Phillip Huang

US Citizen

philliphjhuang@gmail.com | github.com/philliphjhuang | linkedin.com/in/phillip-hj-huang | philliphjhuang.github.io

EDUCATION

Master of Science in Computer Science, **University of Southern California** **Aug. 2024 – May 2026 (expected)**
Relevant Courses: *Artificial Intelligence, Deep Learning, Natural Language Processing* *Los Angeles, CA*

Bachelor of Science in Computer Science, **Rutgers University** **Sep. 2020 – Jan. 2024**
Relevant Courses: *Data Structures, Algorithms, Software Engineering* *New Brunswick, NJ*

SKILLS

Programming: Python, Java, CUDA, C, C++
Frameworks: PyTorch, TensorFlow, NumPy, Pandas, Matplotlib, MITRE Caldera
Machine Learning: NLP, Large Language Models (LLMs), Prompt Engineering, Model Training, Evaluation
Production Tools: AWS, Docker
Spoken Languages: English, Mandarin Chinese

WORK EXPERIENCE

NSIP Masters Intern, **Pacific Northwest National Laboratory** | *Richland, WA (Remote)* **May 2025 – Aug. 2025**

- Developed an end-to-end LLM agent using Claude 3.7 Sonnet and MITRE Caldera on AWS, translating natural language into executable cyber-attacks which reduced implementation time for cybersecurity engineers from weeks to minutes.
- Led evaluation benchmark design by automating dataset generation and validating the agent's output (generated cyber-attacks), where the agent achieved a 73% task completion rate and 80% self-evaluation accuracy in simulated cyber environments.

AI Intern, **Sorenson Communications** | *Salt Lake City, UT (Remote)* **Jun. 2024 – Dec. 2024**

- Implemented DistilBERT and Thutmose Tagger to add punctuation/capitalization and inverse text normalization to ASR datasets to establish a baseline accuracy of 50%.
- Improved punctuation/capitalization and ITN accuracy to 65% on ASR datasets using LLMs (Llama3.1 8B, Gemma 2 9B), contributing to real-time transcription post-processing.

AI Engineer Intern, **Reality AI Lab** | *New York, NY (Remote)* **Feb. 2024 – May 2024**

- Developed a RAG-based chatbot in Streamlit integrated with Google's Gemini LLM to enable real-time LLM interaction for users.
- Built an AI-generated quiz tool for user-provided PDFs with instant feedback/explanations using Gemini to reinforce understanding.

Algorithm Intern, **OmniVision Technologies Inc.** | *Santa Clara, CA* **May 2022 – Aug. 2022**

- Implemented Gaussian blur, sharpenening filters, and demosaic algorithms on Bayer images to improve color reconstruction and reduce artifacts.
- Enhanced image brightness using gamma correction and auto white balance to ensure accurate color reproduction.

RESEARCH EXPERIENCE

Graduate Student Researcher, **USC FORTIS Lab** | *Los Angeles, CA* **Apr. 2025 – Present**

- Used AWS Bedrock to prompt a range of hosted LLMs — Meta's Llama 3.2 (1B, 3B, 11B, 90B), Llama 3.0 (8B, 70B), Mistral 2, and Mixtral 2 — for large-scale simulation of political survey responses using persona-driven prompts.
- Executed controlled experiments to test the effects of question order and demographic context on response bias, using LLM personas to simulate realistic human behavior in survey settings.
- Benchmarked model outputs against real survey data using the inconsistency rate as the primary metric, comparing performance across demographic groups (sex, age, education, race, ethnicity, and income).

PROJECTS

Named Entity Recognition with Bidirectional LSTM **Apr. 2025**

- Trained a bidirectional LSTM model in PyTorch for Named Entity Recognition on the CoNLL-2003 dataset.
- Achieved 77% F1 score on the dev dataset using standard word embeddings.
- Enhanced model with case-sensitive GloVe embeddings, resolving capitalization conflicts and improving F1 score to 88%.