Vivek Verma

vivekverma@berkeley.edu | vivek3141 | vivek3141 | vcubingx | vcubingx

Education

University of California, Berkeley — August 2024 - May 2025
M.S. Computer Science — GPA: 4.00/4.00

University of California, Berkeley — August 2021 - May 2024
B.A. Computer Science, B.A. Applied Mathematics — GPA: 3.93/4.00

Relevant Coursework: Deep Reinforcement Learning†, Natural Language Processing†, LLM Agents†, Computer Vision†, Combinatorial Algorithms†, Randomized Algorithms†, Operating Systems, Machine Learning, Computer Graphics, Probability Theory, Numerical Analysis, Abstract Algebra, Abstract Linear Algebra, Real Analysis, Complex Analysis. Graduate Level†
Teaching: CS 162 (Operating Systems, 10-hour TA), DS-GA 1008 (Deep Learning @ NYU), Math 198 (Speedcubing).
Activities and Societies: Rubik's Cube Club @ Berkeley, Competitive Programming @ Berkeley, Cal Badminton.

Experience

Berkeley Artificial Intelligence Research (BAIR) — January 2022 – Present
Graduate Researcher at Berkeley NLP Group, Advisors: Nicholas Tomlin, Dan Klein — Berkeley, CA
- Designed state-of-the-art structured search classifier for AI-generated text, achieving 99.1 F1, improving baselines by 5.9 F1.
- Implemented and created full-stack web application demo for classifier, grew to 40,000+ users and published BAIR article.
- Researched document-level entropy trends using LLMs, implementing sliding window algorithm for GPT-2 evaluation.

Databricks — May 2024 – August 2024
Software Engineering Intern — San Francisco, CA
- Designed and implemented LLM-based search overview, integrating with the Databricks Assistant to support follow-ups.
- Formulated synthetic dataset loop for automated continuous improvement of Assistant responses based off user feedback.
- Created production-scale endpoints for query classification, Assistant response generation and suggesting follow-up questions.

Waymo — May 2023 – August 2023
Software Engineering Intern — Mountain View, CA
- Designed and implemented trajectory-based optimization method for computing probability of an event being safety-relevant.
- Used method to design classifier that automates human annotation at 91% accuracy, reducing costs by up to 500,000%.
- Implemented algorithm and analyzer in C++, allowing deployment to large simulation jobs with 2,500,000+ miles of data.

Google — August 2022 – December 2022
Software Engineering Intern — Sunnyvale, CA
- Designed variance-weighted linear regression algorithm for multi-source clock synchronization in distributed systems.
- Achieved 2x speedup over traditional averaging methods, wrote simulations and implemented algorithm to verify results.
- To be integrated into Google Cloud clusters to improve clock latency helping with expansion to large Hospitals/Banks.

3blue1brown — June 2021 – August 2021
Content Intern — San Mateo, CA
- Created interactive math lessons on Fourier Series, Partial Differential Equations, Riemann Hypothesis on 3blue1brown.com.

Projects

Math Content Creator on YouTube | *Python, GLSL, OpenGL, PyTorch, NumPy, Cairo, Manim*
- Programmatically created 25+ explanatory math videos in OpenGL/Python that visualize concepts from Deep Learning, Natural Language Processing, Complex Analysis, Graph Theory, etc. with 90,000 subscribers and 3,500,000 views.
- Videos utilized by courses at UC Berkeley, Stanford and NYU; 30+ universities across 7 countries.

ML-Python | *Python, C++, TensorFlow, Keras, NumPy, Matplotlib, PyPi*
- Created high-level python library with 100,000+ downloads for common ML algorithms such as CNNs and Deep Q-Learning.
- Implemented visualizations for training process and optimized gradient descent with C++ extensions for Python.

Publications

- **Vivek Verma**, Eve Fleisig, Nicholas Tomlin and Dan Klein. Ghostbuster: Detecting Text Ghostwritten by Large Language Models. *In Proceedings of the Association of Computational Linguistics: NAACL 2024*. https://arxiv.org/abs/2305.15047
- **Vivek Verma***, Nicholas Tomlin*, and Dan Klein. Revisiting Entropy Rate Constancy in Text. *In Findings of the Association of Computational Linguistics: EMNLP 2023*. https://arxiv.org/abs/2305.12084

Miscellaneous

Awards: Google Code Jam (Top 500), International Collegiate Programming Contest/ACM ICPC (1st Place Division 2 2022, Top 15 Division 1 2023, Top 15 Division 1 2024), Steven H. Strogatz Prize for Math Communication.
Languages/Frameworks/Tools: Python, C/C++, Java, Scala, SQL, PyTorch, Huggingface, Docker, Google Cloud.