



COLLEGE OF PHARMACY  
**PHARMACEUTICAL SCIENCES**  
UNIVERSITY OF MICHIGAN

## Pharmaceutical Sciences Seminar Series

Wednesday, June 28, 2023  
4:00pm  
NCRC Building 10 South Atrium  
[Zoom](#)

### **“CoGT: Ensemble machine learning method and its application on JAK inhibitor discovery”**

Presented by:



**Yingzi Bu**

Ph.D. Candidate, Pharmaceutical Sciences  
University of Michigan

**Abstract:** The discovery of new drug candidate to inhibit its intended target is a complex and resource-consuming process. Machine learning (ML) method for predicting drug-target interactions (DTI) is a potential solution to improve the efficiency. However, traditional ML approaches have limitations in accuracy. In this study, we developed a novel ensemble model CoGT for DTI prediction using multilayer perceptron (MLP), which integrated graph-based models to extract non-Euclidean molecular structure and large pretrained models, specifically chemBERTa, to process simplified molecular input line entry system (SMILES) input. The performance of CoGT was evaluated using compounds inhibiting four Janus kinases (JAKs). Results showed that the large pretrained model, chemBERTa, was better than other conventional ML models in predicting DTI across multiple evaluation metrics, while the Graph Neural Network (GNN) was effective for prediction on imbalanced data sets. To take full advantage of the strengths of these different models, we developed an ensemble model, CoGT, which outperformed other individual ML models in predicting compounds' inhibition on different isoforms of JAKs. Our data suggest that the ensemble model CoGT has the potential to accelerate the process of drug discovery.