



DISSERTATION DEFENSE



## Donna Tjandra

Predicting the Timing of Clinical Outcomes in the Presence of Multiple Events and Noisy Labels

Wednesday, June 5, 2024

9:30am – 11:30am

3901 Beyster

Hybrid – [Zoom](#) Passcode: 235217

**ABSTRACT:** In survival analysis one aims to predict the probability of a new event occurring over time, given time-to-event (TTE) training data. Unlike standard regression, outcomes for individuals in the training set may be censored. While standard regression approaches exclude such individuals, survival analysis approaches account for censorship. However, current approaches often make limiting assumptions during training, such as 1) the timing of different events are independent, and 2) TTEs are labeled accurately. In this thesis, we tackle the limitations of these assumptions and propose novel approaches that specifically address 1) potential dependencies among events and 2) label noise. We first tackle the challenge of accounting for potential dependencies among TTEs by proposing a novel hierarchical approach that models multi-event TTE prediction at different time granularities. We then tackle noisy labels with respect to event occurrence, considering scenarios where some individuals are labeled as experiencing the event when they did not and vice versa by modeling the underlying pattern of label noise. Finally, we tackle label noise with respect to the timing of the event in the context of multiple labelers. Again, by learning to model the noise, we show that we can map noisy predictions to a prediction of the ground truth. Overall, our approaches improve modeling accuracy and expand the applicability of survival analysis to more realistic datasets.

**CHAIR:** Prof. Jenna Wiens