1 (SAR) and Sentinel-2 (Optical) satellite imagery, achieving a 91.5% accuracy and 0.979 AUC.
- Engineered a custom preprocessing pipeline to handle disparate data distributions, implementing log-scaling for SAR and Min-Max scaling for Optical bands to ensure robust detection in cloud-covered regions.
- Optimized a 3-layer CNN architecture with dropout regularization and Adam optimization, resulting in a 92.7% recall rate for identifying high-risk disaster zones in the SEN12-FLOOD dataset.

WriteLight — Hackathon Project | GitHub | Live Demo Summer 2025

- Designed and implemented backend architecture for coin-based reward tracking and in-app item purchases using MongoDB schemas and RESTful API routes for user authentication.
- Collaborated with 10 teammates to ensure seamless API integration, resolved frontend/backend data flow issues, and contributed to project-wide decisions on structure, responsiveness, and developer tooling.

AI Summarizer & Translator | GitHub | Live Demo Winter 2024

- Developed a Chrome extension with GPT-4o API to summarize and translate web content across 10 languages, enhancing accessibility and readability for diverse users.
- Optimized performance cost by implementing secure token handling, client-side rate limiting, and a responsive Tailwind CSS interface.

Professional, Leadership & Service Experience

Calculus Help Tutor *Sept 2025 – Present*

Union College Calculus Help Center Schenectady, NY

- Led weekly study sessions for 200+ students across pre-calculus algebra, single-variable calculus, and vector calculus, tailoring explanations to individual learning styles.
- Developed Socratic questioning techniques to guide students toward independent problem-solving rather than providing direct answers, strengthening their mathematical intuition and conceptual understanding.
- Created visual aids and step-by-step breakdowns for complex topics like integration techniques, partial derivatives, and multivariable optimization to improve student retention.

ICPC Competitive Programming Organizer *Jan 2025 – Apr 2025*

Union College Schenectady, NY

- Organized ICPC (International Collegiate Programming Contest) weekly workshops for an 8-member team, focusing on solving specific algorithmic problems for each session.
- Designed practice sets with 3 in-session problems and 3 take-home problems, and facilitated group debugging sessions to strengthen contest readiness.

President, Union College Badminton Club *Jan 2025 – Present*

Union College Athletics Schenectady, NY

- Expanded the club to 30+ active members by organizing weekly practices and coordinating termly sign-up events.
- Planned and ran on-campus tournaments and mixed-doubles games, fostering an inclusive and competitive environment for players of all levels.

Co-Organizer & Volunteer *Jul 2024 – Aug 2024*

My Body is My Body Kathmandu, Nepal

- Led body safety workshops for 1,100+ students across 19 schools, using interactive songs, skits, and discussions to teach boundaries and prevention.
- Partnered with local teachers and staff to adapt materials culturally and foster student engagement on sensitive topics.

Web Development and Leadership Intern *June 2024 – Aug 2024*

Nobel Navigators Remote

- Co-Developed a WordPress site on the AI revolution, presented during a bi-weekly EXPO showcase to 65+ members.
- Created and managed web content, ensuring accessibility, responsive design, and presentation for a diverse audience.
- Co-facilitated 12 training sessions for 40+ interns on leadership and introductory web development, improving peer collaboration and presentation skills - resulting in each intern presenting their own website at the end of the program.