

# Vivek Verma

✉ vivekverma@berkeley.edu | [in](#) vivek3141 | [G](#) vivek3141 | [v](#) vcubingx | [t](#) vcubingx

## EDUCATION

### University of California, Berkeley

August 2021 - May 2025

B.A. Computer Science, B.A. Applied Mathematics

GPA: 4.00/4.00 (Major), 3.92/4.00 (Overall)

**Relevant Coursework:** Combinatorial Algorithms and Data Structures<sup>†</sup>, Deep Reinforcement Learning<sup>†\*</sup>, Natural Language Processing<sup>†</sup>, Operating Systems, Probability Theory, Abstract Algebra, Complex Analysis\* In Progress\* Graduate Level<sup>†</sup>

**Honors:** Upsilon Pi Epsilon (top 33% of CS students), Dean's List (top 10% of undergraduates).

**Teaching:** CS 162 (Operating Systems), DS-GA 1008 (Deep Learning @ NYU), Math 198 (Speedcubing).

**Activities and Societies:** Rubik's Cube Club @ Berkeley, Competitive Programming @ Berkeley, Cal Badminton.

## EXPERIENCE

### Waymo

May 2023 – August 2023

*Software Engineering Intern, Simulation*

Mountain View, CA

- Designed and implemented trajectory-based optimization method for computing probability of an event being safety-relevant.
- Implemented algorithm and analyzer in C++, to be used in large-scale simulation jobs with 2,500,000+ miles of data.
- Used classifier to automate human annotation, to be deployed in production to save costs, presented method and results.

### Berkeley Artificial Intelligence Research (BAIR)

January 2022 – Present

*Undergraduate 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 32.7 F1.
- Implemented and created production-scale web application demo for classifier, published BAIR article with 100,000+ viewers.
- Researched document-level entropy trends in LLMs, implementing algorithm for 32x speedup of GPT-2 text generation.

### Google

August 2022 – December 2022

*Software Engineering Intern, Google Cloud*

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 using next.js, p5.js and react.js on Fourier Series, Partial Differential Equations, Riemann Zeta Function and Differentials, published on 3blue1brown.com, with 5,000,000 subscribers and 300,000,000 views.

## 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 Complex Analysis, Multivariable Calculus, Fractional Calculus, Measure Theory and Graph Theory.
- 75,000 subscribers, 2,500,000 views, 125,000 hours of watch time and 25,000,000 impressions.
- 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. *Pending, International Conference on Learning Representations (ICLR) 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>

## HONORS/AWARDS

**1st Place (out of ~85 teams), ACM ICPC Pacific Northwest, Division 2**

2022

**Round 2 Qualifier (top 500 out of ~20,000), Google Code Jam**

2022

**Top 15, ACM ICPC Pacific Northwest, Division 1**

2023

## SKILLS

**Areas:** NLP, Reinforcement Learning, Deep Learning, Machine Learning, Back-end Web Development, Statistics.

**Languages/Frameworks/Tools:** Python, C/C++, Java, GLSL, SQL, PyTorch, Huggingface, TensorFlow, Flask, NumPy, OpenGL, OpenCV, Matplotlib, Manim, Jupyter, Git, Docker, Vim, Google Cloud.