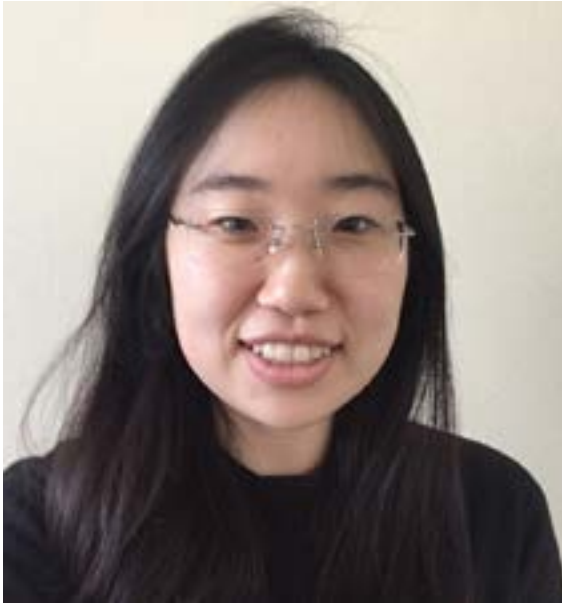




## DISSERTATION DEFENSE



# JEEHEH OH

## Improving the Generalization Performance of Deep Learning Models Applied to Clinical Time Series by Addressing Task Specific Structure

Wednesday, February 24, 2021

10:00 AM

[Virtual](#)

**ABSTRACT:** As the adoption of electronic health records (EHRs) increases, so do the opportunities to improve patient care using these data. Applied to EHR data, machine learning techniques can help identify complex patterns between patient covariates and outcomes. However, in order to augment clinical care these models must generalize. In healthcare settings, generalization performance is often hindered by limited training data driven by low incidence rates of the outcomes of interest. To address this challenge, we develop and evaluate methods that combine deep learning techniques with knowledge about the task structure to improve sample efficiency.

We augment learning algorithms by exploiting known task structure pertaining to i) invariances, ii) signal dynamics, and address challenges associated with iii) class imbalance driven low homophily. First, different types of temporal invariances are present in clinical time-series tasks but may vary across tasks. We propose a novel approach, 'Sequence Transformer Networks' that learns to recognize and exploit task-specific invariances. Second, risk factors for a given adverse outcome may change as a patient's admission progresses. We propose a novel RNN-based architecture in which we relax weight sharing over time to capture such time-varying relationships. Finally, GNNs are a popular method for learning feature representations from graphs but are often evaluated on tasks with graphs presenting high homophily. In clinical tasks, high class imbalance leads to asymmetrical homophily: high homophily with respect to the majority class and on average, but low homophily with respect to the minority class of interest. We address class imbalance driven low homophily by evaluating an attention-based mechanism against recently proposed methods for dealing with low homophily in general. By adapting techniques to better leverage task structure, we are able to improve sample efficiency and predictive performance on tasks such as estimating the risk for adverse patient outcomes.

**CHAIR:** Prof. Jenna Wiens