
Few Shot Learning for Rare Disease Diagnosis

Emily Alsentzer

Postdoctoral Fellow

Brigham and Women's Hospital and Harvard Medical School

February 29, 2024

1:30PM-2:50PM

MSOB X303

Abstract:

Rare diseases affect 300-400 million people worldwide, yet each of the over 7000 unique diseases has very low prevalence. Many patients with rare genetic conditions remain undiagnosed due to clinicians' lack of experience with the individual diseases and the heterogeneity of clinical presentations. Deep learning has demonstrated success in aiding the diagnosis of more common diseases. However, rare disease diagnosis necessitates that approaches learn from limited data and extrapolate to novel genetic conditions. In this talk, I will describe our work to develop few-shot methods to diagnose patients with rare genetic conditions. To address the issue of data scarcity, we develop a framework for simulating realistic rare disease patients. To enrich the model with external knowledge, we frame diagnosis as a subgraph prediction task, developing a graph neural network to represent a patient's phenotypic and genetic data in relation to a knowledge graph of known phenotype, gene, and disease associations. With our large simulated patient cohort and knowledge-infused patient representations, we can perform multi-faceted diagnosis of patients in the Undiagnosed Diseases Network, performing causal gene discovery, retrieving "patients-like-me", and providing interpretable characterizations of novel disease presentations. This line of work demonstrates the impact of incorporating medical knowledge for model generalizability and illustrates the potential for deep learning methods to accelerate the molecular diagnosis of rare disease patients.

Reading list:

- (1) Primer on Graph Neural Networks: <https://distill.pub/2021/gnn-intro/>
- (2) Introduction to the Undiagnosed Diseases Network:
<https://www.nejm.org/doi/full/10.1056/NEJMoa1714458>