THEO ZHANG

(408) 713-8400 | theo.zhang@ucla.edu | https://theozhangg.github.io/ | https://www.linkedin.com/in/theo-zhang/

EDUCATION

University of California, Los Angeles (UCLA)
B.S. in Computer Science, GPA: 3.73/4.00 *Expected June 2025*

- Upsilon Pi Epsilon Computer Science Honor Society Member (2022-present)

Cornell University
Online Certificate in Machine Learning Foundations *August 2023*

EXPERIENCE

Speech and Cognitive Development Lab at UCLA **Los Angeles, CA**
Research Assistant | Advisor: Assistant Professor Meg Cychosz *January 2024 – Present*

- Processed and curated large-scale speech datasets (53K+ samples) from a novel 200K+ sample corpus, ensuring balanced and de-aggregated training, development, and testing sets across key factors (labels, ages, languages, environments)
- Developed deep learning classification models (PyTorch, TensorFlow) to analyze infant speech development and assess speech maturity levels
- Evaluated the effectiveness of different model architectures for classifying children's speech maturity level
- Co-first author on a paper submitted to *Interspeech 2025*, currently under review

UCSF Archives and Special Collections **San Francisco, CA**
Senior Data Science Intern and Summer Research Fellow | Advisor: Lisa Nguyen *June 2024 – September 2024*

- Evaluated the impact of different optical character recognition (OCR) tools on the digital archival process, including potential biases and silences introduced into datasets that were created with OCR
- Independently authored paper on the impact of OCR biases specific to handwritten documents
- Developed pipelines and datasets for the UCSF Digital Health Humanities Institute to be used in workshops and classes

GumGum **Los Angeles, CA**
ML/AI Engineering Intern *August 2023 – December 2023*

- Designed sound classification deep learning models (PyTorch) for context-driven ad placement in video-based media
- Expanded GumGum's existing audio processing pipeline to increase product use cases beyond pure speech classification

Athenic AI **San Francisco, CA**
Software Engineering Intern *June 2022 – September 2022*

- Developed an SMS/email notification and file upload system (React.js), improving customer insights in AI-driven analytics
- Refactored frontend codebase, optimizing backend performance by 50% and increasing readability for documentation
- Resolved over 30 frontend and backend tickets to improve user experience and efficiency of product development

LEADERSHIP

IDEA Hacks 2024 Director **Los Angeles, CA**

- Director of the largest hardware hackathon on the West Coast with an attendance of 230+ participants
- Led a team of 20 to secure $40k in grants and corporate sponsors, manage parts and participants database, create website, and manage event logistics such as venue, judging, catering, workshops, and overseeing over 50 volunteers

IEEE at UCLA **Los Angeles, CA**

- **President 2024-2025**: Chair of the largest IEEE student branch in the U.S. with 400+ members, directed a team of 28 officers, raised $50k (75% increase from the previous year) in funding from UCLA, alumni, and corporations, managed long-term goals and new initiatives, and served as the point of contact for school-wide and IEEE Regional and National
- **External Vice President 2023-2024**: Interfaced with IEEE nationally and managed external club relations
- **Secretary 2022-2023**: Published weekly newsletter to ~850 members, led the Alumni Mentorship Program for 66 members

PUBLICATIONS

Employing self-supervised learning models for cross-linguistic child speech maturity classification *Under Review, 2025*
*Madurya Suresh, **Theo Zhang**, Anne Warlaumont, Kasia Hitczenko, Alejandrina Cristia, Margaret Cychosz*

Silence in OCR: What could handwritten documents tell us? *UCSF eScholarship, 2025*
Theo Zhang

SKILLS

- **Languages**: C/C++, Python, JavaScript
- **ML Tools**: scikit-learn, NumPy, pandas, PyTorch, TensorFlow, Keras, matplotlib
- **Other**: Git, Linux, React.js, Arduino