Kyi Lei Aye

Alhambra, CA, 91801 | (626)-558-5766 | ayekyilei2003@gmail.com | LinkedIn | Github

RESEARCH INTEREST

I am a junior Computer Science student at University of California, Irvine. I have a strong interest in advancing multisensory agentic AI systems and interpreting their complex reasoning with linguistic tools to improve model intelligence and support reliable decision-making in high-stakes domains.

EDUCATION

University of California, Irvine **Sept 2025 - June 2027**

B.S. in Computer Science, Specialization in Intelligent Systems | *GPA* - 4.0/4.0

College of San Mateo **Aug 2022 - May 2025**

A.S. in Computer Science, Math, Physics | *GPA* - 3.94/4.0

RELEVANT COURSES

Object Oriented Programming, Data Structures & Algorithms, Algorithm Designs & Analysis, Information Retrieval, Intro to AI, Machine Learning Specialization, Multivariable Calculus, Linear Algebra, Differential Equations, Discrete Math, Probability & Statistics, Mechanics, Electricity & Magnetism, Thermodynamics, Modern Physics.

RESEARCH EXPERIENCE

Stanford University School of Medicine | Machine Learning Research Intern **Sept 2024 - Aug 2025**

- Fine-tuned a BEiT-BERT based pathology vision-language model on **~6K** image pairs with Pytorch and knowledge distillation, increasing model accuracy rate by **72%** in tissue classification in grayscale cancer microscopy image domain.
- Utilized text and image embeddings with OpenAI, Claude AI, DeepSeek AI, CONCH, and MUSK (Pathology vision-language) models to analyze pathology images, achieving **~85%** accuracy rate in skin tissue recognition.
- Implemented a UNI self-supervised image encoder pipeline in Tensorflow with a cosine similarity technique for pixel-level cancer classification, achieving **70%** precision in identifying cancer areas in pathology images.

Stanford Inclusive Mentoring in Data Science | Research Scholar **Dec 2024 - March 2025**

- Built a potential architecture that includes pseudo-labeling techniques and weakly supervised learning models for research pipeline to analyze label-free pathology images.
- Utilized GPT-4V model with In Context Learning technique to conduct pseudo-labeling of images and evaluated its consistency and accuracy.

College of San Mateo Honors Project Program | Honors Research Scholar **Aug 2023 - Dec 2024**

- Collected survey data from **100+** students on campus and conducted data cleaning and preprocessing using Pandas, Numpy, and Matplotlib.
- Performed analysis on relationship between community college students' job searches, academic activities, and internship attainments using a Supervised Regression Machine Learning model, exploring effective job searching techniques.

JOB EXPERIENCE

Cadence Design System | Machine Learning Engineer Fellow **Sept 2025 - Dec 2025**

- Developed an end-to-end preprocessing pipeline, including tokenization and noise reduction, for **48M+** Amazon product reviews using NLTK and Pandas, enhancing data quality for model training.
- Fine-tuned a BERT-Base Cased model on cleaned reviews with Hugging Face Transformers to extract actionable insights for products, achieving **87%** accuracy in multi-class sentiment analysis (positive, neutral, and negative).
- Developed an Aspect-Based Sentiment Analysis pipeline with PyABSA to evaluate sentiment analysis on product reviews at feature level, revealing which features in products are critical to redesign.

Uber | Software Engineer Fellow **Jan 2025 - Aug 2025**

- Integrated a real-time mobile object detection feature in an inventory app using Tensorflow.js and Firebase, reducing manual data entry workload by **20%** and enhancing user satisfaction and experience.
- Built and tested a Retrieval Augmented Generation (RAG) pipeline with Pinecone database and Langchain to upgrade AI-powered chatbot, improving response relevancy and accuracy rate by **35%**.

Mission Cloud | Machine Learning Engineer Intern **Jan 2025 - Aug 2025**

- Built a custom Convoluted Neural Network (CNN) model on **3000** images using Tqdm, Tensorflow, Keras, Pycocotools, and Ffmpeg frameworks.
- Integrated 2D convolutional layers and dense neural network layers in the CNN model, enhancing precision and accuracy of the model by **15%**.

LEADERSHIP

Microsoft Teals Student Chapter | Founder **July 2024 - May 2025**

- Mentored **200+** K-12 students with AP science courses, enhancing students' average test scores by **15%**.
- Delivered 1-1 and small group tutoring sessions, which increased students' assignment completion rate by **20%**.
- Raised funds for the chapter by conducting outreach events and leading community engagement initiatives.

Girls Who Code Club, **San Mateo County** | Vice President **Jan 2024 - May 2025**

- Grew club membership by **60+** through social media contents, flyers, tech speaker events, and club activities.
- Led group projects, increasing members' participation in projects by **30%** through effective progress tracking.

PUBLICATIONS

- D Barash, Y Winetraub, E Manning, A Van Vleck, O Hirsch, **Kyi Lei Aye**, "Leveraging Computational Pathology AI for Noninvasive Optical Imaging Analysis Without Retraining", Nov 2024. [arXiv] | [Code]

CONFERENCE AWARDS & PRESENTATION

- Presentation Award in Computational & System Biology: "Analyzing Pathology Images with Large Language Models & Imaging Models", the American Society for Microbiology's (ASM) conference (Nov 2025).
- Oral Presentation: "Developing a Vision-Language Model for Optical Coherence Tomography Grayscale Images in Non-Invasive Cancer Diagnosis", Southern California Undergraduate Research Conference (Nov 2025).
- Poster Presentation: "Developing a Vision-Language Model for Optical Coherence Tomography Grayscale Images in Non-Invasive Cancer Diagnosis", MIT IEEE Research Technology Conference (Oct 2025).
- Poster Presentation: "Developing a Vision-Language Model for Optical Coherence Tomography Grayscale Images in Non-Invasive Cancer Diagnosis", Stanford Bio-X Interdisciplinary Initiatives Seed Grants Symposium (May 2025).
- Oral Presentation: "Analyzing Pathology Images with Large Language Models & Imaging Models", Northern California Undergraduate Research Conference at San Jose State University (May 2025).
- Third Prize in Physical & Computational Sciences: "Analyzing Pathology Images with Large Language Models & Imaging Models", Stanford Undergraduate Research Conference (April, 2025).
- Oral Presentation: "Analyzing Pathology Images with Large Language Models & Imaging Models", UC Berkeley Research Symposium (April 2025).
- Oral Presentation: "A Multiple Linear Regression Analysis: Identifying the Most Influential Factors in Landing Internship Opportunities for Community College Students", College of San Mateo Honors Seminar (May 2024).

HONORS & AWARDS

- Google Generation Scholarship Recipient (selected as 1 out of 50 in North America) - Google (Oct 2025)
- Undergraduate Transfer Scholarship Recipient - Pursuit of Excellence Foundation (May 2025)
- Undergraduate National Transfer Scholarship Semi-Finalist - Jack Kent Cooke Foundation (March 2025)