Phi Nguyen

US Citizen | (224) 435 8972 | phnguyen@andrew.cmu.edu | linkedin.com/in/phi-nguyen-h | github.com/Ph1so

EDUCATION

Carnegie Mellon University Pittsburgh, PA
B.S., Robotics - GPA: 4.0/4.0 Expected: May 2028

WORK EXPERIENCE

Sanofi Framingham, MA
Machine Learning Engineer Intern *Jun. 2025 - Present*

- Developing **machine learning** model for drug response prediction in **pharmacology** using **PyTorch**
- Cut drug manufacturing costs by **15%** by training a closed-loop multi-robot **Bayes Opt (BoTorch)** system
- Saved **20+ hours** and **$10K+** in reagent cost by storing checkpoints in **SQL** to resume **Bayes Opt** training
- Built REST API **(Spring Boot)** to interface robot **Java** software and **Python** script; tested API via **Postman**

NASA Merritt Island, FL
Software Engineer Intern *Jan. 2025 - May 2025*

- Deployed fine-tuned computer vision model with **95%** accuracy on **HoloLens** using **Azure** for **AR** tours
- Boosted model accuracy by **10%** by building image augmentation scripts with **OpenCV** for data diversity
- Cut prototyping costs by **$100K+** annually by developing custom **Unreal Engine (UE5)** software for hand object manipulation in virtual simulations; now standardized at **10+ VR** simulation teams across NASA
- Co-inventor on provisional patent for software framework; authored peer-reviewed paper at **IEEE SMC-IT**
- Networked **Vicon** server to stream real-time data to **UE** clients for simulation; constructed mocap volume

RESEARCH EXPERIENCE

MIT - Lippman Lab Cambridge, MA
Undergraduate Research Intern *May 2025 – Present*

- Improved forecasting model accuracy by **10%** by revamping scraping pipeline with **Selenium** and **OpenAI**
- Partnered with **3+** firms (e.g., Deloitte, HP) using ML to identify data overlaps and optimize monetization
- Saved **10+** hours weekly across team by containerizing dev/prod environments with **Docker**

Stanford University - Amin Lab Stanford, CA
Undergraduate Research Intern *Nov. 2024 - Jan. 2025*

- Designed and trained **CUDA**-accelerated deep neural network inspired by SpliceAI using **Tensorflow**; achieved **99.7%** accuracy and deployed in drug discovery research
- Collaborated with engineers to build a scalable data processing pipeline on **AWS Batch** using **S3** and **EC2**
- Cut resource costs by **20%** by refactoring **Nextflow** pipelines to parallelize processing across **3PB+** of data
- Developed **R** scripts to interface **Ensembl's MySQL** server and construct mappings for data quantification

MIT - Esvelt Lab Cambridge, MA
Undergraduate Research Intern *Apr. 2024 - Dec. 2024*

- Built agentic AI system with **LangChain** for natural language control of robots; presented at **MIT BioMAN**
- Built **open-source Python** software used by **2K+** researchers at **10+** companies to control **wet-lab robots**
- Cut latency by **30%** by reverse-engineering robot firmware via **Wireshark**; built Python SDK with pylibftdi

PROJECT EXPERIENCE

Feedback Categorizer for Church: Text, Speech, and Image Recognition - GitHub Jul. 2024
Technologies: React, Next.js, Firebase, OpenAI API, Whisper, Hugging Face

- Built a full-stack app to help my church automatically sort community questions and suggestions (text, voice, or images) into custom categories; saves volunteers' time each week by reducing manual sorting

TECHNICAL SKILLS

Languages: Python, C++, C, Java, JavaScript, HTML/CSS, SQL, R, Nextflow
Software: Git, Linux, AWS, TensorFlow, BoTorch, LangChain, CUDA, Azure, Docker, Unreal Engine, React, Next.js, Spring Boot, Flask, PostgreSQL, MySQL, Postman, Blender, Wireshark
Hardware: Vicon, HoloLens, Oculus, HTC Vive