

# ETHAN VERTAL

✉ [ethan.vertal@gmail.com](mailto:ethan.vertal@gmail.com) [in](#) [ethan-vertical](#) ☎ 571.455.0383 📁 Portfolio: [ethanvert.github.io](https://ethanvert.github.io) 📍 Boston, MA

## EDUCATION

### B.S. in Statistics and Machine Learning

Carnegie Mellon University

Aug 2020 - May 2024

Pittsburgh, PA

- **Graduated with University Honors**
- **Relevant Coursework:** Generative AI, Machine Learning with Large Datasets, Intermediate Deep Learning, Statistical Computing, Statistical Methods in Finance, Advanced Methods of Data Analysis, Algorithms & Advanced Data Structures, AI Problem Solving & Representation

## EXPERIENCE

### Outlier

Reviewer (AI Model Training)

Remote

Jun 2024 - Present

- Review and validate complex mathematics prompts for correctness, complexity, and LaTeX formatting, ensuring strict adherence to defined guidelines.
- Contribute to AI model training as part of Reinforcement Learning with Human Feedback (RLHF), grading AI agents' responses to math problems ranging from basic algebra to discrete mathematics.
- Enhance technical communication by reviewing and formatting prompts, playing a key role in model training for state-of-the-art AI solutions.

Attempter (AI Model Training)

Mar - Jun 2024

- Created and tested collegiate-level math problems and answers, evaluating model responses for accuracy and logical reasoning.
- Identified gaps in LLMs' mathematical reasoning, contributing to the development of next-generation reasoning models.

### Carnegie Mellon University

Research Apprentice

Pittsburgh, PA

Jun - Aug 2023

- Led team to design a feasibility report and scalable plan for parallelizing a simulation software on a customized AWS-backed cluster.
- Researched deploying NetLogo simulations on AWS using EC2, S3 Buckets, and Slurm.

### CrafterCMS

Machine Learning Intern

Reston, VA

Jun - Aug 2022

- Developed a TensorFlow-based recommendation system web app, enhancing machine learning model performance to align with company product goals.

Software Engineering Intern

Jun - Aug 2021

- Delivered an Alexa API web application in Groovy, increasing interest and improving customer engagement for enterprise clients such as Papa John's, Marriott, and more.
- Brought product from proof-of-concept to deployment in Ubuntu, using Git/GitHub for version control.

PROJECTS

---

Presidential RAG

- Ideated, developed, and deployed a Retrieval-Augmented Generation (RAG) Agent, powered by LLaMa-3.2-1B Instruct, that references a ChromaDB vector database of a selection of speeches, articles, and literature delivered by or about the Presidents of the United States to answer user questions.
- Web-scraped all data with BeautifulSoup and Selenium for Chrome interactions.
- Explored data in Pandas and Seaborn; cleaned data with regular expressions and string manipulation.
- Experimented with models and storage options on AWS SageMaker, utilized AWS Bedrock model hub to build foundation, and hosted frontend on AWS EC2.

A New Method of Generating Images

- Spearheaded a project team in designing and developing a novel score-based diffusion algorithm for image generation using PyTorch, focusing on innovating machine learning infrastructure and algorithms.
- Documented comprehensive project details and milestones, ensuring clear communication across all phases of the development life cycle.

Analyzing Movie Reviews with GPT2

- Enhanced the sentiment analysis capabilities of GPT2-medium using HuggingFace and PyTorch APIs, improving model performance for natural language processing (NLP) tasks.
- Implemented a custom fine-tuning loop and Low-Rank Adaptor (LoRA) class, demonstrating a mathematical understanding of LLMs and NLP.

AWARDS & ACTIVITIES

---

- **Carnegie Mellon Men’s Soccer Team (2020-2023):** Dedicated 20+ hours/week while balancing a full academic workload. Earned UAA All-Academic and UAA Offensive Player of the Week Awards.
- **Student Athlete Academic Award (2022-2023):** Achieved highest GPA in respective college as a rising fourth-year in the Athletic Department.

SKILLS

---

Languages	Python, R, PostgreSQL, C, Java, Groovy
Tools & Frameworks	PyTorch, LangChain (TextSplitters, Chains, LCEL), PySpark, AWS (SageMaker, Bedrock, EC2, S3, EMR, Alexa, Lambda), Git/GitHub, HuggingFace, NumPy, Matplotlib, Pandas, Distributed Training (DataParallel, ModelParallel), Model Quantization (Quanto, BitsAndBytes), Hyperparameter Tuning (Optuna, GridSearchCV, Hyperopt), Weights & Biases, Linux/Unix