

Vishruth Veerendranath

[linkedin.com/in/vishruth-v](https://www.linkedin.com/in/vishruth-v) | vishruth@cmu.edu | (878) 834-9199 | [vishruth-v.github.io](https://github.com/vishruth-v) | [Google Scholar](#)

EDUCATION

Carnegie Mellon University – School of Computer Science

Master of Science in Intelligent Information Systems (Program Focus on **ML & NLP**). **GPA: 4.15/4.0**
Coursework: Advanced NLP (PhD), On-Device ML (PhD), Visual Learning (PhD), Multimodal ML (PhD)

Pittsburgh, PA
Dec 2024

PES University

Bachelor of Technology in Computer Science and Engineering; **CGPA: 9.81/10 (Rank 5/1000)**

Bangalore, India
May 2023

SKILLS

Programming Languages: Python (Advanced), C/C++ (Proficient), JavaScript (Moderate), Java (Moderate), R (Basic)
Machine Learning: PyTorch, TensorFlow, HuggingFace, Keras, vLLM, NLTK, Scikit-Learn, llama.cpp, LLMFoundry
Cloud/Web/Big Data: PySpark, FAISS, Dagster, Flyte, MongoDB, Express, React, Node.js, Flask, SQL, AWS, Azure, Hadoop

EXPERIENCE

Datology AI

Research Intern

Redwood City, CA
May 2022 – Jul 2023

- Curating high-quality datasets by filtering noisy data using small models, leading to **30%** training speedup (tokens).
- Sped up LLM training & evaluation stack by **40%** & improved iteration time in **pretraining 1-3B** parameter LLMs.

Intel Corporation

Research Engineer Intern

Bangalore, India
Dec 2022 – Jul 2023

- Accelerated EfficientNet model training by **2x** & inference by **5x** by integrating **FAISS** Nearest Neighbor Search optimizations with **TensorFlow** to optimize softmax classification heads with **50k+** classes (tailored to CPU).
- Boosted accuracy **10%** by designing experiments and data augmentation strategies during pilot phase in first store.

Indian Institute of Science

Research Intern

Bangalore, India
Jun 2022 – Aug 2022

- Designed novel rewards to train **RL Agents** for autonomous driving. Improved comfort & safety of by **51%** (**paper**)

Microsoft

Student Developer

Bangalore, India
Jun 2021 – Aug 2021

- Leveraged Azure ML to automate bank's onboarding processes, reducing errors & customer journey time by **60%**.

SELECTED PUBLICATIONS

ECCO: Can we improve model-generated code efficiency without sacrificing functional correctness? **EMNLP 2024*** [\[LINK\]](#)

- Benchmark to evaluate LLMs' ability to generate code with a focus on *runtime, space efficiency + correctness*.
- Adapting LLMs with iterative code revision and finetuning using NL and execution feedback of all three parameters.

Pre-Calc: Learning to Use the Calculator Improves Numeracy in Language Models. **ICML 2024**, AI4Math workshop [\[LINK\]](#)

- Improving numerical reasoning abilities of encoder-based LMs using a tool-use (calculator) training objective.

ScriptONES: Sentiment-Conditioned Music Generation for Movie Scripts. [\[LINK\]](#)

AIMLSys 2023. Presented at ML for Audio workshop, **NeurIPS 2023**

XLDA: Linear Discriminant Analysis for Scaling Continual Learning to Extreme Classification at the Edge. [\[LINK\]](#)

PAC-Bayes meets Interactive Learning workshop, **ICML 2023** (Work at Intel)

C3PO: A Lightweight Copying Mechanism for Translating Pseudocode to Code. **ACL-IJCNLP 2022** (SRW) [\[LINK\]](#)

GraphCoReg: Co-training for Regression on Temporal Graphs. [\[LINK\]](#)

Best Student Paper Award at the 18th Mining and Learning on Graphs (MLG) workshop, **ECML-PKDD 2022**.

Multivariate Covid-19 Forecasting with Vaccinations as a Factor: the case of India and USA. [\[LINK\]](#)

IEEE TENSYP 2022. Presented at the Healthcare AI and Covid-19 workshop, **ICML 2022**

PROJECTS

Multimodal Web Navigation Agent: Used image-code captioning to improve visual web navigation abilities of LLM agents.

On-Device Math Chatbot: Quantized LLM inference, improved math reasoning & reduced hallucination with tool-use.

Multi-Task Diffusion Model: Jointly training on diffusion + discriminative pixel-level image tasks using special timesteps.