

Electronic medical record phenotyping using the anchor and learn framework

Electronic medical records (EMRs) hold a tremendous amount of information about patients that is relevant to determining the optimal approach to patient care. As medicine becomes increasingly precise, a patient's electronic medical record phenotype will play an important role in triggering clinical decision support systems that can deliver personalized recommendations in real time. In this talk, I introduce our recently developed "anchor and learn" framework for efficiently learning statistically driven phenotypes with minimal manual intervention. Using this approach, we developed a phenotype library that uses both structured and unstructured data from the EMR to represent patients for real-time clinical decision support. The resulting phenotypes are interpretable and fast to build. Evaluated in an emergency department setting, we find that our semi-supervised learning approach (which uses no manually labeled data) performs comparably to supervised learning.

Based on joint work with Yoni Halpern and Steven Horng.