Name: Eric Xing Email: epxing@cs.cmu.edu

FM-Association Analysis: A New Paradigm For Multi-Omic Association Mapping to Complex Clinical Phenotypes

Eric Xing and Eran Segal

Carnegie Mellon University
Weitzmann Institute of Science

Recent advancements in multi-omics and clinical profiling technologies have enabled the development of large-scale, deeply phenotyped human cohorts. However, translating this data into actionable healthcare insights remains a significant challenge. The rapid evolution of AI provides a promising approach to address this challenge, especially through foundation models (FMs), which are trained on massive datasets and adaptable to multiple tasks. However, while FMs achieve state-of-the-art performance across various domains, their adoption in medical AI has been limited.

We established the Human Phenotype Project (HPP), a longitudinal cohort of over 13,000 subjects with diverse diseases such as metabolic, neurological, and cardiovascular disorders. This cohort provides a unique opportunity to develop AI FMs using a multi-modal dataset that includes imaging, time-series, multi-omics, tabular data, and lifestyle information.

We present a suite of AI FMs that we developed which provide novel representations of molecular sequences (DNA/RNA/Protein) and of the diverse HPP modalities. We show that fine-tuning the learned representations predicts future onset of disease and trajectories of disease risk factors. We further integrated these FMs into a more holistic view of human health, by developing a generative FM model that creates a time series from the data of each subject and learns to predict the subject's future data. Our model predicts future health trajectories of subjects based on their historical data and predicts how lifestyle and pharmaceutical interventions can affect these trajectories. Our approach has far-reaching implications, including the development of diagnostic tools, augmentation in clinical decision-making, and new avenues for health informatics research.