KWAN HO RYAN CHAN

ryanckh@seas.upenn.edu ⋄ (626) 400-3159 ⋄ ryanchankh.github.io ⋄ Philadelphia, PA

EDUCATION

Doctor of Philosophy in Electrical and Systems Engineering Expected: May '26
University of Pennsylvania

- Awards: National Science Foundation Graduate Research Fellow (NSF GRFP), UPenn Dean's Fellowship
- Advisors: Professor René Vidal, Professor Donald Geman
- Research Interests: Multi-modal Generative Models, Trustworthy ML, Explainable AI, Bioinfomatics

Bachelor of Arts in Applied Mathematics Class of '19
University of California, Berkeley

PROFESSIONAL EXPERIENCE

Apple Inc. May '24 - Sept '24
AI/ML Research Intern Seattle, WA

- Led the project that collaborated with researches across product areas and prepared for submission as first-author.
- Built real-world applications of our method with industrial Fitness and App Store datasets.

Research Project: Scalable and Collaborative Concept-based Recommendation System

- Proposed a scalable and editable recommendation system that leverages Large Language Models to extract interpretable and representative concepts for large-scale collaborative filtering and recommendations.
- Out-performed thorough comparisons with ID-based recommendation systems and zero-shot Large Language Models

RESEARCH HIGHLIGHTS

Uncertainty-aware and Interpretable-by-Design Machine Learning May '24 - Current

- Proposed an interpretable machine learning method that leverages techniques from Conformal Prediction and Prediction-powered inference to improve the interpretability and performance of sequential predictions.
- Organized a large-scale study that focuses on improving test performance via efficient human evaluation and editing.

Interpretable and Multi-modal Predictions for Cancer Metastasis Aug '24 - Current

- Applied interpretable machine learning methods to multi-modal cancer data such as Gene Alterations (DNA), Gene Expressions (scRNA) and clinical features.
- Led project on joint collaboration with Weill Cornell Medical College and Memorial Sloan Kettering Cancer Center.
- Collaborated with pathologists from local per-sample explanation to global analysis of distributions and predictions.

Controllable Multi-modal Generation and Detoxification of Large Language Models Jan '24 - May '24

- Collaborated on a framework for controlling LLM generations via Representation Engineering, Retrieval-augmented Generation and Sparse Decompositions using human-interpretable concepts.
- Released two open-source datasets involving representation of concepts in LLM activation space in open-source LLMs.
- Performed evaluation of question-answering and trustworthy benchmarks on Foundational Models such as multi-modal generative models (StableDiffusion) and Vision-Language Models (GPT-4, LLaVA).
- Co-authored two-joint works, titled "PaCE: Parsimonious Concept Engineering for Large Language Models" [NeurIPS'24] and "Knowledge Pursuit Prompting for Zero-Shot Multi-modal Synthesis" [ECCV'24].

Scalable Interpretable-by-Design Image Classification via Sequential Predictions Feb '22 - Nov '23

- Proposed an interpretable image classification method for large-scale visual datasets with variable-length and open-ended sets of concepts using Large Language Models (GPT-3, LLaMA) and Visual Language models (LLaVA, BLIP-2).
- Conducted a human evaluation on the faithfulness of the concept answering model with over 10k image-concept pairs.
- Published as first author titled "Bootstrapping Variational Information Pursuit with Foundation Models for Interpretable Image Classification" [ICLR'24] and "Variational Information Pursuit for Interpretable Predictions" [ICLR'23].

SKILLS

Programming and Natural Languages: Python, Java, Matlab, LaTeX, English, Chinese (Mandarin and Cantonese)
AI/ML Development: PyTorch, PyTorchLightning, HuggingFace, Sklearn, DGL, Git, Weight & Biases